Linear isoelastic stochastic control problems and backward stochastic differential equations of Riccati type

Dissertation

zur Erlangung des akademischen Grades
des Doktors der Naturwissenschaften an der

Universität Konstanz

Fachbereich Mathematik und Statistik

vorgelegt von

Volker Bürkel

Tag der mündlichen Prüfung: 19. November 2004

Referent: Prof. Dr. Michael Kohlmann, Universität Konstanz
Referent: Prof. Dr. Shanjian Tang, Fudan University, Shanghai

Zusammenfassung

Die vorliegende Arbeit befasst sich mit linear isoelastischen stochastischen Kontrollproblemen. Es handelt sich hierbei um die Aufgabe, für ein festes $q > 1$ das *Kostenfunktional*

$$J(u) = \frac{1}{q}E[\int_{\tau}^{T} Q(s)|x(s)|^q + N(s)|u(s)|^q ds + M|x(T)|^q]$$

über u zu minimieren, wobei die *Kontrollvariable* u aus einem Vektorraum $\mathcal{U}$ stochastischer Prozesse stammt und x die eindeutige, starke Lösung ist der stochastischen Differentialgleichung

$$\begin{aligned} dx(t) &= \{A(s)x(s) + B(s)u(s)\}\, ds + \sum_{i=1}^{d} \left\{C^i(s)x(s) + D^i(s)u(s)\right\} dw^i(s), \\ x(\tau) &= h. \end{aligned}$$

Diese Gleichung bezeichnet man als *Zustandsgleichung*. Man sagt das Problem habe eine Lösung, wenn es ein $\bar{u} \in \mathcal{U}$ gibt sodass

$$J(\bar{u}) = \min_{u \in \mathcal{U}} J(u).$$

Man nennt $\bar{u}$ dann eine *optimale Kontrolle*.
Die linear isoelastischen stochastischen Kontrollprobleme verallgemeinern die linear quadratischen stochastischen Kontrollprobleme bei denen $q = 2$ ist. Es zeigt sich, dass sich die linear isoelastischen genau wie die linear quadratischen Probleme mit Hilfe so genannter *rückwärts-stochastischer Riccati Differentialgleichungen* (*BSRDEs, Backward Stochastic Riccati Differential Equations*) vollständig lösen lassen. Wir leiten die BSRDE im nicht-quadratischen Fall für skalare Zustandsgleichungen ab und zeigen, dass sie unter geeigneten Voraussetzungen eindeutig lösbar ist. Unsere Voraussetzungen sind hierbei im wesentlichen Standard-Annahmen, die mithin die eindeutige Lösbarkeit des entsprechenden Kontrollproblems garantieren. Wir müssen der Messbarkeit der Koeffizienten keine Einschränkungen ausser der üblichen Adaptiertheit auferlegen.
Die BSRDE mit den Unbekannten K und L für den nicht-quadratischen Fall lautet

$$dK = \Bigg\{-q'AK - \sum_{i=1}^{d}(C^i)^2K - 2\sum_{i=1}^{d} C^iL^i - \frac{q-2}{q-1}KBG(K,L)$$

$$
\begin{aligned}
& -\left(\frac{1}{q-1}BK + \sum_{i=1}^{d} D^i\left(C^iK + L^i\right)\right)G(K,L) \\
& -\frac{1}{q-1}QK^{2-q} + \frac{2-q}{2}\frac{1}{K}\sum_{i=1}^{d}\left(L^i + KC^i + KD^iG(K,L)\right)^2\Bigg\} ds + \sum_{i=1}^{d} L^i dw^i, \\
K(T) \quad = \quad & f(M),
\end{aligned}
$$

wobei G eine geeignete, implizit definierte Funktion ist. Die Abbildung f ist gegeben durch $f(v) = |v|^{\frac{1}{q-1}-1}v$ für $v \neq 0$, $f(v) = 0$. Ist (K, L) die eindeutige Lösung der BSRDE, so ist die optimale Kontrolle $\bar{u}$ gegeben durch

$$\bar{u} = G(K, L)\bar{x},$$

wobei $\bar{x}$ die zu $\bar{u}$ gehörige Lösung der Zustandsgleichung ist, d.h. $\bar{x}$ ist der *optimale Zustand*. Die *optimalen Kosten* sind gegeben durch

$$J(\bar{u}) = \frac{1}{q}E[K^{q-1}(\tau)|h|^q].$$

Wir charakterisieren die Lösung des stochastischen Kontrollproblems durch die Lösung eines geeigneten vorwärts-rückwärts stochastischen Systems von Differentialgleichungen, und gewinnen die Lösung (K, L) der BSRDE aus der Lösung dieses Systems. Der technische Kern dieser Methode besteht darin zu zeigen, dass der optimale Zustand $\bar{x}$ zum Anfangswert $h = 1$ fast sicher nie den Wert 0 erreicht. Dies entspricht dem Vorgehen von Tang in [T:GLQO]. Die Methode wurde unabhängig voneinander von Tang und dem Autor entwickelt.
Als Anwendungen betrachten wir Hedging-Probleme in Kapitalmärkten. Wir stellen einen Dualitätszugang her, bei dem sich das duale Problem in den meisten Fällen ebenfalls mit Hilfe von BSRDEs lösen lässt. Darüberhinaus führt der Dualitätszugang zu einem Typ linear isoelastischer Probleme, bei dem sich die üblichen Optimalitätsbedingungen auf unterschiedliche Teile einer zusammengesetzten Kontrollvariablen beziehen.

Abstract

This work deals with a generalization of so called linear quadratic stochastic control problems. This is a type of optimization problem that consists of minimizing the *cost functional*

$$J(u) := \frac{1}{2}E[\int_{\tau}^{T} x(s)'Q(s)x(s) + u(s)'N(s)u(s)ds + x(T)Mx(T)'],$$

where x is the solution of the stochastic differential equation (SDE), the *state equation,*

$$\begin{aligned} dx(t) &= \{A(s)x(s) + B(s)u(s)\}\, ds + \sum_{i=1}^{d} \left\{C^i(s)x(s) + D^i(s)u(s)\right\} dw^i(s), \\ x(\tau) &= h, \end{aligned}$$

and u belongs to some linear space $\mathcal{U}$ of stochastic processes. This is an important class of control problems because they can be used to model a broad variety of problems arising in applications, and because of the good analytic tractability of these problems. The key to this tractability is the so called *Backward Stochastic Riccati Differential Equation (BSRDE).*

We will consider *linear isoelastic control problems,* where the one-dimensional state processes x also follow a linear SDE as above, but the cost functional is given by

$$J(u) = \frac{1}{q}E[\int_{\tau}^{T} Q(s)|x(s)|^q + N(s)|u(s)|^q ds + M|x(T)|^q],$$

for some $q > 1$.

After giving some criteria for the solvability of linear isoelastic problems, we characterize their optimal state x and their optimal control u as part of the solution of the *Forward Backward Stochastic Differential Equation (FBSDE)* with auxiliary condition

$$\begin{aligned} dx(t) &= \{A(s)x(s) + B(s)u(s)\}\, ds + \sum_{i=1}^{d} \left\{C^i(s)x(s) + D^i(s)u(s)\right\} dw^i(s), \\ dy(t) &= \left\{-A(s)y(s) - \sum_{i=1}^{d} C^i(s)z^i(s) - Q(s)\varphi(x(s))\right\} ds + \sum_{i=1}^{d} z^i(s)dw^i(s), \\ & x(\tau) = h, \qquad y(T) = M\varphi(x(T)), \\ & B'y + \sum_{i=1}^{d}(D^i)'z^i + N\varphi(u) = 0, \end{aligned}$$

that arises from the study of the Gâteaux-derivative of J. Here, $\varphi(v) := |v|^{q-2}v$ for $v \neq 0$, $\varphi(v) = 0$. Note that this is a linear system of equations if $q = 2$. It turns out that this system of equations can be *decoupled.* Under our assumptions, there is a uniformly bounded, adapted family of positive random variables $K(t \vee \tau)_{t\in[0,T]}$ such that

$$f\left(y(t \vee \tau)\right) = K(t \vee \tau)x(t \vee \tau), \quad \text{for all} \ \ t \in [0, T],$$

where x and y belong to the solution of the FBSDE and f is the inverse function of φ. This family K will turn out to be a semimartingale and part of a solution of the newly established BSRDE for linear isoelastic stochastic control problems

$$\begin{aligned} dK &= \Bigg\{ -q'AK - \sum_{i=1}^{d}(C^i)^2K - 2\sum_{i=1}^{d} C^iL^i - \frac{q-2}{q-1}KBG(K,L) \\ &\quad - \left(\frac{1}{q-1}BK + \sum_{i=1}^{d} D^i\left(C^iK + L^i\right)\right)G(K,L) \\ &\quad - \frac{1}{q-1}QK^{2-q} + \frac{2-q}{2}\frac{1}{K}\sum_{i=1}^{d}\left(L^i + KC^i + KD^iG(K,L)\right)^2\Bigg\} ds + \sum_{i=1}^{d} L^i dw^i, \\ K(T) &= f(M), \end{aligned}$$

where G is some implicitly defined function. We can show, without imposing any restrictions on the measurability of the coefficients, except adaptedness, that this BSRDE is uniquely solvable and that the solution part L satisfies some strong a-priori estimate. We have the representation

$$K := \frac{f(y)}{x}, \ \ L^i := \frac{f'(y)z^i}{x} - C^i\frac{f(y)}{x} - D^i\frac{f(y)}{x^2}u, \ \ i = 1, \ldots, d,$$

where (x, u, y, z) is the solution of the FBSDE for the initial value $h = 1$. In particular, we will show that x does not attain zero. This can be regarded as the essential element of our method, that was independently developed by the author and Tang, see [T:GLQO]. Given the solution (K, L) of the BSRDE, the optimal state $\bar{x}$ and the optimal control $\bar{u}$ of the corresponding control problem (with initial value h) are related by

$$\bar{u} = G(K, L)\bar{x},$$

and the optimal cost is given by

$$J(\bar{u}) = \frac{1}{q}E[K(\tau)^{q-1}|h|^q].$$

Finally, we apply our results to some financial market hedging problems. A formulation as problems of minimum norm allows us to introduce a dual problem that can also be treated via the BSRDE approach. Hence, one has a choice to pick the one of the two problems that seems to be tractable best. The duality approach also leads to a new, interesting type of linear isoelastic problems.

Acknowledgements

I would like to thank Prof. Dr. Michael Kohlmann for his support, for many suggestions and comments. My thanks also go to Christian Bender, Johannes Leitner and Bernhard Peisl for many fruitful discussions. I'm indebted to Andrew P. Smith, Christoph Safferling and Christina Niethammer for finding many mistakes in writing. Financial support by a grant according to the Landesgraduiertenförderungsgesetz Baden-Württemberg is gratefully acknowledged. I want to thank my family and in particular my parents who supported me in many ways. Finally I want to express my gratitude to the German taxpayers for their readiness to support mathematical research.

Contents

1 Introduction **1**
1.1 The basic problem . . . 1
1.2 Definitions, problem formulation and assumptions . . . 6
1.2.1 Definitions . . . 6
1.2.2 Problem formulation and assumptions . . . 8
1.3 A short survey on known results about stochastic BSRDEs . . . 10

2 Some auxiliary results **13**
2.1 The dual space of $\mathbf{H_q(0, T; \mathbb{R}^m)}$ . . . 13
2.2 Solvability of linear SDEs . . . 16

3 Existence, the FBSDE **19**
3.1 Existence . . . 19
3.2 The FBSDE . . . 24
3.3 Basic properties of the solution . . . 29

4 The Riccati equation **37**
4.1 A feedback representation for the adjoint process . . . 37
4.2 The differential of K . . . 42
4.3 Represention of $\frac{1}{x}u$, the function G . . . 46
4.4 The equation, inherent properties . . . 52

5 Unique solvability, the optimal control **65**
5.1 Integrability of L . . . 66
5.2 Solvability of the BSRDE . . . 68
5.3 Uniqueness and representation of the optimal control . . . 70

6 A financial application and a duality **81**
6.1 The financial market model and problem formulation . . . 82
6.1.1 The market model . . . 82
6.1.2 Two financial market problems . . . 86
6.2 A duality . . . 89
6.2.1 The framework . . . 89

6.2.2 The adjoint operators and orthogonal complements 91
6.3 Solutions to the problems 99
6.3.1 Terminal cost, complete market 100
6.3.2 Terminal cost, incomplete market 102
6.3.3 Running and terminal cost, complete market 104
6.3.4 Running and terminal cost, incomplete market 106

A Some measurable selection arguments **115**
A.1 Measurable Selections 115
A.2 To the proof of Lemma 3.4. 116
A.3 To the proof of Lemma 6.10-3. 117

Chapter 1

Introduction

1.1 The basic problem

Among the issues of stochastic control theory, some of the most intensively studied problems are of the following form: Find some $\bar{u}$ that minimizes the *cost functional*

$$J(u) = E[\int_0^T R(t, x(t), u(t))dt + \Psi(x(T))], \tag{1.1}$$

where

- the *state process* x is the strong solution of the stochastic differential equation (SDE)

$$\begin{aligned} dx(t) &= a(s, x(s), u(s))ds + b(s, x(s), u(s))\, dw(s), & (1.2) \\ x(0) &= h, & (1.3) \end{aligned}$$

  and

- the *control variable* u ranges over some specified set $\mathcal{U}$ of stochastic processes, called the *admissible controls,*

and w is a d-dimensional standard Brownian motion.
The intuitive meaning of this formalism (from the point of view of applications) is the following: The process x is meant to describe some dynamical system in a random environment that can be influenced by a *controller* that implements the control strategy u. The dynamical system is subject to random perturbations.
Consider for example an electric dc motor. Let x be the angular velocity of the shaft of the motor and u be input voltage of the motor. In a simple model, these two quantities are related by the deterministic differential equation (see [KwS:LO], Section 3.3)

$$dx(t) = \{-\alpha x(s) + \kappa u(s)\}\, ds, \quad x(0) = x_0,$$

for some suitable constants α and κ. Here, the input voltage u is the control variable. x_0 is the initial angular velocity. Now suppose the motor responds in a “noisy” way to the input voltage u. We model this effect by introducing a stochastic integral in the above equation,

$$dx(t) = \{-\alpha x(s) + \kappa u(s)\}\, ds + \lambda(s)u(s)dw(s), \quad x(0) = x_0,$$

for some adapted process λ. Assume that the "controller" of the motor wants his device to run at a speed that is close to the constant angular velocity ξ, at least in an average sense, in the time interval $[0, T]$. So he may chose his control variable u (represented by a stochastic process) such that e.g.

$$J(u) = E[\int_0^T \exp\{|x(s) - \xi|\}ds],$$

is minimal. Besides, he may select as an integrand any other function of $|x(s) - \xi|$ that represents his ideas of "being close". Suppose that the controller wants to avoid that the input voltage becomes too high over a long time. So he may change the above cost criterion by introducing an additional term to

$$J(u) = E[\int_0^T \exp\{|x(s) - \xi|\} + p(u(s))\, ds],$$

where p is a function that reflects the controllers's grade of disapproval for high input voltages.
However, one must make sure that for every $u \in \mathcal{U}$ the quantity $J(u)$ is well defined, in particular, that the SDE possesses a unique (strong) solution. In general, i.e. in the absence of further assumptions on the functions a, b, R and Ψ (let us call them, somewhat inaccurate, the "coefficients" of the problem) and the set $\mathcal{U}$, the problem (1.1)-(1.3) will not be well posed. This means that there may be no *optimal control* $\bar{u}$ that minimizes J.

So, let us assume that the problem is correctly defined and well posed. The next, and crucial, step is to determine an optimal control $\bar{u}$, along with the optimal cost $J(\bar{u})$ and the optimal state $\bar{x}$. It is natural that here we will encounter a tradeoff between the generality of the coefficients and $\mathcal{U}$, and the explicitness with which we can describe $\bar{u}$. On the "explicit" end of this scale there are the so called *linear quadratic* (LQ) stochastic control problems."Linear" refers to the linear state equation of this type of problem, "quadratic" to its quadratic cost functional. Stated explicitly, the linear quadratic (stochastic control) problems are of the form

$$\begin{aligned} dx(t) &= \{A(s)x(s) + B(s)u(s) + \gamma(s)\}\, ds \\ &\quad + \sum_{i=1}^{d} \left\{C^i(s)x(s) + D^i(s)u(s) + \Gamma^i(s)\right\} dw^i(s), && (1.4) \\ x(\tau) &= h, && (1.5) \\ J(u) &= \frac{1}{2} E[\int_\tau^T (x(t) - \rho(t))'\, Q(t)\, (x(t) - \rho(t)) + u'(t)N(t)u(t)dt \\ &\quad +(x(T) - \vartheta)'M(x(T) - \vartheta)] && (1.6) \\ &= \min_{u\in\mathcal{U}}! \end{aligned}$$

For the sake of technical orientation we list the usual assumptions imposed on this problem. $(\mathcal{F}_t)_{t\geq 0}$ is the augmentation of the filtration generated by the standard Brownian motion w. τ is a $(\mathcal{F}_t)_t$-stopping time with $\tau < T$.
The processes γ, $(\Gamma^i)_{i\leq d}$ and ρ, as well as the random variable (r.v.) ϑ belong to the data of the problem and are assumed to be square integrable and adapted respectively measurable with respect to the σ-algebra $\mathcal{F}_T$; the coefficients $A, B, (C^i)_{i\leq d}, (D^i)_{i\leq d}, Q$ and N are essentially bounded and also adapted. The essentially bounded r.v. M is $\mathcal{F}_T$- measurable. The initial value h is square integrable. $\mathcal{U}$ is the real linear space of m-dimensional, adapted, square integrable processes. Existence of an optimal control is typically (not exclusively) ensured by one of the following two conditions:

1. N is uniformly positive, Q and M are non-negative.

2. $\sum_{i=1}^d D^i(D^i)'$ and M are uniformly positive, Q and N are non-negative.

Under the above assumptions, the LQ problem can be completely solved if the state process x is one dimensional. For n-dimensional state it can be solved if the first assumption holds and in many particular cases if the second assumption holds, depending, for example, on the measureability of the coefficients or the dimension of w. The key tool for the far-reaching analytic tractability of these problems is the so called ***Backward Stochastic Riccati Differential Equation (BSRDE)***. Solvability of a special subtype of problems depends more or less on the solvability of the corresponding BSRDE.
The equation bears (as with every Backward Stochastic Differential Equation) two unknown processes, K and L, and is given by (suppressing the time argument t)

$$\begin{aligned} dK &= -\Bigg\{A'K + K'A + Q + \sum_{i=1}^d (C^i)KC^i + \sum_{i=1}^d \left((C^i)'L^i + L^iC^i\right) \\ &\quad - \left(KB + \sum_{i=1}^d (C^i)'KD^i + \sum_{i=1}^d L^iD^i\right) \times \left(N + \sum_{i=1}^d (D^i)'KD^i\right)^{-1} \\ &\quad \times \left(KB + \sum_{i=1}^d (C^i)'KD^i + \sum_{i=1}^d L^iD^i\right)'\Bigg\} dt + \sum_{i=1}^d L^i dw^i, \qquad (1.7) \\ K(T) &= M. \qquad (1.8) \end{aligned}$$

K and L^i are $\mathbb{R}^{n\times n}$-valued processes (n is the dimension of the state process x).
Due to the the heavily non-lipschitz right hand side of this equation, its (unique) solvability is quite hard to prove. First existence and uniqueness results for deterministic coefficients $A, B, \ldots$ go back to Wonham [W:MRE]. In this case, the equation becomes a matrix-valued ordinary differential equation (ODE) with $L = 0$. Bismut introduced the equation with stochastic coefficients in [B:LQOC] and proved the existence of a solution for the case that the randomness of the coefficients comes from some smaller filtration than $(\mathcal{F}_t)$. It was not until 2001 that a general (i.e. without restrictions on the measurability of the

coefficients) existence and uniqueness result was published for one-dimensional BSRDE (i.e. for one-dimensional state equations), see [KT:GAS]. In 2003, the papers [T:GLQO] and [KT:MBSR] covered the case of multi-dimensional state-equations. In these papers there are also surveys on the development of BSRDE-theory since the seminal work of Bismut.

Once the problem of solvability is overcome, we may enjoy the benefits derived from the (abstract) knowledge of the process (K, L). Note that we do not address numerical questions and take a naive point of view: if a stochastic differential equation is solvable, the solution is available for us. For simplicity, we assume that $\gamma, (\Gamma^i)_{i\leq d}, \rho$ and ϑ are equal to zero. In this case, the optimal state $\bar{x}$ and the optimal control $\bar{u}$ are linked by the relation

$$\begin{aligned}\bar{u}(t) = &-\left(N(t) + \sum_{i=1}^{d}(D^i(t))'K(t)D^i(t)\right)^{-1}\\ &\times\left(K(t)B(t) + \sum_{i=1}^{d}(C^i(t))'K(t)D^i(t) + \sum_{i=1}^{d} L^i(t)D^i(t)\right)'\bar{x}(t), \end{aligned} \tag{1.9}$$

and the optimal cost is given by

$$J(\bar{u}) = E[h'K(\tau)h], \tag{1.10}$$

where h is the initial value of the state equation.
Such an explicit representation for the optimal control and the optimal cost is very desirable, but its availability for LQ problems depends strongly on the special structure of these problems.
This work is concerned with an extension of linear quadratic problems that we call *linear isoelastic problems*. We consider one-dimensional state equations without inhomogeneous parts

$$\begin{aligned} dx(t) &= \{A(s)x(s) + B(s)u(s)\}\,ds + \sum_{i=1}^{d}\left\{C^i(s)x(s) + D^i(s)u(s)\right\}dw^i(s),\\ x(\tau) &= h, \end{aligned}$$

and *isoelastic* cost functions:

$$J(u) = \frac{1}{q}E[\int_0^T Q(t)|x(t)|^q + N(t)|u(t)|^q dt + M|x(T)|^q],$$

with $q \in (1, \infty)$. The naming “isoelastic” comes from economics. There, one considers the concept of *elasticity*. For a function $x \mapsto p(x)$, its elasticity is formally defined as $\frac{dp}{dx}\frac{x}{p}$. Clearly, the mappings $[0, \infty) \longrightarrow \mathbb{R}$, $x \mapsto \frac{1}{q}x^q$, have a constant elasticity, i.e. an elasticity that is independent of x.
It turns out that, like in the quadratic case, this problem can be solved with the help of the solution of a BSRDE-type equation, and that this newly introduced Riccati-type equation

is uniquely solvable. The conditions we must impose ensure the existence of an optimal control and do not involve restrictions on the measurability of the coefficients.
The text is organized as follows:

The rest of Chapter 1 introduces notation and a statement of the problem we are concerned with. It also contains the basic assumptions that we impose on these problems.

Chapter 2 contains two results. The first one is about the reflexivity of $H_q(\tau, T; \mathbb{R}^m)$. The second result is concerned with the solvability of linear stochastic differential equations and the continuous dependence (in a q-th mean sense) of their solution on some parameters. Both results are not claimed to be new, but are included for the reader's convenience.

Chapter 3 includes existence results for the linear isoelastic control problems and a deterministic counterexample. The solution of the control problem is completely characterized in terms of the solution of a forward backward stochastic differential equation (FBSDE). It is shown that the optimal cost can be represented with the help of the adjoint process. We derive some properties of the solution of the FBSDE, in particular the linearity of $h \mapsto \left(\bar{x}^{\tau,h}, \bar{u}^{\tau,h}, f(\bar{y}^{\tau,h})\right)$, where $\bar{x}^{\tau,h}, \bar{u}^{\tau,h}$ and $\bar{y}^{\tau,h}$ belongs to the solution of the FBSDE for initial time τ and initial value h.

Chapter 4 introduces a feedback representation for the adjoint process, namely the representation $f\left(\bar{y}(t \vee \tau)\right) = K(t \vee \tau)\bar{x}(t \vee \tau)$, $t \in [0, T]$, for some family of random variables $(K(t \vee \tau))_{t\in[0,T]}$. This family is shown to be uniformly bounded and strictly or uniformly positive, depending on the assumption in force. We derive the BSRDE for linear isoelastic problems by differentiating K and introduce the function G. Generalizing a method from [T:GLQO] we derive a-priori estimates for the solution part L of the BSRDE.

Chapter 5 shows that the BSRDE is solvable, essentially by showing that the optimal state process x for the initial value 1 never reaches zero. We show that the solution of the BSRDE is unique and that the optimal control, the optimal state and the optimal cost of the linear isoelastic problem can be derived from the solution (K, L) of the BSRDE.

Chapter 6 describes a financial market model and states two hedging problems. We consider a duality approach for these problems and set up dual problems. The financial market problems and most of the dual problems can be solved with the theory we developed so far. In a special case, the duality approach leads to an new type of linear isoelastic problem where our standard optimality assumptions apply to separate parts of a compound control variable. We establish a BSRDE for this new problem and show that this BSRDE is solvable.

1.2 Definitions, problem formulation and assumptions

1.2.1 Definitions

Throughout this work we make the following assumptions and use the following notations or conventions.
Let be $T > 0$. w is a d-dimensional standard Brownian motion, starting in 0 and defined on the filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F})_{0\le t\le T}, P)$. The filtration $(\mathcal{F})_{0\le t\le T}$ is taken to be the augmentation of the filtration generated by w, see [KS:BM], Chap. 2.7, Def. 7.2. Unless stated otherwise, the linear space $\mathbb{R}^n$ is equipped with the euclidean norm $|\cdot|$. The space $\mathbb{R}^{m\times n}$ of real $m \times n$ -matrices is equipped with the operator norm induced by the euclidean norm. A vector or a matrix with a prime as superscript, i.e. v' or A', denotes the transpose of that vector or matrix, however, for a real number $q \in (1,\infty)$, q' denotes the conjugate exponent of q, $\frac{1}{q} + \frac{1}{q'} = 1$ respectively $q' = \frac{q}{q-1}$. When A is a $m \times n$-matrix with entries a_{ij}, we may write $A = [a_{ij}]$. We also use the notation $[A]_{ij}$ for the (i, j)- entry a_{ij} of A. If $A \in \mathbb{R}^{m\times m}$ is uniformly positive, i.e. if there is a $k > 0$ such that $\lambda' A\lambda \ge k|\lambda|^2$ for all $\lambda \in \mathbb{R}^m$, we write $A \gg 0$.
A *stopping time* always means a stopping time with respect to the filtration $(\mathcal{F}_t)$. When considering a real, finite-dimensional, normed linear space V as a measurable space, V is endowed with the Borel σ-algebra induced by the norm. $\mathcal{B}$ denotes the Borel-σ-algebra on $[0, T]$, completed with respect to the Lebesgue-measure. The Lebesgue-measure is denoted by *Leb*. Let τ_1 and τ_2 be two stopping times with values in $[0, T]$; equip the stochastic interval $[\tau_1, \tau_2]$ (or $[\tau_1, \tau_2)$, etc.) with the trace-σ- algebra derived from $\mathcal{B} \otimes \mathcal{F}$.
Let V be a real, finite dimensional, normed vector space. A V-valued stochastic process on $[\tau_1, \tau_2]$ is a measurable mapping from $[\tau_1, \tau_2]$ to V.
Let $q \in (1,\infty)$ and a stopping time $\tau \in [0, T]$ be given. The spaces $L^q_{\mathcal{F}}(\tau, T; V)$ respectively $H_q(\tau, T; V)$ consist of all V-valued, $(\mathcal{F}_{\tau\vee t})_t$-adapted processes z on $[\tau, T]$ that satisfy

$$\|z\|_{L^q_{\mathcal{F}}} := \left(E[\int_\tau^T |z(s)|^q ds]\right)^{\frac{1}{q}} < \infty,$$

respectively

$$\|z\|_{H_q} := \left(E[\left(\int_\tau^T |z(s)|^2 ds\right)^{\frac{q}{2}}]\right)^{\frac{1}{q}} < \infty.$$

By $R_q(\tau, T; V)$ we denote the space of all V-valued, $(\mathcal{F}_{\tau\vee t})_t$-adapted processes z with paths that are right-continuous and have left-side limits (i.e. are RCLL), and that satisfy

$$\|z\|_{R_q} := \left(E[\sup_{\tau\le t\le T} |z(t)|^q]\right)^{\frac{1}{q}} < \infty.$$

$L^q_{\mathcal{F}}(\Omega, C([\tau, T]; V))$ is the subspace of the processes in $R_q(\tau, T; V)$ whose paths are $P-a.s.$ continuous, endowed with the restriction of the norm $\|\cdot\|_{R_q}$. We will denote this restriction

by $\|\cdot\|_{L_c^q}$.
$L^q_{\mathcal{F}_\tau}(V)$ is the space of V-valued, $\mathcal{F}_\tau$-measurable random variables ζ with finite norm

$$|\zeta|_{L^q} := (E[|\zeta|^q])^{\frac{1}{q}}\,.$$

Accordingly, $L^\infty_{\mathcal{F}_\tau}(V)$ consists of all V-valued, $\mathcal{F}_\tau$-measurable random variables ζ with finite norm

$$|\zeta|_{L^\infty} := \text{ess.sup}_{\omega\in\Omega}|\zeta(\omega)|.$$

Further, by $L^\infty_{\mathcal{F}}(\tau, T; V)$ we denote the space of all V-valued, $(\mathcal{F}_{\tau\vee t})_t$-adapted, essentially bounded processes z, endowed with the norm

$$\|z\|_{L^\infty_{\mathcal{F}}} := \text{ess.sup}_{(t,\omega)\in[\tau,T]}|z(t,\omega)|.$$

Finally, $L^\infty_{\mathcal{F}}(\Omega, C([\tau, T]; V))$ is the subspace of processes z in $L^\infty_{\mathcal{F}}(\tau, T; V)$ whose paths are continuous. We equip the smaller space with the norm of the larger space and denote this restriction by $\|\cdot\|_{L_c^\infty}$.
All these spaces are complete in their respective norms; for $R_q(\tau, T; V)$ see [DM:PPB], Chap. VII, § 3, no. 64 .
In the notation of stochastic processes we will often skip the time variable, and the argument ω is, as usual, completely suppressed. The mutual variation process of two continuous semimartingales m, n is denoted by $< m, n >$. We will use the the following acronyms:

BSDE	for	backward stochastic differential equation
BSRDE	for	backward stochastic Riccati differential equation
ODE	for	ordinary differential equation
r.v.	for	random variable
SDE	for	stochastic differential equation

For a normed linear space W, W^* denotes the dual space of W. The indicator of a set S, i.e. the function that takes the value 1 on S and that equals zero on the complement of S, is denoted by $\mathbf{1}_S$.
Bearing some basic results of convex analysis in mind, it is not surprising that the derivative of the function $\mathbb{R}^m \longrightarrow \mathbb{R}$, $u \mapsto |u|^q$, plays an important role, as well as the derivative's inverse. We will use a fixed notation for these two functions.

Definition 1.1 *For a given $q > 1$ define the functions $\varphi, f : \mathbb{R}^n \longrightarrow \mathbb{R}^n$ by*

$$\varphi(v) = |v|^{q-2}v, \qquad f(v) = |v|^{\frac{1}{q-1}-1}v,$$

for $v \neq 0$ and

$$\varphi(0) = f(0) = 0.$$

We will in general not specify n. If e.g. the argument of f is real, then f is meant to be a mapping $\mathbb{R} \longrightarrow \mathbb{R}$. Using this convention, we may note that φ and f are *multiplicative* in the following sense: For $\alpha \in \mathbb{R}$ and $v \in \mathbb{R}^n$ we have $\varphi(\alpha v) = \varphi(\alpha)\varphi(v)$, $f(\alpha v) = f(\alpha)f(v)$ (this will turn out to be a crucial property). φ and f are continuous on $\mathbb{R}^n$ (note that $q-2 > -1$, $\frac{1}{q-1} - 1 > -1$ for $q > 1$) and mutually inverse, i.e. $\varphi(f(v)) = f(\varphi(v)) = v$ for all $v \in \mathbb{R}^n$. Note that for stopping times τ, γ with $\tau < T, \gamma \le T$, the following mappings are well defined and continuous:

$$\begin{aligned}
L^q_{\mathcal{F}}(\tau,T;\mathbb{R}^m) &\longrightarrow L^{q'}_{\mathcal{F}}(\tau,T;\mathbb{R}^m),\\
u &\mapsto \varphi(u);\\
L^q_{\mathcal{F}}(\Omega, C([\tau,T];\mathbb{R})) &\longrightarrow L^{q'}_{\mathcal{F}}(\Omega, C([\tau,T];\mathbb{R})),\\
x &\mapsto \varphi(x);\\
L^{q'}_{\mathcal{F}}(\tau,T;\mathbb{R}^m) &\longrightarrow L^{q}_{\mathcal{F}}(\tau,T;\mathbb{R}^m),\\
v &\mapsto f(v);\\
L^{q'}_{\mathcal{F}}(\Omega, C([\tau,T];\mathbb{R})) &\longrightarrow L^{q}_{\mathcal{F}}(\Omega, C([\tau,T];\mathbb{R})),\\
y &\mapsto f(y);\\
L^q_{\mathcal{F}_\gamma}(\mathbb{R}) &\longrightarrow L^{q'}_{\mathcal{F}_\gamma}(\mathbb{R}),\\
x &\mapsto \varphi(x);\\
L^{q'}_{\mathcal{F}_\gamma}(\mathbb{R}) &\longrightarrow L^{q}_{\mathcal{F}_\gamma}(\mathbb{R}),\\
y &\mapsto f(y).
\end{aligned}$$

1.2.2 Problem formulation and assumptions

As indicated in the previous section, this work is concerned with a particular type of control problem. We wish to formulate the specific form of this problem and a corresponding framework in which it is considered. We will allow for random initial times and values.

Definition 1.2 : *Problem $\mathcal{P}(\tau, h)$*
Fix some $q \in (1,\infty)$. Let τ be a stopping time with $\tau < T$. Assume that we are given stochastic processes
$A \in L^\infty_{\mathcal{F}}(\tau,T;\mathbb{R})$, $B \in L^\infty_{\mathcal{F}}(\tau,T;\mathbb{R}^{1\times m})$, $C^i \in L^\infty_{\mathcal{F}}(\tau,T;\mathbb{R})$ and
$D^i \in L^\infty_{\mathcal{F}}(\tau,T;\mathbb{R}^{1\times m})$ respectively for $i \in \{1,\ldots,d\}$, Q, $N \in L^\infty_{\mathcal{F}}(\tau,T;\mathbb{R})$ and a r.v. $M \in L^\infty_{\mathcal{F}_T}(\mathbb{R})$. For some $h \in L^q_{\mathcal{F}_\tau}(\mathbb{R})$ we denote by $\mathcal{P}(\tau,h)$ the problem

$$J(u) = \frac{1}{q}E[\int_\tau^T Q(t)|x(t)|^q + N(t)|u(t)|^q dt + M|x(T)|^q] = \min_{u\in\mathcal{U}}! \tag{1.11}$$

where

$$\mathcal{U} = L^q_{\mathcal{F}}(\tau,T;\mathbb{R}^m) \cap H_q(\tau,T;\mathbb{R}^m) \text{ if } N \neq 0, \tag{1.12}$$

respectively

$$\mathcal{U} = H_q(\tau,T;\mathbb{R}^m) \text{ if } q \ge 2 \text{ and } N = 0, \tag{1.13}$$

and x is the unique strong solution of

$$dx(t) = \{A(s)x(s) + B(s)u(s)\}\,ds + \sum_{i=1}^{d}\left\{C^i(s)x(s) + D^i(s)u(s)\right\}dw^i(s), \tag{1.14}$$

$$x(\tau) = h. \tag{1.15}$$

A solution *of this problem is a process $\bar{u} \in \mathcal{U}$ in which the cost functional in (1.11) attains its minimum, $J(\bar{u}) = \min_{u\in\mathcal{U}} J(u)$. $\bar{u}$ is called the* optimal control *for the problem, the process $\bar{x}$ corresponding to $\bar{u}$ is called the* optimal state *for the problem.*
The collection of processes respectively r.v. $A, B, (C^i)_{1\le 1\le d}, (D^i)_{1\le i\le d}, Q, N,$ and M are called the coefficients *of problem $\mathcal{P}(\tau, h)$.*

Some comments on this definition are in order. It claims that the SDE possesses a unique strong solution. This may be clear from the literature, but will also be proved in the next section. For the cost functional to be well defined, the control u and the corresponding state process x in general (if Q or N do not vanish) should be q-integrable i.e. we should require $\|u\|_{L^q_{\mathcal{F}}} < \infty$ and $\|x\|_{L^q_{\mathcal{F}}} < \infty$. The latter is achieved by taking u from $H_q(\tau, T; \mathbb{R}^m)$ (by the Burkholder-Gundy-Davis inequality), the former by taking u from $L^q_{\mathcal{F}}(\tau, T; \mathbb{R}^m)$, so the domain $\mathcal{U}$ is just the intersection of these two spaces. Yet, for $q \le 2$ we have (by Jensens's inequality) $H_q(\tau, T; \mathbb{R}^m) \subseteq L^q_{\mathcal{F}}(\tau, T; \mathbb{R}^m)$, and vice versa for $q \ge 2$ (by Hölder's inequality). So the minimization is performed on one of these two normed linear spaces.

So far, nothing has been said about the existence of an optimal control $\bar{u}$. Given the reflexivity of $\mathcal{U}$, this existence assertion (stated in Section 3.1) will follow from classical results of convex analysis, provided that we can make sure that J is coercive, i.e. $J(u) \to \infty$ for $\|u\|_{\mathcal{U}} \to \infty$, where $\|\cdot\|_{\mathcal{U}}$ is the norm of $L^q_{\mathcal{F}}(\tau, T; \mathbb{R}^m)$ or $H_q(\tau, T; \mathbb{R}^m)$, depending on the value of q. Each of the following three assumptions will guarantee coercivity, the first one in the case $q \le 2$, the second and third in the case $q \ge 2$.

Assumption A1 *The r.v. M and the stochastic process $\sum_{i=1}^d (D^i)'D^i$ are uniformly positive, i.e. there is an $\epsilon > 0$ such that $M \ge \epsilon$, $P-a.s.$, and for all $v \in \mathbb{R}^m$ we have $v'\left(\sum_{i=1}^d (D^i)'D^i\right)v \ge \epsilon|v|^2$, $Leb \otimes P-a.s.$. The processes Q and N are non-negative. q belongs to* $(1, 2]$.

This assumption implicitly requires $m \le d$. To see this consider the $\mathbb{R}^{m\times d}$-valued process σ whose i-th column is $(D^i)'$, $[\sigma]_{ji} = D^i_j$, i.e.

$$\sigma := \left[(D^1)', \ldots, (D^i)', \ldots, (D^d)'\right]. \tag{1.16}$$

We have the representation $\sum_{i=1}^d (D^i)'D^i = \sigma\sigma'$, hence the sum cannot be a regular matrix unless the kernel of $\sigma' \in \mathbb{R}^{d\times m}$ is trivial. This yields $m \le d$. In a financial market model, when the SDE (1.14), (1.15) describes the value process generated by the initial endowment

h and the portfolio u, Assumption A1 would imply that the underlying market is arbitrage free.
The above assumption allows the weight process N of the immediate control cost to be identical to zero. In order to ensure the existence of an optimal control, one may even allow the process N to become negative, see Lemma 3.3 in Chapter 3. This is a particular feature of stochastic control theory. In the deterministic case, i.e. when all coefficients are deterministic and $C^i = D^i = 0$, $i = 1, \ldots, d$, the minimization problem $\mathcal{P}(0, h)$ will in general be ill posed if N is not positive, see Remark 3.3 in Chapter 3, too.
But also in the stochastic case the uniform positivity of N will help us to ensure that an optimal control exists.

Assumption A2 *The process N is uniformly positive ($Leb \otimes P - a.s.$). The process Q and the r.v. M are non-negative (i.e. $Q \geq 0$, $Leb \otimes P - a.s$ respectively $M \geq 0, P - a.s.$). q belongs to $[2, \infty)$.*

As pointed out, for $q \geq 2$ the presence of N in the cost functional forces us to choose u from $L^q_{\mathcal{F}}(\tau, T; \mathbb{R}^m)$. If N vanishes, we may also cast problem $\mathcal{P}(\tau, h)$ as a minimization problem over $H_q(\tau, T; \mathbb{R}^m)$. The following assumption gives the framework within which this case will be considered.

Assumption A3 *The r.v. M and the stochastic process $\sum_{i=1}^d (D^i)' D^i$ are uniformly positive. The process Q is non-negative, $Q \geq 0$, $Leb \otimes P - a.s.$, and N is zero, $N = 0$. q belongs to $[2, \infty)$. The minimization is performed over $\mathcal{U} = H_q(\tau, T; \mathbb{R}^m)$.*

Later on, we will have to strengthen Assumption A2 to Assumption A4, see page 40.

Assume that we are given the coefficients of a problem $\mathcal{P}(\tau, h)$. If γ is a stopping time with $\tau \leq \gamma < T$, we can construct a new problem (or "subproblem") $\mathcal{P}(\gamma, h_\gamma)$ whose coefficients are given by the restriction of the coefficients of problem $\mathcal{P}(\tau, h)$.

Definition 1.3 *(Subproblem)*
Assume we are given a problem $\mathcal{P}(\tau, h)$, and a stopping time γ with $\tau \leq \gamma < T$. Let h_γ be in $L^q_{\mathcal{F}_\gamma}(\mathbb{R})$. Unless otherwise stated, the coefficients of problem $\mathcal{P}(\gamma, h_\gamma)$ are meant to be the restrictions
$A|_{[\gamma,T]}, B|_{[\gamma,T]}, (C^i|_{[\gamma,T]})_{1 \leq i \leq d}, (D^i|_{[\gamma,T]})_{1 \leq i \leq d}, Q|_{[\gamma,T]}, N|_{[\gamma,T]},$ *and* M.

Note that if Assumption A1, A2 or A3 holds for problem $\mathcal{P}(\tau, h)$, then the respective assumption also holds for the subproblems $\mathcal{P}(\gamma, h_\gamma)$.

1.3 A short survey on known results about stochastic BSRDEs

The typical approach to linear quadratic stochastic control problems is the BSRDE. However, there are alternatives. For example, linear quadratic problems can be tackled by the

Maximum Principle in the framework of linear state equations and convex cost functionals, see e.g. [CK:TSMP]. Besides, one may try to handle these problems via Dynamic Programming, see for example [YZ:SC]. Nevertheless, for linear quadratic problems both these methods lead quite naturally to the consideration of a BSRDE, see the two cited references ([CK:TSMP], Section 3.6 and [YZ:SC], Chapter 6) for the case of deterministic coefficients. In this work, we will treat the linear isoelastic problem with BSRDEs, too, and the purpose of this section is to give a brief survey on results for BSRDEs corresponding to linear quadratic problems. See also the surveys in [KT:GAS] and [T:GLQO] from which we take much of the information presented in this section.
Consider a linear quadratic stochastic control problem with deterministic coefficients. The corresponding BSRDE reduces to the ordinary matrix-valued differential equation

$$\begin{aligned}\frac{dK}{dt} &= -\Bigg\{A'K + K'A + Q + \sum_{i=1}^{d}(C^i)KC^i \\ &\quad - \left(KB + \sum_{i=1}^{d}(C^i)'KD^i\right)\left(N + \sum_{i=1}^{d}(D^i)'KD^i\right)^{-1}\left(KB + \sum_{i=1}^{d}(C^i)'KD^i\right)'\Bigg\}, \\ K(T) &= M.\end{aligned}$$

This equation was solved by Wonham (1968), see [W:MRE]. It is an ordinary, matrix valued Riccati differential equation that gave the name *Riccati* to the more general equations. To the best of our knowledge, the first one who introduced the BSRDE (1.7), (1.8) with stochastic coefficients $A, B, (C^i)_{1\le i\le d}, (D^i)_{1\le i\le d}, Q, N$ and M was Bismut (1976) in [B:LQOC]. Bismut also proved the solvability of the BSRDE in the case of a uniformly positive N, but he had to impose some restrictions on the coefficients. Assume $d_0 < d$ and let $(\mathcal{F}_t^{(2)})_{t\in[0,T]}$ be the augmentation of the filtration generated by the Brownian motion $w^{d_0+1}, \ldots, w^d$. In [B:LQOC], Bismut assumed that the coefficients are adapted to $(\mathcal{F}_t^{(2)})_{t\in[0,T]}$ and that $D^{d_0+1} = \ldots = D^d = 0$, $C^{d_0+1} = \ldots = C^d = 0$. The restriction on the measurability entails for the solution part L of the BSRDE that $L^1 = \ldots = L^{d_0} = 0$. The restriction on the C^i and D^i leads to an equation where the remaining $L^{d_0+1}, \ldots, L^d$ do not appear in the drift term. In the paper [B:CDSL] Bismut kept the measurability assumption for the coefficients, but he only required that $D^{d_0+1} = \ldots = D^d = 0$.
In 1990, the paper [PP:ASBS] initiated an intense research on the subject of BSDEs. This seminal article contains an existence result for solutions of BSDEs with a uniformly Lipschitz-continuous driver. Yet, the BSRDE is an example for a BSDE with a heavily non-Lipschitz driver. There are methods to overcome the Lipschitz assumption, see e.g. [K:REU] and [LSM:EBS], but no one precisely matches the needs of the BSRDE. However, some arguments developed in last two cited papers were adapted by Kohlmann and Tang in their paper [KT:GAS], which contains an existence result for the one-dimensional BSRDE without any restrictions on the measurability of the coefficients, except, of course, adaptedness. Apart from the boundedness of the coefficients, the essential assumption in this paper is that either N or M and $\sum_{i=1}^{d}(D^i)'D^i$ are uniformly positive. The latter

of these conditions, $M \gg 0$, $\sum_{i=1}^{d} (D^i)' D^i \gg 0$, points to a quite interesting direction of research. This condition can make sure that the corresponding control problem has a solution even if N becomes negative (i.e. not too negative). This is a distinct difference between deterministic and stochastic control theory. All the papers cited so far assume N to be non-negative. The papers [CLZ: SLQ], [CZ:SLQ2] and [HZ:ISRE] deal with BSRDEs arising from linear quadratic stochastic control problems with a possibly negative N, including existence results for some special cases.

Concerning the case of a non-negative N, the theory made important progresses in the last two years. In [T:GLQO], Tang proved a global existence and uniqueness result for the multidimensional BSRDE without further restrictions on the measurability of the coefficients if N is uniformly positive. It is a distinct feature of this article that it makes no use of an iteration in order to approximate the solution. In [KT:MBSR], some special cases of the multidimensional BSRDE is tackled when $\sum_{i=1}^{d} (D^i)' D^i$ and M is uniformly positive.

Chapter 2

Some auxiliary results

This chapter contains two results. The first one is concerned with the dual space of $H_q(0, T; \mathbb{R}^m)$, the second one with the solvability of SDEs such as (1.14), (1.15). Both results are not claimed to be new, although we have not found any reference for the first one in the literature. The second result can be derived from much more general statements, see for example [P:SIDE], Chap. V. We include this second result on SDEs because in our specific setting they are somewhat more immediately applicable.

2.1 The dual space of $\mathbf{H_q(0, T; \mathbb{R}^m)}$

The following lemma establishes an isomorphism between the dual space of $H_q(0, T; \mathbb{R}^m)$ and $H_{q'}(0, T; \mathbb{R}^m)$, that unfortunately is not shown to be isometric. The lemma also states that the H_q-spaces are reflexive, which will be needed to ensure the solvability of problem $\mathcal{P}(\tau, h)$. For simplicity we will write H_q instead of $H_q(0, T; \mathbb{R}^m)$.

Lemma 2.1 *For $q > 1$, H_q^* is isomorphic to $H_{q'}$, more precisely the application $i : H_{q'} \longrightarrow H_q^*$,*

$$v \mapsto \left(u \mapsto E[\int_0^T v'(s)u(s)ds] \right) := i(v), \tag{2.1}$$

is an isomorphism. H_q is reflexive.

Proof: First, by using Hölder's inequality twice, the linear and injective application (2.1) is well defined and continuous. We start by showing the assertion on the reflexivity of H_q. Set

$$Y := \left\{ z : [0, T] \longrightarrow \mathbb{R}^m : z \text{ is measurable and } |z|_Y := \left(\int_0^T |z(t)|^2 dt \right)^{\frac{1}{2}} < \infty \right\}.$$

We endow Y with the Sigma-algebra induced by $|\cdot|_Y$. According to [DS:LO], Lemma III.11.16 (b) and Thm. III.11.17, the normed linear space

$$L_q(Y) := \left\{ a : \Omega \longrightarrow Y : a \text{ is } \mathcal{F}_T - \text{measurable and } \|a\|_q := (E[|a|_Y^q])^{\frac{1}{q}} < \infty \right\}$$

can be identified with the normed linear space

$$\Big\{ b : [0,T] \times \Omega \longrightarrow \mathbb{R}^m : \\ b \text{ is } \mathcal{B} \otimes \mathcal{F}_T - \text{measurable and } \left(E[\left(\int_0^T |b(t)|^2 dt\right)^{\frac{q}{2}}]\right)^{\frac{1}{q}} < \infty \Big\}.$$

The identification means that for every a there is a $Leb \otimes P - a.s.$ unique b such that $a(\omega) = b(\cdot, \omega)$. Conversely, for every b, the mapping $\Omega \longrightarrow Y$, $\omega \mapsto h(\cdot, \omega)$ is measurable. Hence, H_q can be regarded as a closed subspace of

$$\begin{aligned} W_q &:= \Big\{ c : [0,T] \times \Omega \longrightarrow \mathbb{R}^m : c \text{ is} \\ &\qquad \mathcal{B} \otimes \mathcal{F}_T - \text{measurable and } \|c\|_{W_q} := \left(E[\left(\int_0^T |c(t)|^2 dt\right)^{\frac{q}{2}}]\right)^{\frac{1}{q}} < \infty \Big\} \\ &= L_q(\Omega, \mathcal{F}_T, Y). \end{aligned}$$

Thus, by [D:VM], Chap. II, § 13.5, Cor. 1, $W_{q'}$ is isometrically isomorphic to W_q^*, and the isomorphism is of the form (2.1) as an application of $W_{q'} \longrightarrow W_q^*$. In particular, W_q is reflexive, hence H_q, being a closed subspace of W_q, is also reflexive.

To show that i is an isomorphism we will consider its dual application i^*. We claim that i is the dual of the application $j : H_q \longrightarrow H_{q'}^*$, $u \mapsto \Big(v \mapsto E[\int_0^T u'vds]\Big)$(j is the mapping i after interchange of q and q'). The dual mapping $j^* : H_{q'}^{**} \longrightarrow H_q^*$ is defined by

$$\begin{aligned} v^{**} \in H_{q'}^{**} \quad \mapsto \quad & (H_q \longrightarrow \mathbb{R}) \\ & u \mapsto v^{**}(j(u)). \end{aligned}$$

As $H_{q'}$ is reflexive, for every $v^{**} \in H_{q'}^{**}$ there is a $v \in H_{q'}$ such that $v^{**}(\phi^*) = \phi^*(v)$ for all $\phi^* \in H_{q'}^*$. If we (canonically) identify $H_{q'}$ in this way with its second dual $H_{q'}^{**}$, the application j^* takes the form

$$\begin{aligned} j^* : H_{q'} \quad &\longrightarrow \quad H_q^* \\ v \quad &\mapsto \quad u \mapsto (j(u))\,(v) \\ &\qquad = u \mapsto \; E[\int_0^T v'u\; ds], \end{aligned}$$

hence, $i = j^*$ and $i^* = j$.

Next, we show that there is a constant $C > 0$ such that $\|j(u)\|_{H_{q'}^*} \geq C\,\|u\|_{H_q}$ for all $u \in H_q$. To demonstrate this, for a given $u = (u_1, \ldots, u_m)'$ define the process $v = (v_1, \ldots, v_m)'$ as

$$v_k(t) := u_k(t) \left(\int_0^t |u(s)|^2 ds \right)^{\frac{q}{2}-1}, \tag{2.2}$$

$k = 1,\ldots,m$. By convention, we set $0\cdot\infty = 0$ if the integral is zero and $q < 2$ (i.e. $\frac{q}{2} - 1 < 0$). The process v is $(\mathcal{F}_t)$- adapted and we have

$$\begin{aligned}
&\|v\|_{H_{q'}}^{q'} \\
&= E[\left(\int_0^T |v(t)|^2 dt\right)^{\frac{q'}{2}}] \\
&= E[\left(\int_0^T |u(t)|^2 \left(\int_0^t |u(s)|^2 ds\right)^{q-2} dt\right)^{\frac{q'}{2}}] \\
&\qquad \text{and with } h(t) := \textstyle\int_0^t |u|^2 ds \text{ because of } \int_0^T \frac{dh}{dt}\, h^{q-2} dt = \frac{1}{q-1}\left[h^{q-1}\right]_0^T \\
&= \left(\frac{1}{q-1}\right)^{\frac{q'}{2}} E[\left(\int_0^T |u(t)|^2 dt\right)^{(q-1)\frac{q'}{2}}] \\
&= \left(\frac{1}{q-1}\right)^{\frac{q'}{2}} E[\left(\int_0^T |u(t)|^2 dt\right)^{\frac{q}{2}}] \\
&= \left(\frac{1}{q-1}\right)^{\frac{q'}{2}} \|u\|_{H_q}^q .
\end{aligned}$$

We may calculate that for arbitrary u and for v as above

$$\begin{aligned}
&\frac{1}{\|v\|_{H_{q'}}} |j(u)(v)| \\
&= \frac{1}{\|v\|_{H_{q'}}} E[\int_0^T v'(t)u(t)dt] \\
&= \frac{1}{\|v\|_{H_{q'}}} E[\int_0^T |u(t)|^2 \left(\int_0^t |u(s)|^2 ds\right)^{\frac{q'}{2}-1} dt] \\
&= \frac{1}{\|v\|_{H_{q'}}} \frac{2}{q} E[\left(\int_0^T |u(t)|^2 dt\right)^{\frac{q}{2}}] \\
&\qquad \text{and with } \|v\|_{H_{q'}} = \left(\tfrac{1}{q-1}\right)^{\frac{1}{2}} \|u\|_{H_q}^{\frac{q}{q'}} \\
&= \frac{2}{q}(q-1)^{\frac{1}{2}} \|u\|_{H_q}^{q-\frac{q}{q'}} \\
&= \frac{2}{q}(q-1)^{\frac{1}{2}} \|u\|_{H_q} .
\end{aligned}$$

Thus,

$$\|j(u)\|_{H_{q'}^*} \quad = \sup_{\|v\|_{H_{q'}}=1} \{|j(u)(v)|\}$$

$$\geq \frac{2}{q}(q-1)^{\frac{1}{2}} \|u\|_{H_q}. \quad (2.3)$$

As $j = i^*$, this implies that i is onto, see [R:FA], Chap. 4, Thm. 4.15 . As all spaces involved are Banach spaces, i is an isomorphism and the Lemma is proven. ■

The spaces $H_q(\tau, T; \mathbb{R}^m)$ are closed subspaces of $H_q(0, T; \mathbb{R}^m)$, so they are also reflexive. If we replace in the above proof the integrals $\int_0^T$ and $\int_0^t$ by $\int_{\tau(\omega)}^T$ and $\int_{\tau(\omega)}^t$ we see that the same arguments apply for the mapping $i : H_q(\tau, T; \mathbb{R}^m) \longrightarrow H^*_{q'}(\tau, T; \mathbb{R}^m)$. We omitted the stopping times for simplicity; in the first instance it is somewhat easier to consider processes defined on the product space $[0, T] \times \Omega$, rather than on the stochastic interval $[\tau, T]$. So we get the following Corollary

Corollary 2.2 *The assertions of the preceding lemma also hold for the spaces $H_q(\tau, T; \mathbb{R}^m)$.*

2.2 Solvability of linear SDEs

The following lemma and its corollary show that (1.14), (1.15), the SDE for the problem $\mathcal{P}(0, h)$, has a well behaved solution that depends continuously on its initial value. The transition to the problems with random initial time will be straightforward. The method of the proof is to show that the integral operator induced by the SDE is a contraction on small time intervals. It is taken from [YZ:SC], Chap. 1, Thm. 6.3.

Lemma 2.3 *Assume $a \in L^\infty_{\mathcal{F}}(0, T; \mathbb{R})$ and $c \in L^\infty_{\mathcal{F}}(0, T; \mathbb{R}^{1\times d})$. For $q > 1$ and $h \in R_q(0, T; \mathbb{R})$, the SDE*

$$x(t) = h(t) + \int_0^t a(s)x(s)ds + \int_0^t x(s)c(s)dw(s) \quad (2.4)$$

possesses a unique solution $x \in R_q(0, T; \mathbb{R})$.
The linear mapping $R_q(0, T; \mathbb{R}) \longrightarrow R_q(0, T; \mathbb{R})$, $h \mapsto x$, is continuous.

Proof: The proof works with a contraction argument on subintervals of $[0, T]$. Let β be a common bound for $|a|$ and $|c|$. By the Burkholder-Gundy-Davis-inequaliy, there is a universal constant k such that for all $T_0 \in (0, T]$ and all $x \in R_q(0, T; \mathbb{R})$

$$\begin{aligned} E[\sup_{0\leq t\leq T_0} \left|\int_0^t x(s)c(s)dw(s)\right|^q] &\leq kE[\left(\int_0^{T_0} |x(s)|^2|c(s)|^2 ds\right)^{\frac{q}{2}}] \\ &\leq k\beta^q E[\left(T_0 \sup_{0\leq t\leq T_0} |x(s)|^2\right)^{\frac{q}{2}}] \\ &\leq k\beta^q T_0^{\frac{q}{2}} E[\sup_{0\leq t\leq T_0} |x(s)|^q] \\ &= k\beta^q T_0^{\frac{q}{2}} \|x\|^q_{R_q} < \infty. \quad (2.5) \end{aligned}$$

For the drift-term of equation (2.4) we have the estimate

$$E[\sup_{0\le t\le T_0}\left|\int_0^t a(s)x(s)ds\right|^q] \le T_0^q\beta^q\,\|x\|_{R_q} < \infty, \tag{2.6}$$

for T_0 and x as above. Hence, for arbitrary $T_0 \in (0,T]$, the application $R_q(0,T_0;\mathbb{R}) \longrightarrow R_q(0,T_0;\mathbb{R})$, $x \mapsto h(\cdot) + \int_0^\cdot axds + \int_0^\cdot xcdw$, is well defined. Using the inequality $|\xi+\eta|^q \le 2^{q-1}(|\xi|^q + |\eta|^q)$, $\xi,\eta\in\mathbb{R}$, and the estimate (2.5) and (2.6), we find that this mapping is a contraction, provided we choose T_0 sufficiently small such that

$$2^{q-1}\left(k\beta^q T_0^{\frac{q}{2}} + \beta^q T_0^q\right) < 1.$$

Hence, for sufficiently small T_0 we get by the Contraction Theorem a unique solution on $[0,T_0]$. This solution may be continued by the same procedure on the intervals $[T_0,2T_0],[2T_0,3T_0],\dots$. By this iteration we find a unique solution on $[0,T]$.

Hence, the application $h\mapsto x$ is well defined and, from the uniqueness of the solution, also linear. To show its continuity we may apply the Closed Graph Theorem: let (h_n) be a converging sequence in $R_q(0,T;\mathbb{R})$ with limit h. Assume that the sequence of corresponding solutions (x_n) converges to some x in $R_q(0,T;\mathbb{R})$. The calculations which led to (2.5) and (2.6) now show, that

$$\begin{aligned} x_n(\cdot) = h_n(\cdot) + \int_0^\cdot ax_nds + \int_0^\cdot x_ncdw \\ \to h(\cdot) + \int_0^\cdot axds + \int_0^\cdot xcdw, \end{aligned}$$

$n\to\infty$, in $R_q(0,T;\mathbb{R})$. Thus, x is the solution belonging to h, and the Closed Graph Theorem gives the desired continuity. ■

This, of course, easily carries over to the framework of equation (1.14), (1.15).

Corollary 2.4 *For $q>1$ and a given stopping time τ with $0\le\tau<T$ consider the SDE*

$$\begin{aligned} dx &= \{a(s)x(s)+\alpha(s)\}\,ds + \{x(s)c(s)+\beta(s)\}\,dw(s) \\ x(\tau) &= h, \end{aligned}$$

where, $a\in L^\infty_{\mathcal{F}}(0,T;\mathbb{R})$, $c\in L^\infty_{\mathcal{F}}(0,T;\mathbb{R}^{1\times d})$, $\alpha\in L^q_{\mathcal{F}}(0,T;\mathbb{R})$, $\beta\in H_q(0,T;\mathbb{R}^{1\times d})$ and $h\in L^q_{\mathcal{F}_\tau}(\mathbb{R})$.
This SDE possesses a unique solution $x\in L^q_{\mathcal{F}}(\Omega, C([\tau,T];\mathbb{R}))$, and there is a constant k, independent of τ,h,α and β such that

$$\|x\|_{L^q_c} \le k\left(|h|_{L^q} + \|\alpha\|_{L^q_{\mathcal{F}}} + \|\beta\|_{H_q}\right).$$

The same is true if $\alpha\in H_q(\tau,T;\mathbb{R})$ (in this case, on the right hand side of the above estimate, the norm of α is replaced by $\|\alpha\|_{H_q}$).

Proof: For a given τ and a triple (h, f, g), define the process $h(\cdot)$ by $h(t) = \mathbf{1}_{[\tau,T]}(t)h + \int_0^t \mathbf{1}_{[\tau,T]}(s)\alpha(s)ds + \int_0^t \mathbf{1}_{[\tau,T]}(s)\beta(s)dw(s)$. The desired solution is now given by that of (2.4) for this particular $h(\cdot)$. As $h(\cdot)$ depends linearly and continuously on (h, α, β), the existence of a k with the claimed property follows from the continuity statement in the above Lemma.

To proof the last assertion it suffices to show that $H_q(\tau, T; \mathbb{R}) \ni \alpha \mapsto \int_0^\cdot \mathbf{1}_{[\tau,T]}(s)\alpha(s)ds \in R_q(0, T; \mathbb{R})$ is also well defined and continuous. We have

$$
\begin{aligned}
E[\sup_{\tau \le t \le T} \left| \int_0^t \mathbf{1}_{[\tau,T]}(s)\alpha(s)ds \right|^q] &\le E[\left(\int_0^T \mathbf{1}_{[\tau,T]}(s)|\alpha(s)|ds \right)^q] \\
&\le E[\left(T^{\frac{1}{2}} \left(\int_0^T \mathbf{1}_{[\tau,T]}(s)|\alpha(s)|^2 ds \right)^{\frac{1}{2}} \right)^q], \\
&\quad \text{by the Cauchy-Schwarz-inequality,} \\
&= T^{\frac{q}{2}} \|\alpha\|_{H_q},
\end{aligned}
$$

hence we may take α from $H_q(\tau, T; \mathbb{R})$. ■

One may wonder why we insist on the requirement $\alpha \in H_q$. Let us place ourselves in the situation of Assumption A3, when $N = 0$ and $q \geq 2$ (i.e. $L^q \subset H_q$). Here, we perform the minimization of J over $H_q(\tau, T; \mathbb{R}^m)$ instead of $L^q_{\mathcal{F}}(\tau, T; \mathbb{R}^m)$. In this case, in (1.14), the term $B(s)u(s)$, playing the role of α, will belong to $H_q(\tau, T; \mathbb{R})$ instead of $L^q_{\mathcal{F}}(\tau, T; \mathbb{R})$.

So far, we have established that problem $\mathcal{P}(\tau, h)$ is at least well defined. Existence of an optimal control is now easily seen. This is deferred to the next chapter, where also an initial characterization of the optimal state is given.

Chapter 3

Existence, the FBSDE, and basic properties of solutions

It is not hard to see that under Assumptions A1 - A3 the problem $\mathcal{P}(\tau, h)$ is uniquely solvable. Yet, the proof is not constructive and gives us no idea what the optimal control $\bar{u}$ may look like. If we want to get more information, we may consider - just like in calculus - the derivative $J'(u)$ and try to resolve $J'(u) = 0$, which turns out to be hard (that is what may happen in calculus, too). In fact, this equation will provide a characterization of the optimal state in terms of a so called *Forward Backward Stochastic Differential Equation* (FBSDE), with an auxiliary condition. This is a system of two stochastic differential equations where the first equation has a specified initial value and the second equation a specified terminal value that depends on the terminal value of the first equation. Unfortunately, the "driver" of the first equation depends on the solution of the second and vice-versa. In short, it's tricky. And although we stated in the Introduction that we are not concerned about numerical questions, we will allow ourselves to consider the FBSDE as a mere theoretical tool. However, in this role it provides excellent help in settling, for example, questions of continuous and linear dependence.

The first section of this chapter contains an existence result for the problems $\mathcal{P}(\tau, h)$ and some counterexamples. Section 2 will establish the FBSDE and Section 3 collects some properties of the solution of the FBSDE.

3.1 Existence

Let us cite for convenience a basic result from convex analysis. We take it from [ET:CA], Chap. II, Prop. 1.2.

Proposition 3.1 *Let U be a real reflexive Banach-space and $\mathcal{C}$ be a closed, convex set of U. Assume that $J : \mathcal{C} \longrightarrow \mathbb{R}$ is convex and lower semicontinuous. If $\mathcal{C}$ is bounded or if J is coercive over $\mathcal{C}$ (i.e. $J(u) \to +\infty$ if $\|u\|_U \to \infty, u \in \mathcal{C}$), then*

$$J(\bar{u}) = \min_{u \in \mathcal{C}} J(u)$$

has at least one solution. The solution is unique if J is strictly convex over $\mathcal{C}$. ■

In the problem $\mathcal{P}(\tau, h)$, the minimization is performed over the whole linear space $\mathcal{U}$, so we will essentially have to check the coercivity. In the cases of A1 and A3, coercivity will be due to the *terminal cost* part of J, i.e. to $\frac{1}{q}E[M|x(T)|^q]$. We will make some use of BSDE-theory to exploit the fact that, to some extend, the "coercivity of a forward solution" translates into "continuous dependence of a backward solution".

Lemma 3.2 *Under the Assumptions A1, A2 or A3, the problem $\mathcal{P}(\tau, h)$ has a unique solution $\bar{u} \in \mathcal{U}$.*

Proof: Let us begin with A2, when $\mathcal{U} = L^q_{\mathcal{F}}(\tau, T; \mathbb{R}^m)$ and N is uniformly positive. The coercivity follows directly from $qJ(u) \geq E[\int_\tau^T N|u|^q ds] \geq \epsilon \|u\|^q_{L^q_{\mathcal{F}}}$ for some $\epsilon > 0$. The strict convexity of J is clear.

If A1 or A3 are in force, we have $\mathcal{U} = H_q(\tau, T; \mathbb{R}^m)$ and (recalling the definition of σ from (1.16)) $\sigma\sigma' \gg 0$, $M \gg 0$. Set

$$\theta := \sigma'(\sigma\sigma')^{-1}B'. \tag{3.1}$$

Then,

$$B' = \sigma\theta,$$

and due to $\sigma\sigma' \gg 0$ we have $\theta \in L^\infty_{\mathcal{F}}(\tau, T\ \mathbb{R}^d)$. Define the vector C by $C := [C^1, \ldots, C^i, \ldots, C^d] \in L^\infty_{\mathcal{F}}(\tau, T\ \mathbb{R}^d)$. With these notations the SDE (1.14) reads as

$$dx = \{Ax + \theta'\sigma' u\}\, ds + \{xC + u'\sigma\}\, dw. \tag{3.2}$$

At a preliminary stage, let us assume that

$$h = 0.$$

Now, suppose that J is not coercive, i.e. that there is a sequence $(v_n)_n$ in $H_q(\tau, T; \mathbb{R}^m)$ and a $c > 0$ such that $J(v_n) \leq c$ and $\|v_n\|_{H_q} \to \infty$, $n \to \infty$. We pass to the sequence $u_n := \frac{1}{\|v_n\|_{H_q}} v_n$. As $\|v_n\|_{H_q}$ tends to infinity and $h = 0$ we get $J(u_n) \to 0$, $n \to \infty$. Denote the solution of (3.2) that corresponds to u_n by x_n. Because of $M \gg 0$ we may estimate $qJ(u_n) \geq E[M|x_n(T)|^q] \geq \epsilon E[|x_n(T)|^q]$ for some $\epsilon > 0$, independent of n, hence $|x_n(T)|_{L^q} \to 0$, $n \to \infty$. We want to show that this contradicts $\|u_n\|_{H_q} \equiv 1$. To this end, we consider (3.2) as a BSDE. Define

$$R_n := x_n, \qquad\qquad S_n := x_n C + u_n'\sigma.$$

Note that $\sigma' u_n = S_n' - C' x_n$. (R_n, S_n) then solves the BSDE

$$\begin{aligned} dR_n &= \{AR_n + \theta'\sigma' u_n\}\, ds + S_n dw \\ &= \{AR_n + \theta'(S_n' - C'x_n)\}\, ds + S_n dw \\ &= \{(A - C\theta)R_n + \theta' S_n'\}\, ds + S_n dw, & (3.3) \\ R_n(T) &= x_n(T). & (3.4) \end{aligned}$$

(R_n, S_n) is the unique solution of this BSDE in $L^q_{\mathcal{F}}(\Omega, C([\tau, T]; \mathbb{R})) \times H_q(\tau, T; \mathbb{R}^m)$, see Thm. 5.1 in [EPQ:BSDE]. Prop. 5.1 in the same article yields that there is a $k > 0$, independent of n, such that

$$\|R_n\|_{L^q_c} + \|S_n\|_{H_q} \le k\,|x_n(T)|_{L^q} \to 0,\ n \to \infty, \tag{3.5}$$

because the terminal value of the equation tends to zero, $|x_n(T)|_{L^q} \to 0$, $n \to \infty$ (continuous dependence on the terminal value). As $\|x_n\|_{L^q_c} \to 0$ implies $\|x_n\|_{H_q} \to 0$, hence $\|x_n C\|_{H_q} \to 0$, it follows that $\|u_n'\sigma\|_{H_q} = \|S_n - x_n C\|_{H_q} \to 0$, $n \to \infty$. But from the uniform positivity of $\sigma\sigma'$ and the normalization of the u_n we get

$$\begin{aligned}
\|u_n'\sigma\|^q_{H_q} &= E[\left(\int_\tau^T u_n'\sigma\sigma' u_n ds\right)^{\frac{q}{2}}] \\
&\ge E[\left(\epsilon\int_\tau^T |u_n|^2 ds\right)^{\frac{q}{2}}] \\
&= \epsilon^{\frac{q}{2}}\,\|u_n\|_{H_q} \\
&= \epsilon^{\frac{q}{2}},
\end{aligned}$$

hence, the assumption $|x_n(T)|_{L^q} \to 0$ contradicts $\|u_n\|_{H_q} \equiv 1$, and J is coercive if $h = 0$. Now assume h to be arbitrary and assume $(v_n)_n$ is a sequence with $\|v_n\|_{H_q} \to \infty$, $n \to \infty$. Denote by x_n the corresponding solutions of (3.2) and by x_n^0 the solutions for the control v_n and the initial value $h = 0$. Let Γ solve the SDE $d\Gamma = A\Gamma ds + \Gamma C dw$, $\Gamma(\tau) = h$; then $x_n = x_n^0 + \Gamma$, and in particular we have $J(v_n) \ge \frac{1}{q}E[M|x_n(T)|^q] \ge \epsilon E[|x_n^0(T) + \Gamma(T)|^q] = \epsilon\,|x_n^0(T) + \Gamma(T)|^q_{L^q}$, with some $\epsilon > 0$ independent of n. As we have seen, $|x_n^0(T)|_{L^q}$ tends to infinity, hence $J(v_n) \to \infty$, $n \to \infty$ also if $h \ne 0$.
Finally, in order to show that J is strictly convex, choose two control processes $u_1 \ne u_2$ and denote by x_1, x_2 the corresponding solutions of the state equation (1.14). By the strict convexity of the function $\mathbb{R} \longrightarrow \mathbb{R}$, $r \mapsto \frac{1}{q}|r|^q$, and the uniform positivity of M, it suffices to show, that $x_1(T) \ne x_2(T)$ (in $L^q_{\mathcal{F}_T}(\mathbb{R})$). Set $R = x_1 - x_2$ and $S = RC + (u_1 - u_2)'\sigma$. Again, (R, S) satisfies the BSDE

$$dR = \{(A - C\theta)R + \theta' S'\}\,ds + Sdw,\quad R(T) = x_1(T) - x_2(T).$$

Now suppose that $x_1(T) - x_2(T) = 0$. The already mentioned Prop. 5.1 of [EPQ:BSDE] then entails $R = 0$, $S = 0$, hence $\|(u_1 - u_2)'\sigma\|_{H_q} = 0$. As σ has maximal rank, this is not possible unless $u_1 = u_2$, hence the assumption $x_1(T) - x_2(T) = 0$ leads to a contradiction and the lemma is proved. ∎

Instead of the particular θ used above we may have chosen any essentially bounded, adapted process that satisfies $B' = \sigma\theta$. These processes play a prominent role in mathematical finance. There, we have $C = 0$ and the processes θ help to parametrize the so called (signed) *martingale measures,* i.e. the (signed) measures dominated by P under which the

semimartingale $\int_0^\cdot B' ds + \int_0^\cdot \sigma dw$ is a martingale. These measures (i.e. their P-densities) turn out to be the dual space of the linear space of "attainable terminal values" $x(T)$ when $h = 0$ (when $h \neq 0$, the set of all possible terminal values $x(T)$ is of course an affine space). We will return to this parametrization in Chapter 6 when we discuss some financial market problems.
We have already mentioned that, apart from Assumptions 1-3, there are other situations where an optimal control may exist. The following remark gives an example of such a situation that is of interest because it exhibits a basic difference between the stochastic and deterministic setting of linear isoelastic problems: the stochastic problem may have a solution, even if the weighting process N is uniformly negative. This is impossible in the deterministic case. For linear quadratic problems, a discussion of this observation and examples are given in [YZ:SC].

Remark 3.3 *1. In the deterministic setting, i.e. for deterministic (but time dependent) coefficients with $C^i = 0$, $D^i = 0$, $i = 1, \dots, d$, and deterministic control processes u, the problem $\mathcal{P}(0,1)$ cannot have a solution if N is strictly negative on a set S with $Leb(S) > 0$.*

2. Assume that $q < 2$, $C^i = 0$, $i = 1, \dots, d$, and M as well as $\sum_{i=1}^d (D^i)' D^i$ are uniformly positive. Then there is a constant $\delta > 0$ such that the problem $\mathcal{P}(0,1)$ is uniquely solvable if $-\delta \leq N$, $Leb \otimes P - a.s.$.

Proof:

1. We will exploit the fact that L^1- and L^q-norms are not equivalent to show that there is a sequence (u_n) such that $J(u_n) \to -\infty$, $n \to \infty$.
The state equation reduces to $dx = \{Ax + Bu\}ds$, $x(0) = 1$, where A and B are $\mathbb{R}$- resp. $\mathbb{R}^m$-valued, measurable and essentially bounded functions on $[0,T]$. The set $\mathcal{U}$ of admissible control consists of of the $\mathbb{R}^m$-valued, measurable and q-integrable functions on $[0,T]$. Let us define the essentially bounded and uniformly positive function $E(t) := \exp\{\int_0^t A(s)ds\}$. Then the solution of the state equation is given by $x(t) = E(t) + E(t)\int_0^t \frac{1}{E(s)} B(s)u(s)ds$.
Now let us assume that there is a set $S \subset [0,T]$ with $N < 0$ on S and $Leb(S) > 0$. We will show that this entails $\inf_u J(u) = -\infty$. Without loss of generality we may assume that there is an $\epsilon > 0$ such that $N(s) \leq -\epsilon$ for $s \in S$ (otherwise we may quit to a subset S_0 of S with $Leb(S_0) > 0$). As the completed Borel-field is atom-free there is a measurable function $v : S \longrightarrow \mathbb{R}$ and a sequence S_n of measurable subsets of S such that $\int_S |v(s)|ds < \infty$, $\int_{S_n} |v(s)|^q ds < \infty$ for all n, but $\int_{S_n} |v(s)|^q ds \to \infty$, $n \to \infty$ (compare [M:IP], Thm 6.4.2; it suffices to show that such v and S_n exit if $S = [0,1]$, and for this particular S take $S_n = [0, 1 - \frac{1}{n}]$, $v(s) = (1-s)^\alpha$ with $-1 < \alpha < -\frac{1}{q}$). Define $u_n = \mathbf{1}_{S_n} v$ on S and $u_n = 0$ outside S. Denote by x_n the solutions of the state equations corresponding to u_n. As the functions N an Q are essentially bounded (M is a scalar), there are constants $c_j > 0$ such that for all $t \in [0,T]$ and all n

$$|x_n(t)| = \left| E(t) + E(t)\int_0^t \frac{1}{E(s)} B(s)u_n(s)ds \right|$$

$$
\begin{aligned}
&\leq |E(t)|\left(1+\int_0^t \left|\frac{1}{E(s)}B(s)\right| |u_n(s)|ds\right)\\
&\leq c_1\left(1+\int_0^T |u_n(s)|ds\right)\\
&\leq c_2,
\end{aligned}
$$

since $\int_0^T |u_n(s)|ds \leq \int_S |v(s)|ds < \infty$. Hence, $\int_0^T |Q(s)||x_n(s)|^q ds + |M||x_n(T)| < c_3$ for all n. It follows, that , $J(u) \leq c_3 + \int_0^T N(s)|u_n(s)|^q ds = c_3 + \int_{S_n} N(s)|v(s)|^q ds \leq c_3 - \epsilon \int_{S_n} |v(s)|^q ds$. As we have chosen v and S_n such that the last integral tends to ∞, we get $J(u_n) \to -\infty$, $n \to \infty$, and there exists no minimal value of the problem that can be attained.

2. We will show that J is coercive if N is not "too much negative". We proceed as in Lemma 3.2 and will use BSDE theory. Choose a $\theta \in L^\infty_{\mathcal{F}}(0, T\ \mathbb{R}^d)$ such that $\sigma\theta = B'$. The state equation then reads as $dx = \{Ax + \theta'\sigma' u\}ds + u'\sigma dw$, $x(0) = 1$. Consider this as a BSDE $dR = \{AR + \theta' S'\}ds + Sdw$, $R(T) = x(T)$, with $x = R$ and $S = u'\sigma$. By Prop. 5.1 in [EPQ:BSDE], there is a $c_1 > 0$ independent of u such that $\|S\|^q_{H_q} \leq c_1 \, |x(T)|^q_{L^q}$. As $\sigma\sigma' \gg \epsilon$ and $M \gg \epsilon$ we have

$$
\begin{aligned}
q\,J(u) &= E[\int_0^T Q|x|^q + N|u|^q ds + M|x(T)|^q]\\
&\geq E[\int_0^T N|u|^q ds + \epsilon|x(T)|^q]\\
&\geq E[\int_0^T N|u|^q ds] + \frac{\epsilon}{c_1}\,\|\sigma' u\|^q_{H_q}\\
&= E[\int_0^T N|u|^q ds] + \frac{\epsilon}{c_1}E[\left(\int_0^T u'\sigma\sigma' u\, ds\right)^{\frac{q}{2}}]\\
&\geq E[\int_0^T N|u|^q ds] + \frac{\epsilon^2}{c_1}\,\|u\|^q_{H_q}\\
&\geq \left(\text{ess.inf}_{(s,\omega)} N\right) \|u\|^q_{L^q_{\mathcal{F}}} + \frac{\epsilon^2}{c_1}\,\|u\|^q_{H_q}\,.
\end{aligned}
$$

Let us assume that ess.inf $N \leq 0$ (the case of a positive N is covered by Lemma 3.2). As $\|u\|^q_{L^q_{\mathcal{F}}} \leq T^{\frac{q}{2}}\,\|u\|_{H_q}$, then there is a $c_2 > 0$ independent of u such that

$$
q\,J(u) \geq \left(T^{\frac{q}{2}}\text{ess.inf}_{(s,\omega)} N + c_2\right) \|u\|^q_{H_q}\,.
$$

So, if $0 < \delta < c_2$ and $T^{\frac{q}{2}}$ess.inf $N \geq -\delta$, we get $J(u) \to \infty$ if $\|u\|_{H_q} \to \infty$, i.e, J is coercive, and the assertion of the remark follows.

■

From the preceding proofs of existence we do not get further knowledge on the optimal control or the optimal state process. The next section will give a characterization that is not really very explicit, but that will e.g. help to decide if a certain process is or is not the optimal control for a given problem.

3.2 The FBSDE

The *Forward Backward Stochastic Differential Equation* arises directly from the study of the Gâteaux-derivative ("directional derivative", see [ET:CA], Def. I.5.2) of the functional $u \mapsto J(u)$. So, we first have to determine this derivative. We imitate the procedure of [CK:TSMP], Lemma 1.1.

Lemma 3.4 *The Gâteaux-derivative J' of the cost functional $J : \mathcal{U} \longrightarrow \mathbb{R}$ in problem $\mathcal{P}(\tau, h)$ is given by the following:*
For $v \in \mathcal{U}$ denote by ξ the solution of the SDE

$$d\xi(t) = \{A(s)\xi(s) + B(s)v(s)\}\, ds + \sum_{i=1}^{d} \left\{C^i(s)\xi(s) + D^i(s)v(s)\right\} dw^i(s), \tag{3.6}$$

$$\xi(\tau) = 0. \tag{3.7}$$

Then, for all $u, v \in \mathcal{U}$

$$J'(u) \cdot v = E[\int_\tau^T Q(s)\varphi(x(s))\xi(s) + N(s)v'(s)\varphi(u(s))ds + M\varphi(x(T))\xi(T)].$$

Here, the dot " $\cdot$ " denotes the duality product between $\mathcal{U}^*$ and $\mathcal{U}$.
The derivative (as a mapping $J' : \mathcal{U} \longrightarrow \mathcal{U}^*, u \mapsto J'(u)$) is continuous.

Proof: We will consider the three parts of the cost functional separately and set $J_1(u) := \frac{1}{q}E[\int_\tau^T N|u|^q ds]$, $J_2(u) := \frac{1}{q}E[\int_\tau^T Q|x|^q ds]$, $J_3(u) := \frac{1}{q}E[M|x(T)|^q]$. For $u, v \in \mathcal{U}$ we must evaluate the limit $\frac{1}{\lambda}(J_k(u + \lambda v) - J(u))$, $\lambda \to 0$, $\lambda \geq 0$, $k = 1, 2, 3$.
Let us start with J_1. We have

$$\frac{1}{\lambda}(J_1(u + \lambda v) - J_1(u)) = \frac{1}{\lambda}E[\int_\tau^T N\, \frac{1}{q}(|u + \lambda v|^q - |u|^q)ds]. \tag{3.8}$$

Recall that φ is the derivative of $\mathbb{R}^m \longrightarrow \mathbb{R}$, $u \mapsto \frac{1}{q}|u|^q$. From the Mean Value Theorem, for every $\lambda > 0$ there is a mapping $m_\lambda : [0, T] \times \Omega \supset [\tau, T] \longrightarrow [0, 1]$ such that

$$\frac{1}{q}(|u(s) + \lambda v(s)|^q - |u(s)|^q) = \lambda v'(s)\varphi(u(s) + \lambda m_\lambda(s)v(s)), \tag{3.9}$$

for all $(s,\omega) \in [\tau, T]$. We may assume that m_λ is adapted, see Appendix A. The differential quotient (3.8) now reads as

$$\frac{1}{\lambda}\left(J_1(u+\lambda v) - J_1(u)\right) = E[\int_\tau^T N\, v'\varphi(u+\lambda m_\lambda v)ds], \tag{3.10}$$

and we want to take take the limit $\lambda \searrow 0$ in the right hand side integral. As m_λ is bounded, it is clear that $u + \lambda m_\lambda v \to u$, $\lambda \searrow 0$, (s,ω)-pointwise on $[\tau, T]$. As $u \mapsto \frac{1}{q}|u|^q$ is convex, its derivative is monotone, i.e. for all $z_1, z_2 \in \mathbb{R}^m$ we have $(z_1 - z_2)'(\varphi(z_1) - \varphi(z_2)) \geq 0$. For arbitrary scalars $\alpha, \beta \in [0,1]$ with $\alpha \leq \beta$ we may set $z_1 = u + \beta v$ and $z_2 = u + \alpha v$. The monotonicity of φ yields $(\beta - \alpha)v'\left(\varphi(u+\beta v) - \varphi(u+\alpha v)\right) \geq 0$, hence $v'\varphi(u+\beta v) \geq v'\varphi(u+\alpha v)$. For $\lambda \in (0,1)$ we have $\lambda m_\lambda \in (0,1)$. Thus,

$$v'\varphi(u) = v'\varphi(u + 0\cdot v) \leq v'\varphi(u+\lambda m_\lambda \cdot v) \leq v'\varphi(u + 1\cdot v),$$

$Leb \otimes P - a.s.$ for $\lambda \in (0,1)$, so we can apply the Dominated Convergence Theorem in (3.10) and get

$$\frac{1}{\lambda}\left(J_1(u+\lambda v) - J_1(u)\right) \to E[\int_\tau^T N\, v'\varphi(u)ds] = J_1'(u)\cdot v,$$

$\lambda \searrow 0$.

For the functionals J_2 and J_3 we may proceed very similarly. Let us denote by x^u the solution of (1.14), (1.15) that corresponds to u and by ξ^v the solution of (3.6), (3.7) that corresponds to v. Then, $x^{u+\lambda v} = x^u + \lambda \xi^v$. Choose $u, v \in \mathcal{U}$ and $\lambda > 0$. Again, by the Mean Value Theorem there is a (adapted) process $m_\lambda^{(2)}$ and a r.v. $m_\lambda^{(3)}$ with values in $[0,1]$ such that

$$\begin{aligned}
\frac{1}{\lambda}\left(J_2(u+\lambda v) - J_2(u)\right) &= \frac{1}{q}E[\int_\tau^T Q\,\left(|x^u + \lambda\xi^v|^q - |x^u|^q\right)ds] \\
&= E[\int_\tau^T Q\,\xi^v\varphi(x^u + \lambda m_\lambda^{(2)}\xi^v)ds], \\
\frac{1}{\lambda}\left(J_3(u+\lambda v) - J_3(u)\right) &= \frac{1}{q}E[M\left(|x^u(T) + \lambda\xi^v(T)|^q - |x^u(T)|^q\right)] \\
&= E[M\,\xi^v(T)\varphi\left(x^u(T) + \lambda m_\lambda^{(3)}\xi^v(T)\right)].
\end{aligned}$$

Just as in the calculation for J_1, the monotonicity of φ entails for $\lambda \in (0,1)$

$$\begin{aligned}
\xi^v\varphi(x^u) \leq \xi^v\varphi(x^u + \lambda m_\lambda^{(2)}\xi^v) \leq \xi^v\varphi(x^u + \xi^v), \\
\xi^v(T)\varphi(x^u(T)) \leq \xi^v(T)\varphi(x^u(T) + \lambda m_\lambda^{(3)}\xi^v(T)) \leq \xi^v(T)\varphi(x^u(T) + \xi^v(T)),
\end{aligned}$$

$Leb \otimes P - a.s.$ respectively $P - a.s.$. The Dominated Convergence Theorem yields

$$\begin{aligned}
J_2'(u)\cdot v &= E[\int_\tau^T Q(s)\,\xi^v(s)\varphi(x^u(s))ds], \\
J_3'(u)\cdot v &= E[M\,\xi^v(T)\varphi(x^u(T))],
\end{aligned}$$

hence, $J' = J_1' + J_2' + J_3'$ has the desired form.
From Corollary 2.4 we know that the mappings $\mathcal{U} \longrightarrow L^q_{\mathcal{F}}(\tau, T; \mathbb{R})$, $u \mapsto x^u$, $\mathcal{U} \longrightarrow L^q_{\mathcal{F}_T}(\mathbb{R})$, $u \mapsto x^u(T)$ are continuous. Along with the continuity of $L^q_{\mathcal{F}}(\tau, T; \mathbb{R}^m) \longrightarrow L^{q'}_{\mathcal{F}}(\tau, T; \mathbb{R}^m)$, $u \mapsto \varphi(u)$, (accordingly for $x \mapsto \varphi(x)$ and $x(T) \mapsto \varphi(x(T))$), we get, that $\mathcal{U} \longrightarrow L^{q'}_{\mathcal{F}}(\tau, T; \mathbb{R}) \times L^{q'}_{\mathcal{F}}(\tau, T; \mathbb{R}^m) \times L^{q'}_{\mathcal{F}_T}(\mathbb{R})$, $u \mapsto (\varphi(x^u), \varphi(u), \varphi(x^u(T))$ is continuous. The claimed continuity of J' now follows from the boundedness of the linear operator $\mathcal{U}^* \longrightarrow L^{q'}_{\mathcal{F}}(\tau, T; \mathbb{R}) \times L^{q'}_{\mathcal{F}_T}(\mathbb{R})$, $v \mapsto (\xi^v, \xi^v(T))$. ■

Yet, we can state that problem $\mathcal{P}(\tau, h)$ has a solution $u \in \mathcal{U}$ if and only if $J'(u) \cdot v = 0$ for all $v \in \mathcal{U}^*$ (the sufficiency part relies on the continuity of J'), see [ET:CA], Prop. II.2.1. The FBSDE may (in our context) be seen as some paraphrase of this condition that makes it more tractable. FBSDEs are a familiar tool in stochastic control theory, in particular in the theory built around the so called *Maximum Principle.* There, a relation is established between the optimal control on one side and the optimal state process as well as the *first- and second-order adjoint processes* on the other side. The link is given by the fact that the optimal control maximizes (pointwise!) the *Hamiltonian* function of the control problem. However, in our FBSDE only the first-order adjoint process will appear, and it doesn't explicitly include a Hamiltonian or a pointwise maximization thereof. Implicitly, though, this maximization is incorporated in the *auxiliary condition* (3.14). Yet, we will not try to give an introduction to the Maximum Principle and recommend [YZ:SC] for further reading.

Proposition 3.5 *Given the coefficients of problem $\mathcal{P}(\tau, h)$, consider the FBSDE (more precisely: FBSDE with auxiliary condition)*

$$dx(t) = \{A(s)x(s) + B(s)u(s)\}\, ds + \sum_{i=1}^{d} \left\{C^i(s)x(s) + D^i(s)u(s)\right\} dw^i(s), \tag{3.11}$$

$$dy(t) = \left\{-A(s)y(s) - \sum_{i=1}^{d} C^i(s)z^i(s) - Q(s)\varphi\left(x(s)\right)\right\} ds + \sum_{i=1}^{d} z^i(s)dw^i(s), \tag{3.12}$$

$$x(\tau) = h, \qquad y(T) = M\varphi(x(T)), \tag{3.13}$$

$$B'y + \sum_{i=1}^{d}(D^i)'z^i + N\varphi(u) = 0, \quad Leb \otimes P - a.s.. \tag{3.14}$$

The problem $\mathcal{P}(\tau, h)$ is solvable if and only if this FBSDE has a solution

$$(x, u, y, z) \in L^q_{\mathcal{F}}(\Omega, C([\tau, T]; \mathbb{R})) \times \mathcal{U} \times L^{q'}_{\mathcal{F}}(\Omega, C([\tau, T]; \mathbb{R})) \times H_{q'}(\tau, T; \mathbb{R}^d).$$

If the FBSDE has a solution (x, u, y, z), then u is an optimal control with corresponding state x. If x and u are an optimal state and control for problem $\mathcal{P}(\tau, h)$, then (x, u) belongs to a solution of the FBSDE.

Proof: For $u, v \in \mathcal{U}$, we use the notation x^u and ξ^v from Lemma 3.4 for the solution of the state equation with initial value h and control u respectively initial value 0 and control v. Let $(y^u, z_u) \in L^{q'}_{\mathcal{F}}(\Omega, C([\tau, T]; \mathbb{R})) \times H_{q'}(\tau, T; \mathbb{R}^m)$ be the unique solution of the BSDE

$$\begin{aligned} dy^u &= \left\{ -Ay^u - \sum_{i=1}^d C^i z_u^i - Q\varphi(x^u) \right\} ds + \sum_{i=1}^d z_u^i dw^i, \\ y^u(T) &= M\varphi(x^u(T)). \end{aligned}$$

Applying Itô's formula to $y^u \xi^v$ yields

$$\begin{aligned} d(y^u \xi^v) &= \left\{ -Q\varphi(x^u)\xi^v + y^u Bv + \sum_{i=1}^d D^i z_u^i v \right\} ds \\ &\quad + \sum_{i=1}^d \left\{ \xi^v z_u^i + y^u C^i \xi^v + y^u D^i v \right\} dw^i, \end{aligned}$$

hence

$$E[y^u(T)\xi^v(T) - y^u(0)\xi^v(0)] = E[\int_\tau^T -Q\varphi(x^u)\xi^v + y^u Bv + \sum_{i=1}^d D^i z_u^i v ds].$$

Note that $y^u(T) = M\varphi(x^u(T))$ and $\xi^v(0) = 0$. Adding $E[\int_\tau^T Q\varphi(x^u)\xi^v + N(\varphi(u))'vds]$ on both sides of the last equality then yields

$$\begin{aligned} & E[\int_\tau^T y^u Bv + \sum_{i=1}^d D^i z_u^i v + N(\varphi(u))'vds] \\ &= E[\int_\tau^T Q\varphi(x^u)\xi^v + N(\varphi(u))'vds + M\varphi(x^u(T))\xi^v(T)] \\ &= J'(u) \cdot v, \end{aligned}$$

by Lemma 3.4. The control $u \in \mathcal{U}$ is a solution for problem $\mathcal{P}(\tau, h)$ if and only if $J'(u)\cdot v = 0$ for all $v \in \mathcal{U}$, (see [ET:CA], Prop. II.2.1). Thus, $u \in \mathcal{U}$ solves problem $\mathcal{P}(\tau, h)$ if and only if

$$E[\int_\tau^T \left(y^u B + \sum_{i=1}^d D^i z_u^i + N(\varphi(u))' \right) vds] = 0$$

for all $v \in \mathcal{U}$, i.e. if and only if

$$B'y^u + \sum_{i=1}^d (D^i)' z_u^i + N\varphi(u) = 0, \quad Leb \otimes P - a.s.,$$

which implies the assertion of the proposition. ■

From Lemma 3.2 we can immediately deduce the following corollary.

Corollary 3.6 *Under Assumptions A1-A3, the FBSDE (3.11)-(3.14) has a unique solution* $(\bar{x}, \bar{u}, \bar{y}, \bar{z}) \in L^q_{\mathcal{F}}(\Omega, C([\tau, T]; \mathbb{R})) \times \mathcal{U} \times L^{q'}_{\mathcal{F}}(\Omega, C([\tau, T]; \mathbb{R})) \times H_{q'}(\tau, T; \mathbb{R}^m)$. ■

The equation (3.12) is in general called the *adjoint equation* (for problem $\mathcal{P}(\tau, h)$) and its solution part y is called the *adjoint process.* Being doubly inaccurate we will continue to use the terminology "FBSDE" for the system of equations (3.11) - (3.13) and the auxiliary condition (3.14). In the standard definition of a FBSDE, no unknown process u occurs and there's no auxiliary condition, see [MY:FBSDE]. Actually, if $N > 0$, one could replace u by $f(\frac{1}{N}(B'y + \sum_{i=1}^d (D^i)'z^i))$ in (3.11) and skip (3.13). The result would be a "standard" FBSDE.
A direct study of the FBSDE in order to find an optimal control is not very encouraging. Though, if one has a guess of what the optimal control might be, this could be checked with the help of the FBSDE. As a fairly simple example, consider the problem $dx = u'dw$, $x(0) = 1$, $J(u) = \frac{1}{q}E[|x(T)|^q]$, i.e. the problem of finding an integrand in $H_q(0, T; \mathbb{R}^d)$, such that the corresponding stochastic integral is close to 1 (in the q-th mean) at time T. The existence of an optimal control is assured by Lemma 3.2; depending on q the situation is covered by Assumptions A1 or A3. In the quadratic case, i.e. for $q = 2$, the solution is immediate:

$$\begin{aligned} E[|x(T)|^2] &= E[\left(1 + \int_0^T u'dw\right)^2] \\ &= E[1 + 2\int_0^T u'dw + \left(\int_0^T u'dw\right)^2] \\ &= 1 + E[\int_0^T |u|^2 ds], \end{aligned}$$

by the isometry property. Hence, J is minimized for $u = 0$. If $1 < q < 2$, one can set $\bar{u} = 0$ in the FBSDE and see that it's also optimal for these q. Of course, in this situation it would not have been hard to get the same result by differentiating $|x|^q$, although one must take some care about the zeros of x.

Note that for $q = 2$ the FBSDE is *linear.* In a setting where the FBSDE is uniquely solvable, this entails that the solution $(\bar{x}, \bar{u}, \bar{y}, \bar{z})$ - in particular $\bar{y}$ - depends *linearly* on the initial value h. This is a simple observation, but it permitted J.-M. Bismut in his seminal paper [B:LQOC] to introduce a BSDE, depending only on the coefficients of the problem - the stochastic Riccati-equation (1.7), (1.8) - whose solution (K, L) allows a *decoupling* of the FBSDE. This means, if one knows (K, L), one can arrange things in such a way that the FBSDE falls apart into a SDE and a BSDE that can both be solved *independently.* More precisely, there is a function G (which depends on the coefficients of the control problem) such that the ansatz $\bar{u} = G(K, L)\bar{x}$, plugged into (3.11), transfers this equation into a simple SDE and indeed yields the optimal state and the optimal control. If $\bar{x}$ is known, equation (3.12) reduces to a linear BSDE.

From (1.9) one sees that in the quadratic case we have

$$G(K,L) = -\left(N + \sum_{i=1}^{d}(D^i)'KD^i\right)^{-1}\left(KB + \sum_{i=1}^{d}(C^i)'KD^i + \sum_{i=1}^{d}L^iD^i\right)',$$

where, of course, we have to make sure that $\left(N + \sum_{i=1}^{d}(D^i)'KD^i\right)^{-1}$ exists.
Our aim is to carry over this method to the non-quadratic case. To this end, we should know a bit more about how, for example $\bar{y}$, depends on h (looking at the FBSDE (3.11)-(3.14), a linear dependence such as in the quadratic case can not be expected). In the next section we collect some properties of the solution of the FBSDE that can be derived quite immediately.

3.3 Basic properties of the solution

Consider a problem $\mathcal{P}(\tau, h)$ and assume, that $\bar{x}$ is an optimal state for this problem. Assume that $\bar{x}$ happens to attain zero. Intuitively, it is clear that $\bar{u}$ vanishes after this time and that $\bar{x}$ stays at 0. We want to make this precise and therefore introduce the following:

Definition 3.7 *(stopping time τ_0) Suppose that we are given a problem $\mathcal{P}(\tau, h)$ and an optimal state $\bar{x}$ for this problem. We define τ_0 as the first time that $\bar{x}$ attains 0, i.e. τ_0 is the stopping time*

$$\tau_0 := \inf\{t \in [0,T] : \bar{x}(t \vee \tau) = 0\} \wedge T.$$

Then, given one of our standard assumptions, we have the following Lemma.

Lemma 3.8 *Consider problem $\mathcal{P}(\tau, h)$ and let one of the Assumptions A1-A3 hold. Let $(\bar{x}, \bar{u}, \bar{y}, \bar{z})$ be the corresponding solution of the FBSDE (3.11) - (3.14). Then $(\bar{x}, \bar{u}, \bar{y}, \bar{z})$ vanishes after τ_0. More precisely*

$$(\bar{x}, \bar{u}, \bar{y}, \bar{z}) = \mathbf{1}_{[\tau,\tau_0]}(\bar{x}, \bar{u}, \bar{y}, \bar{z}).$$

Proof: Set $(\bar{x}_0, \bar{u}_0, \bar{y}_0, \bar{z}_0) = \mathbf{1}_{[\tau,\tau_0]}(\bar{x}, \bar{u}, \bar{y}, \bar{z})$. From the definition of τ_0 we get $\bar{x}(\cdot \wedge \tau_0) = \mathbf{1}_{[\tau,\tau_0]}\bar{x}$. This yields that $\bar{x}_0$ is the solution of the state equation (1.14) for $u = \bar{u}_0$. We have $|\bar{u}_0| \leq |\bar{u}|$, $|\bar{x}_0| \leq |\bar{x}|$, and since our assumptions demand Q, N and M to be non-negative, we get $J(\bar{u}_0) \leq J(\bar{u})$. As the optimal control is unique, it follows that

$$\bar{u}_0 = \bar{u}, \qquad \bar{x}_0 = \bar{x}.$$

We will now show that $(\bar{x}_0, \bar{u}_0, \bar{y}_0, \bar{z}_0)$ solves the FBSDE (3.11) - (3.14). The initial and terminal values (3.13), as well as the auxiliary condition (3.14) are easily checked. Let us consider the adjoint equation (3.12). Let Γ be the solution of the SDE $d\Gamma = A\Gamma dt +$

$\Gamma\sum_{i=1}^d C^i dw^i$, $\Gamma(\tau) = 1$. For a stopping time γ with $\tau \le \gamma \le T$ we then have the representation (compare [EM:BSDE] Prop 1.2.4.)

$$\begin{aligned}\bar{y}(\gamma) &= \Gamma(\gamma)^{-1}E[\Gamma(T)M\varphi(\bar{x}(T)) + \int_\gamma^T \Gamma(s)Q(s)\varphi(\bar{x}(s))ds|\mathcal{F}_\gamma] \\ &= \Gamma(\gamma)^{-1}E[\Gamma(T)M\varphi(\bar{x}_0(T)) + \int_\gamma^T \Gamma(s)Q(s)\varphi(\bar{x}_0(s))ds|\mathcal{F}_\gamma],\end{aligned} \tag{3.15}$$

where we replaced $\bar{x}$ by $\bar{x}_0$ in the second line. Let us show that $\bar{y}(\gamma) = 0$ on the set $\{\tau_0 < \gamma\}$ (this implies that $\bar{y} = \mathbf{1}_{[\tau,\tau_0]}\bar{y}$). On the set $\{\tau_0 < \gamma\}$ we have $\{\tau_0 < T\}$, hence for all $\mathcal{F}_\gamma-$ measurable subsets $S \subset \{\tau_0 < \gamma\}$

$$\begin{aligned}\mathbf{1}_S E[\Gamma(T)M\varphi(\bar{x}_0(T))|\mathcal{F}_\gamma] &= E[\mathbf{1}_S\Gamma(T)M\varphi(\bar{x}_0(T))|\mathcal{F}_\gamma] \\ &= E[\underbrace{\mathbf{1}_S\mathbf{1}_{[\tau,\tau_0]}(T)}_{=0}\Gamma(T)M\varphi(\ \bar{x}(T))|\mathcal{F}_\gamma] \\ &= 0,\end{aligned}$$

similarly,

$$\begin{aligned}\mathbf{1}_S E[\int_\gamma^T \Gamma(s)Q(s)\varphi(\mathbf{1}_{[\tau,\tau_0]}(s)\bar{x}(s))ds|\mathcal{F}_\gamma] &= E[\mathbf{1}_S\int_\gamma^T \Gamma(s)Q(s)\varphi(\mathbf{1}_{[\tau,\tau_0]}(s)\bar{x}(s))ds|\mathcal{F}_\gamma] \\ &= E[\int_\gamma^T \Gamma(s)Q(s)\varphi(\mathbf{1}_S\mathbf{1}_{[\tau,\tau_0]}(s)\bar{x}(s))ds|\mathcal{F}_\gamma] \\ &= 0,\end{aligned}$$

thus,

$$\begin{aligned}\bar{y}(\gamma) &= \Gamma(\gamma)^{-1}E[\Gamma(T)M\varphi(\bar{x}_0(T)) + \int_\gamma^T \Gamma(s)Q(s)\varphi(\bar{x}_0(s))ds|\mathcal{F}_\gamma] \\ &= 0,\ P-a.s. \text{ on } \{\tau_0 < \gamma\}.\end{aligned}$$

Finally, we consider the "diffusion"-part z. Since $\bar{y} = \bar{y}_0$, we can deduce that $(\bar{y}_0, \bar{z})$ is the unique solution of the BSDE (3.12) corresponding to the coefficients of the problem $\mathcal{P}(\tau, h)$. Let $\gamma \ge \tau$ again be a stopping time. Itô's formula, applied to $\Gamma\bar{y}_0$, then yields

$$\sum_{i=1}^d \int_\tau^\gamma \Gamma\bar{z}^i dw^i = \Gamma(\gamma)\bar{y}_0(\gamma) - \bar{y}_0(\tau) + \int_\tau^\gamma \Gamma Q\varphi(\bar{x}_0)ds - \sum_{i=1}^d \int_\tau^\gamma \Gamma C^i\bar{y}_0 dw^i.$$

On the right hand side of this equality, each term that depends on γ contains $\bar{y}_0$ or $\bar{x}_0$, hence $\sum_{i=1}^d \int_{\tau_0}^\gamma \Gamma\bar{z}^i dw^i = 0$ for all γ with $\gamma > \tau_0$. Since $\Gamma \neq 0$, this implies that $\bar{z} = 0$ on $(\tau_0, T]$, i.e. $\bar{z} = \bar{z}_0$. ■

Yet, this lemma is redundant to the extent that $\bar{x}$ never reaches 0, except on the set $\{h = 0\}$ of the zeros of the initial value. This is due to the fact that the optimal state will turn out to be a stochastic exponential. However, the previous lemma will be useful in proving this exponential representation for the optimal state.
The next lemma relates the initial value of $\bar{y}$ to the optimal cost $J(\bar{u})$.

Lemma 3.9 *If $\bar{u}$ is optimal for problem $\mathcal{P}(\tau, h)$, then*

$$J(\bar{u}) = \frac{1}{q} E[\bar{y}(\tau)h],$$

where $\bar{y}$ is the adjoint process from the corresponding solution of (3.11) - (3.14).

Proof: Let $\bar{x}$ be an optimal state for problem $\mathcal{P}(\tau, h)$. Note that $Q\varphi(\bar{x})\bar{x} = Q|\bar{x}|^q$, $M\varphi(\bar{x}(T))\bar{x}(T) = M|\bar{x}(T)|^q$. Applying Itô's formula to $\bar{y}\bar{x}$ and integrating yields

$$\begin{aligned} E[\bar{y}(T)\bar{x}(T) - \bar{y}(\tau)\bar{x}(\tau)] &= E[M|\bar{x}(T)|^q - \bar{y}(\tau)h] \\ &= E[\int_\tau^T -Q|\bar{x}|^q + \bar{u}'\left(B'y + \sum_{i=1}^d (D^i)'z^i\right) ds] \\ &= E[\int_\tau^T -Q|\bar{x}|^q - \bar{u}'N\varphi(\bar{u})ds] \\ &= E[\int_\tau^T -Q|\bar{x}|^q - N|\bar{u}|^q], \end{aligned}$$

where in the third line we used the auxiliary condition (3.14). Rearranging the last equation (as an equality of the right hand side of the first and the last line) yields the assertion. ■

The next proposition may be viewed as one of the central observations of this work. Almost all the statements of this section are similar in the quadratic case $q = 2$. Concerning the dependence of $\bar{y}$ on h, there's a difference. In the quadratic case, $h \mapsto \bar{y}$ is linear and bounded. It turns out that in the "general" case, $h \mapsto f(\bar{y})$ still is linear and bounded. It may not be seen why this is of particular importance. This will become clear in the next chapter. For the moment, we will give a very rough outline of the opportunities this linearity offers: the mapping $h \mapsto f(\bar{y}(\tau))$ will be continuous, too, and this will imply the existence of a random variable $K(\tau)$ such that $f(\bar{y}(\tau)) = K(\tau)h = K(\tau)\bar{x}(\tau)$. Varying over the initial time will yield a family of random variables (indeed, a process) $(K(t \vee \tau))_{t \le T}$, that allows us to represent the optimal state in terms of the adjoint process and vice-versa by

$$f(\bar{y}(t \vee \tau)) = K(t \vee \tau)\bar{x}(t \vee \tau), \quad t \le T.$$

From this representation it is clear that K possesses a differential at least on the stochastic interval $[\tau, \tau_0)$ (where τ_0 is the first vanishing time for $\bar{x}$). Further, K is part of the solution of a BSDE, namely the Riccati-equation. This equation will only depend on the coefficients

of the problem $\mathcal{P}(\tau,h)$, i.e. it does not involve $\bar{x},\bar{u},\bar{y}$ or $\bar{z}$ - but its solution (K,L) will enable us to explicitly construct $\bar{x}$ and $\bar{u}$.
The differential of K will be found simply by differentiating $\frac{f(\bar{y})}{\bar{x}}$ (of course we must take care about the zeros of $\bar{x}$ - this is a crucial point). The differential of this quotient may be considered independently and without any knowledge that the quotient represents some particular process K. But it will actually be of great importance for us that we can derive some properties of K (essentially concerning boundedness and positivity) that could not be proved easily for $\frac{f(\bar{y})}{\bar{x}}$.
For convenience let us introduce the following notation.

Notation 3.10 *(solution of the FBSDE depending on τ and h) For a stopping time $\tau < T$ and an $h \in L^q_{\mathcal{F}_\tau}(\mathbb{R})$ we denote by*

$$(\bar{x}^{\tau,h}, \bar{u}^{\tau,h}, \bar{y}^{\tau,h}, \bar{z}^{\tau,h})$$

the solution of (3.11)-(3.14) corresponding to problem $\mathcal{P}(\tau,h)$, given that this FBSDE has a unique solution.

With this notation, we have

Proposition 3.11 *Let one of the Assumptions A1-A3 hold. Then, for a given stopping time $\tau < T$ and a given $h_0 \in L^q_{\mathcal{F}_\tau}(\mathbb{R})$ we have for all $b \in L^\infty_{\mathcal{F}_\tau}(\mathbb{R})$*

$$(\bar{x}^{\tau,bh_0}, \bar{u}^{\tau,bh_0}, f(\bar{y}^{\tau,bh_0})) = b(\bar{x}^{\tau,h_0}, \bar{u}^{\tau,h_0}, f(\bar{y}^{\tau,h_0})).$$

In particular, the mapping

$$\begin{array}{rcl} L^q_{\mathcal{F}_\tau}(\mathbb{R}) & \longrightarrow & L^q_{\mathcal{F}}(\Omega, C([\tau,T];\mathbb{R})) \times \mathcal{U} \times L^q_{\mathcal{F}}(\Omega, C([\tau,T];\mathbb{R})), \\ h & \mapsto & (\bar{x}^{\tau,h}, \bar{u}^{\tau,h}, f(\bar{y}^{\tau,h})), \end{array}$$

is linear.

Proof: We start with the first assertion. Set $(\bar{x},\bar{u},\bar{y},\bar{z}) := (\bar{x}^{\tau,h_0}, \bar{u}^{\tau,h_0}, \bar{y}^{\tau,h_0}, \bar{z}^{\tau,h_0})$ and for $b \in L^\infty_{\mathcal{F}_\tau}(\mathbb{R})$ define $(x_0,u_0,f(y_0)) := b\,(\bar{x},\bar{u},f(\bar{y}))$, in particular $y_0 = \varphi(b)\bar{y}$. Additionally, set $z_0 = \varphi(b)\bar{z}$. We will check that (x_0,u_0,y_0,z_0) satisfies (3.11)-(3.14). Note that these processes inherit their integrability from $\bar{x}$, etc., due to the essential boundedness of b. We use the fact that (stochastic and deterministic) integration, starting with τ, commutes with multiplication with an $\mathcal{F}_\tau$-measurable random variable. Thus, from (3.11) we get for all stopping times $\gamma \geq \tau$

$$\begin{aligned} b\,\bar{x}(\gamma) &= b\bar{x}(\tau) + b\int_\tau^\gamma A\bar{x} + B\bar{u}ds + \sum_{i=1}^d b\int_\tau^\gamma \left\{C^i\bar{x} + D^i\bar{u}\right\} dw^i \\ &= b\bar{x}(\tau) + \int_\tau^\gamma A(b\bar{x}) + B(b\bar{u})ds + \sum_{i=1}^d \int_\tau^\gamma \left\{C^i(b\bar{x}) + D^i(b\bar{u})\right\} dw^i, \end{aligned}$$

and

$$\begin{aligned}\varphi(b)\bar{y}(\gamma) &= \varphi(b)\bar{y}(\tau) + \varphi(b)\int_\tau^\gamma A\bar{y} - \sum_{i=1}^d C^i\bar{z} - Q\varphi(\bar{x})ds + \sum_{i=1}^d \varphi(b)\int_\tau^\gamma \bar{z}^i dw^i \\ &= \varphi(b)\bar{y}(\tau) + \int_\tau^\gamma A(\varphi(b)\bar{y}) - \sum_{i=1}^d C^i(\varphi(b)\bar{z}) - Q\varphi(b)\varphi(\bar{x})ds \\ &\quad + \sum_{i=1}^d \int_\tau^\gamma (\varphi(b)\bar{z}^i)dw^i.\end{aligned}$$

Since $\varphi(b)\varphi(\bar{x}) = \varphi(b\bar{x})$, this implies that (x_0, u_0, y_0, z_0) satisfies the differential equations (3.11) and (3.12). We have $y_0(T) = \varphi(b)\bar{y}(T) = \varphi(b)M\varphi(\bar{x}) = M\varphi(b\bar{x}) = M\varphi(x_0)$, hence the terminal condition in (3.13) holds, and the initial condition is obvious. Using the compatibility of φ with multiplication again shows that the auxiliary condition (3.14) also holds for u_0, y_0 and z_0. Hence, the first assertion of the proposition is proved.
For the assertion on linearity choose $h_1, h_2 \in L^q_{\mathcal{F}_\tau}(\mathbb{R})$ and set $H = 1 + |h_1| + |h_2| \in L^q_{\mathcal{F}_\tau}(\mathbb{R})$. Define $b_i := \frac{h_i}{H}$, $i = 1, 2$. The b_i are essentially bounded, $\mathcal{F}_\tau$-measurable, and we have $h_1 + h_2 = (b_1 + b_2)H = b_1 H + b_2 H$. Thus,

$$\begin{aligned}&(\bar{x}^{\tau,h_1+h_2}, \bar{u}^{\tau,h_1+h_2}, f(\bar{y}^{\tau,h_1+h_2})) \\ &= (\bar{x}^{\tau,(b_1+b_2)H}, \bar{u}^{\tau,(b_1+b_2)H}, f(\bar{y}^{\tau,(b_1+b_2)H})) \\ &= (b_1 + b_2)(\bar{x}^{\tau,H}, \bar{u}^{\tau,H}, f(\bar{y}^{\tau,H})) \\ &= b_1(\bar{x}^{\tau,H}, \bar{u}^{\tau,H}, f(\bar{y}^{\tau,H})) + b_2(\bar{x}^{\tau,H}, \bar{u}^{\tau,H}, f(\bar{y}^{\tau,H})) \\ &= (\bar{x}^{\tau,b_1H}, \bar{u}^{\tau,b_1H}, f(\bar{y}^{\tau,b_1H})) + (\bar{x}^{\tau,b_2H}, \bar{u}^{\tau,b_2H}, f(\bar{y}^{\tau,b_2H})) \\ &= (\bar{x}^{\tau,h_1}, \bar{u}^{\tau,h_1}, f(\bar{y}^{\tau,h_1})) + (\bar{x}^{\tau,h_2}, \bar{u}^{\tau,h_2}, f(\bar{y}^{\tau,h_2})).\end{aligned}$$

Finally, for $\alpha \in \mathbb{R}$, $\alpha(\bar{x}^{\tau,h}, \bar{u}^{\tau,h}, f(\bar{y}^{\tau,h})) = (\bar{x}^{\tau,\alpha h}, \bar{u}^{\tau,\alpha h}, f(\bar{y}^{\tau,\alpha h}))$ follows from the first assertion, and linearity is shown. ■

In the above proposition we did not explicitly mention what happens to z if h varies. However, from the proof it is clear, that $\bar{z}^{\tau,bh_0} = \varphi(b)\bar{z}^{\tau,h_0}$ (where we used the notation of the proof). Yet, for $q < 2$ the process $f(\bar{z}^{\tau,h})$ does not in general belong to $H_q(\tau, T; \mathbb{R}^m)$, and a statement about $\bar{z}^{\tau,h}$ may be misleading in this respect. Nevertheless, it is correct that $h \mapsto f(\bar{z}^{\tau,h})$ is linear as a mapping from $L^q_{\mathcal{F}_\tau}(\mathbb{R})$ to the linear space of all d-dimensional adapted processes.

In the next lemma, the linearity eases the proof of continuous dependence of $(\bar{x}^{\tau,h}, \bar{u}^{\tau,h})$ on h.

Lemma 3.12 *Under each of the the Assumptions A1-A3, the mapping*

$$\begin{aligned}L^q_{\mathcal{F}_\tau}(\mathbb{R}) &\longrightarrow L^q_{\mathcal{F}}(\Omega, C([\tau, T]; \mathbb{R})) \times \mathcal{U} \times L^{q'}_{\mathcal{F}}(\Omega, C([\tau, T]; \mathbb{R})) \times H_{q'}(\tau, T; \mathbb{R}^d), \\ h &\mapsto (\bar{x}^{\tau,h}, \bar{u}^{\tau,h}, \bar{y}^{\tau,h}, \bar{z}^{\tau,h}),\end{aligned}$$

is continuous. The optimal cost depends continuously on the initial value, i.e. $L^q_{\mathcal{F}_\tau}(\mathbb{R}) \longrightarrow \mathbb{R}$, $h \mapsto J(\bar{u}^{\tau,h})$*, is continuous.*

Proof: First, let us show that

$$h \mapsto (\bar{x}^{\tau,h}, \bar{u}^{\tau,h})$$

is continuous. By linearity, it suffices to show continuity in $h = 0$. Let $(h_n)_n$ be a sequence in $L^q_{\mathcal{F}_\tau}(\mathbb{R})$ with $h_n \to 0$, $n \to \infty$. Set $(\bar{x}_n, \bar{u}_n) := (\bar{x}^{\tau,h_n}, \bar{u}^{\tau,h_n})$. Plugging the control $u = 0$ (with initial value h_n) in J yields $J(\bar{u}^{\tau,h_n}) \leq J(0) \to 0$, $n \to \infty$, by Corollary 2.4. The next step depends on the assumption that is in force.

If Assumption A2 holds, the process N is uniformly positive, hence $J(\bar{u}_n) \to 0$ implies $\|\bar{u}_n\|_{L^q_{\mathcal{F}}} \to 0$, $n \to \infty$. From Corollary 2.4 it follows that $\bar{x}_n \to 0$, $n \to \infty$.

If Assumption A1 or A3 holds, the r.v. M is uniformly positive, hence from $J(\bar{u}_n) \to 0$ we get $E[|\bar{x}_n(T)|^q] \to 0$, $n \to \infty$. We write the SDE (3.11) again as in (3.2) as

$$dx = \{Ax + \theta'\sigma'u\}\, ds + \{xC + u'\sigma\}\, dw$$

and consider it as a BSDE as in (3.3), (3.4),

$$\begin{aligned} dR_n &= \{(A - \theta C)R_n + \theta' S_n'\}ds + S_n dw, \\ R_n(T) &= \bar{x}_n(T), \end{aligned}$$

with $R_n = \bar{x}_n$ and $S_n = \bar{x}_n C + \bar{u}_n \sigma$. From [EPQ:BSDE] it follows that $(R_n, S_n) \to 0$, $n \to \infty$, in $L^q_{\mathcal{F}}(\Omega, C([\tau,T];\mathbb{R})) \times H_q(\tau,T;\mathbb{R}^d)$. The uniform positivity of $\sigma\sigma'$ then gives, that $\bar{u}_n \to 0$ in $H_q(\tau,T;\mathbb{R}^m)$, $n \to \infty$.

So, if one of the Assumptions A1-A3 holds, the mapping $h \mapsto (\bar{x}^{\tau,h}, \bar{u}^{\tau,h})$ is continuous.

We will now have to show that $h \mapsto (\bar{y}^{\tau,h}, \bar{z}^{\tau,h})$ is also continuous. Let $(h_n)_n$ be a sequence that converges in $L^q_{\mathcal{F}_\tau}(\mathbb{R})$ to h and set $(\bar{x}_n, \bar{u}_n, \bar{y}_n, \bar{z}_n) = (\bar{x}^{\tau,h_n}, \bar{u}^{\tau,h_n}, \bar{y}^{\tau,h_n}, \bar{z}^{\tau,h_n})$ and $(\bar{x}, \bar{z}, \bar{y}, \bar{z}) =: (\bar{x}^{\tau,h}, \bar{u}^{\tau,h}, \bar{y}^{\tau,h}, \bar{z}^{\tau,h})$. We already know that $\bar{x}_n \to \bar{x}$ in $L^q_{\mathcal{F}}(\Omega, C([\tau,T];\mathbb{R}))$, hence $Q\varphi(\bar{x}_n) \to Q\varphi(\bar{x})$ in $L^{q'}_{\mathcal{F}}(\Omega, C([\tau,T];\mathbb{R}))$, and $M\varphi(\bar{x}_n(T)) \to M\varphi(\bar{x}(T))$ in $L^{q'}_{\mathcal{F}_T}(\mathbb{R})$, $n \to \infty$. It follows from [EPQ:BSDE], Prop. 5.1 (or the Closed Graph Theorem), that $(\bar{y}_n, \bar{z}_n) \to (\bar{y}, \bar{z})$ in $L^{q'}_{\mathcal{F}}(\Omega, C([\tau,T];\mathbb{R})) \times H_{q'}(\tau,T;\mathbb{R}^d)$, $n \to \infty$. The continuity of $h \mapsto J(\bar{u}^{\tau,h})$ is clear from the continuity of $h \mapsto (\bar{x}^{\tau,h}, \bar{u}^{\tau,h})$, and the lemma is proved. ∎

Given a fixed initial time τ, it turns out that we need not be overly concerned about the initial value h if we are only interested in $\bar{x}^{\tau,h}$ and $\bar{u}^{\tau,h}$. To some extent, it suffices to consider $h = 1$.

Lemma 3.13 *Let one of the Assumptions A1-A3 hold. Then, for* $h \in L^q_{\mathcal{F}_\tau}(\mathbb{R})$ *we have*

$$(\bar{x}^{\tau,h}, \bar{u}^{\tau,h}) = h(\bar{x}^{\tau,1}, \bar{u}^{\tau,1}).$$

Further, $y^{\tau,h} = \varphi(h) y^{\tau,1}$.

Proof: For $h \in L^q_{\mathcal{F}_\tau}(\mathbb{R})$, choose a sequence (h_n) with $h_n \in L^\infty_{\mathcal{F}_\tau}(\mathbb{R})$, such that $|h - h_n|_{L^q} \to 0$, $n \to \infty$. From Lemma 3.12 we have that

$$(\bar{x}^{\tau,h_n}, \bar{u}^{\tau,h_n}) \to (\bar{x}^{\tau,h}, \bar{u}^{\tau,h}),$$

$n \to \infty$, in $L^q_{\mathcal{F}}(\Omega, C([\tau,T];\mathbb{R})) \times \mathcal{U}$; from Proposition 3.11 we get $(\bar{x}^{\tau,h_n}, \bar{u}^{\tau,h_n}) = h_n(\bar{x}^{\tau,1}, \bar{u}^{\tau,1})$. There is a subsequence $(n_k)_k$ such that $P-a.s.$ $h_{n_k} \to h,\ k \to \infty$, hence we have $P-a.s.$

$$h_{n_k}\bar{x}^{\tau,1}(t \vee \tau) = \bar{x}^{\tau,h_{n_k}}(t \vee \tau) \to h\bar{x}^{\tau,1}(t \vee \tau),$$

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

$$h_{n_k}\bar{u}^{\tau,1} = \bar{u}^{\tau,h_{n_k}} \to h\bar{u}^{\tau,1},$$

$Leb \otimes P - a.s.$, $k \to \infty$. Besides, the norm-convergence $(\bar{x}^{\tau,h_{n_k}}, ^{\tau,h_{n_k}}) \to (\bar{x}^{\tau,h}, \bar{u}^{\tau,h})$, $k \to \infty$, in $L^q_{\mathcal{F}}(\Omega, C([\tau,T];\mathbb{R})) \times \mathcal{U}$ implies the existence of a sub-sub-sequence $(n_{k_j})_j$ such that $P-a.s.$

$$\bar{x}^{\tau,h_{n_{k_j}}}(t \vee \tau) \to \bar{x}^{\tau,h}(t \vee \tau),$$

for all $t \in [0,T]$ and $\bar{u}^{\tau,h_{n_{k_j}}} \to \bar{u}^{\tau,h}$, $Leb \otimes P - a.s.$, $j \to \infty$. As the limits are unique, this now implies that $P-a.s.$

$$\bar{x}^{\tau,h}(t \vee \tau) = h\bar{x}^{\tau,1}(t \vee \tau),$$

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

$$\bar{u}^{\tau,h} = h\bar{u}^{\tau,1}$$

$Leb \otimes P - a.s.$, which proves the first assertion of the lemma.
Now recall the representation (3.15) for the adjoint process from Lemma 3.8. Using $x^{\tau,h} = hx^{\tau,1}$, this representation yields for every stopping time γ with $\tau \le \gamma \le T$

$$\begin{aligned}
&\bar{y}^{\tau,h}(\gamma)\\
&= \Gamma(\gamma)^{-1}E[\Gamma(T)M\varphi(\bar{x}^{\tau,h}(T)) + \int_\gamma^T \Gamma(s)Q(s)\varphi(\bar{x}^{\tau,h}(s))ds|\mathcal{F}_\gamma]\\
&= \Gamma(\gamma)^{-1}E[\Gamma(T)M\varphi(h)\varphi(\bar{x}^{\tau,1}(T)) + \int_\gamma^T \Gamma(s)Q(s)\varphi(h)\varphi(\bar{x}^{\tau,1}(s))ds|\mathcal{F}_\gamma]\\
&= \varphi(h)\Gamma(\gamma)^{-1}E[\Gamma(T)M\varphi(\bar{x}^{\tau,1}(T)) + \int_\gamma^T \Gamma(s)Q(s)\varphi(\bar{x}^{\tau,1}(s))ds|\mathcal{F}_\gamma].
\end{aligned}$$

In the fourth line we used the fact that h is $\mathcal{F}_\gamma-$measurable. Reading the representation (3.15) "from the right to the left" now shows that

$$\bar{y}^{\tau,1}(\gamma) = \Gamma(\gamma)^{-1}E[\Gamma(T)M\varphi(\bar{x}^{\tau,1}(T)) + \int_\gamma^T \Gamma Q\varphi(\bar{x}^{\tau,1})ds|\mathcal{F}_\gamma],$$

hence $y^{\tau,h} = \varphi(h)\bar{y}^{\tau,1}$, as required. ■

Remarks on Chapter 3 This chapter was very much in the spirit of [B:LQOC]. There,

in the quadratic case $q = 2$, the solution of problem $\mathcal{P}(\tau, h)$ is also characterized in terms of a FBSDE. Bismut himself states that his method is taken from the theory of controlled partial differential equations, especially from [L:OCSG], Chapter 4. When the above cited article of Bismut was published, there was not much known about BSDEs, not even linear BSDEs; hence Bismut had to develop the necessary tools on his own - with the goal of deriving another BSDE that is very much harder to treat than linear BSDEs.
Nowadays, there's a vast literature on BSDEs, and the situation of "Lipschitz-drivers" is covered very thoroughly. In this chapter, one of the benefits we extracted from this "ready-to-use" theory is the assertion that Assumptions A1, A3 are sufficient to guarantee the existence of an optimal control. This type of condition (i.e. uniform positivity of $\sigma\sigma'$ and M) is well known in quadratic theory, see for example [CLZ: SLQ]. In this article, one also finds examples where the quadratic problem is well posed even if the control weighting process is indefinite. In [KT:GAS], the global solvability of the Riccati equation (1.7), (1.8) is shown to be solvable under Assumption A1 for $q = 2$.
The statement that the solution of a problem $\mathcal{P}(\tau, h)$ can be characterized in terms of the FBSDE (3.11) -(3.14) is not surprising. It may be more interesting to note that the isoelastic cost functions allow for an explicit representation of the optimal cost in terms of the adjoint process, as stated in Lemma 3.9. Though this observation is immediate, it seems to be new. Also, the assertion on the linear and continuous dependence $h \mapsto (x^{\tau,h}, u^{\tau,h}, f(y^{\tau,h}))$ seems to be new. The proof of this fact relies heavily on the special form of the cost functions. As the reader may notice, the last two results mentioned were not really hard to produce, although, they will turn out to be crucial (especially the linearity of $h \mapsto f(y^{\tau,h})$) in our attempt to find and solve a Riccati-equation for problem $\mathcal{P}(\tau, h)$.

Chapter 4

The Riccati equation

In the introduction, we announced that we wanted to solve problems of type $\mathcal{P}(\tau, h)$ with the help of Backward Stochastic Riccati Differential Equations. But so far, we have not yet defined this equation (except in the Introduction). The reason is that we need some implicitly defined function G for establishing the BSRDE. We will investigate this function G in the third section of this chapter.
In the first section we will make an initial attempt to decouple the FBSDE (3.11)-(3.14). This is achieved by introducing a process (in the first instance rather a family of random variables) K that allows us to represent the adjoint process $\bar{y}$ in terms of the optimal state $\bar{x}$. Actually, we will have $f(\bar{y}) = K\bar{x}$. This K will be the first part of the solution (K, L) of the Riccati-equation. In order to derive the BSRDE, we will consider the differential of $\frac{f(\bar{y})}{\bar{x}}$ in Section 4.2.
It will turn out that we can get quite strong a-priori knowledge about the boundedness of K. In Section 4.4 we will investigate what we can deduce about the second component L of a solution (K, L) of the Riccati-equation - given that first component K actually satisfies the mentioned boundedness condition.

4.1 A feedback representation for the adjoint process

One of our principal aims is to "decouple" the FBSDE (3.11)- (3.14). Basically, this means that we want to specify a stochastic process that follows some (backward) stochastic differential equation and that allows us to construct the optimal control $\bar{u}$. The specification of this decoupling process should be "exogenous" in the sense that it does not make use of the optimal state or the optimal control - the quantities we are actually looking for. Thus, the specification, i.e. the stochastic differential equation, is expected to involve only the coefficients of the problem $\mathcal{P}(\tau, h)$, but not $\bar{x}$ or $\bar{u}$. The next proposition is a first step in this direction, although there is no concern about the optimal control. The proposition introduces the family of r.v. K that allows the construction of $\bar{y}$ (the adjoint process) from $\bar{x}$ (the optimal state) by $f(\bar{y}) = K\bar{x}$. For non-vanishing $\bar{x}$ this of course yields $K = \frac{f(\bar{y})}{\bar{x}}$,

and differentiating this fraction will make us suspect (in Section 4.2) that K is a candidate for an "exogenously definable object" that is strongly related to the solution of problem $\mathcal{P}(\tau, h)$.
The techniques of the following proposition are those from [B:LQOC], Proposition 4.2. There, in the quadratic case ($q = 2$) we have $\bar{y} = K\bar{x}$. The key insight that allows us to make use of these techniques in the isoelastic case is the linear and continuous dependence of $f(\bar{y}^{\tau,h})$ on h.

Proposition 4.1 *Let one of the Assumptions A1-A3 hold and let τ be a stopping time with $0 \leq \tau < T$. Then, for every $t \in [0, T]$ there is a $P - a.s.$-unique $K(t \vee \tau) \in L^{\infty}_{\mathcal{F}_{t\vee\tau}}(\mathbb{R})$ such that*

$$f(\bar{y}^{\tau,h}(t \vee \tau)) = K(t \vee \tau)\, \bar{x}^{\tau,h}(t \vee \tau) \tag{4.1}$$

for every $h \in L^q_{\mathcal{F}_\tau}(\mathbb{R})$.

Proof: We start by considering $t = 0$, i.e. we are looking for a $K(\tau)$ such that $f(\bar{y}^{\tau,h}(\tau)) = K(\tau)\bar{x}^{\tau,h}(\tau) = K(\tau)h$. From Proposition 3.11, we know that the application $L^q_{\mathcal{F}_\tau}(\mathbb{R}) \longrightarrow L^q_{\mathcal{F}_\tau}(\mathbb{R})$, $h \mapsto f(\bar{y}^{\tau,h}(\tau))$ is linear; from Lemma 3.12 we see that it is also continuous. Let A, B be two disjoint sets in $\mathcal{F}_\tau$. As $\mathbf{1}_{A\cup B} = \mathbf{1}_A + \mathbf{1}_B$, linearity entails that $f(\bar{y}^{\tau,\mathbf{1}_{A\cup B}})(\tau) = f(\bar{y}^{\tau,\mathbf{1}_A})(\tau) + f(\bar{y}^{\tau,\mathbf{1}_B})(\tau)$, hence the mapping $\mathcal{F}_\tau \longrightarrow \mathbb{R}$, $A \mapsto E[f\,(\bar{y}^{\tau,\mathbf{1}_A}(\tau))]$ is a signed measure. It is clear that sets with $P(A) = 0$ are mapped to 0. By the Radon-Nikodym-Theorem, there is a $P - a.s.$-unique, integrable, $\mathcal{F}_\tau$-measurable r.v. $K(\tau)$ such that $E[f\,(\bar{y}^{\tau,\mathbf{1}_A}(\tau))] = E[\mathbf{1}_A K(\tau)]$, compare [H:MT], Sect. 31, Thm B. As $f\,(\bar{y}^{\tau,\mathbf{1}_A})\,(\tau) = \mathbf{1}_A f\,(\bar{y}^{\tau,1})\,(\tau)$, see Proposition 3.11, we get $E[\mathbf{1}_A\,(f\,(\bar{y}^{\tau,1}(\tau)) - K(\tau))] = 0$ for all $A \in \mathcal{F}_\tau$. This entails $K(\tau) = f\,(\bar{y}^{\tau,1}(\tau))$, $P - a.s.$. By Proposition 3.11 again, this gives $hf\,(\bar{y}^{\tau,1}(\tau)) = f\left(\bar{y}^{\tau,h}(\tau)\right) = K(\tau)h$ if $h = \mathbf{1}_S$, $S \in \mathcal{F}_\tau$. By linearity, this immediately extends to initial values h that are finitely valued, simple r.v., i.e. to the initial values $h = \sum_{j=1}^{k} \alpha_j \mathbf{1}_{S_j}$, where the S_j are disjoint, $\mathcal{F}_\tau$-measurable sets, and $\alpha_j \in \mathbb{R}$.
As $h \mapsto f(\bar{y}^{\tau,h})$ is linear and continuous, there is a $k > 0$ such that $\left|f(\bar{y}^{\tau,h}(\tau))\right|^q_{L^q} \leq k\,|h|^q_{L^q}$ for all $h \in L^q_{\mathcal{F}_\tau}(\mathbb{R})$. Consequently, for all simple, finitely valued and $\mathcal{F}_\tau$-measurable h we have $E[|f(\bar{y}^{\tau,h}(\tau))|^q] = E[|K(\tau)|^q|h|^q] \leq kE[|h|^q]$; this entails that $K(\tau)$ is essentially bounded. Hence, the mapping $L^q_{\mathcal{F}_\tau}(\mathbb{R}) \longrightarrow L^q_{\mathcal{F}_\tau}(\mathbb{R})$, $h \mapsto K(\tau)h$ is well defined, linear and continuous. It coincides with $h \mapsto f(\bar{y}^{\tau,h}(\tau))$ for simple r.vs. h, and as the simple r.vs. are dense in $L^q_{\mathcal{F}_\tau}(\mathbb{R})$, it follows that $f(\bar{y}^{\tau,h}(\tau)) = K(\tau)\,h$ for all $h \in L^q_{\mathcal{F}_\tau}(\mathbb{R})$.
Now recall Definition 1.3 and consider the family of problems $\mathcal{P}(t\vee\tau, h_{t\vee\tau})$, $t \in [0, T)$, whose coefficients are those of problem $\mathcal{P}(\tau, h)$, restricted to the subinterval $[t \vee \tau, T]$, i.e the coefficients of $\mathcal{P}(t \vee \tau, h_{t\vee\tau})$ are $A|_{[t\vee\tau,T]}, B|_{[t\vee\tau,T]}, \ldots, N|_{[t\vee\tau,T]}$ and M. If Assumption A1, A2 or A3 holds for problem $\mathcal{P}(\tau, h)$, this is also true for problem $\mathcal{P}(t \vee \tau, h_{t\vee\tau})$. With the same procedure as above, we can now construct a $P - a.s.$-unique, essentially bounded, $\mathcal{F}_{t\vee\tau}$-measurable r.v. $K(t\vee\tau)$ such that $f(\bar{y}^{t\vee\tau,h_{t\vee\tau}}(t\vee\tau)) = K(t\vee\tau)h_{t\vee\tau} = K(t\vee\tau)\bar{x}^{t\vee\tau,h_{t\vee\tau}}(t\vee\tau)$ for all $h_{t\vee\tau} \in L^q_{\mathcal{F}_{t\vee\tau}}(\mathbb{R})$.
In this way we have constructed a family of r.v. $K(t \vee \tau)$, $t \in [0, T)$, to which we add $K(T) := f(M)$; let us show that this family indeed satisfies (4.1). For $t = T$ this is clear from the terminal condition on $\bar{y}^{\tau,h}$. Now fix some arbitrary $h \in L^q_{\mathcal{F}_\tau}(\mathbb{R})$ and consider, as

above, for $t \in [0,T)$ the problem $\mathcal{P}(t \vee \tau, \bar{x}^{\tau,h}(t \vee \tau))$, where we take the optimal state of problem $\mathcal{P}(\tau, h)$ at time $t \vee \tau$ as the initial value of a new problem with initial time $t \vee \tau$. Set $h_{t\vee\tau} := \bar{x}^{\tau,h}(t \vee \tau)$. From the uniqueness of the optimal control it is clear that the solution of these problems, arising from a "premature halt" of $\bar{x}^{\tau,h}$, are the restriction of $\bar{x}^{\tau,h}$, $\bar{u}^{\tau,h}$ to $[t \vee \tau, T]$, $\bar{x}^{t\vee\tau,h_{t\vee\tau}} = \bar{x}^{\tau,h}|_{[t\vee\tau,T]}$, $\bar{u}^{t\vee\tau,h_{t\vee\tau}} = \bar{u}^{\tau,h}|_{[t\vee\tau,T]}$. Clearly, $\bar{y}^{\tau,h}|_{[t\vee\tau,T]}$ then satisfies (3.12)-(3.14) of the FBSDE for problem $\mathcal{P}(t \vee \tau, h_{t\vee\tau})$, thus $\bar{y}^{t\vee\tau,h_{t\vee\tau}} = \bar{y}^{\tau,h}|_{[t\vee\tau,T]}$. Hence, from the construction of $K(\cdot \vee \tau)$ we get for all $t \in [0,T]$, that

$$f(\bar{y}^{\tau,h}(t \vee \tau)) = f(\bar{y}^{t\vee\tau,h_{t\vee\tau}}(t \vee \tau)) = K(t \vee \tau)h_{t\vee\tau},$$

which is the desired representation, and the proposition is shown. ■

The independence of K on h reflects what we have seen in Lemma 3.13, i.e. the fact that $\bar{x}^{\tau,h} = h\bar{x}^{\tau,1}$ and $f(\bar{y}^{\tau,h}) = hf(\bar{y}^{\tau,1})$. So far, we do not know if the "family of r.vs." $K(t \vee \tau)$, $t \in [0,T]$, is actually a stochastic process, i.e. we do not know if there is any measurability with respect to t. However, if $\tau_0 := \inf\{t \in [\tau, T] : \bar{x}^{\tau,1}(t) = 0\} \wedge T$, then we have $K(t \vee \tau) = \frac{f(\bar{y}^{\tau,1}(t\vee\tau))}{\bar{x}^{\tau,1}(t\vee\tau)}$ on the stochastic interval $[\tau, \tau_0)$, and due to the continuity of paths of $\bar{x}^{\tau,1}, \bar{y}^{\tau,1}$, at least the restriction $K : [\tau, \tau_0) \longrightarrow \mathbb{R}$ is a continuous stochastic process (note that $\tau < \tau_0$).
Next to this t-measurability/continuity of the restricted family K, there are two further properties of K that will help us (and we are going to show next): its uniform boundedness and its positivity. To get the latter in the form we will need, unfortunately we will have to strengthen Assumption 2. The boundedness property is readily available:

Lemma 4.2 *Let one of the Assumptions A1, A2 or A3 hold. Then the family $K(t \vee \tau)$, $t \in [0,T]$, of Proposition 4.1 is uniformly essentially bounded, i.e. there is some $k > 0$ such that*

$$\sup\{|K(t \vee \tau)|_{L^\infty} : \ t \in [0,T]\} \leq k.$$

Proof: Recall the representation of the optimal cost from Lemma 3.9, $J(\bar{u}^{\tau,h}) = \frac{1}{q}E[\bar{y}^{\tau,h}(\tau)h]$. Applying this to the problems $\mathcal{P}(t \vee \tau, h_{t\vee\tau})$, using Proposition 4.1 and the optimality of $\bar{u}^{t\vee\tau,h_{t\vee\tau}}$, yields for all $t \in [0,T]$

$$\begin{aligned}
J_{t\vee\tau}(\bar{u}^{t\vee\tau,h_{t\vee\tau}}) &= \frac{1}{q}E[\bar{y}^{t\vee\tau,h_{t\vee\tau}}(t \vee \tau)h_{t\vee\tau}] \\
&= \frac{1}{q}E[\varphi(K(t \vee \tau)h_{t\vee\tau})h_{t\vee\tau}] \\
&= \frac{1}{q}E[\varphi(K(t \vee \tau))|h_{t\vee\tau}|^q] \\
&\leq J_{t\vee\tau}(0),
\end{aligned}$$

where $J_{t\vee\tau}$ denotes the cost functional for initial time $t \vee \tau$. Denote by $\tilde{x}_{t\vee\tau}^{0,h_{t\vee\tau}}$ the solution of the state equation for problem $\mathcal{P}(t \vee \tau, h_{t\vee\tau})$ (i.e. equation (1.14), (1.15) with initial time $t \vee \tau$ instead of τ) with $u = 0$. By Corollary 2.4, there is a $k_1 > 0$, *independent of* *t and* $h_{t\vee\tau}$, such that $\left\|\tilde{x}_{t\vee\tau}^{0,h_{t\vee\tau}}\right\|_{L_c^q}^q \leq k_1 \left|h_{t\vee\tau}\right|_{L^q}^q$. From the essential boundedness of Q and M there is a k_2 independent of t and h such that $J_{t\vee\tau}(0) \leq k_2 \left\|\tilde{x}_{t\vee\tau}^{0,h_{t\vee\tau}}\right\|_{L_c^q}^q$. Putting this together, we get for all $t \in [0,T]$ and all $h_{t\vee\tau} \in L^q_{\mathcal{F}_{t\vee\tau}}(\mathbb{R})$

$$\begin{aligned} \frac{1}{q}E[\varphi(K(t\vee\tau))|h_{t\vee\tau}|^q] &\leq J_{t\vee\tau}(0) \\ &\leq k_2 \left\|\tilde{x}_{t\vee\tau}^{0,h_{t\vee\tau}}\right\|_{L_c^q}^q \\ &\leq k_2 k_1 E[|h_{t\vee\tau}|^q], \end{aligned}$$

with k_1, k_2 independent of t and $h_{t\vee\tau}$. This entails the uniform boundedness of $\varphi(K(\cdot\vee\tau))$, hence that of $K(\cdot\vee\tau)$, and the lemma is proved. ∎

In the previous lemma, our main argument relied on the relationship between the optimal cost and the adjoint process (stated in Lemma 3.9) for the problems $\mathcal{P}(\tau,h)$. We want to exploit this connection again in order to prove that K is strictly positive, respectively uniformly positive (depending on the assumption in force). This positivity, respectively uniform positivity, is a technical requirement for the following. Unfortunately, under Assumption A2 the aforementioned link between the optimal cost and K shows us that $K(\tau)$ cannot always be strictly positive. Consider problem $\mathcal{P}(\tau,h)$ with, say, $N\equiv 1$, $Q=0$ and $M=0$. This matches the requirements of Assumption A2. It is clear that the optimal control is identically zero, $\bar{u}^{\tau,h}=0$, since the only contribution to the cost functional is the immediate control cost. For all $h \in L^q_{\mathcal{F}_\tau}(\mathbb{R})$, the optimal cost is zero, thus, by Lemma 3.9

$$0 = J(\bar{u}^{\tau,h}) = \frac{1}{q}E[\bar{y}^{\tau,h}(\tau)h] = \frac{1}{q}E[\varphi(K(\tau))|h|^q],$$

for all $h \in L^q_{\mathcal{F}_\tau}(\mathbb{R})$, hence $K(\tau)=0$. With the same argument applied to the problems $\mathcal{P}(t\vee\tau, h_{t\vee\tau})$, it follows that $K\equiv 0$.
This example is, of course, pathological, but it is not excluded by Assumption A2. However, the reasoning shows that strict positivity of K could be expected if the following condition holds: For all $t\in[0,T]$ and all $h_{t\vee\tau} \in L^q_{\mathcal{F}_{t\vee\tau}}(\mathbb{R})$ we have

$$J_{t\vee\tau}(\bar{u}^{t\vee\tau,h_{t\vee\tau}}) > 0 \text{ if } h_{t\vee\tau} \neq 0, \tag{4.2}$$

where $J_{t\vee\tau}$ is the cost functional and $\bar{u}^{t\vee\tau,h_{t\vee\tau}}$ is the optimal control for problem $\mathcal{P}(t\vee\tau, h_{t\vee\tau})$. The following Assumption is stronger than A2 and will imply (4.2).

Assumption A4 *In addition to Assumption A2 we have* $M>0$, $P-a.s.$.

This yields what we desire.

Remark 4.3 *Under Assumption A4, (4.2) holds.*

Proof: Suppose that there is a $h_{t\vee\tau} \in L^q_{\mathcal{F}_{t\vee\tau}}(\mathbb{R})$ such that $J(\bar{u}^{t\vee\tau,h_{t\vee\tau}}) = 0$. We have to show that $h_{t\vee\tau} = 0$. From the uniform positivity of N it follows that $\bar{u}^{t\vee\tau,h_{t\vee\tau}} = 0$, thus $\bar{x}^{t\vee\tau,h_{t\vee\tau}}$ satisfies the SDE $dx = Axds + \sum_{i=1}^d C^i x dw^i$, $x(t\vee\tau) = h_{t\vee\tau}$. Therefore $\bar{x}^{t\vee\tau,h_{t\vee\tau}}(T) = h_{t\vee\tau}\exp\{\int_{t\vee\tau}^T A - \frac{1}{2}\sum_{i=1}^d |C^i|^2 ds + \sum_{i=1}^d \int_{t\vee\tau}^T C^i dw^i\}$. Since $J(\bar{u}^{t\vee\tau,h_{t\vee\tau}}) = 0$, the expectation $E[M|\bar{x}^{t\vee\tau,h_{t\vee\tau}}(T)|^q]$ equals zero. The strict positivity of M and of the exponential part of $\bar{x}^{t\vee\tau,h_{t\vee\tau}}$ now implies $h_{t\vee\tau} = 0$, $P-a.s.$, and since $J(u) \geq 0$ for all $u \in \mathcal{U}$, the remark is proved. ■

In the case of Assumption A1 and A3, the cost functional is even uniformly positive.

Remark 4.4 *If Assumptions A1 or A3 hold, there is a $\delta > 0$ such that for all $t \in [0,t]$ and all $h_{t\vee\tau} \in L^q_{\mathcal{F}_{t\vee\tau}}(\mathbb{R})$*

$$J(\bar{u}^{t\vee\tau,h_{t\vee\tau}}) \geq \delta\, |h_{t\vee\tau}|^q_{L^q}.$$

Proof: From the representation of (1.14) as a BSDE as in (3.3) and (3.5), it follows with Prop. 5.1. in [EPQ:BSDE] that there is a $k > 0$ such that $\left\|\bar{x}^{t\vee\tau,h_{t\vee\tau}}\right\|_{L^q_{\mathcal{F}}} + \left\|C\bar{x}^{t\vee\tau,h_{t\vee\tau}} + (\bar{u}^{t\vee\tau,h_{t\vee\tau}})'\sigma\right\|_{H_q} \leq k\left|\bar{x}^{t\vee\tau,h_{t\vee\tau}}(T)\right|_{L^q}$. The constant k can be chosen independently of t and $h_{t\vee\tau}$; it depends on the the Lipschitz-constant of the "driver" of the BSDE and of T. Since $\bar{x}^{t\vee\tau,h_{t\vee\tau}}(t\vee\tau) = h_{t\vee\tau}$, this yields, using $M \gg \epsilon$ for some $\epsilon > 0$ in the third line,

$$\begin{aligned}
E[|h_{t\vee\tau}|^q] &\leq \left\|\bar{x}^{t\vee\tau,h_{t\vee\tau}}\right\|^q_{L^q_c} \\
&\leq k^q\left|\bar{x}^{t\vee\tau,h_{t\vee\tau}}\right|^q_{L^q} \\
&\leq k^q\frac{1}{\epsilon}E[M|\bar{x}^{\tau,h}(T)|^q] \\
&\leq k^q\frac{1}{\epsilon}J(\bar{u}^{t\vee\tau,h_{t\vee\tau}})
\end{aligned}$$

for all $t \in [0,T]$ and all $h_{t\vee\tau} \in L^q_{\mathcal{F}_{t\vee\tau}}(\mathbb{R})$. This proves the remark. ■

These positivity properties of the the cost functional carry over to K.

Lemma 4.5 *Let one of the Assumptions A1, A3 or A4 hold. Consider the family of r.v. $K(t\vee\tau), t \in [0,T]$, from Proposition 4.1.*

a) If Assumption A1 or A3 holds, then there is a $\delta_0 > 0$ such that $K(t\vee\tau) \geq \delta_0$, $P-a.s.$ for all $t \in [0,T]$.

b) If Assumption A4 holds, then $K(t\vee\tau) > 0$, $P-a.s.$ for all $t \in [0,T]$.

Proof: Let $t \in [0,T)$ be arbitrary. Applying Lemma 3.9 to problem $\mathcal{P}(t\vee\tau, h_{t\vee\tau})$ for some $h_{t\vee\tau} \in L^q_{\mathcal{F}_{t\vee\tau}}(\mathbb{R})$ yields

$$J(\bar{u}^{t\vee\tau,h_{t\vee\tau}}) = \frac{1}{q}E[\bar{y}^{t\vee\tau,h_{t\vee\tau}}(t\vee\tau)h_{t\vee\tau}] = \frac{1}{q}E[\varphi(K(t\vee\tau))|h_{t\vee\tau}|^q].$$

Hence, by the previous remarks, we have for all $h_{t\vee\tau} \in L^q_{\mathcal{F}_{t\vee\tau}}(\mathbb{R})$ that

$$\frac{1}{q}E[\varphi(K(t\vee\tau))|h_{t\vee\tau}|^q] > 0$$

if Assumption A4 holds, and that

$$\frac{1}{q}E[\varphi(K(t\vee\tau))|h_{t\vee\tau}|^q] > \delta E[|h_{t\vee\tau}|^q]$$

(δ independent of t and $h_{t\vee\tau}$) if Assumption A1 or A3 holds. This implies the assertions of the lemma for $t < T$. Noting $K(T) = f(M)$ completes the proof of the lemma. ■

So far in this section, we have often considered problems $\mathcal{P}(t\vee\tau, h_{t\vee\tau})$. In the sequel, we will focus again on the initial time τ. In view of Lemma 3.13 we will also restrict ourselves to the initial value $h = 1$. Recall Definition 3.7 and let us fix the following notation.

Notation 4.6 *Let τ be a stopping time with $\tau < T$ and let one of the Assumptions A1, A3 or A4 hold. Until further notice, we will use the notation*

$$x := \bar{x}^{\tau,1},\ u := \bar{u}^{\tau,1},\ y := \bar{y}^{\tau,1},\ z := \bar{z}^{\tau,1}.$$

The definition of the stopping time τ_0 introduced in Definition 3.7 then applies to the optimal state process x,

$$\tau_0 := \inf\{s \in [\tau, T] : x(s) = 0\} \wedge T.$$

K will denote the family $K(t\vee\tau)$, $t \in [0,T]$, from Proposition 4.1.

Note that, due to the continuity of x, we have $\tau < \tau_0$ $P-a.s.$.
The properties of K we encountered in this section rely on the connection between K and the optimal cost of problem $\mathcal{P}(\tau, h)$ (via the adjoint process y). The next step is to investigate the differential of K. This will ultimately lead to to the Riccati-equation for the linear isoelastic problem.

4.2 The differential of K

The heading of this section implicitly suggests that K is a semimartingale. Yet, this is far from obvious. By construction we have $f(y) = Kx$. Hence, on the stochastic interval $[\tau, \tau_0)$, the representation

$$K = \frac{f(y)}{x} \tag{4.3}$$

is available. Noting that f is twice differentiable on $\mathbb{R}\backslash\{0\}$, we could establish that K is indeed a semimartingale on $[\tau, \tau_0)$, if we knew that y does not vanish on this interval. As $y = \varphi(Kx)$ this is clear (given Assumption A1, A3 or A4) from the positivity of K and

the definition of τ_0 (note that $x(\tau) = 1$).
Then, on $[\tau, \tau_0)$, we have (denoting the derivative with a prime $'$)

$$\begin{aligned} d\left(\frac{f(y)}{x}\right) &= \frac{1}{x}d(f(y)) + d(\frac{1}{x})(f(y)) + d < \frac{1}{x}, f(y) > \\ &= \left\{-A\frac{f'(y)y}{x} - \sum_{i=1}^{d} C^i \frac{f'(y)z^i}{x} - Q\frac{f'(y)\varphi(x)}{x} + \frac{1}{2}\sum_{i=1}^{d}\frac{f''(y)(z^i)^2}{x}\right\} ds \\ &\quad + \sum_{i=1}^{d}\frac{f'(y)z^i}{x}dw^i \\ &\quad + \left\{-A\frac{f(y)}{x} - B\frac{f(y)}{x^2}u + \sum_{i=1}^{d}\frac{f(y)}{x^3}\left(C^i x + D^i u\right)^2\right\} ds \\ &\quad + \sum_{i=1}^{d}\left\{-C^i\frac{f(y)}{x} - D^i\frac{f(y)}{x^2}u\right\} dw^i \\ &\quad + \sum_{i=1}^{d}\left\{-C^i\frac{1}{x} - D^i\frac{1}{x^2}u\right\} f'(y)z^i ds. \end{aligned} \tag{4.4}$$

The rest of this and the next section is dedicated to manipulate (4.4) in such a way that it results in a Backward Stochastic Riccati Differential Equation. In other words, our aim is to eliminate the processes x, u, y, z in order to get a (backward) differential equation whose *driver* only depends on the coefficients of the problem, of K and of some process L.
The typical structure of a BSDE does not admit much choice of what L might be. The general form of these equations is $dK = Driver(K, L, s, \omega)ds + \sum_{i=1}^{d} L^i dw^i,\ K(T) = terminal\ \ value$. Thus we are led to set, on $[\tau, \tau_0)$,

$$L^i := \frac{f'(y)}{x}z^i - C^i\frac{f(y)}{x} - D^i\frac{f(y)}{x^2}u, \quad i = 1, \ldots, d. \tag{4.5}$$

Note that $f'(y) = \frac{1}{q-1}|y|^{\frac{1}{q-1}-1}$, hence $yf'(y) = \frac{1}{q-1}f(y)$. Replacing $\frac{f(y)}{x}$ by K we can write (4.4) as

$$\begin{aligned} dK &= \left\{-q'AK - \sum_{i=1}^{d} C^i\frac{f'(y)z^i}{x} - \frac{1}{q-1}Q\frac{f'(y)\varphi(x)}{x} + \frac{1}{2}\sum_{i=1}^{d}\frac{f''(y)(z^i)^2}{x}\right. \\ &\quad \left. -BK\frac{u}{x} + \sum_{i=1}^{d} K\left(C^i + D^i\frac{u}{x}\right)^2 + \sum_{i=1}^{d}\left(-C^i - D^i\frac{u}{x}\right)\frac{f'(y)z^i}{x}\right\} ds \\ &\quad + \sum_{i=1}^{d} L^i dw^i. \end{aligned} \tag{4.6}$$

Next, we have that $f'(y)\varphi(x) = \frac{1}{q-1}xK^{2-q}$, hence $Q\frac{f'(y)\varphi(x)}{x} = \frac{1}{q-1}QK^{q-2}$. We will replace the expression $\frac{f''(y)(z^i)^2}{x}$. Note that $f''(y)f(y) = (2-q)f'(y)^2$, hence

$$\frac{1}{2}\frac{f''(y)(z^i)^2}{x} = \frac{2-q}{2}\frac{1}{K}\left(\frac{f'(y)z^i}{x}\right)^2.$$

Plugging this into (4.6) yields

$$\begin{aligned} dK &= \Bigg\{-q'AK - \sum_{i=1}^d C^i\frac{f'(y)z^i}{x} - \frac{1}{q-1}QK^{2-q} \\ &\quad + \frac{2-q}{2}\frac{1}{K}\sum_{i=1}^d\left(\frac{f'(y)z^i}{x}\right)^2 \\ &\quad - BK\frac{1}{x}u + \sum_{i=1}^d K\left(C^i + D^i\frac{u}{x}\right)^2 + \sum_{i=1}^d\left(-C^i - D^i\frac{1}{x}u\right)\frac{f'(y)z^i}{x}\Bigg\}ds \\ &\quad + \sum_{i=1}^d L^i dw^i, \end{aligned} \tag{4.7}$$

and with $\frac{f'(y)z^i}{x} = L^i + KC^i + KD^i\frac{u}{x}$

$$\begin{aligned} dK &= \Bigg\{-q'AK - \sum_{i=1}^d C^i(L^i + KC^i + KD^i\frac{1}{x}u) - \frac{1}{q-1}QK^{2-q} \\ &\quad + \frac{2-q}{2}\frac{1}{K}\sum_{i=1}^d\left(L^i + KC^i + KD^i\frac{u}{x}\right)^2 - BK\frac{1}{x}u \\ &\quad + \sum_{i=1}^d K\left(C^i + D^i\frac{1}{x}u\right)^2 + \sum_{i=1}^d\left(-C^i - D^i\frac{1}{x}u\right)\left(L^i + KC^i + KD^i + \frac{1}{x}u\right)\Bigg\}ds \\ &\quad + \sum_{i=1}^d L^i dw^i. \end{aligned}$$

This simplifies to

$$\begin{aligned} dK &= \Bigg\{-q'AK - \sum_{i=1}^d (C^i)^2K - 2\sum_{i=1}^d C^iL^i - \left(BK + \sum_{i=1}^d D^i\left(C^iK + L^i\right)\right)\frac{1}{x}u \\ &\quad - \frac{1}{q-1}QK^{2-q} + \frac{2-q}{2}\frac{1}{K}\sum_{i=1}^d\left(L^i + KC^i + KD^i\frac{1}{x}u\right)^2\Bigg\}ds + \sum_{i=1}^d L^i dw^i. \end{aligned} \tag{4.8}$$

Note that all the above calculations are justified on the stochastic interval $[\tau, \tau_0)$ (since $x, y, K > 0$ on this interval).
In order to get a stochastic differential equation where the unknown processes are solely K and L, we must eliminate the expression $\frac{1}{x}u$. Note that in the quadratic case $q = 2$ this is not hard. There, we have $L^i = \frac{1}{x}z^i - KC^i - KD^i\frac{1}{x}u$, hence $z^i = x\left(L^i + KC^i + KD^i\frac{1}{x}u\right)$. Substituting z^i in (3.14) then yields

$$B'y + x\sum_{i=1}^{d}(D^i)'(L^i + KC^i + KD^i\frac{1}{x}u) \ + \ Nu \ = \ 0,$$

and dividing by x

$$KB' + \sum_{i=1}^{d}(D^i)'(L^i + KC^i + KD^i\frac{1}{x}u) \ + \ N\frac{1}{x}u \ = \ 0,$$

hence

$$\frac{1}{x}u = -\left(N + K\sum_{i=1}^{d}(D^i)'D^i\right)^{-1}\left(KB' + \sum_{i=1}^{d}(D^i)'(L^i + KC^i)\right),$$

at least on $[\tau, \tau_0)$.
In the non-quadratic case, the auxiliary condition (3.14) does not reduce to a linear equation. The representation of $\frac{1}{x}u$ in terms of K, L and the coefficients of the problem will still be possible, but it will involve an implicitly defined function G. In the next section we will introduce this function. Let us prepare for this by rephrasing the auxiliary condition. The subsequent calculations hold $Leb \otimes P - a.s.$ on $[\tau, \tau_0)$. From (4.5) we get

$$z^i = \frac{x}{f'(y)}\left(L^i + KC^i + KD^i\frac{1}{x}u\right).$$

Let us replace z^i in (3.14) by this expression. This yields

$$B'y + \sum_{i=1}^{d}(D^i)'\frac{x}{f'(y)}\left(L^i + KC^i + KD^i\frac{1}{x}u\right) \ + \ N\varphi(u) = 0.$$

Dividing by $\varphi(x)$ leads to

$$B'\frac{y}{\varphi(x)} + \sum_{i=1}^{d}(D^i)'\frac{x}{f'(y)\varphi(x)}\left(L^i + KC^i + KD^i\frac{1}{x}u\right) \ + \ N\varphi(\frac{1}{x}u) = 0.$$

Since $f'(y)\varphi(x) = \frac{1}{q-1}xK^{2-q}$ and $\frac{y}{\varphi(x)} = \varphi(K)$ we get

$$\varphi(K)B' + (q-1)\varphi(K)\sum_{i=1}^{d}(D^i)'\left(\frac{L^i}{K} + C^i + D^i\frac{1}{x}u\right) \ + \ N\varphi(\frac{1}{x}u) = 0,$$

i.e.

$$\varphi(K)B' + (q-1)\varphi(K)\sum_{i=1}^{d}(D^i)'\left(\frac{L^i}{K}+C^i\right)$$
$$+(q-1)\varphi(K)\left(\sum_{i=1}^{d}(D^i)'D^i\right)\frac{1}{x}u + N\varphi(\frac{1}{x}u) = 0, \quad Leb\otimes P - a.s., \tag{4.9}$$

on $[\tau, \tau_0)$. This last equation is the one we are going to resolve for $\frac{1}{x}u$.

4.3 Represention of $\frac{1}{x}u$, the function G

We wish to represent $\frac{1}{x}u$ in terms of $B, (C^i)_{1\le i\le d}, (D^i)_{1\le i\le d}, N, K$ and L, i.e. we are looking for a function $G = G(B, (C^i)_{1\le i\le d}, (D^i)_{1\le i\le d}, N, K, L)$ such that

$$\varphi(K)B' + (q-1)\varphi(K)\sum_{i=1}^{d}(D^i)'\left(\frac{L^i}{K}+C^i\right)$$
$$+(q-1)\varphi(K)\left(\sum_{i=1}^{d}(D^i)'D^i\right)G(\cdot) + N\varphi(G(\cdot)) = 0. \tag{4.10}$$

Showing that such a function G exists is a real-analytic problem. Therefore, in this section the quantities $B, (C^i)_{1\le i\le d}, \ldots$ are meant to be real variables or vectors in $\mathbb{R}^m, \mathbb{R}^d, \ldots$, rather than stochastic processes. We hope that here no confusion arises for the reader. Of course, once G is well defined we want to plug in the stochastic processes. Hence we will choose the range of these processes as the domain of G. Here some distinction with respect to the assumption in force is necessary. First, note that under Assumption A3 things are quite easy, due to the condition $N = 0$. There we have

$$\begin{aligned} &G(B, (C^i)_{1\le i\le d}, (D^i)_{1\le i\le d}, N, K, L)\\ &= -\left[\sum_{i=1}^{d}(D^i)'D^i\right]^{-1}\left(\frac{1}{q-1}B' + \sum_{i=1}^{d}(D^i)'\left(\frac{L^i}{K}+C^i\right)\right)\\ &= -\left[K\sum_{i=1}^{d}(D^i)'D^i\right]^{-1}\left(\frac{1}{q-1}B'K + \sum_{i=1}^{d}(D^i)'\left(L^i + C^iK\right)\right). \end{aligned} \tag{4.11}$$

Thus, in the case of Assumption A3 it is clear what G is (likewise in the case of Assumption A1 if $N = 0$). We will mainly consider Assumptions A1 and A4.
To ease notation we will use the following convention: $\frac{1}{n} \le \sum_{i=1}^{d}(D^i)'D^i$ means that for all $v \in \mathbb{R}^m$ we have $\frac{1}{n}|v|^2 \le v'(\sum_{i=1}^{d}(D^i)'D^i)v$ (or, alternatively, that all eigenvalues of $\sum_{i=1}^{d}(D^i)'D^i$ are greater than or equal to $\frac{1}{n}$).

Notation 4.7 *For every $n \in \mathbb{N}_{\geq 1}$ define*

$$\begin{aligned}\mathcal{D}_n^{(1)} &:= \{(B,(C^i)_{1\leq i\leq d},(D^i)_{1\leq i\leq d},N,K,L) \in \mathbb{R}\times\mathbb{R}^d\times\mathbb{R}^{m\times d}\times\mathbb{R}_{\geq 0}\times\mathbb{R}_{>0}\times\mathbb{R}^d : \\ &\quad |B|,|C^i|,|D^i| \leq n,\ \frac{1}{n} \leq \sum_{i=1}^d (D^i)'D^i,\ 0\leq N\leq n,\ \frac{1}{n}\leq K\leq n\},\end{aligned}$$

and

$$\begin{aligned}\mathcal{D}_n^{(4)} &:= \{(B,(C^i)_{1\leq i\leq d},(D^i)_{1\leq i\leq d},N,K,L) \in \mathbb{R}\times\mathbb{R}^d\times\mathbb{R}^{m\times d}\times\mathbb{R}_{\geq 0}\times\mathbb{R}_{>0}\times\mathbb{R}^d : \\ &\quad |B|,|C^i|,|D^i| \leq n,\ \frac{1}{n} \leq N\leq n,\ \ 0< K\leq n\}.\end{aligned}$$

If Assumption A1, respectively A4, holds, then we know that there is a n such that the range of the *processes*[1] $(B,(C^i)_{1\leq i\leq d},(D^i)_{1\leq i\leq d},N,K,L)$ is in $\mathcal{D}_n^{(1)}$, respectively $\mathcal{D}_n^{(4)}$. We will encounter some statements involving constants that depend on the (upper and lower) bounds of $B,\dots,K,L$, where the vector of these variables is valued in $\mathcal{D}_n^{(1)}$ or $\mathcal{D}_n^{(4)}$. The use of the index n will allow us simply to say that the constants depend on n.
We now can state precisely what we are looking for. For every $n \in \mathbb{N}_{\geq 1}$ let us introduce the mapping

$$F : \left(\mathcal{D}_n^{(1)}\times\mathbb{R}^m\right)\cup\left(\mathcal{D}_n^{(4)}\times\mathbb{R}^m\right) \longrightarrow \mathbb{R}, \tag{4.12}$$

that is given by

$$\begin{aligned}(B,(C^i)_{1\leq i\leq d},(D^i)_{1\leq i\leq d},N,K,L,G) \;\mapsto\;& \varphi(K)B' + (q-1)\varphi(K)\sum_{i=1}^d (D^i)'\left(\frac{L^i}{K}+C^i\right) \\ &+(q-1)\varphi(K)\left(\sum_{i=1}^d (D^i)'D^i\right)G + N\varphi(G).\end{aligned} \tag{4.13}$$

Note that F is well defined. For simplicity of notation set

$$v := (B,(C^i)_{1\leq i\leq d},(D^i)_{1\leq i\leq d},N,K,L). \tag{4.14}$$

Bearing (4.9) in mind, the problem we address in the following lemma is to properly define a (continuous) function $G : \mathcal{D}_n^{(1)}\cup\mathcal{D}_n^{(4)} \longrightarrow \mathbb{R}^m$ such that $F(v,G(v)) = 0$ for all v.

Lemma 4.8 *For some $n \in \mathbb{N}_{\geq 1}$, let F be given by (4.12), (4.13). Then there is a unique function $G : \mathcal{D}_n^{(1)}\cup\mathcal{D}_n^{(4)} \longrightarrow \mathbb{R}^m$ such that $F(v,G(v)) = 0$ for all $v \in \mathcal{D}_n^{(1)}\cup\mathcal{D}_n^{(4)}$ (where v is defined in (4.14)). G is continuous.*

[1] K will be seen to be a process

Proof: We use the notation (4.14). For $v \in \mathcal{D}_n^{(1)} \cup \mathcal{D}_n^{(4)}$ set

$$r(v) := \varphi(K)B' + (q-1)\varphi(K)\sum_{i=1}^{d}(D^i)'\left(\frac{L^i}{K}+C^i\right) \in \mathbb{R}^m$$

and

$$\mathcal{A}(v) := (q-1)\varphi(K)\left(\sum_{i=1}^{d}(D^i)'D^i\right) \in \mathbb{R}^{m\times m}.$$

The function F then reads as $F(v,G) = r(v) + \mathcal{A}(v)G + N\varphi(G)$, with $\varphi(G) = |G|^{q-2}G$. The only non-linear term in the equation $F(v,G) = 0$ is $|G|^{q-2}$, so we will first try to determine the modulus of G.

Let us fix some $v \in \mathcal{D}_n^{(1)} \cup \mathcal{D}_n^{(4)}$ with

$$r(v) \neq 0.$$

For $\alpha \in \mathbb{R}_{>0}$ set

$$G_\alpha := -\left(\mathcal{A}(v) + \alpha^{q-2}\mathrm{diag}(N)\right)^{-1} r(v).$$

Note that the inverse is well defined. If $v \in \mathcal{D}_n^{(1)}$, $\mathcal{A}(v)$ is uniformly positive and $\alpha^{q-2}\mathrm{diag}(N)$ is positive-semidefinite, and vice versa if $v \in \mathcal{D}_n^{(4)}$. G_α is the solution of $r(v) + \mathcal{A}(v)G + \alpha^{q-2}\mathrm{diag}(N)G = 0$. Hence, if we could find an $\bar{\alpha}$ such that $\bar{\alpha} = |G_{\bar{\alpha}}|$, then $G_{\bar{\alpha}}$ would solve $F(v,G) = 0$ for our fixed v. For $\alpha \in \mathbb{R}_{>0}$ set

$$h(v,\alpha) := \frac{|G_\alpha|}{\alpha} = |(\alpha\mathcal{A}(v) + \alpha^{q-1}\mathrm{diag}(N))^{-1}r(v)|.$$

As $r(v) \neq 0$ and either $\mathcal{A}(v)$ or $\mathrm{diag}(N)$ are positive definite, h is strictly decreasing in α with $h(v,\alpha) \to \infty$, $\alpha \to 0$, $h(v,\alpha) \to 0$ $\alpha \to \infty$. Hence there is a unique $\bar{\alpha} > 0$ with $h(v,\bar{\alpha}) = 1$. Obviously, $|G_{\bar{\alpha}}| = \bar{\alpha}$ and hence $F(v,G_{\bar{\alpha}}) = 0$. So we may define $G(v) := G_{\bar{\alpha}}$.

If

$$r(v) = 0$$

we set $G(v) = 0$; moreover, due to the regularity of $\mathcal{A}(v) + |G|^{q-2}\mathrm{diag}(N)$ for $G \neq 0$, this is the only possible choice. Besides, $G(v) = 0$ is only possible if $r(v) = 0$, as we have seen above (recall that we had $\bar{\alpha} > 0$). Now note that for every solution G of $F(v,G) = 0$ with $r(v) \neq 0$ we have $h(v,|G|) = 1$, hence $|G| = \bar{\alpha}$ and consequently $G = G_{\bar{\alpha}}$. Thus we have defined an application $G : \mathcal{D}_n^{(1)} \cup \mathcal{D}_n^{(4)} \longrightarrow \mathbb{R}^m$, and if $F(v,G^0) = 0$ for some $v \in \mathcal{D}_n^{(1)} \cup \mathcal{D}_n^{(4)}$ and some $G^0 \in \mathbb{R}^m$, then $G^0 = G(v)$. Hence, G is well defined and unique.

We now turn to the question of continuity. In the first instance we will show that G is continuous on the the set

$$S := (\mathcal{D}_n^{(1)} \cup \mathcal{D}_n^{(4)})\backslash\{r(v) = 0\}.$$

Consider the function h as a mapping $h : S \times \mathbb{R}_{>0} \longrightarrow \mathbb{R}$, $(v,\alpha) \mapsto h(v,\alpha)$; we have seen that there is a function $\bar{\alpha} = \bar{\alpha}(v) > 0$ such that $h(v,\bar{\alpha}(v)) - 1 = 0$. As $\alpha\mathcal{A}(v) + \alpha^{q-1}\mathrm{diag}(N)$

is regular in a small neighborhood of $(v, \bar{\alpha}(v))$ for $(v, \bar{\alpha}(v)) \in S \times \mathbb{R}_{>0}$, it follows that $(\partial_\alpha h)(v, \bar{\alpha}(v)) \neq 0$ for $v \in S$. The Implicit Function Theorem now yields that $v \mapsto \bar{\alpha}(v)$ is continuously differentiable on the interior of S, and it is easily seen that this extends to continuity on S. Now note that $G(v) = -(\mathcal{A}(v) + \bar{\alpha}(v)^{q-2}\text{diag}(N))^{-1}r(v)$, for $v \in S$; this entails continuity of G on S.
To show continuity on $\{r(v) = 0\}$ consider a sequence (v_n) in $(\mathcal{D}_n^{(1)} \cup \mathcal{D}_n^{(4)})\backslash\{r(v) = 0\}$ with $v_j \to v \in \mathcal{D}_n^{(1)} \cup \mathcal{D}_n^{(4)}$, $j \to \infty$, and $r(v) = 0$. We must show that $G(v_j) \to 0$, $j \to \infty$. Denote by N_j the N-component of v_j and set $\alpha_j = |G(v_j)|$ (note that $\alpha_j \neq 0$ for all j). Then we have $G(v_j) = -(\mathcal{A}(v_j) + \alpha_j^{q-2}\text{diag}(N_j))^{-1}r(v_j)$. Taking the modulus on both sides and dividing by α_j gives, as above,

$$\begin{aligned} 1 &= \left|\left(\alpha_j\mathcal{A}(v_j) + \alpha_j^{q-1}\text{diag}(N_j)\right)^{-1} r(v_j)\right| \\ &\leq \left|\left(\alpha_j\mathcal{A}(v_j) + \alpha_j^{q-1}\text{diag}(N_j)\right)^{-1}\right| \underbrace{|r(v_j)|}_{\to 0}, \end{aligned}$$

hence $\left|\left(\alpha_j\mathcal{A}(v_j) + \alpha_j^{q-1}\text{diag}(N_j)\right)^{-1}\right| \to \infty$, $j \to \infty$. As n is fixed, there is a common lower bound, depending on n and independent of j, for the eigenvalues of $\mathcal{A}(v_j)$ (if $v \in \mathcal{D}_n^{(1)}$) respectively for N_j (if $v \in \mathcal{D}_n^{(4)}$). Thus, $\left|\left(\alpha_j\mathcal{A}(v_j) + \alpha_j^{q-1}\text{diag}(N_j)\right)^{-1}\right| \to \infty$ is only possible if $\alpha_j \to 0$, $j \to \infty$. This implies $G(v_j) \to 0$, $j \to \infty$, and the lemma is shown. ∎

Now all ingredients are in place to finally state the Riccati-equation - just replace $\frac{1}{x}u$ by $G(\cdot)$ in (4.8). We delay this statement to the beginning of the next section and turn our attention first to some properties of G - being implicitly defined we know very little so far about this function. These properties, given in the next two lemmas, will be useful when investigating some a-priori estimates for solutions of the Riccati-equation.

Lemma 4.9 *Consider the function G defined in Lemma 4.8.*

1. *For all $n \in \mathbb{N}_{\geq 1}$ and all $(B, (C^i)_i, (D^i)_i, N, K, L) \in \mathcal{D}_n^{(1)} \cup \mathcal{D}_n^{(4)}$ we have*

$$\left(\frac{1}{q-1}KB + \sum_{i=1}^{d} D^i(KC^i + L^i)\right) G(B, (C^i)_i, (D^i)_i, N, K, L) \leq 0.$$

2. *Assume $q \leq 2$. Then, for all $n \in \mathbb{N}_{\geq 1}$ there are constants $a, b > 0$, depending only on n and q, such that*

$$|G(B, (C^i)_i, (D^i)_i, N, K, L)| \leq a + b|L|.$$

for all $(B, (C^i)_i, (D^i)_i, N, K, L) \in \mathcal{D}_n^{(1)}$. The same is true if $q \geq 2$ and G is given by (4.11).

3. *Assume $q \geq 2$. Then, for all $n \in \mathbb{N}_{\geq 1}$ there are constants $a, b > 0$, depending only on n and q, such that*
$$|G(B,(C^i)_i,(D^i)_i,N,K,L)| \leq a + b|L|^{\frac{1}{q-1}}$$
for all $(B,(C^i)_i,(D^i)_i,N,K,L) \in \mathcal{D}_n^{(4)}$.

Proof: We omit the arguments of G.

1. From the definition we have
$$\varphi(K)B' + (q-1)\varphi(K)\sum_{i=1}^d (D^i)'\left(\frac{L^i}{K} + C^i\right)$$
$$= -(q-1)\varphi(K)\left(\sum_{i=1}^d (D^i)'D^i\right)G - N\varphi(G).$$
Multiplying with $\frac{1}{q-1}K^{2-q} \in (0,\infty)$ yields after transposition
$$\frac{1}{q-1}BK + \sum_{i=1}^d D^i\left(L^i + C^iK\right)$$
$$= -KG'\left(\sum_{i=1}^d (D^i)'D^i\right) - \frac{1}{q-1}K^{2-q}N(\varphi(G))',$$
and multiplication with G from the right gives
$$\left(\frac{1}{q-1}BK + \sum_{i=1}^d D^i\left(L^i + C^iK\right)\right)G$$
$$= -KG'\left(\sum_{i=1}^d (D^i)'D^i\right)G - \frac{1}{q-1}K^{2-q}N|G|^q, \tag{4.15}$$
where the right hand side is equal to or less than zero.

2. If G is defined by (4.11) the linear growth of G in L is clear. For the general case, note that on $\mathcal{D}_n^{(1)}$ we have $\frac{1}{n} \leq K \leq n$ and $\sum_{i=1}^d (D^i)'D^i \geq \frac{1}{n}$. From equality (4.15) we get
$$\left|\left(\frac{1}{q-1}BK + \sum_{i=1}^d D^i\left(L^i + C^iK\right)\right)G\right|$$
$$= KG'\left(\sum_{i=1}^d (D^i)'D^i\right)G + \frac{1}{q-1}K^{2-q}N|G|^q$$
$$\geq KG'\left(\sum_{i=1}^d (D^i)'D^i\right)G$$
$$\geq \frac{1}{n^2}|G|^2,$$

for all $(B, (C^i)_i, (D^i)_i, N, K, L) \in \mathcal{D}_n^{(1)}$. From the boundedness of $B, K, (C^i)_i, (D^i)_i$ and K there are constants a_0, b_0 depending only on n and q such that

$$a_0 + b_0|L||G| \geq |G|^2$$

for all $(B, (C^i)_i, (D^i)_i, N, K, L) \in \mathcal{D}_n^{(1)}$. This yields the assertion that is to be shown.

3. The arguments are the same as above. Again, from (4.15) we have

$$\begin{aligned}
&\left|\left(\frac{1}{q-1}BK + \sum_{i=1}^{d} D^i\left(L^i + C^iK\right)\right)G\right| \\
&= KG'\left(\sum_{i=1}^{d}(D^i)'D^i\right)G + \frac{1}{q-1}K^{2-q}N|G|^q \\
&\geq \frac{1}{q-1}K^{2-q}N|G|^q \\
&\geq \frac{1}{q-1}\frac{1}{n^{q-1}}|G|^q
\end{aligned}$$

for all $(B, (C^i)_i, (D^i)_i, N, K, L) \in \mathcal{D}_n^{(4)}$, since $K^{2-q} \geq \frac{1}{n^{q-2}}$ and $N \geq \frac{1}{n}$. On the left hand side, every term except L and G is bounded, hence there are constants $a_0, b_0 > 0$ depending only on n and q such that

$$a_0 + b_0|L||G| \geq |G|^q,$$

which entails the assertion (note that $\frac{1}{q-1} \leq 1$).

■

The following corollary is only of technical relevance and is essentially a simple calculation. We will need it when dealing with the Riccati-equation in the case of Assumption A4.

Corollary 4.10 *Suppose $q > 2$ and set $l = \frac{1}{q-1}$. For every $n \in \mathbb{N}_{\geq 1}$ there is a $k_0 > 0$ depending only on n and q such that for all $(B, (C^i)_i, (D^i)_i, N, K, L) \in \mathcal{D}_n^{(4)}$*

$$\left|\left(lBK + \sum_{i=1}^{d} D^i(C^iK + L^i)\right)G(B, (C^i)_i, (D^i)_i, N, K, L))K^{l-1}\right| \leq k_0\left(1 + \left(K^{l-1}|L|\right)^{l+1}\right).$$

Proof: From Lemma 4.9-3 there are constants k_i depending only on n and q such that

$$\left|\left(lBK + \sum_{i=1}^{d} D^i(C^iK + L^i)\right)G(B, (C^i)_i, (D^i)_i, N, K, L)K^{l-1}\right|$$

$$
\begin{aligned}
&\leq k_1 \left| \left(lBK + \sum_{i=1}^{d} D^i (C^i K + L^i) \right) \right| \left(1 + |L|^l\right) K^{l-1} \\
&\leq k_2 \left(K + |L|\right) \left(1 + |L|^l\right) K^{l-1} \\
&\leq k_3 \left(K^l + K^l |L|^l + K^{l-1} \left(|L| + |L|^{l+1}\right)\right). \qquad (4.16)
\end{aligned}
$$

We have $K^l|L|^l = K^{l(2-l)}\left(K^{l-1}|L|\right)^l \leq n^{l(2-l)}\left(K^{l-1}|L|\right)^l$ and $K^{l-1}|L|^{l+1} = K^{-l(l-1)}\left(K^{l-1}|L|\right)^{l+1} \leq n^{-l(l-1)}\left(K^{l-1}|L|\right)^{l+1}$. So, (4.16) can be continued by

$$
\begin{aligned}
&\left| \left(lBK + \sum_{i=1}^{d} D^i (C^i K + L^i) \right) G(B, (C^i)_i, (D^i)_i, N, K, L) K^{l-1} \right| \\
&\leq k_4 \left(1 + \left(K^{l-1}|L|\right)^l + K^{l-1}|L| + \left(K^{l-1}|L|\right)^{l+1}\right) \\
&\leq k_5 \left(1 + \left(K^{l-1}|L|\right)^{l+1}\right),
\end{aligned}
$$

with some k_5 depending only on n and q. This proves the lemma. ■

In the following, we will omit the variables $B, (C^i)_i, (D^i)_i$ and N of G.

Notation 4.11 *G will always denote the function introduced in Lemma 4.8 respectively in (4.11). We will suppress the arguments $B, (C^i)_i, (D^i)_i$ and N and write $G = G(K, L)$.*

We are now ready to define the Riccati-equation as an object in its own right.

4.4 The equation, inherent properties

By replacing $\frac{1}{x}u$ by $G(K, L)$ in (4.8) we are led to the Riccati-equation. To the best of our knowledge, this generalization of the "conventional" Backward Stochastic Riccati Differential Equations is new.

Definition 4.12 *Let $A, B, (C^i)_i, (D^i)_i, N, Q$ and M be the coefficients of a problem $\mathcal{P}(\tau, 1)$ which either satisfy Assumption A1, A3 or A4. The* Backward Stochastic Riccati Differential Equation *(*BSRDE*) for these coefficients is given by*

$$
\begin{aligned}
dK &= \Bigg\{ -q'AK - \sum_{i=1}^{d} (C^i)^2 K - 2 \sum_{i=1}^{d} C^i L^i - \frac{q-2}{q-1} KBG(K, L) \\
&\quad - \left(\frac{1}{q-1} BK + \sum_{i=1}^{d} D^i \left(C^i K + L^i\right) \right) G(K, L) \\
&\quad - \frac{1}{q-1} QK^{2-q} + \frac{2-q}{2} \frac{1}{K} \sum_{i=1}^{d} \left(L^i + KC^i + KD^i G(K, L)\right)^2 \Bigg\} ds + \sum_{i=1}^{d} L^i dw^i, \qquad (4.17) \\
K(T) &= f(M). \qquad (4.18)
\end{aligned}
$$

A solution *of this equation is a pair of adapted processes* K *and* $L = (L^1, \dots, L^d)$*, the* L^i *being real valued, such that*

1. $K \in L^\infty_{\mathcal{F}}(\tau, T; \mathbb{R}) \cap L^\infty_{\mathcal{F}}(\Omega, C([\tau, T]; \mathbb{R}))$ *and in addition*
   If Assumption A1 or A3 holds: there is a $c > 0$ *such that* $P - a.s.$ $K(t \vee \tau) \geq c$ *for all* $t \in [0, T]$*;*
   If Assumptions A4 holds: $P - a.s.$ $K(t \vee \tau) > 0$ *for all* $t \in [0, T]$*;*
2. $\int_\tau^T |L(s)|^2 ds < \infty$, $P - a.s.$
3. (K, L) *satisfies the BSDE (4.17), (4.18), i.e.*

$$
\begin{aligned}
K(t \vee \tau) = f(M) - \int_{t\vee\tau}^T \Bigg\{ & - q'A(s)K(s) \\
& - \sum_{i=1}^d (C^i(s))^2 K(s) - 2\sum_{i=1}^d C^i(s)L^i(s) - \frac{q-2}{q-1}K(s)B(s)G(K(s), L(s)) \\
& - \left(\frac{1}{q-1}B(s)K(s) + \sum_{i=1}^d D^i(s)\left(C^i(s)K(s) + L^i(s)\right) \right) G(K(s), L(s)) \\
& - \frac{1}{q-1}Q(s)K^{2-q}(s) \\
& + \frac{2-q}{2}\frac{1}{K}\sum_{i=1}^d \left(L^i(s) + K(s)C^i(s) + K(s)D^i(s)G(K(s), L(s))\right)^2 \Bigg\} ds \\
& - \sum_{i=1}^d \int_{t\vee\tau}^T L^i(s)dw^i(s),
\end{aligned}
$$

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

Note that we are not completely free in the choice of the coefficients for which we may formulate the Riccati-equation. We must make sure that the function G is properly defined for these coefficients. As the definition of G involves K in a critical way, the definition of a solution must incorporate requirements on K such that the expression $G(K, L)$ is well defined. Note that in the quadratic case this requirement that guarantees the existence of G would typically be formulated by demanding that $\left(\mathrm{diag}(N) + K\sum_{i=1}^d (D^i)'D^i\right)^{-1}$ exists. The main goal now is to show that $K = \frac{f(y)}{x}$ and L, given by (4.5), is a solution of the BSRDE. We straightaway encounter two major obstacles. First, the expression $\frac{f(y)}{x}$ is properly defined only on $[\tau, \tau_0)$, and there is no evidence that K, as constructed in Proposition 4.1, possesses a differential outside this interval. Secondly, we also do not know whether L is $P-a.s.$ pathwise square-integrable. These obstacles are closely related. Let us sketch the main ideas without going into details too much. Suppose that we were able to show that L is pathwise square integrable on $[\tau, \tau_0)$, i.e. $\int_\tau^{\tau_0} |L|^2 ds < \infty$ $P - a.s.$, where L is given by

(4.5). From Lemma 4.9-2,3 we then get $\int_\tau^{\tau_0} |G(K,L)|^2 ds < \infty \;\; P-a.s.$ Now note that on $[\tau,\tau_0)$ the equality (4.9) holds. Yet, the function G resolves this equality for given quantities $B,(C^i),(D^i),N,K$ and L. As G is unique among the mappings with this property, it follows that $\frac{1}{x}u = G(K,L), \;\; Leb \otimes P - a.s.$ on $[\tau,\tau_0)$ (here the uniqueness of G becomes important). Consequently, we have $u = G(K,L)x$ on $[\tau,\tau_0)$, and from Lemma 3.8 we know that $u = 0$ on $(\tau_0,T]$. We may extend G by $G = 0$ on $[\tau_0,T]$ and hence get $u = G(K,L)x$, $Leb \otimes P - a.s.$, on all of $[\tau,T]$. This is very good news, because it shows that x is actually a stochastic exponential, $dx = \{A + BG(K,L)\}xds + \sum_{i=1}^d \{C^i + D^i G(K,L)\}xdw^i$, $x(\tau) = 1$, *and the coefficients of this SDE are pathwise square integrable.* From this integrability we can deduce that $P-a.s.$ we have $x(t \vee \tau) > 0$ for all t, i.e. $\tau_0 = T$ - this is the link to the first obstacle mentioned above, the integrability of L.

The preceding discussion is, of course, not completely rigorous, but it may have justified that it would be worth investigating the integrability of L. The subsequent three theorems are crucial in this respect. They deal with a-priori estimates of the integrals of L, if L is known to be part of a solution of a Backward Riccati Equation (the reader will hopefully not be confused that we will denote generic solutions of the BSRDE by (K,L); we will try to make clear in the respective context if we mean such a solution or the particular processes defined in Section 4.2).

Besides that, the following theorems are very helpful for proving the solvability of the BSDRE, they are a key insight and interesting in their own right. They exhibit a surprising property of these equations. Roughly speaking, they show that the strong requirements imposed on K entail strong integrability properties of L. The method we use is taken from [T:GLQO]; there, the author considers the differential of K^2, whereas we will investigate the differential of $K^{\frac{1}{q-1}}$ or K^{-r} for some $r > 0$. We will omit the case $q = 2$ for technical reasons (the quadratic case is meanwhile well covered in the literature, for example by the aforementioned article of Tang and in [KT:GAS]).

Theorem 4.13 *Assume $q < 2$. Let (K,L) satisfy the differential equation (4.17) on $[\tau,T]$, where the coefficients satisfy Assumption A1. Assume that $n \in \mathbb{N}_{\geq 1}$ is such that $(B,(C^i)_i,(D^i)_i,N,K,L) \in \mathcal{D}_n^{(1)}$, $|A|,|Q| \leq n$, $Leb \otimes P - a.s.$, and $|M| \leq n$, $P-a.s.$. Then, for every $p \in [1,\infty)$ there is a $k > 0$ depending only on n,q,p and T such that*

$$E[\left(\int_\tau^T |L(s)|^2 ds\right)^p] \leq k.$$

Proof: Set $l = \frac{1}{q-1}$. For $j \in \mathbb{N}_{\geq 1}$ introduce the stopping times

$$\gamma_j := \inf\{t \geq \tau : \int_\tau^t |L|^2 ds \geq j\} \wedge T, \; \inf \emptyset := \infty.$$

By Itô's formula, (writing simply G instead of $G(K,L)$)

$$d(K^l(t \wedge \gamma_j)) \;\; = \;\; l\mathbf{1}_{[\tau,\gamma_j]}K^{l-1}dK + \frac{1}{2}l(l-1)\mathbf{1}_{[\tau,\gamma_j]}K^{l-2}d < K >$$

$$
\begin{aligned}
= \ & \mathbf{1}_{[\tau,\gamma_j]}\Bigg\{-l(l+1)AK^l - l\sum_{i=1}^d (C^i)^2K^l - 2l\sum_{i=1}^d C^iK^{l-1}L^i \\
& -l\left(lBK+\sum_{i=1}^d D^i(C^iK+L^i)\right)GK^{l-1} + l(l-1)K^lBG \\
& -l^2QK^{2-q+l-1} + \frac{1}{2}(l-1)\sum_{i=1}^d (L^i+C^iK+KD^iG)^2K^{l-1}\Bigg\}\,ds \\
& +\mathbf{1}_{[\tau,\gamma_j]}lK^{l-1}Ldw + \ \frac{1}{2}l(l-1)\mathbf{1}_{[\tau,\gamma_j]}K^{l-2}|L|^2ds. \qquad (4.19)
\end{aligned}
$$

Setting $t=T$, we get for all j

$$
\begin{aligned}
& K^l(\gamma_j)-K^l(\tau)+\int_\tau^T \mathbf{1}_{[\tau,\gamma_j]}\Bigg\{l(l+1)AK^l + 2l\sum_{i=1}^d C^iK^{l-1}L^i + l\sum_{i=1}^d (C^i)^2K^l \\
& \quad +l\left(lBK+\sum_{i=1}^d D^i(C^iK+L^i)\right)GK^{l-1} - l(l-1)BK^lG + l^2QK^{2-q+l-1}\Bigg\}\,ds \\
& \quad -l\int_\tau^T \mathbf{1}_{[\tau,\gamma_j]}K^{l-1}Ldw \\
= \ & (l-1)\int_\tau^T \mathbf{1}_{[\tau,\gamma_j]}\left\{\frac{1}{2}\sum_{i=1}^d (L^i+C^iK+D^iKG)^2K^{l-1} + \frac{l}{2}K^{l-2}|L|^2\right\}ds \\
\geq \ & 0,
\end{aligned}
$$

where the last inequality is due to the assumption $q<2$, i.e. $l>1$. By Lemma 4.9-1,

$$
l\left(lBK+\sum_{i=1}^d D^i(C^iK+L^i)\right)GK^{l-1}\leq 0,
$$

therefore

$$
\begin{aligned}
& K^l(\gamma_j)-K^l(\tau)+\int_\tau^T \mathbf{1}_{[\tau,\gamma_j]}\Bigg\{l(l+1)AK^l + 2l\sum_{i=1}^d C^iK^{l-1}L^i \\
& \quad +l\sum_{i=1}^d (C^i)^2K^l - l(l-1)BK^lG + l^2QK^{1-q+l}\Bigg\}\,ds \\
& \quad -l\int_\tau^T \mathbf{1}_{[\tau,\gamma_j]}K^{l-1}Ldw \\
\geq \ & (l-1)\int_\tau^T \mathbf{1}_{[\tau,\gamma_j]}\left\{\frac{1}{2}\sum_{i=1}^d (L^i+C^iK+D^iKG)^2K^{l-1} + \frac{1}{2}lK^{l-2}|L|^2\right\}ds.
\end{aligned}
$$

Taking into account that the coefficients are bounded by n and that $\frac{1}{n} \leq K \leq n$, from this last inequality it follows that there is a $k_1 > 0$ depending only on n, q and T such that $P - a.s.$ for all j

$$\begin{aligned} & k_1 + k_1 \left(\int_\tau^T \mathbf{1}_{[\tau,\gamma_j]} \{|G| + |L|\}\, ds + \left| \int_\tau^T \mathbf{1}_{[\tau,\gamma_j]} K^{l-1} L dw \right| \right) \\ \geq & \int_\tau^T \mathbf{1}_{[\tau,\gamma_j]} |L|^2 ds. \end{aligned} \tag{4.20}$$

Now choose a $p > 1$. By the Burkholder-Gundy-Davis inequalities there is a constant $c > 0$ depending only on p and T such that

$$\begin{aligned} E[\left| \int_\tau^T \mathbf{1}_{[\tau,\gamma_j]} K^{l-1} L dw \right|^p] & \leq cE[\left(\int_\tau^T \mathbf{1}_{[\tau,\gamma_j]} K^{2l-2} |L|^2 ds \right)^{\frac{p}{2}}] \\ & \leq cn^{p(l-1)} E[\left(\int_\tau^T \mathbf{1}_{[\tau,\gamma_j]} |L|^2 ds \right)^{\frac{p}{2}}]. \end{aligned}$$

Lemma 4.9-2 asserts the existence of constants a, b depending only on n and q such that $|G| \leq a + b|L|$. Using the relation $|\xi + \eta|^p \leq 2^{p-1}(|\xi|^p + |\eta|^p)$, $\xi, \eta \in \mathbb{R}$, we can conclude that there are some $k_i > 0$ depending only on n, p, q and T such that, starting from (4.20), the following estimates hold for all j:

$$\begin{aligned} & E[\left(\int_\tau^T \mathbf{1}_{[\tau,\gamma_j]} |L|^2 ds \right)^p] \\ \leq & k_2 \left(1 + E[\left(\int_\tau^T \mathbf{1}_{[\tau,\gamma_j]} \{|G| + |L|\} ds \right)^p] + E[\left| \int_\tau^T \mathbf{1}_{[\tau,\gamma_j]} K^{l-1} L dw \right|^p] \right) \\ \leq & k_3 \left(1 + E[\left(\int_\tau^T \mathbf{1}_{[\tau,\gamma_j]} |L| ds \right)^p] + E[\left| \int_\tau^T \mathbf{1}_{[\tau,\gamma_j]} K^{l-1} L dw \right|^p] \right) \\ \leq & k_4 \left(1 + E[\left(\int_\tau^T \mathbf{1}_{[\tau,\gamma_j]} |L| ds \right)^p] + E[\left(\int_\tau^T \mathbf{1}_{[\tau,\gamma_j]} |L|^2 ds \right)^{\frac{p}{2}}] \right). \end{aligned} \tag{4.21}$$

The Cauchy-Schwarz inequality gives

$$\int_\tau^T \mathbf{1}_{[\tau,\gamma_j]} |L| ds \leq T^{\frac{1}{2}} \left(\int_\tau^T \mathbf{1}_{[\tau,\gamma_j]} |L|^2 ds \right)^{\frac{1}{2}},$$

and we may continue (4.21) by

$$E[\left(\int_\tau^T \mathbf{1}_{[\tau,\gamma_j]} |L|^2 ds \right)^p] \leq k_5 \left(1 + E[\left(\int_\tau^T \mathbf{1}_{[\tau,\gamma_j]} |L|^2 ds \right)^{\frac{p}{2}}] \right).$$

Set $\alpha := \left(\int_\tau^T \mathbf{1}_{[\tau,\gamma_j]} |L|^2 ds \right)^{\frac{p}{2}}$. Since $\xi\eta \leq \frac{1}{2}(\xi^2 + \eta^2)$, $\xi, \eta \in \mathbb{R}$, the last inequality entails that

$$E[\alpha^2] \leq E[k_5 + k_5 \alpha] \leq E[k_5 + \frac{1}{2}(\alpha^2 + k_5^2)],$$

thus

$$\frac{1}{2}E[\alpha^2] \leq k_5 + \frac{1}{2}k_5^2.$$

Hence, with $k := 2k_5 + k_5^2$, for all j we have

$$E[\Big(\int_\tau^T \mathbf{1}_{[\tau,\gamma_j]}|L|^2 ds\Big)^p] \leq k,$$

and by the Monotone Convergence Theorem

$$\begin{aligned} E[\Big(\int_\tau^T |L|^2 ds\Big)^p] &= E[\lim_{j\to\infty}\Big(\int_\tau^T \mathbf{1}_{[\tau,\gamma_j]}|L|^2 ds\Big)^p] \\ &= \lim_{j\to\infty} E[\Big(\int_\tau^T \mathbf{1}_{[\tau,\gamma_j]}|L|^2 ds\Big)^p] \\ &\leq k. \end{aligned}$$

Since k only depends on n, p, q and T, the theorem is proved. ∎

The following theorem will enable us to establish the integrability of L if Assumption A4 is in force.

Theorem 4.14 *Assume $q > 2$ and set $l := \frac{1}{q-1}$. Let (K, L) be a solution of BSRDE (4.17) on $[\tau, T]$, where the coefficients satisfy Assumption A4. Assume that $n \in \mathbb{N}_{\geq 1}$ is such that $(B, (C^i)_i, (D^i)_i, N, K, L) \in \mathcal{D}_n^{(4)}$, $|A|, |Q| \leq n$, $Leb \otimes P-a.s.$, and $|M| \leq n$, $P-a.s.$. Then, for every $p \in (1, \infty)$ there is a $k > 0$ depending only on n, q, p and T such that*

$$E[\Big(\int_\tau^T \left(K^{l-1}(s)|L(s)|\right)^2 ds\Big)^p] \leq k.$$

Proof: Similarly as in Theorem 4.13, define the stopping times γ_j by

$$\gamma_j := \inf\{t \geq \tau : \int_\tau^t \left(K^{l-1}|L|\right)^2 ds \geq j\} \wedge T,\ \inf\emptyset := \infty.$$

From (4.19) we get (the special form of γ_j does not affect the working up to this point), since $l < 1$,

$$\begin{aligned} 0 \leq (1-l)\int_\tau^T \mathbf{1}_{[\tau,\gamma_j]}&\left\{\frac{l}{2}K^{l-2}|L|^2 + \frac{1}{2}\sum_{i=1}^d (L^i + C^iK + D^iKG)^2 K^{l-1}\right\} ds \\ = \quad -K^l(\gamma_j) + K^l(\tau) &+ \int_\tau^T \mathbf{1}_{[\tau,\gamma_j]}\Bigg\{-l(l+1)AK^l - 2l\sum_{i=1}^d C^iK^{l-1}L^i \\ -l\sum_{i=1}^d (C^i)^2K^l &+ l(l-1)BK^lG - l^2QK^{2-q+l-1} \end{aligned}$$

$$-l\left(lBK+\sum_{i=1}^{d}D^i(C^iK+L^i)\right)GK^{l-1}\Bigg\}ds$$
$$+l\int_\tau^T \mathbf{1}_{[\tau,\gamma_j]}K^{l-1}Ldw,$$

and, since Q is non-negative,

$$\begin{aligned}0\le(1-l)&\int_\tau^T \mathbf{1}_{[\tau,\gamma_j]}\frac{l}{2}K^{l-2}|L|^2ds\\ \le\;& -K^l(\gamma_j)+K^l(\tau)+\int_\tau^T \mathbf{1}_{[\tau,\gamma_j]}\Bigg\{-l(l+1)AK^l-2l\sum_{i=1}^{d}C^iK^{l-1}L^i\\ &+l(l-1)BK^lG-l\sum_{i=1}^{d}(C^i)^2K^l-l\left(lBK+\sum_{i=1}^{d}D^i(C^iK+L^i)\right)GK^{l-1}\Bigg\}ds\\ &+l\int_\tau^T \mathbf{1}_{[\tau,\gamma_j]}K^{l-1}Ldw. \qquad (4.22)\end{aligned}$$

Hence, according to the Lemma 4.9-3 and Corollary 4.10, there are k_i depending only on n, q and T such that

$$\begin{aligned}0\le&\int_\tau^T \mathbf{1}_{[\tau,\gamma_j]}K^{l-2}|L|^2ds=\int_\tau^T \mathbf{1}_{[\tau,\gamma_j]}K^{-l}\left(K^{l-1}|L|\right)^2ds\\ \le\;& k_1+\int_\tau^T \mathbf{1}_{[\tau,\gamma_j]}\Big\{k_1+k_1K^{l-1}|L|+k_1K^l\left(1+|L|^l\right)+k_1\\ &+k_1\left(1+K^{l-1}|L|^{l+1}\right)\Big\}ds+\left|l\int_\tau^T \mathbf{1}_{[\tau,\gamma_j]}K^{l-1}Ldw\right|\\ \le\;& k_2\left(1+\int_\tau^T \mathbf{1}_{[\tau,\gamma_j]}\left\{K^{l-1}|L|+K^l|L|^l+K^{l-1}|L|^{l+1}\right\}ds\right)\\ &+\left|l\int_\tau^T \mathbf{1}_{[\tau,\gamma_j]}K^{l-1}Ldw\right|. \qquad (4.23)\end{aligned}$$

We want the expressions in the fourth line of the preceding inequalities to be dominated by $\left(K^{l-1}|L|\right)^{l+1}$ and some constants. For the first expression it is clear that

$$K^{l-1}|L|\le const.\left(1+\left(K^{l-1}|L|\right)^{l+1}\right),$$

for all K, L, with a constant depending only on l. Next, we have

$$\begin{aligned}K^l|L|^l&=\left(K^{2-l}K^{l-1}|L|\right)^l\\ &\le n^{(2-l)l}\left(K^{l-1}|L|\right)^l\\ &\le n^{(2-l)l}const.\left(1+\left(K^{l-1}|L|\right)^{l+1}\right),\end{aligned}$$

for all K, L and a constant depending only on l. Finally, set $r := \frac{l-l^2}{l+1} > 0$. Then, $\frac{l-1}{l+1} = l - 1 + r$, hence

$$\begin{aligned} K^{l-1}|L|^{l+1} &= \left(K^{\frac{l-1}{l+1}}|L|\right)^{l+1} \\ &= K^{r(l+1)}\left(K^{l-1}|L|\right)^{l+1} \\ &\leq n^{r(l+1)}\left(K^{l-1}|L|\right)^{l+1}. \end{aligned}$$

Now examine the last expression in the first line of (4.23). There, we have

$$\frac{1}{n^l}\left(K^{l-1}|L|\right)^2 \leq K^{-l}\left(K^{l-1}|L|\right)^2.$$

Putting all this together, from (4.23) we get that there is a k_3 independent of j such that

$$\begin{aligned} &\int_\tau^T \mathbf{1}_{[\tau,\gamma_j]}(K^{l-1}|L|)^2 ds \\ &\leq k_3\left(1 + \int_\tau^T \mathbf{1}_{[\tau,\gamma_j]}\left(K^{l-1}|L|\right)^{l+1} ds + \left|\int_\tau^T \mathbf{1}_{[\tau,\gamma_j]}K^{l-1}L dw\right|\right), \end{aligned}$$

hence, for $p > 1$,

$$\begin{aligned} &\left(\int_\tau^T \mathbf{1}_{[\tau,\gamma_j]}(K^{l-1}|L|)^2 ds\right)^p \\ &\leq k_4\left(1 + \left(\int_\tau^T \mathbf{1}_{[\tau,\gamma_j]}\left(K^{l-1}|L|\right)^{l+1} ds\right)^p + \left|\int_\tau^T \mathbf{1}_{[\tau,\gamma_j]}K^{l-1}L dw\right|^p\right), \end{aligned}$$

With a k_4 independent of j. Taking expectations, the Burkholder-Gundy-Davis-inequality yields

$$\begin{aligned} &E[\left(\int_\tau^T \mathbf{1}_{[\tau,\gamma_j]}(K^{l-1}|L|)^2 ds\right)^p] \\ &\leq k_5 + k_5 E[\left(\int_\tau^T \mathbf{1}_{[\tau,\gamma_j]}\left(K^{l-1}|L|\right)^{l+1} ds\right)^p] \\ &\quad + k_5 E[\left(\int_\tau^T \mathbf{1}_{[\tau,\gamma_j]}(K^{l-1}|L|)^2 ds\right)^{\frac{p}{2}}], \end{aligned}$$

and since $\frac{l+1}{2} < 1$ we can apply Jensen's inequality for concave functions and get

$$\begin{aligned} &E[\left(\int_\tau^T \mathbf{1}_{[\tau,\gamma_j]}(K^{l-1}|L|)^2 ds\right)^p] \\ &\leq k_6 + k_6 E[\left(\int_\tau^T \mathbf{1}_{[\tau,\gamma_j]}\left(K^{l-1}|L|\right)^2 ds\right)^{\frac{(l+1)p}{2}}] \\ &\quad + k_6 E[\left(\int_\tau^T \mathbf{1}_{[\tau,\gamma_j]}(K^{l-1}|L|)^2 ds\right)^{\frac{p}{2}}], \end{aligned} \tag{4.24}$$

where k_6 is independent of j. Since $l<1$, there is a constant $a>0$ depending only on l,p and k_6 such that

$$k_6 + k_6\eta^{\frac{(l+1)p}{2}} + k_6\eta^{\frac{p}{2}} \le a + \frac{1}{2}\eta^p,$$

for all $\eta \ge 0$. Hence, from (4.24) we get (setting $\eta := \int_\tau^T \mathbf{1}_{[\tau,\gamma_j]}(K^{l-1}|L|)^2 ds$)

$$E[\Big(\int_\tau^T \mathbf{1}_{[\tau,\gamma_j]}(K^{l-1}|L|)^2 ds\Big)^p] \le$$
$$a + \frac{1}{2}E[\Big(\int_\tau^T \mathbf{1}_{[\tau,\gamma_j]}(K^{l-1}|L|)^2 ds\Big)^p] \tag{4.25}$$

hence

$$E[\Big(\int_\tau^T \mathbf{1}_{[\tau,\gamma_j]}(K^{l-1}|L|)^2 ds\Big)^p] \le 2a, \tag{4.26}$$

and the constant on the right hand side depends on n,T,q and p only - in particular, it is independent of j. As $\mathbf{1}_{[\tau,\gamma_j]}(K^{l-1}|L|)^2 \uparrow (K^{l-1}|L|)^2$, $Leb \otimes P - a.s.$, $j\to\infty$, the lemma is now proved by applying the Monotone Convergence Theorem. ∎

Corollary 4.15 *Suppose that the assumptions of Theorem 4.14 hold. Then, for all $p>1$ there is a constant $k>0$ depending only on n,q,T and p such that*

$$E[\Big(\int_\tau^T |L|^2 ds\Big)^p] \le k.$$

Proof: Since $l<1$ we have $K^{l-1}|L| \ge n^{l-1}|L|$ and the assertion follows from the preceding theorem. ∎

Let us turn to the case of Assumption A3. In this case G is given explicitly. Set $C = (C^1,\dots,C^d)'$ and let σ be given by (1.16). Then, under Assumption A3 we have

$$\begin{aligned} G(K,L) &= -\left[\sum_{i=1}^d (D^i)'D^i\right]^{-1}\left(\frac{1}{q-1}B' + \sum_{i=1}^d (D^i)'\left(\frac{L^i}{K} + C^i\right)\right) \\ &= -(\sigma\sigma')^{-1}\left(\frac{1}{q-1}B' + \frac{1}{K}\sigma L + \sigma C\right). \end{aligned}$$

We may enter this into the BSRDE (4.17) and firstly calculate that

$$-\left(\frac{1}{q-1}BK + \sum_{i=1}^d D^i\left(C^iK + L^i\right)\right)G(K,L)$$

$$
\begin{aligned}
&= \left(\frac{1}{q-1}BK+\sigma L+\sigma CK\right)'(\sigma\sigma')^{-1}\left(\frac{1}{q-1}B'+\frac{1}{K}\sigma L+\sigma C\right)\\
&= \frac{1}{(q-1)^2}B(\sigma\sigma')^{-1}B'K+\frac{2}{q-1}B(\sigma\sigma')^{-1}\sigma CK+C'\sigma'(\sigma\sigma')^{-1}\sigma CK\\
&\quad+\frac{2}{q-1}B(\sigma\sigma')^{-1}\sigma L+2C'\sigma'(\sigma\sigma')^{-1}\sigma L+\frac{1}{K}L'\sigma'(\sigma\sigma')^{-1}\sigma L,
\end{aligned}
$$

and

$$
\begin{aligned}
&\sum_{i=1}^{d}\left(L^i+KC^i+KD^iG(K,L)\right)^2\\
&\quad= (L+KC+K\sigma'G(K,L))'(L+KC+K\sigma'G(K,L))\\
&\quad= K^2|C|^2+\frac{1}{(q-1)^2}K^2B(\sigma\sigma')^{-1}B'-K^2C'\sigma'(\sigma\sigma')^{-1}\sigma C\\
&\qquad -2KC'\sigma'(\sigma\sigma')^{-1}\sigma L+2KC'L-L'\sigma'(\sigma\sigma')^{-1}\sigma L+|L|^2.
\end{aligned}
$$

If we replace the respective expressions in (4.17) we get, after some simplification, the following form of the BSRDE under Assumption A3:

$$
\begin{aligned}
dK &= \Big\{-q'AK-\frac{q}{2}|C|^2K+\frac{q}{2(q-1)^2}B(\sigma\sigma')^{-1}B'K+q'B(\sigma\sigma')^{-1}\sigma CK\\
&\quad+\frac{q}{2}C'\sigma'(\sigma\sigma')^{-1}\sigma CK-qC'L+q'B(\sigma\sigma')^{-1}\sigma L+qC'\sigma'(\sigma\sigma')^{-1}\sigma L\\
&\quad+\frac{q}{2}\frac{1}{K}L'\sigma'(\sigma\sigma')^{-1}\sigma L+\frac{2-q}{2}\frac{1}{K}|L|^2-\frac{1}{q-1}QK^{2-q}\Big\}\,ds\\
&\quad+L'dw, &(4.27)\\
K(T) &= f(M). &(4.28)
\end{aligned}
$$

Note that we made here no particular use of the fact that $q>2$. Hence, (4.27) also represents the BSRDE for a problem that satisfies Assumption A1 with $N=0$. For (4.27) an analogous statement to Theorem 4.13 holds.

Theorem 4.16 *Assume $q>2$. Let (K,L) satisfy the differential equation (4.27) on $[\tau,T]$, where the coefficients satisfy Assumption A3. Assume that $n\in\mathbb{N}_{\geq 1}$ is such that $(B,(C^i)_i,(D^i)_i,N,K,L)\in\mathcal{D}_n^{(1)}$, $|A|,|Q|\leq n$, $Leb\otimes P-a.s.$, and $|M|\leq n$, $P-a.s.$. Then, for every $p\in[1,\infty)$ there is a $k>0$ depending only on m,d,n,q,p and T such that*

$$E[\left(\int_\tau^T|L(s)|^2ds\right)^p]\leq k.$$

Proof: Let n be such that $Leb\otimes P-a.s.$ $(B,(C^i)_{1\leq i\leq d},(D^i)_{1\leq i\leq d},N,K,L)\in\mathcal{D}_n^{(1)}$. Define γ_j as in the proof of Theorem 4.13. Since $\sigma\sigma'$ is uniformly positive and σ is essentially bounded, there are constants $c_1, c_2 > 0$ such that $\lambda' (\sigma\sigma')^{-1} \lambda \leq c_1|\lambda|^2$ and $|\sigma\mu| \leq c_2|\mu|$, $Leb \otimes P - a.s.$ for all $\lambda \in \mathbb{R}^m$ and $\mu \in \mathbb{R}^d$. Now choose an $r > 0$ such that

$$r^2 + r - c_1 c_2^2 q r \geq 0,$$

and consider the differential of K^{-r}. By Itô's formula we get

$$\begin{aligned} d\left(K^{-r}(t \wedge \gamma_j)\right) &= \mathbf{1}_{[\tau,\gamma_j]} \frac{-r}{K^{r+1}} dK + \mathbf{1}_{[\tau,\gamma_j]} \frac{-r(-r-1)}{2\,K^{r+2}} d < K > \\ &\mathbf{1}_{[\tau,\gamma_j]} \Big\{ rq' A K^{-r} + \frac{rq}{2}|C|^2 K^{-r} - \frac{rq}{2(q-1)^2} B\left(\sigma\sigma'\right)^{-1} B' K^{-r} - rq' B \left(\sigma\sigma'\right)^{-1} \sigma C K^{-r} \\ &- \frac{rq}{2} C'\sigma' \left(\sigma\sigma'\right)^{-1} \sigma C K^{-r} + \frac{r}{q-1} Q K^{1-q-r} + rqK^{-r-1} C' L - rq' K^{-r-1} B \left(\sigma\sigma'\right)^{-1} \sigma L \\ &- rqK^{-r-1} C'\sigma' \left(\sigma\sigma'\right)^{-1} \sigma L - r\frac{q}{2} K^{-r-2} L'\sigma' \left(\sigma\sigma'\right)^{-1} \sigma L - \frac{r(2-q)}{2} K^{-r-2}|L|^2 \Big\}\, ds \\ &- r\mathbf{1}_{[\tau,\gamma_j]} K^{-r-1} L' dw + \frac{1}{2}(-r)(-r-1)\mathbf{1}_{[\tau,\gamma_j]} K^{-r-2}|L|^2 ds. \end{aligned}$$

Setting $t = T$, this yields, for all j

$$\begin{aligned} &\int_\tau^T \mathbf{1}_{[\tau,\gamma_j]} \Big\{ \frac{1}{2}(-r)(-r-1)\mathbf{1}_{[\tau,\gamma_j]} K^{-r-2}|L|^2 - r\frac{q}{2} K^{-r-2} L'\sigma' \left(\sigma\sigma'\right)^{-1} \sigma L \\ &\quad - \frac{r(2-q)}{2} K^{-r-2}|L|^2 \Big\}\, ds \\ = &\quad K^{-r}(\gamma_j \wedge T) - K^{-r}(\tau) - \int_\tau^T \mathbf{1}_{[\tau,\gamma_j]} \Big\{ rq' A K^{-r} + \frac{rq}{2}|C|^2 K^{-r} \\ &\quad - \frac{rq}{2(q-1)^2} B\left(\sigma\sigma'\right)^{-1} B' K^{-r} - rq' B\left(\sigma\sigma'\right)^{-1} \sigma C K^{-r} \\ &\quad - \frac{rq}{2} C'\sigma' \left(\sigma\sigma'\right)^{-1} \sigma C K^{-r} + \frac{r}{q-1} Q K^{1-q-r} + rqK^{-r-1} C' L - rq' K^{-r-1} B\left(\sigma\sigma'\right)^{-1} \sigma L \\ &\quad - rqK^{-r-1} C'\sigma' \left(\sigma\sigma'\right)^{-1} \sigma L \Big\}\, ds - r\int_\tau^T \mathbf{1}_{[\tau,\gamma_j]} K^{-r-1} L' dw. \end{aligned} \tag{4.29}$$

Let us examine the very first line of the above expression, more precisely the integrand. By the definition of c_1 and c_2, we have $Leb \otimes P - a.s.$ for all j

$$\begin{aligned} &\frac{1}{2}(-r)(-r-1)K^{-r-2}|L|^2 - r\frac{q}{2} K^{-r-2} L'\sigma' \left(\sigma\sigma'\right)^{-1} \sigma L \\ &\geq \frac{1}{2}(-r)(-r-1)K^{-r-2}|L|^2 - c_1 r\frac{q}{2} K^{-r-2} L'\sigma'\sigma L \\ &\geq \frac{1}{2}(-r)(-r-1)\mathbf{1}_{[\tau,\gamma_j]} K^{-r-2}|L|^2 - c_1 c_2^2 r\frac{q}{2} K^{-r-2}|L|^2. \end{aligned} \tag{4.30}$$

Taking into account how how r was chosen, it follows that $P - a.s.$

$$\int_\tau^T \mathbf{1}_{[\tau,\gamma_j]} \left\{ \frac{1}{2}(-r)(-r-1)\mathbf{1}_{[\tau,\gamma_j]} K^{-r-2}|L|^2 - r(q-1)K^{-r-2}L'\sigma'\left(\sigma\sigma'\right)^{-1}\sigma L \right) ds \geq 0,$$

for all j. Thus, we get from (4.29) the estimate (note that $q > 2$)

$$\begin{aligned} 0 \leq & \int_\tau^T \mathbf{1}_{[\tau,\gamma_j]} \frac{-r(2-q)}{2} K^{-r-2}|L|^2 ds \\ \leq & \; K^{-r}(\gamma_j \wedge T) - K^{-r}(\tau) - \int_\tau^T \mathbf{1}_{[\tau,\gamma_j]} \left\{ rq'AK^{-r} + \ldots \right\} ds \\ & -r \int_\tau^T \mathbf{1}_{[\tau,\gamma_j]} K^{-r-1}L'dw. \end{aligned} \tag{4.31}$$

Now recall that all the coefficients of the problem, as well as K, are essentially bounded, and that K and $\sigma\sigma'$ are uniformly positive. All upper and lower bounds depend on n. Given that r is chosen depending on n and q, from (4.31) it now follows that there are constants $k_i > 0$ depending only on n, q and T such that for all j we have $P - a.s.$ (note that K is uniformly positive)

$$\begin{aligned} & k_1 \left(1 + \int_\tau^T \mathbf{1}_{[\tau,\gamma_j]} |L| ds + \left| \int_\tau^T \mathbf{1}_{[\tau,\gamma_j]} K^{-r-1}L'dw \right| \right) \\ \geq & \; \frac{-r(2-q)}{2} \int_\tau^T \mathbf{1}_{[\tau,\gamma_j]} K^{-r-2}|L|^2 ds \\ \geq & \; k_2 \int_\tau^T \mathbf{1}_{[\tau,\gamma_j]} |L|^2 ds. \end{aligned}$$

The proof now follows exactly the same pattern as the proof of Theorem 4.13 after (4.20). ■

Remarks on Chapter 4 In our presentation we have chosen not to postulate the BSRDE, but to develop it through the calculation of the differential of $\frac{f(y)}{x}$. The equation might also have been derived via the Dynamic Programming Principle; see [YZ:SC], Chapter 6, for an application of this principle to linear quadratic problems with deterministic coefficients. We have done this for some special cases. But as our approach completely relies on the FBSDE (as a special form of the Maximum Principle) we have not included our (heuristic) calculations using Hamilton-Jacobi-Bellman equations.
The results of Section 4.1 were necessary to make the calculation of the differential of $\frac{f(y)}{x}$ rigorous on the interval $[\tau, \tau_0)$. Yet, these results are an example of the mutual benefits that a parallel investigation of stochastic linear isoelastic control problems and stochastic Riccati equations may yield. In our context it may seem quite artificial, in some sense, to oppose control problems and Riccati-equations, since in our presentation the control problems are the essential reason for studying the Riccati-equation. However, in the quadratic

case, the long-standing open question of the solvability of the Riccati equation led to the development of highly sophisticated analytic techniques for treating Riccati equations, and so these equations developed “a life of their own right”, and (to avoid misunderstandings) we do not deny this right in any way. In the quadratic case, there are of course further applications of Riccati equations, for example in the theory of stabilizing systems and filter theory (the latter one is quite a “Hilbert-space-theory”, so Riccati equations for $q \neq 2$ are not expected to be useful there).

The results for Section 4.1 are an extension of the “quadratic” theory, as found in [B:LQOC]. The new result is essentially to consider the representation $y = \varphi(Kx)$ instead of $y = Kx$. The statement of a “non-quadratic” BSRDE (i.e. a BSRDE for non-linear-quadratic problems) is new. As already mentioned, the Theorems 4.13, 4.14 and 4.16 are non-trivial generalizations to a non-quadratic BSRDE of a result of Tang, see Theorem 5.1 in [T:GLQO].

Chapter 5

Unique solvability, representation of the optimal control and the optimal cost

In this chapter we will show that the BSRDE is uniquely solvable under Assumptions A1, A3 and A4. For quadratic equations ($q = 2$) there have been various approaches to handle the problem of solvability via successive approximation. Due to the high non-linearity of the equation, this methods generally involve very demanding estimates for the approximating sequence. The method used here is, to the best of our knowledge, new and was developed independently by Tang (see [T:GLQO]) and the author (Tang considered linear quadratic problems with a n-dimensional state equation).
The method is based on the following observation: given the solution of a BSRDE, one can construct the optimal state x and the optimal control u, as well as the solution (y, z) of the adjoint equation. But one can also reverse this construction: Given the solution (x, u, y, z) of the FBSDE (that is known to exist and to satisfy the auxiliary condition (3.14)), one *defines* the processes K and L by

$$K := \frac{f(y)}{x}, \quad L^i := \frac{f'(y)z^i}{x} - C^i \frac{f(y)}{x} - D^i \frac{f(y)}{x^2} u, \quad i = 1, \ldots, d, \tag{5.1}$$

and tries to show that (K, L) actually is "the" solution of the BSRDE. As already mentioned, one main problem here is to show that x does not vanish - given this, it would be clear from the continuity of x and y, that L as constructed above is pathwise square integrable. Along with the results of Chapter 4, this would immediately show that (K, L) is a solution of the BSRDE (4.17), (4.18).
Yet, we proceed slightly different and first show in Section 5.1 that L is "H_p-integrable" on $[\tau, \tau_0)$. This will entitle us to proof $\tau_0 = T$ and hence solvabilty of the BSRDE in Section 5.2. Section 5.3 will address the question whether the solution of the BSRDE is unique. It will turn out that this is question intimately related to the problem of constructing a solution for our control problem $\mathcal{P}(\tau, 1)$ out of an *arbitrary* solution of the BSRDE.

5.1 Integrability of L

Recall the discussion preceding Theorem 4.13. There, we argued why it would be of much importance to know that $\int_\tau^{\tau_0} |L|^2 ds < \infty,\ P-a.s.$, where L is given by (5.1). In this section we will get an even stronger integrability of L with the help of the a-priori estimates of Theorem 4.13, Corollary 4.15 and Theorem 4.16. This is done by a truncation procedure that we take from [T:GLQO] and that carries over straightforward to the BSRDE with $q \neq 2$.
The method is to "stop" the process x when he reaches the level $\frac{1}{j}$. It turns out that one can construct a family of control problems whose optimal state actually is this stopped process. The optimal state of these modified problems never reaches zero. Roughly speaking, this means that everything we proved so far to hold "locally" on $[\tau, \tau_0)$ for the original problem holds on $[\tau, T]$ for the modified problem.
Consider (as "original" problem) problem $\mathcal{P}(\tau, 1)$ and let one of the Assumptions A1, A3 or A4 hold. Remember that (x, u, y, z) is the solution of the FBSDE (3.11)-(3.14) for problem $\mathcal{P}(\tau, 1)$. In particular we have $x(\tau) = 1$.
For $j \in \mathbb{N}_{>1}$ let us introduce the stopping times

$$\tau_j := \inf\{t > \tau : |x(t)| \leq \frac{1}{j}\} \wedge T,\ \inf \emptyset := \infty. \tag{5.2}$$

Further, for $j \in \mathbb{N}_{>1}$ and $t \in [0, T]$ set

$$\begin{aligned} x_j(t \vee \tau) &:= x(\tau \vee (t \wedge \tau_j)),\ u_j := \mathbf{1}_{[\tau,\tau_j]}, \\ y_j(t \vee \tau) &:= y(\tau \vee (t \wedge \tau_j)),\ z_j := \mathbf{1}_{[\tau,\tau_j]}, \end{aligned}$$

as well as on $[\tau, T]$

$$\begin{aligned} &A_j := \mathbf{1}_{[\tau,\tau_j]} A,\ B_j := \mathbf{1}_{[\tau,\tau_j]} B, \\ &C_j := \mathbf{1}_{[\tau,\tau_j]} C = \mathbf{1}_{[\tau,\tau_j]} (C^1, \dots, C^d) =: (C_j^1, \dots, C_j^d), \\ &Q_j := \mathbf{1}_{[\tau,\tau_j]} Q, M_j := \varphi(K_j(T))\ (K_j \text{ see below.}) \end{aligned}$$

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

$$\begin{aligned} &K_j(t \vee \tau) := \frac{f(y_j(t \vee \tau))}{x_j(t \vee \tau)}, \\ &L_j^i(t \vee \tau) := \frac{f'(y_j(t \vee \tau))}{x_j(t \vee \tau)} z_j^i - C_j^i(t \vee \tau) K_j(t \vee \tau) \\ &-D^i(t \vee \tau) K_j(t \vee \tau) \frac{1}{x_j(t \vee \tau)} u_j(t \vee \tau), \\ &L_j(t \vee \tau) := (L_j^1(t \vee \tau), \dots, L_j^d(t \vee \tau)). \end{aligned}$$

Note that due to the definition of τ_j these quantities are well defined, i.e. $\frac{1}{x_j}$ and $f'(y_j)$ are well defined on $[\tau, T]$ for all j. We want to show that (K_j, L_j) actually is a solution of the BSRDE (4.17), (4.18) with coefficients $A_j, B_j, (C_j^i)_{1 \leq i \leq d}, (D^i)_{1 \leq i \leq d}, Q_j, N$ and M_j. Let us introduce the control problems corresponding to these coefficients.

Definition 5.1 *(Problem $\mathcal{P}_j(\tau,1)$)*
The problem $\mathcal{P}_j(\tau,1)$ is the linear isoelastic stochastic control problem with coefficients $A_j, B_j, (C_j^i)_{1\leq i\leq d}, (D^i)_{1\leq i\leq d}, Q_j, N$ and M_j.

From Lemma 4.2 we see that K_j is essentially bounded, Lemma 4.5 gives, that K_j is uniformly positive if Assumption A1 or A3 holds respectively strictly positive if Assumption A4 holds. As a consequence, M_j is uniformly respectively strictly positive if Assumption A1 or A3 respectively A4 holds. Note that the coefficients $(D^i)_{1\leq i\leq d}$ and N remain unchanged in problem $\mathcal{P}_j(\tau,1)$. It follows that the coefficients of problem $\mathcal{P}_j(\tau,1)$ satisfy Assumption A1 or A3 respectively A4 if those of problem $\mathcal{P}(\tau,1)$ do. Stopping the processes x and y shows that x_j and y_j have the differentials

$$\begin{aligned} dx_j &= \{A_j x_j + B_j u_j\}\, ds + \sum_{i=1}^d \{C_j^i x_j + D^i u_j\}\, dw^i, \\ dy_j &= \left\{-A_j y_j - \sum_{i=1}^d C_j^i z_j^i - Q_j\varphi(x_j)\right\} ds + \sum_{i=1}^d z_j^i dw^i, \\ x_j(\tau) = 1, &\qquad y_j(T) = M_j\varphi(x_j(T)). \end{aligned}$$

In addition, multiplying (3.14) with $\mathbf{1}_{[\tau,\tau_j]}$ gives

$$B_j y_j + \sum_{i=1}^d D^i z_j^i + N\varphi(u_j) = 0, \quad Leb\otimes P - a.s., \tag{5.3}$$

hence (x_j, u_j, y_j, z_j) turns out to be the solution of the FBSDE (3.11)-(3.14) corresponding to problem $\mathcal{P}_j(\tau,1)$, i.e x_j is the optimal state and u_j the optimal control for this problem, see Proposition 3.5 and its Corollary. The "trick" of this truncation now is, of course, that x_j does not vanish, hence all calculations and statements that hold on $[\tau,\tau_0)$ actually hold on $[\tau,T)$.
As $x_j(t)\geq \frac{1}{j}$ we get that

$$\int_\tau^T |L_j(s)|^2 ds < \infty, \;\; P - a.s. \tag{5.4}$$

for all j.
It is clear, that for all j there is a **n independent of j** such that

$$(B_j, (C_j^i)_{1\leq i\leq d}, (D^i)_{1\leq i\leq d}, N, K_j, L_j) \in \mathcal{D}_n^{(1)}, \quad Leb\otimes P - a.s. \text{ on } [\tau,T] \tag{5.5}$$

if Assumption A1 or A3 holds and

$$(B_j, (C_j^i)_{1\leq i\leq d}, (D^i)_{1\leq i\leq d}, N, K_j, L_j) \in \mathcal{D}_n^{(4)}, \quad Leb\otimes P - a.s \text{ on } [\tau,T] \tag{5.6}$$

if Assumption A4 holds. Let us write
$G_j(K_j,L_j)$ for $G(B_j,(C_j^i)_{1\leq i\leq d},(D^i)_{1\leq i\leq d},N,K_j,L_j)$. As in Section 4.2 we may cast (5.3)

in the form of (4.9) (with $B, (C^i), \ldots$ replaced by $B_j, (C^i_j), \ldots$), and Lemma 4.8 then shows, that $\frac{1}{x_j}u_j = G_j(K_j, L_j)$, $Leb \otimes P - a.s.$ on $[\tau, T]$ (note the uniqueness statement on G). Calculating dK_j yields

Lemma 5.2 *For every j, (K_j, L_j) is a solution of the BSRDE (4.17), (4.18) for the coefficients $A, B_j, (C^i_j)_{1\le i\le d}, (D^i)_{1\le i\le d}, Q_j, N$ and M_j of problem $\mathcal{P}_j(\tau, 1)$.*

Proof: The proof is just a summing-up of the preceding arguments. K_j is essentially bounded and uniformly positive respectively strictly positive if Assumption A1, A3 respectively A4 holds; the process L_j is $P - a.s.$ pathwise square integrable. Hence properties 1 and 2 of Definition 4.12 are satisfied, and we have to look at the differential of K_j. The calculations of Section 4.2 (applied to the situation of problem $\mathcal{P}_j(\tau, 1)$) show at first instance, that the differential of K_j is given on $[\tau, T]$ by a modification of (4.8), where the modification consists in substituting $A, B, (C^i), Q, K, L, u$ and x with $A_j, B_j, (C^i_j), Q_j, K_j, L_j, u_j$ and x_j. The terminal condition $K_j(T) = f(M_j)$ is obvious. By Lemma 4.8, condition (5.3) yields that $\frac{1}{x_j}u_j = G(K_j, L_j)$, $Leb \otimes P - a.s.$. Substituting this into the modified (4.8) shows that K_j follows the desired differential equation, and the proof is finished. ■

This now yields what we wanted to know about the unmodified L. Note that we have defined L only on $[\tau, \tau_0)$

Theorem 5.3 *Let problem $\mathcal{P}(\tau, 1)$ be given, and let Assumption A1, A3 or A4 hold. Consider L as defined in (4.5) and extend it arbitrarily (but measurable and adapted) to $[\tau, \tau_0]$. For every $p > 1$ we have*

$$E\Big[\Big(\int_\tau^{\tau_0} |L|^2 ds\Big)^p\Big] < \infty.$$

Proof: Fix a $p > 1$. Consider the stopping times τ_j from (5.2), the problems $\mathcal{P}_j(\tau, 1)$ and the processes (K_j, L_j) introduced in this section. Lemma 5.2 shows that the (K_j, L_j) satisfy a BSRDE. From (5.5) respectively (5.6) we know that we can apply Theorems 4.13, 4.16 respectively Corollary 4.15 to (K_j, L_j), hence there is a $k > 0$, independent of j, such that $E[(\int_\tau^{\tau_0} |L_j|^2 ds)^p] < k$. As $\tau_j \uparrow \tau_0$, $j \to \infty$, we have $|L_j| \uparrow |L|$, $Leb \otimes P - a.s.$ on $[\tau, \tau_0]$, since every term of the sum that defines L_j contains an indicator of $[\tau, \tau_j]$. The assertion now follows from the Monotone Convergence Theorem. ■

5.2 Solvability of the BSRDE

Almost all ingredients for the proof of solvability are available now. Yet, before combining them in the proof of Theorem 5.5, let us recall a well known result on the solution of linear, homogeneous SDE (see for example [WW:SI], Thm. 8.2.1).

Remark 5.4 *Let a and b^i, $i = 1, \ldots, d$, be adapted, real valued processes with $\int_\tau^T |a(s)|ds < \infty$, $P-a.s.$, and $\int_\tau^T |b^i(s)|^2 ds < \infty$, $P-a.s.$, $i = 1, \ldots, d$. Suppose $\zeta_\tau \in \mathbb{R}$, $\zeta_\tau \neq 0$. Then, the SDE*

$$\begin{aligned} d\zeta &= a\zeta ds + \sum_{i=1}^d b^i \zeta dw^i, \\ \zeta(\tau) &= \zeta_\tau \end{aligned}$$

possesses a unique strong solution ζ on $[\tau, T]$ such that $P-a.s.$ $\zeta(t) \neq 0$ for all $t \in [\tau, T]$.

Proof: Set

$$\zeta_1(t \vee \tau) := \zeta_\tau \exp\left\{ \int_\tau^{t\vee\tau} a - \frac{1}{2}|b|^2 ds + \sum_{i=1}^d \int_\tau^{t\vee\tau} b^i dw^i \right\}.$$

This obviously is a solution for the SDE. Its inverse $\zeta_0 = \frac{1}{\zeta_1}$ satisfies the differential equation

$$d\zeta_0 = (-a + |b|^2)\zeta_0 ds - b\zeta_0 dw, \quad \zeta_0(\tau) = \frac{1}{\zeta_\tau}.$$

Let ζ_2 be another solution for the equation in the statement of the remark. By Itô's formula, $d(\zeta_2\zeta_0) = 0$; as $(\zeta_2\zeta_0)(\tau) = 1$ we have $P-a.s.$ $\zeta_2\zeta_0 = 1$ for all t, hence $P-a.s.$ $\zeta_1(t) = \zeta_2(t)$ for all t. This proves the remark. ■

This yields one of our central results. We must exclude the "conventional" case $q = 2$, since it is not covered by Theorem 4.13, Theorem 4.14 or Theorem 4.16. However, the papers [T:GLQO] and [KT:MBSR] show that in the quadratic case the *multidimensional* BSRDE is solvable in case of Assumption A1 respectively in a special case of Assumption A4.

Theorem 5.5 *Let Assumption A1, A3 or A4 hold with $q \neq 2$.*

1. *We have $\tau_0 = T$, $P-a.s.$, and $x(T) \neq 0$, $P-a.s.$.*
2. *The BSRDE (4.17), (4.18) has a solution (K, L), such that $L \in H_p(\tau, T; \mathbb{R}^d)$ for every $p > 1$.*

Proof: Let K and L be given by (4.3) and (4.5).

1. The proof follows the plan briefly outlined after Definition 4.12.
   Recall equality (4.9), i.e. the representation of $\frac{1}{x}u$ in terms of $A, B, \ldots, K$ and L. Along with Lemma 4.8 this shows that $\frac{1}{x}u = G(K, L)$ respectively $u = G(K, L)x$ on $[\tau, \tau_0)$ (keeping Notation 4.11 in mind). Also recall that from Lemma 3.8 we get, that $u = 0$ on $(\tau_0, T]$ and $x = 0$ on $[\tau_0, T]\backslash\{(T, \omega) : x(T) \neq 0\}$.
   Extend (K, L) to $[\tau, \tau_0]$ such that the pair remains adapted and K remains positive

and uniformly bounded. Define the processes a and b by $a = A+BG(K,L)$ on $[\tau,\tau_0]$, $a = 0$ on $(\tau_0,T]$ and $b^i = C^i + D^iG(K,L)$ on $[\tau,\tau_0]$, $b^i = 0$ on $(\tau_0,T]$. The SDE for the optimal state thus reads as

$$dx = axds + bxdw, \quad x(\tau) = 1.$$

We want to apply the above remark and thus have to check the integrability of a and the b^i. Combining Lemma 4.9-2 respectively -3 with Theorem 5.3 shows that $\int_\tau^T |a|ds < \infty$, $P-a.s.$, and $\int_\tau^T |b|^2ds < \infty$, $P-a.s.$. From the above remark it now follows that x is invertible, hence $\tau_0 = T$, $P-a.s.$, and $x(T) > 0$, $P-a.s.$.

2. From the first part we know that $\tau_0 = T$ and $x(T) > 0$, $P-a.s.$. Hence, $K(T) = \frac{f(y(T))}{x(T)}$ is well defined, and by (3.13) it equals $f(M)$, i.e. (4.18) is satisfied. From the calculations of Subsection 4.2 we know, that (4.8) holds on $[\tau,T)$. This, by continuity of K, of course extends to $[\tau,T]$. Now we may replace $\frac{1}{x}u$ by $G(K,L)$ in (4.8) (the reasoning is the same as in the first part of the proof). This yields (4.17). The statement about L follows from Theorem 5.3, and the theorem is shown.

■

The above result says nothing whether the solution of the BSRDE is unique. This question is tackled in the next section. There, in the interplay between the control problem and the BSDRE, we will somehow change direction: in the existence part we found that one may construct a solution of the BSRDE out the solution of the control problem. Uniqueness will follow from the fact that a solution of the control problem can be constructed out of an arbitrary solution of the BSRDE.
But we do not want to suggest that we proved an equivalence between the solvability of the BSRDE and the corresponding control problem; as we could guarantee the existence the function G only if one of the Assumptions A1, A3 or A4 holds, we automatically had to restrict our considerations about BSRDE to settings in which the solvability of the control problem is granted.

5.3 Uniqueness and representation of the optimal control

In the quadratic case $q = 2$, the knowledge of a solution of the BSRDE entitles us to completely solve the corresponding problem of control. The core motivation for our investigation on "non-quadratic" BSRDE was that we hoped to find a similar situation in the case of a linear isoelastic control problem. What we have seen so far is more or less the other way round - knowing the solution (x,u,y,z) of the FBSDE (3.11)-(3.13), we can construct a solution of the BSRDE (4.17), (4.18) via (4.3), (4.5). The question now is: when we are given a solution (K_0,L_0) of the BSRDE, can we construct (x,u,y,z) out of this? Since we are mainly interested in the optimal state x and the optimal control u this

question may seem overdone; yet, if we just know u, then x is given automatically through (3.11), and (y, z) is easily derived from (3.12); hence, concerning the theory, it makes no big difference whether we look for u or (x, u, y, z).
If we knew that the solution of the BSRDE was unique, the answer would be easy: as we have identified the unique solution to be (K, L) given by (4.3), (4.5), we get $(K_0, L_0) = (K, L)$, hence $u = G(K_0, L_0)x$, what gives (x, u). The pair (y, z) then is given through (3.12) (along with its terminal condition) - alternatively one may set $y = \varphi(K_0 x)$ and then finally resolve the representation (4.5) to get the z^i.
Yet, as we have not proved uniqueness of the solution of the BSRDE so far, we must look out for another possibility. One way is to plug in the ansatz

$$X := G(K_0, L_0)U \tag{5.7}$$

into the state equation (3.11) and to check out (via the FBSDE), whether the resulting pair (X, U) is the optimal state and the optimal control for the control problem. If it indeed is, one can easily see that this fact entails the unique solvability of the BSRDE (the details are carried out at the end of this section). The technical problem of this approach is that at first instance there is no evidence that U constructed in this way actually belongs to $\mathcal{U}$. Hence, we are lead to proceed in two steps. First, we *assume* that the process U we get from (5.7) is in $\mathcal{U}$. In the second step we introduce a family of "localized" problems to which we may apply the result from step one (the localization is quite similar to that of Section 5.1). Approximating U by the optimal controls of the localized problems (whose norms turn out to be uniformly bounded) then yields $U \in \mathcal{U}$. The idea of such an localization is inspired by [KT:MLQ]. So, let us start with the case that (5.7) yields an $U \in \mathcal{U}$.

Lemma 5.6 *Given Assumption A1, A3 or A4, consider a problem $\mathcal{P}(\tau, 1)$. Let (K_0, L_0) be a solution of the corresponding BSRDE (4.17), (4.18). Define the process X as the unique strong solution of*

$$dX = \{A + BG(K_0, L_0)\} X ds + \sum_{i=1}^{d} \left\{C^i + D^i G(K_0, L_0)\right\} X dw^i, \quad X(\tau) = 1,$$

as well as

$$U := G(K_0, L_0)X, \; Y := \varphi(K_0 X), \; Z^i := (q-1)\varphi(X)K_0^{q-2} \left(C^i K_0 + D^i K_0 G(K_0, L_0) + L_0^i\right).$$

If $U \in \mathcal{U}$ (and consequently $X \in L^q_{\mathcal{F}}(\Omega, C([\tau, T]; \mathbb{R}))$), then U is the optimal control and X is the optimal state for problem $\mathcal{P}(\tau, 1)$. $(X, U, Y, (Z^i)_{1 \le i \le d})$ satisfies the FBSDE (3.11)-(3.13).

Proof: First note that Theorems 4.13, 4.16 and Corollary 4.15, along with Lemma 4.9 show that the SDE for X satisfies the hypothesis of Remark 5.4. Hence this equation has a unique strong solution, namely

$$X(t \vee \tau) = \exp\left(\int_\tau^{t \vee \tau} \left\{A + BG(K_0, L_0) - \frac{1}{2}\sum_{i=1}^{d} \left(C^i + D^i G(K_0, L_0)\right)^2\right\} ds\right.$$

$$+\sum_{i=1}^{d}\int_{\tau}^{t\vee\tau}\left\{C^i+D^iG(K_0,L_0)\right\}dw^i\Bigg). \tag{5.8}$$

We will show that the quadruple (X,U,Y,Z) is the solution of the FBSDE (3.11)-(3.13) corresponding to problem $\mathcal{P}(\tau,1)$ (with $Z:=(Z^1,\ldots,Z^d)$). This, of course, then implies (by Proposition 3.5), that (X,U) is the optimal state and and control for problem $\mathcal{P}(\tau,1)$. To start, we will check the regularity of X,U and Y. By hypothesis, we have $U\in\mathcal{U}$, hence $X\in L^q_{\mathcal{F}}(\Omega,C([\tau,T];\mathbb{R}))$. As K_0 is continuous and essentially bounded, we get $Y=\varphi(K_0X)\in L^{q'}_{\mathcal{F}}(\Omega,C([\tau,T];\mathbb{R}))$. Let us turn to Z. From Lemma 4.9, Theorem 4.13, Corollary 4.15 and Theorem 4.16 we see that $D^iK_0G(K_0,L_0)+L_0^i$ belongs to $H_p(\tau,T;\mathbb{R})$ for every $p\in[1,\infty)$. As K_0^{q-2} is essentially bounded (K_0 is required to be uniformly positive if $q<2$) and $\varphi(X)\in L^{q'}_{\mathcal{F}}(\Omega,C([\tau,T];\mathbb{R}))$, we may so far conclude by Hölder's inequality that $Z\in H_{p_1}(\tau,T;\mathbb{R}^d)$ for any p_1 with $1\le p_1<q'$.
Next we show that (Y,Z) solves BSDE (3.12) with the terminal condition specified in (3.13). As $K_0X>0$ we may apply Itô's formula to $\varphi(K_0X)$ also for $q<3$. We will use the representations $(\partial\varphi)(K_0X)=(q-1)\frac{\varphi(K_0X)}{K_0X}$, $(\partial^2\varphi)(K_0X)=(q-1)(q-2)\frac{\varphi(K_0X)}{(K_0X)^2}$, along with the abbreviations $G_0:=G(K_0,L_0)$ and $l=\frac{1}{q-1}$.

$$\begin{aligned}
&d\left(\varphi(K_0X)\right)\\
&= \Bigg\{-A\varphi(K_0X)-QK_0^{2-q}\frac{\varphi(K_0X)}{K_0}-\frac{1}{l}\sum_{i=1}^{d}(C^i)^2\varphi(K_0X)\\
&\quad -\frac{2}{l}\sum_{i=1}^{d}C^iL_0^i\frac{\varphi(K_0X)}{K_0}+BG_0\varphi(K_0X)+\frac{\varphi(K_0X)}{lK_0}\sum_{i=1}^{d}L_0^i\left(C^i+D^iG_0\right)\\
&\quad -\frac{1}{l}\left(lBG_0\varphi\left(K_0X\right)+\frac{\varphi(K_0X)}{lK_0}\sum_{i=1}^{d}D^i\left(C^iK_0+L_0^i\right)G_0\right)\Bigg\}ds\\
&\quad +\frac{1}{l}\sum_{i=1}^{d}\left\{\left(C^i+D^iG_0\right)\varphi(K_0X)+\frac{\varphi(K_0X)}{K_0}L_0^i\right\}dw^i\\
&= \left\{-A\varphi(K_0X)-\frac{\varphi(K_0X)}{l}\sum_{i=1}^{d}C^i\left(C^i+D^iG_0+\frac{L_0^i}{K_0}\right)-QK_0^{2-q}\frac{\varphi(K_0X)}{K_0}\right\}ds\\
&\quad +\sum_{i=1}^{d}Z^idw^i\\
&= \left\{-AY-\sum_{i=1}^{d}C^iZ^i-Q\varphi(X)\right\}ds+\sum_{i=1}^{d}Z^idw^i,
\end{aligned} \tag{5.9}$$

and Y satisfies the terminal condition

$$Y(T)=M\varphi(X(T)). \tag{5.10}$$

Note that (5.9) is just (3.12) and (5.10) is the corresponding terminal condition, but still we have to show, that $Z \in H_{q'}(\tau, T; \mathbb{R}^d)$. Consider (5.9), (5.10) as an "isolated" BSDE. As $Q\varphi(X) \in L^{q'}_{\mathcal{F}}(\Omega, C([\tau, T]; \mathbb{R}))$, $M\varphi(X(T)) \in L^{q'}_{\mathcal{F}_T}(\mathbb{R})$, and the coefficients are uniformly bounded, this BSDE has a unique solution $(\eta, \zeta) \in L^{q'}_{\mathcal{F}}(\Omega, C([\tau, T]; \mathbb{R})) \times H_{q'}(\tau, T; \mathbb{R}^d)$, see [EPQ:BSDE], Thm. 5.1. For any $p_1 \in [1, q')$ we have $L^{q'}_{\mathcal{F}}(\Omega, C([\tau, T]; \mathbb{R})) \subset L^{p_1}_{\mathcal{F}}(\Omega, C([\tau, T]; \mathbb{R}))$ and $L^{q'}_{\mathcal{F}_T}(\mathbb{R}) \subset L^{p_1}_{\mathcal{F}_T}(\mathbb{R})$ as inclusion of sets. Hence there is a unique solution in $L^{p_1}_{\mathcal{F}}(\Omega, C([\tau, T]; \mathbb{R})) \times H_{p_1}(\tau, T; \mathbb{R}^d)$, which must coincide with (η, ζ), as this pair also belongs to $L^{p_1}_{\mathcal{F}}(\Omega, C([\tau, T]; \mathbb{R})) \times H_{p_1}(\tau, T; \mathbb{R}^d)$. Now note that we know (Y, Z) is in $L^{p_1}_{\mathcal{F}}(\Omega, C([\tau, T]; \mathbb{R})) \times H_{p_1}(\tau, T; \mathbb{R}^d)$, and as the above calculation shows, this pair of processes satisfies the BSDE (5.9), (5.10). By this, we get $(\eta, \zeta) = (Y, Z)$, and in particular $Z \in H_{q'}(\tau, T; \mathbb{R}^d)$.
Finally, we have to show that U, Y and Z satisfy the auxiliary condition (3.14) of the FBSDE. Keeping in mind that X is strictly positive we may calculate

$$\begin{aligned}
& B'Y + \sum_{i=1}^{d} (D^i)' Z^i + N\varphi(U) = 0 \\
:\Longleftrightarrow \quad & B'\varphi(K_0 X) + (q-1) \sum_{i=1}^{d} (D^i)' \left(\varphi(X) K_0^{q-2} \left(C^i K_0 + D^i K_0 G(K_0, L_0) + L_0^i \right) \right) \\
& + N\varphi(G(K_0, L_0) X) = 0 \\
\overset{:\varphi(X)}{\Longleftrightarrow} \quad & B'\varphi(K_0) + (q-1) K_0^{q-2} \sum_{i=1}^{d} (D^i)' \left(C^i K_0 + D^i K_0 G(K_0, L_0) + L_0^i \right) \\
& + N\varphi(G(K_0, L_0)) = 0 \\
\Longleftrightarrow \quad & B'\varphi(K_0) + (q-1)\varphi(K_0) \sum_{i=1}^{d} (D^i)' \left(\frac{L_0^i}{K_0} + C^i \right) \\
& + (q-1)\varphi(K_0) \left(\sum_{i=1}^{d} (D^i)' D^i \right) G(K_0, L_0) + N\varphi(G(K_0, L_0)) = 0.
\end{aligned}$$

The statement in the last line is true, as it is the very definition of $G = G(B, (C^i)_{1 \le i \le d}, (D^i)_{1 \le i \le d}, N, K_0, L_0)$, as introduced in Lemma 4.8. Hence U, Y and Z satisfy the auxiliary condition (3.14), and the lemma is shown. ■

The following corollary will be very helpful both in getting rid of the assumption $U \in \mathcal{U}$ and showing uniqueness of the solution of the BSRDE.

Corollary 5.7 *Consider problem $\mathcal{P}(\tau, 1)$ and let the assertions of Lemma 5.6 hold. Then the optimization problem $\mathcal{P}(\tau, h)$ (with the same coefficients as $\mathcal{P}(\tau, 1)$ and an initial value $h \in L^q_{\mathcal{F}_\tau}(\mathbb{R})$) has the optimal cost $\frac{1}{q} E[K_0^{q-1}(\tau) |h|^q]$.*

Proof: The optimal state and the adjoint process for problem $\mathcal{P}(\tau, 1)$ are given by X and Y as defined in the statement of the above Lemma. Let $\bar{y}^{\tau, h}$ be the adjoint process for problem $\mathcal{P}(\tau, h)$. From Lemma 3.13 we have $\bar{y}^{\tau,h} = \varphi(h)Y$. Lemma 3.9 now shows that the optimal cost for problem $\mathcal{P}(\tau, h)$ is given by

$$\frac{1}{q}E[\bar{y}^{\tau,h}(\tau)h] = \frac{1}{q}E[\varphi(h)Y(\tau)h] = \frac{1}{q}E[|h|^q K_0^{q-1}(\tau)],$$

what proves the corollary. ∎

We now want to get rid of the assumption $U \in \mathcal{U}$. For this purpose, let us introduce the following stopping times. This definition is inspired by Lemma 3.10 in [KT:MLQ].

Definition 5.8 *(stopping times T_k)*
Consider a problem $\mathcal{P}(\tau, 1)$ for which one of the Assumptions A1, A3 or A4 holds and let (K_0, L_0) be a solution of the corresponding BSRDE (4.17), (4.18). For $k \in \mathbb{N}_{\geq 1}$ set

$$\begin{aligned} T_k \quad := \quad \inf \Bigg\{ t > \tau : \Bigg| \int_\tau^t A + BG(K_0, L_0) - \frac{1}{2}\sum_{i=1}^d \left(C^i + D^i G(K_0, L_0)\right)^2 ds \\ + \sum_{i=1}^d \int_\tau^t C^i + D^i G(K_0, L_0) dw^i \Bigg| \geq k \Bigg\} \wedge T. \end{aligned} \tag{5.11}$$

Using the T_k, we want to construct a family of control problems whose optimal controls somehow "automatically" happen to belong to $\mathcal{U}$.

Definition 5.9 *(Problem $\mathcal{P}_k(\tau, 1)$, K_k, L_k, G_k, X_k, U_k)*
Consider a problem $\mathcal{P}(\tau, 1)$ for which one of the Assumptions A1, A3 or A4 holds, and let (K_0, L_0) be a solution of the corresponding BSRDE (4.17),(4.18). Let the T_k be defined as in Definition 5.8. Now, depending on the Assumption in force, set

- *If Assumption A1 or A3 holds:*
  $A_k := \mathbf{1}_{[\tau,T_k]}A$, $\quad B_k := \mathbf{1}_{[\tau,T_k]}B$, $\quad C_k^i := \mathbf{1}_{[\tau,T_k]}C^i$, $\quad D_k^i := D^i$,
  $N_k := \mathbf{1}_{[\tau,T_k]}N$, $\quad Q_k := \mathbf{1}_{[\tau,T_k]}Q$, $\quad M_k := \varphi(K_0(T_k))$,

- *If Assumption A4 holds:*
  $A_k := \mathbf{1}_{[\tau,T_k]}A$, $\quad B_k := \mathbf{1}_{[\tau,T_k]}B$, $\quad C_k^i := \mathbf{1}_{[\tau,T_k]}C^i$, $\quad D_k^i := \mathbf{1}_{[\tau,T_k]}D^i$,
  $N_k := N$, $\quad Q_k := \mathbf{1}_{[\tau,T_k]}Q$, $\quad M_k := \varphi(K_0(T_k))$.

The problem $\mathcal{P}_k(\tau, 1)$ is defined as the problem $\mathcal{P}(\tau, 1)$ after replacing the coefficients $A, B, (C^i)_{1\leq i\leq d}, (D^i)_{1\leq i\leq d}, Q, N$ and M by $A_k, B_k, (C_k^i)_{1\leq i\leq d}, (D_k^i)_{1\leq i\leq d}, Q_k, N_k$ and M_k. Further, set

$$K_k := K_0(\cdot \wedge T_k),\ L_k := \mathbf{1}_{[\tau,T_k]}L_0,$$

and

$$G_k := G_k(K_k, L_k) := G(B_k, (C_k^i)_{1\leq i\leq d}, (D_k^i)_{1\leq i\leq d}, N_k, K_k, L_k).$$

Finally, let X_k be the solution of

$$dX_k = \{A_k + B_k G_k\} X_k ds + \sum_{i=1}^{d} \left\{C_k^i + D_k^i G_k\right\} X_k dw^i, \ X_k(\tau) = 1, \tag{5.12}$$

and set $U_k := G_k X_k$.

This definition is quite lengthy and we want to sketch what the plan with these newly defined objects is. We want to show that (K_k, L_k) is a solution for the BSRDE corresponding to problem $\mathcal{P}_k(\tau, 1)$. It will turn out that $X_k = X(\cdot \wedge T_k)$, where X is given by (5.8). This will entail, by the definition of the T_k, that the X_k are essentially bounded. This implies that $U_k \in \mathcal{U}$ for all k, so that we can apply Lemma 5.6, and (X_k, U_k) turns out to be the optimal state and the optimal solution of problem $\mathcal{P}_k(\tau, 1)$. The representation of the optimal cost applied to $\mathcal{P}_k(\tau, 1)$ then yields an upper bound for the H_q- respectively L^q-norms of the U_k. Via a limiting procedure this will lead to $U = G(K_0, L_0)X \in \mathcal{U}$ - the desired result that allows us to apply Lemma 5.6 to problem $\mathcal{P}(\tau, 1)$. The next Lemma makes some of these statements precise and shows that G_k and X_k as introduced above are well defined.

Lemma 5.10 *We place ourselves in the situation of Definition 5.9. For all $k \in \mathbb{N}_{\geq 1}$ we have:*

1. *If Assumption A1 or A3 respectively A4 holds for problem $\mathcal{P}(\tau, 1)$, then*
   - *K_k is uniformly positive respectively strictly positive;*
   - *the same is true for problem $\mathcal{P}_k(\tau, 1)$.*
2. *G_k is well defined and $G_k = \mathbf{1}_{[\tau, T_k]} G(K_0, L_0)$.*
3. *(K_k, L_k) solves the BSRDE (4.17), (4.18) for problem $\mathcal{P}_k(\tau, 1)$.*
4. *Let X be given by (5.6), then $X_k = X(\cdot \wedge T_k)$. In particular, X_k is essentially bounded.*

Proof:

1. - The uniform respectively strict positivity of K_k is inherited from K_0.
   - In the case of Assumption A1 or A3, D_k^i remains unchanged and K_0 is uniformly positive. Hence $\sum_{i=1}^d (D_k^i)'(D_k^i)$ and $\varphi(K_0(T_k))$ are both uniformly positive. Similarly, in the case of Assumption A4, N_k remains unchanged and K_0 is strictly positive, so N_k is uniformly positive and M_k is strictly positive.
2. From 1. we see, that the value of the vector $(B_k, (C_k^i)_{1 \leq i \leq d}, (D_k^i)_{1 \leq i \leq d}, N_k, K_k, L_k)$ is $Leb \otimes P - a.s.$ in the domain of G, hence G_k is well defined. On the stochastic interval $[\tau, T_k]$, $(B_k, (C_k^i)_{1 \leq i \leq d}, (D_k^i)_{1 \leq i \leq d}, N_k, K_k, L_k)$ and $(B, (C^i)_{1 \leq i \leq d}, (D^i)_{1 \leq i \leq d}, N, K, L)$ coincide, hence $G_k = G(K_0, L_0)$ on $[\tau, T_k]$. On $(T_k, T]$, the defining equality for G_k (see (4.10)) reduces to $(q-1)\varphi(K_k)\left(\sum_{i=1}^d (D^i)'D^i\right) G_k = 0$ in case of Assumption A1, and $N\varphi(G_k) = 0$ in the case of Assumption A4, i.e. $G_k = 0 = \mathbf{1}_{[\tau, T_k]} G(K_0, L_0)$ on $(T_k, T]$.

3. We stop K_0 in T_k. Observing part *2* and $A_k K_0 = A_k K_k, \ldots$, yields for every $t \in [0,T]$,

$$\begin{aligned}
&K_0(t \wedge T_k) - K_0(\tau) \\
&= \int_\tau^t \mathbf{1}_{[\tau,T_k]} \Bigg\{ -q' A K_0 - \sum_{i=1}^d (C^i)^2 K_0 - 2\sum_{i=1}^d C^i L_0^i - \frac{q-2}{q-1} B G(K_0,L_0) K_0 \\
&\quad - \left(\frac{1}{q-1} B K_0 + \sum_{i=1}^d D^i \left(C^i K_0 + L_0^i \right) \right) G(K_0,L_0) \\
&\quad - \frac{1}{q-1} Q K_0^{2-q} + \frac{2-q}{2} \frac{1}{K_0} \sum_{i=1}^d \left(L_0^i + K_0 C^i + K_0 D^i G(K_0,L_0) \right)^2 \Bigg\} ds \\
&\quad + \int_\tau^t \mathbf{1}_{[\tau,T_k]} \sum_{i=1}^d L_0^i dw^i \\
&= \int_\tau^t \Bigg\{ -q' A_k K_0 - \sum_{i=1}^d (C_k^i)^2 K_0 - 2\sum_{i=1}^d C_k^i L_k^i - \frac{q-2}{q-1} B_k G_k K_0 \\
&\quad - \left(\frac{1}{q-1} B_k K_0 + \sum_{i=1}^d D_k^i \left(C_k^i K_0 + L_k^i \right) \right) G_k \\
&\quad - \frac{1}{q-1} Q_k K_0^{2-q} + \frac{2-q}{2} \frac{1}{K_0} \sum_{i=1}^d \left(L_k^i + K_0 C_k^i + K_0 D_k^i G_k \right)^2 \Bigg\} ds \\
&\quad + \int_\tau^t \sum_{i=1}^d L_k^i dw^i \\
&= \int_\tau^t \Bigg\{ -q' A_k K_k - \sum_{i=1}^d (C_k^i)^2 K_k - 2\sum_{i=1}^d C_k^i L_k^i - \frac{q-2}{q-1} B_k G_k K_k \\
&\quad - \left(\frac{1}{q-1} B_k K_k + \sum_{i=1}^d D_k^i \left(C_k^i K_k + L_k^i \right) \right) G_k \\
&\quad - \frac{1}{q-1} Q_k K_k^{2-q} + \frac{2-q}{2} \frac{1}{K_k} \sum_{i=1}^d \left(L_k^i + K_k C_k^i + K_k D_k^i G_k \right)^2 \Bigg\} ds \\
&\quad + \int_\tau^t \sum_{i=1}^d L_k^i dw^i.
\end{aligned}$$

As $K_k(T) = M_k$, (K_k, L_k), is thus seen to satisfy the BSRDE for problem $\mathcal{P}_k(\tau, 1)$.

4. As above, let us stop the process under consideration in T_k. Set $X^{T_k} := X(\cdot \wedge T_k)$. This leads for every $t \in [0,T]$ to

$$X^{T_k}(t) - 1$$

$$
\begin{aligned}
&= \int_\tau^t \mathbf{1}_{[\tau,T_k]} \{A + BG(K_0, L_0)\} X ds + \int_\tau^t \mathbf{1}_{[\tau,T_k]} \sum_{i=1}^d \left\{C^i + D^i G(K_0, L_0)\right\} X dw^i \\
&= \int_\tau^t \{A_k + B_k G_k\} X ds + \int_\tau^t \sum_{i=1}^d \left\{C_k^i + D_k^i G_k\right\} X dw^i \\
&= \int_\tau^t \{A_k + B_k G_k\} X^{T_k} ds + \int_\tau^t \sum_{i=1}^d \left\{C_k^i + D_k^i G_k\right\} X^{T_k} dw^i,
\end{aligned}
$$

i.e. X^{T_k} and X_k follow the same SDE. With Remark 5.4 one can see that the SDE (5.12) is uniquely solvable (compare the first step in the proof of Lemma 5.6), hence $X^{T_k} = X_k$. From the representation (5.8) and the definition of T_k we now get that

$$
\begin{aligned}
|X(t \wedge T_k)| \;&\le\; \exp\left\{\left|\int_\tau^{t\wedge T_k} A + BG(K_0, L_0) - \frac{1}{2}\sum_{i=1}^d \left(C^i + D^i G(K_0, L_0)\right)^2 ds \right.\right. \\
&\qquad \left.\left. + \int_\tau^{t\wedge T_k} \sum_{i=1}^d \int_\tau^t C^i + D^i G(K_0, L_0) dw^i \right|\right\} \\
&\le\; \exp\{k\}, \;\; P - a.s.,
\end{aligned}
$$

for all $t \in [0, T]$. This proves the lemma.

■

We now can turn to our second principal result.

Theorem 5.11 *Consider a problem $\mathcal{P}(\tau, 1)$ with $q \neq 2$ and suppose that Assumption A1, A3 or A4 holds. Let (K_0, L_0) be a solution of the corresponding BSRDE (4.17), (4.18), and define X as the unique strong solution of*

$$
dX = \{A + BG(K_0, L_0)\} X ds + \sum_{i=1}^d \left\{C^i + D^i G(K_0, L_0)\right\} X dw^i.
$$

Then X is the optimal state and $U := G(K_0, L_0)X$ is the optimal control for problem $\mathcal{P}(\tau, 1)$. The solution of the corresponding FBSDE (3.11)- (3.14) is given by (X, U, Y, Z) as defined in Lemma 5.6.

Proof: We want to apply Lemma 5.6 and thus have to show that $U \in \mathcal{U}$. The assertions of the theorem then follow.
To this end, consider the problems $\mathcal{P}_k(\tau, 1)$ introduced in Definition 5.9. By Lemma 5.10 - 3, (K_k, L_k) is a solution of the BSRDE for problem $\mathcal{P}_k(\tau, 1)$. Obviously, we have $L_k \in H_p(\tau, T; \mathbb{R}^d)$ for every $p > 1$, since the same is true for L_0 by Theorems 4.13, 4.16 respectively Corollary 4.15, hence $G_k = G(K_k, L_k) \in \mathcal{U}$ by Lemma 4.9. As X_k is essentially

bounded, we get $U_k = G_k X_k \in \mathcal{U}$ and may thus apply Lemma 5.6. In particular, (X_k, U_k) is the optimal state and optimal control for problem $\mathcal{P}_k(\tau, 1)$.
Let J_k be the cost functional (1.11) for problem $\mathcal{P}_k(\tau, 1)$ (i.e. with coefficients Q_k, N_k and M_k). Then we find from Corollary 5.7, that

$$J_k(U_k) = \frac{1}{q} E[K_k^{q-1}(\tau)] = \frac{1}{q} E[K_0^{q-1}(\tau)],$$

where the last equality is due to $\tau < T_k$, $P-a.s.$, i.e. $K_k(\tau) = K_0(\tau \wedge T_k) = K_0(\tau)$, $P-a.s.$. We are now going to show that there is a common bound for $\|U_k\|_{H_q}$ respectively $\|U_k\|_{L^q_{\mathcal{F}}}$ if Assumption A1 or A3 respectively A4 holds.

- Let Assumption A1 or A3 hold. The reasoning of the proof of Lemma 3.2 shows, that there is a $c > 0$, independent of k, such that

$$E[\left(\int_\tau^T |U_k|^2 ds\right)^{\frac{q}{2}}] \leq c\, E[|X_k(T)|^q],$$

  ("coercivity of J"). For the application of this reasoning note that K_0 is uniformly positive, hence there is an $\epsilon > 0$ such that $M_k \geq \epsilon$, $P-a.s.$ for all $k \in \mathbb{N}_{\geq 1}$. Using the same ϵ, we have for all k

$$E[|X_k(T)|^q] \leq \frac{1}{\epsilon} E[M_k |X_k(T)|^q] \leq \frac{q}{\epsilon} J_k(U_k) \leq \frac{1}{\epsilon} E[K_0^{q-1}(\tau).]$$

  Combining the last two inequalities shows, that $\frac{c}{\epsilon} E[K_0^{q-1}(\tau)]$ is an upper bound, independent of k, for $\|U_k\|_{H_q}$.

- Let Assumption A4 hold. Then there is an $\epsilon > 0$ such that $N_k \equiv N \geq \epsilon$, $Leb \otimes P-a.s.$ for all k. Consequently,

$$\epsilon E[\int_\tau^T |U_k|^q ds] \leq E[\int_\tau^T N_k |U_k|^q ds] \leq q\, J_k(U_k) \leq E[K_0^{q-1}(\tau)],$$

  for all $k \in \mathbb{N}_{\geq 1}$, what gives an upper bound, independent of k, for $\|U_k\|_{L^q_{\mathcal{F}}}$.

The final step is a limiting procedure. Note that by Remark 5.10-2. we have $U_k := G_k X_k = \mathbf{1}_{[\tau, T_k]} G(K_0, L_0) X_k = \mathbf{1}_{[\tau, T_k]} G(K_0, L_0) X = \mathbf{1}_{[\tau, T_k]} U$. As $T_k \uparrow T$, $P-a.s.$, $k \to \infty$, this yields $U_k \to U$, $Leb \otimes P-a.s.$, $k \to \infty$. Therefore, the uniform boundedness of the norms of the U_k gives $\|U\|_{H_q} \leq \liminf_k \|U_k\|_{H_q} < \infty$ respectively $\|U\|_{L^q_{\mathcal{F}}} \leq \liminf_k \|U_k\|_{L^q_{\mathcal{F}}} < \infty$ if Assumption A1 or A3 respectively A4 holds. It follows, that $U \in \mathcal{U}$ and the proof is finished. ■

Given this theorem, it is quite easy now to deduce that the BSRDE has a unique solution. We want to point out that the representation of the optimal cost as stated in Corollary 5.7 played an important role in the proof of Theorem 5.11 and also does in the following proof.

Corollary 5.12 *Consider a problem $\mathcal{P}(\tau, 1)$ and let Assumption A1, A3 or A4 hold. Then the solution of the corresponding BSRDE (4.17), (4.18) is unique.*

Proof: Assume that (K_1, L_1) and (K_2, L_2) are two solutions of the BSRDE. Consider the family of subproblems $\mathcal{P}(\gamma, h_\gamma)$, $\tau \leq \gamma < T$, $h_\gamma \in L^q_{\mathcal{F}_\gamma}(\mathbb{R})$, as introduced in Definition 1.3. For these subproblems, Assumption A1, A3 respectively A4 holds if they do for problem $\mathcal{P}(\tau, 1)$. It is clear that the processes (K_1, L_1), (K_2, L_2), restricted to the stochastic interval $[\gamma, T]$, are a solution for the BSRDE (4.17), (4.18) corresponding to subproblem $\mathcal{P}(\gamma, h_\gamma)$. From the above theorem and Corollary 5.7 it follows that the optimal cost for such a problem is given both by $E[K_1^{q-1}(\gamma)|h_\gamma|^q]$ and $E[K_2^{q-1}(\gamma)|h_\gamma|^q]$ for all $h_\gamma \in L^q_{\mathcal{F}_\gamma}(\mathbb{R})$. This entails that $K_1(\gamma) = K_2(\gamma)$, $P - a.s.$, for all stopping times γ with $\tau \leq \gamma < T$. As K_1 and K_2 have continuous paths and $K_1(T) = K_2(T) = f(M)$, we see that $K_2 = K_2 =: K$ in $L^\infty_{\mathcal{F}}(\tau, T; \mathbb{R}) \cap L^\infty_{\mathcal{F}}(\Omega, C([\tau, T]; \mathbb{R}))$.
Let (x, u) be the optimal state and the optimal control for problem $\mathcal{P}(\tau, 1)$. By the above theorem and the uniqueness of (x, u) we get $u = G(K_1, L_1)x = G(K_2, L_2)x$. From Theorem 5.5 we know that x does not vanish, so that we may divide by x and get $\frac{1}{x}u = G(K_1, L_1) = G(K_2, L_2) =: G$. The above theorem states, that the process z belonging to the unique solution of the FBSDE (3.11)-(3.14) can be represented as

$$z^i = (q-1)\varphi(x)K^{q-2}\left(C^iK + D^iKG + L_1^i\right) = (q-1)\varphi(x)K^{q-2}\left(C^iK + D^iKG + L_2^i\right).$$

This gives $L_1 = L_2$, and the corollary is shown. ■

Remarks on Chapter 5 In this chapter again we heavily exploited the interdependence between the BSRDE and the corresponding problem of control, in some kind of pingpong: starting from the problem of control and the FBSDE, the processes K and L that we constructed out of x, u, y and z are a solution of the BSRDE. Going back from the BSRDE to the control problem one sees that any solution (K_0, L_0) of the BSRDE yields an optimal control u. Returning from the problem of control to the BSRDE, this implies that the solution of the BSRDE is unique. One may object that this in somewhat pretty, but not helpful, since the proof of existence is not constructive. Sharper, if the main motivation for the study of the BSRDE is the wish to solve the control problem, where's the sense of solving the BSRDE with the help of the processes we are actually looking for? The answer is that the results of this paper indeed do not help to solve the stochastic control problem *numerically*. But actually it is not clear that a constructive method would do, at least in the case of stochastic coefficients. The constructive methods generally involve some type of Picard-like iteration (if one takes "Picard-like" in a very wide sense), see for example [KT:GAS]. Admittedly, these iteration gives an idea how one may tackle the problem of solving the BSRDE numerically. But in the course of this iteration, some auxiliary BSDEs with Lipschitz driver must be solved, some real functions must be minimized etc. in each iteration. The (constructive) standard proof of the solvability of BSDEs with Lipschitz-driver is based on an iteration itself, and in each of these latter iterations one has to calculate a martingale representation! Hence, to us it is not clear that a constructive method in general also delivers a feasible numerical approach (there are exceptions,

especially in the case of deterministic coefficients, when the BSRDE degenerates to an ordinary differential equation with auxiliary condition; to avoid misunderstandings: to the best of our knowledge no author who was concerned with BSRDE with stochastic coefficients and who used iterations claimed that this method is suitable as a pattern for a numerical scheme). As this point is of importance to us, let us consider the example of ordinary differential equations with Lipschitz-driver. The typical proof for the theorem about the existence of solutions (i.e. the Picard-Lindelöf-Theorem) makes use of the Picard-iteration. But at first instance it is not clear that a numerical version of this method yields an approximation of the solution, since one must take into account the numerical integration that has to performed in each step.
Besides, there are some natural discretization schemes to tackle BSDEs numerically, see for example [EM:BSDE], Chapter 4. Of course, it is not clear (and there's no evidence) whether or not this particular scheme, applied to the BSRDE for a linear isoelastic problem, converges to a solution. But if one knows that a solution *exists* and is unique, it makes sense to ask whether or not it *does*.
Now that the solvability of (4.17), (4.18) is clear one may wonder if such a BSRDE can be stated and solved for more general linear isoelastic control problems. A first generalization would be to consider n-dimensional state equations. If one finds a suitable BSRDE for this situation and tries to apply the method used in the present work and in [T:GLQO], one would rather consider the first degeneration time of the matrix-valued fundamental system for the optimal state than the first time the optimal state hits zero, see [T:GLQO]. A next, and presumably rather difficult step, would be to consider inhomogeneous problems, i.e. problems with state equation $dx = \{Ax + Bu + g\}\, ds + \sum_{i=1}^{d} \{C^i x + D^i u + k\}\, dw^i$ and cost functional $J(u) = \frac{1}{q} E[\int_\tau^T Q|x-\rho|^q + N|u|^q ds + M|x(T) - \vartheta|^q]$. One may also try to investigate the situation with the modified cost functional $J(u) = \frac{1}{q} E[\left(\int_\tau^T (x-\rho)' Q (x-\rho) ds\right)^{\frac{q}{2}} + \left(\int_\tau^T u' N u ds\right)^{\frac{q}{2}} + ((x(T) - \vartheta)' M (x(T) - \vartheta))^{\frac{q}{2}}]$, with matrix-valued processes/r.v. Q, N and M. Here, one could choose $\mathcal{U} = H_q(\tau, T; \mathbb{R}^m)$ independent of q.
The last point we want to mention is the question what happens when N becomes negative and the control problem is still solvable. In [CLZ: SLQ] there are some special cases where BSRDE for linear quadratic problems are solved.
To the best of our knowledge, no BSRDE for linear isoelastic stochastic control problems have occurred so far in the literature. Thus, all results concerning their unique solvability, the representation of the optimal control for the underlying problem and the representation of the optimal cost are new.

Chapter 6

A financial application and a duality

In this final chapter we will have a look at some possible applications of the results we have developed so far. The problems we consider belong to the so called *hedging problems* of mathematical finance. And though they are treated in the very last chapter, they initially gave rise for us to study linear isoelastic stochastic control problems as one possible mathematical framework in which these hedging problems may fit. Yet, we will see that the theory does not fit under all circumstances. We have to assume that the *claims* under consideration are *attainable* (the terminology is explained in Subsection 6.1.1).
Section 6.1 introduces the market model and the problems we are considering. In Section 6.2 we introduce a duality relation that links the financial market problem to a dual problem (that turns out to be essentially another linear isoelastic stochastic control problem). This dual problem can also be treated with the BSRDE methods we developed, with two exceptions. Hence, one has a choice to consider the primal or dual problem, depending on which one seems more easy to handle. In the first of the mentioned exceptions (*terminal cost, complete market,* Subsection 6.3.1), the dual problem is stunningly simple to treat and can be solved "by hand". It is typical for financial market problems concerning terminal wealth that in the case of so called complete markets almost everything finally relies on a proper manipulation of the (unique) so called *equivalent martingale measure.* Subsection 6.3.1 may be seen as an illustration of this.
In the second exception (see Subsection 6.3.4), the linear isoelastic problem to which the dual problem gives rise does not fit in the BSRDE framework we set up so far. There, we encounter some sort of mixing of the assumptions we used until now: the first part of the control variable "behaves" according to Assumption A4, the second part of the control variable "behaves" as is required in Assumption A1 or A3. We can introduce a BSRDE for this type of control problem, and through the duality relation, we can prove that this particular BSRDE is solvable. The treatment of the financial market problems and their dual problems with the help of BSRDEs is done Section 6.3
We skip the convention that (x, u, y, z) is the solution of the FBSDE for problem $\mathcal{P}(\tau, 1)$.

6.1 The financial market model and problem formulation

Let us start with the description of the underlying financial market model.

6.1.1 The market model

A financial market with $m+1$ assets is typically modeled by some filtered probability space $(\Omega, P, \mathcal{F}, (\mathcal{F}_t)_{t\geq 0},)$ and a set of adapted stochastic processes $S_j,\ j=0,\dots,m$, defined on this space. The *r.v.* $S_j(t)$ is the (stochastic) value of one unit of asset j at time t. By abuse of language we will often make no difference between the asset j and its price process S_j. It is usual to denote by S_0 the non-risky asset, the *bond* (which is assumed to pay interests continuously). The $S_1,\dots,S_m$ are supposed to be risky assets, typically *stocks* .
We will use the following "Brownian" model. Let w be a d-dimensional Brownian motion on the filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F})_{0\leq t\leq T}, P)$. The filtration $(\mathcal{F})_{0\leq t\leq T}$ is supposed to be the augmentation of the filtration generated by w.

Assumption (*market model coefficients)*
For the rest of this chapter, let all the processes r, b_j, $j\in\{1,\dots,m\}$, and σ_{ji}, $i\in\{1,\dots,d\}$, $j\in\{1\dots,m\}$, belong to $L^\infty_{\mathcal{F}}(0,T;\mathbb{R})$. Set $\sigma := [\sigma_{ji}]_{\substack{1\leq j\leq m\\ 1\leq i\leq d}}$, and assume, that

$$\sigma\sigma' \gg 0, \tag{6.1}$$

i.e., $\sigma\sigma'$ is uniformly positive.

Let $s_1,\dots,s_m$ be strictly positive real numbers. In our market model, the price processes $S_0,\dots,S_m$ on $[0,T]$ then are given by the following SDEs.

$$dS_0(t) = r(s)ds,\ S_0(0)=1, \tag{6.2}$$

and

$$dS_j(t) = b_j(s)S_j(s)ds + \sum_{i=1}^{d} \sigma_{ji}(s)S_j(s)dw^i(s),\ S_j(0)=s_j, \tag{6.3}$$

for $j\in\{1,\dots,m\}$.
Let us consider an agent that is trading in this financial market, i.e. trading in the securities $S_0, S_1,\dots,S_m$. We impose some standard idealizations on trading: first, we assume that there are no frictions, i.e. no transaction costs or time delay. There is no dividend paying. Besides, the securities can be traded in arbitrary portions of one unit, i.e. the number of shares the agent buys or sells need not be an integer. The model also allows for unlimited short selling, in particular, for unlimited borrowing. The agent is a *small investor,* this means, his trading has no influence on the price of the securities. We will assume that the trading strategies of the investor are *self financing,* i.e. apart form the

agent's initial financial endowment, all money for buying securities comes from the selling of other securities from his portfolio.
Let us have a look at the investment strategies that the agent may implement. If we denote by $\pi_j(t)$, $j \in \{0, \ldots, m\}$, $t \in [0, T]$, the amount of money the agent invests at time t in security S_j, then it is reasonable to allow only for $(\mathcal{F}_t)$-adapted *portfolio processes* $\pi = (\pi_0, \pi_1, \ldots, \pi_m)$ in our model - otherwise, the "investment decision" represented by $\pi_j(t)$ would use information that is not available at time t. So, let π be a self financing portfolio process and let x be the resulting wealth process of the agent, i.e. $x = \pi_0 + \pi_1 + \ldots + \pi_m$. Set $u = (\pi_1, \ldots, \pi_m)'$, $B = (b_1 - r, \ldots, b_m - r)$ and $A = r$. It can be argued (see [KS:MMF]), that the wealth process x of the self financing portfolio process π then follows the SDE

$$dx = \{Ax + Bu\}\, ds + u'\sigma dw, \tag{6.4}$$

$$x(0) = x_0, \tag{6.5}$$

where x_0 is the initial endowment of the agent. We will sometimes indicate the dependence of x on x_0 and u by writing $x^{x_0,u}$. To make the above equation meaningful, we must require, that

$$P(\int_0^T |u(s)|^2 ds < \infty) = 1. \tag{6.6}$$

Besides, given an arbitrary adapted u that satisfies this integrability condition, we can easily find a process π_0 such that (π_0, u) is a self financing trading strategy, just by setting $\pi_0 = x - \sum_{j=1}^m u_j$. So, we will focus on the positions the agents holds in the risky assets $S_1, \ldots, S_m$, and will, by abuse of language, call the allocation $u = (u_1, \ldots, u_m)$ in these assets a *portfolio* or *investment strategy*. It turns out that the requirements that trading strategies u are adapted and satisfy (6.6) are not sufficient to exclude a phenomenon that is considered as highly undesirable in economic theory, that of *arbitrage*. Loosely speaking, a financial market bears an arbitrage opportunity if an agent can make money out of nothing. Formally, the portfolio (technically: the stochastic process) u is an arbitrage opportunity, if for the corresponding solution of (6.4), (6.5) with $x_0 = 0$ (i.e. no initial endowment) we have

$$x(T) \geq 0 \text{ and } P(x(T) > 0) > 0.$$

This would mean that the agent starts with no money ($x_0 = 0$), is not exposed to the risk of losing money ($x(T) \geq 0$) and can gain money with a non-zero probability ($P(x(T) > 0) > 0$). In [KS:MMF], Section 1.2, there are two examples of arbitrage possibilities. There are typically two ways to work around this problem. The first one is only to allow for *tame* portfolios, i.e. portfolios, that satisfy a minimal integrability condition and whose resulting value process x is bounded from below. The second one is to impose some stronger integrability condition on the portfolio to exclude doubling strategies and other pathologies. See [DeS:ACPM] for a discussion of the subject in the much more delicate framework of a financial market that is given by a general, m-dimensional semimartingale. We will follow the second alternative.

Definition 6.1 *As set $\mathcal{V}$ of admissible portfolios we choose*

$$\mathcal{V} := H_q(0, T; \mathbb{R}^m), \tag{6.7}$$

where q, at this instance, may be an arbitrary real with $q > 1$. It is a standard result that, if we restrict the implementable trading strategies to this set, the market allows no arbitrage:

Remark 6.2 *Consider the financial market that is given by the price processes $S_0, \ldots, S_m$ as introduced above, and let q be a real with $q > 1$. Then $\mathcal{V}$ contains no arbitrage opportunity for this market.*

Proof: Let us assume that u is an arbitrage opportunity for this market. We will show that this leads to a contradiction.
Let x be the solution u of (6.4), (6.5) for this particular u with $x_0 = 0$. As in Lemma 3.2, set $\theta := \sigma'(\sigma\sigma')^{-1}B' \in L^\infty_{\mathcal{F}}(0, T\ \mathbb{R}^d)$ and write (6.4), (6.5) as

$$dx = \{Ax + \theta'\sigma' u\}\, ds + u'\sigma dw. \tag{6.8}$$

Set $R := x$ and $S := u'\sigma$ and consider the above SDE as a BSDE, like in (3.3), (3.4):

$$dR = \{AR + \theta' S'\}\, ds + Sdw,\ \ R(T) = x(T).$$

Let Γ be the solution of $d\Gamma = -A\Gamma ds - \theta'\Gamma dw$, $\Gamma(0) = 1$. Then we can represent R by

$$\Gamma(t)R(t) = E[\Gamma(T)x(T)|\mathcal{F}_t],\ t \in [0, T],$$

(compare [EM:BSDE], Prop. 1.2.4), in particular

$$R(0) = x_0 = E[\Gamma(T)x(T)] = 0.$$

Obviously, we have $\Gamma(T) > 0$, $P - a.s.$. As u was supposed to be an arbitrage opportunity, we have $x(T) \geq 0$, thus the last equality cannot hold unless $x(T) = 0$, what contradicts the arbitrage assumption. ■

Note that the above proof would also work if knew that there is *any* essentially bounded process θ such that $\sigma\theta = B'$.
The agent may, of course, pursue different purposes when trading in the financial market. We will address the mathematical problems that arise when the agent is looking for *optimal hedging strategies:* Suppose that the agent has some stochastic financial obligation, a so called *contingent claim*, at time T. This obligation is modeled by an $\mathcal{F}_T$-measurable r.v. ξ. $\xi(\omega)$ is the amount of money the agent has to pay if at time T the "state" ω is realized (if this amount is negative the agent receives money). A typical example for a contingent claim is a (European) call option for a traded security. When selling a European call option for the asset S_j with *strike price* k and *maturity time* T, the agent engages himself to sell one unit of the asset S_j at time T for the price k, if the buyer wishes so. Of course, the buyer of the claim will not do so if the stock exchange price of S_j at time T is smaller than k. Thus, the resulting obligation for the agent can be modeled by the r.v. $\max\{S_j(T) - k, 0\}$. The agent is looking for some self-financing trading strategy u such that, given a fixed

initial endowment x_0, the terminal state $x^{x_0,u}(T)$ of his wealth is as close as possible to ξ. For us, "close to" will mean that the quantity $E[|\xi - x^{x_0,u}(T)|^q]$ becomes small (with a fixed $q > 1$). There are plenty of other criteria, economically perhaps even more intuitive than ours, to measure proximity of the claim and the terminal value. In particular, one may be more interested in criteria that do not punish an overshooting of $x^u(T)$ over ξ, as in this case the agent gains more money from his portfolio than he needs to satisfy his obligation. Hence one may prefer a criteria that depends on $\max\{\xi - x^u(T), 0\}$ rather than on $|\xi - x^u(T)|$.
Unfortunately, we cannot handle the case of general claims $\xi \in L^q_{\mathcal{F}_T}(\mathbb{R})$. We must require that ξ is *attainable* (see below). Given an attainable claim, there is a fixed fair price at time $t = 0$ for ξ. Yet, if ξ is non-negative and attainable, it will turn out that given our optimality criterion the optimal terminal wealth stays below ξ if the initial endowment e_0 of the agent is smaller than the fair price of the claim at time $t = 0$. Hence, it would (a posteriori) make no difference if we considered the cost functional $E[|\xi - x^{e_0,u}(T)|^q]$ or $E[(\max\{\xi - x^{e_0,u}(T), 0\})^q]$. Hedging problems with cost functionals that depend on $\max\{\xi - x^{e_0,u}(T), 0\}$ are extensively studied in [FL:QH] and [FL:EH]. There, the authors use statistical test theory and determine a *test* ψ such that the optimal terminal value $\bar{x}^{e_0,u}(T)$ of the hedging problem is given by $\bar{x}^{e_0,u}(T) = \psi\xi$.

Definition 6.3 *(attainable claims)*
A claim $\xi \in L^q_{\mathcal{F}_T}(\mathbb{R})$ is called attainable if there is an initial endowment $\xi_0 \in \mathbb{R}$ and a portfolio $u_\xi \in \mathcal{V}$ such that the resulting terminal wealth $x^{\xi_0,u_\xi}(T)$ equals ξ,

$$x^{\xi_0,u_\xi}(T) = \xi, \quad P - a.s.. \tag{6.9}$$

One may wonder if it makes much sense to consider the problem of minimizing $E[|\xi - x^{x_0,u}(T)|^q]$ under this assumption, since the choice $(x_0, u) = (\xi_0, u_\xi)$ obviously minimizes this criterion. The idea is that the agent will in general not want to provide the initial endowment ξ_0, but a smaller amount x_0.
There are situations in which (6.9) turns out to be no restriction, namely in *complete markets*. A market is called *complete*, if very claim is attainable, i.e. (6.9) holds for every $\xi \in L^q_{\mathcal{F}_T}(\mathbb{R})$. In our market model, the completeness depends on the relation between d, the dimension of the Brownian Motion, and m, the number of the risky assets respectively the dimension of the control process. We state another standard result:

Remark 6.4 *Consider the financial market as described in this section. Then the following two assertions are equivalent:*

1. *For every $\xi \in L^q_{\mathcal{F}_T}(\mathbb{R})$ there is a $\xi_0 \in \mathbb{R}$ and a $u \in \mathcal{V}$ such that $\xi = x^{\xi_0,u_\xi}(T)$, where x^{ξ_0,u_ξ} is the solution of (6.4), (6.5) for $(x_0, u) = (\xi_0, u_\xi)$.*

2. *$d = m$.*

Proof: Recall that the condition $\sigma\sigma' \gg 0$ entails $d \geq m$. Again, set $\theta = \sigma'(\sigma\sigma')^{-1}B'$.
We first proof the implication 1. $\Rightarrow$ 2. by constructing a contradiction. Suppose $m \leq d-1$,

i.e σ has not full rank. Then, there is a process $\kappa \in H_q(0,T;\mathbb{R}^d)$ such that $u'\sigma \neq \kappa'$ (in $H_q(0,T;\mathbb{R}^d)$) for all $u \in H_q(0,T;\mathbb{R}^m)$ (in particular $\kappa \neq 0$). Choose some $\mu_0 \in \mathbb{R}$ and let μ be the solution of the (forward) SDE

$$d\mu = \{A\mu + \theta'\kappa\} + \kappa' dw, \quad \mu(0) = \mu_0, \tag{6.10}$$

and set $\xi = \mu(T)$. We may consider the above equation as a BSDE

$$dR = \{AR + \theta' S\} + S' dw, \quad R(T) = \xi. \tag{6.11}$$

From the unique solvability of this BSDE we get $R = \mu$ and $S = \kappa$. We claim that ξ is not attainable, i.e. there is no pair $(x_0, u) \in \mathbb{R} \times H_q(0,T;\mathbb{R}^m)$ such that $\xi = x^{\xi_0,u}(T)$. Otherwise, the pair $R = x$, $S := \sigma' u$ would yield the unique solution of the BSDE (6.11) (since $Bu = \theta'\sigma' u$). By this we would have found a u such that $\sigma' u = \kappa$, what contradicts our choice of κ.
To show the implication 2. $\Rightarrow$ 1., assume $d = m$ and chose some $\xi \in L^q_{\mathcal{F}_T}(\mathbb{R})$ and let (R, S) be the corresponding solution of (6.11). As $\sigma\sigma' \gg 0$ we see that σ is $Leb \otimes P - a.s.$ invertible, and $\sigma^{-1} = \sigma'(\sigma\sigma')^{-1}$ shows that σ^{-1} is essentially bounded. Thus, $u := (\sigma')^{-1} S$ is well defined and in $H_q(0,T;\mathbb{R}^m)$. Plugging this into (6.11) shows, that $x := R$ is a solution of (6.4) with initial value $x_0 := R(0)$ and terminal value $x(T) = \xi$. ■

The agent may actually be not interested in the nominal value $x(t)$, $t \in [0,T]$, of his wealth, but in the discounted value $\widetilde{x}(t)$. As we assume continuous payment of interests with interest rate A, the discount factor $\frac{1}{S_0(t)}$ at time t is given by $\exp\left\{-\int_0^t A(s)ds\right\}$. Hence, the present value $\widetilde{x}(t)$ of the agents wealth at time t is given by

$$\widetilde{x}(t) = \exp\left\{-\int_0^t A(s)ds\right\} x(t). \tag{6.12}$$

Set $\tilde{B}(\cdot) := \exp\left\{-\int_0^\cdot A(s)ds\right\} B(\cdot) \in L^\infty_{\mathcal{F}}(0,T;\mathbb{R}^m)$ and $\tilde{\sigma}(\cdot) := \exp\left\{-\int_0^\cdot A(s)ds\right\} \sigma(\cdot) \in L^\infty_{\mathcal{F}}(0,T;\mathbb{R}^{m\times d})$. It is easily checked that the discounted wealth resulting from the initial endowment x_0 and the trading strategy u follows the SDE

$$d\widetilde{x} = \tilde{B} u ds + u'\tilde{\sigma} dw, \quad x(0) = x_0, \tag{6.13}$$

i.e. $\widetilde{x}(t) = x_0 + \int_0^t u'\tilde{B}' ds + \int_0^t u'\tilde{\sigma} dw$. Hence, the discounted wealth process is the stochastic integral of the portfolio u with respect to the special semimartingale $\int_0^\cdot \tilde{B}' ds + \int_0^\cdot \tilde{\sigma} dw$. Both following problems will for technical reasons be stated in a discounted setting. For simplicity of notation we will omit the tilde on the coefficients.

6.1.2 Two financial market problems

We will now specify the two particular problems we address. We work within the framework of the previous subsection. Assume we are given the following data: an initial endowment

$e_0 \in \mathbb{R}$, an attainable contingent claim $\xi \in L^q_{\mathcal{F}_T}(\mathbb{R})$, along with a reproducing initial endowment ξ_0 and portfolio u_ξ (such that $\xi = x^{\xi_0,u_\xi}(T)$). We assume that $\xi_0 \neq e_0$. As already indicated, we ask, which portfolio $\bar{u}$ the agent must choose to drive his terminal wealth close to ξ in the q-th mean. So, for some $q > 1$ we want to minimize $J(u) = E[|x^{e_0,u}(T) - \xi|^q]$ over $u \in \mathcal{V}$. In order to put different weights in this cost criterion on the scenarios that may occur at time T, the agent may choose a r.v. $M \in L^\infty_{\mathcal{F}}(0,T;\mathbb{R})$ with $M \gg 0$ and consider the cost functional

$$J^{(T)}(u) := E[M|x^{e_0,u}(T) - \xi|^q], \tag{6.14}$$

that is to be minimized over $\mathcal{V}$. The superscript "T" refers to *terminal* cost. Along with the state equation (6.4), (6.5) this, of course, constitutes a stochastic control problem. But due to the presence of ξ it is not of the type that we could solve directly with the help of the BSRDE. This is why we had to impose attainability of ξ. We may write

$$x^{e_0,u}(T) - \xi = x^{e_0,u}(T) - x^{\xi_0,u_\xi}(T) = x^{e_0-\xi_0,u-u_\xi}(T).$$

It is clear, that $u - u_\xi$ ranges over all $\mathcal{V}$ if u ranges over $\mathcal{V}$, $\mathcal{V} - u_\xi = \mathcal{V}$. Hence,

$$\min_{u\in\mathcal{V}} E[M\,|x^{e_0,u}(T) - \xi|^q] = \min_{u\in\mathcal{V}} E[M\,|x^{e_0-\xi_0,u-u_\xi}(T)|^q] = \min_{u\in\mathcal{V}} E[M\,|x^{e_0-\xi_0,u}(T)|^q].$$

Under the given assumptions we may thus confine ourselves to the case $\xi = 0$ and tackle the minimization of $J^{(T)}$ by considering the following (abstract) control problem. It is for technical reasons stated in a discounted setting,

Definition 6.5 *(terminal cost hedging problem)*
Let the assumptions on the financial market coefficients hold with $A = 0$. Assume $M \in L^\infty_{\mathcal{F}_T}(\mathbb{R})$ with $M \gg 0$. Fix some $q > 1$. By the terminal cost hedging problem we mean the stochastic control problem

$$J^{(T)}(u) := \frac{1}{q} E[M\,|x(T)|^q] = \min_{u\in\mathcal{V}}! \tag{6.15}$$

where x is the solution of

$$dx(t) = B(s)u(s)ds + \sum_{i=1}^{d} D^i(s)u(s)dw^i(s), \tag{6.16}$$

$$x(0) = x_0, \tag{6.17}$$

with $D^i := (\sigma_{1,i}, \ldots, \sigma_{m,i})$ and $x_0 \in \mathbb{R}$, $x_0 \neq 0$.

Note that for $q = 2$ this amounts to the so called *mean variance hedging problem.* There is a vast literature on mean-variance hedging, at this instance we only cite [GLP:MVH].

The terminal cost hedging problem takes into account only the terminal wealth of the agent at time T. Though, it may occur that the agent wants to get rid of his obligation ξ

before maturity - maybe he found an alternative investment that pays better for bearing risk, or he simply wants to reduce his exposure to uncertainty. Imagine, the agent sold the claim ξ at time 0 for the price ξ_0 and invested the amount x_0 with $x_0 < \xi_0$ to run a hedging strategy for ξ. If he wants to buy back ξ at time t, the money primarily available to him at this time is the amount $x^{x_0,u}(t)$ produced by his hedging strategy. Of course, there is also the remaining amount of $\xi_0 - x_0$, plus interest, but the agent typically does not want to spend this money on hedging. Otherwise he would have done so from the beginning, thus creating a situation where he excluded his risk but also his chance of earning money - the agent sells the claim for ξ_0, but he spends this money immediately for a hedging strategy.

If ξ is an attainable claim with replicating portfolio u_ξ and the initial endowment ξ_0, then it can be argued, that $x^{\xi_0,u_\xi}(t)$ is the fair price for the claim ξ (as $x^{\xi_0,u_\xi}(t)$ is a $\mathcal{F}_t$ - measurable r.v. we suppose that we can identify the "states" ω which turn out to be realized at time t, or, to put it differently, at time t we know where to evaluate $x^{\xi_0,u_\xi}(t)$ in order to get the fair price for ξ).
Hence, if the agents wants to have a chance to settle the claim ξ before maturity without too much of loss, he may try to keep track of the fair price (at least in some average sense) and consider the compound cost functional

$$J^{(RT)}(u) = \frac{1}{q}E[\int_0^T Q(s)|x^{\xi_0,u_\xi}(s) - x^{x_0,u}(s)|^q ds \; + \; M|x^{\xi_0,u_\xi}(T) - x^{x_0,u}(T)|^q] \tag{6.18}$$

and minimize it over all trading strategies $u \in \mathcal{V}$. The superscript "RT" stands for *running and terminal* cost. The r.v. M and the process Q are both required to be uniformly positive, for technical reasons. They may help the agent to model the different emphasis he puts on running and terminal costs, as well to put different weights on the "states" ω.
As above we can "shift controls" via the relation $x^{\xi_0,u_\xi} - x^{x_0,u} = x^{\xi_0 - x_0, u_\xi - u}$ and thus cast the problem of minimizing $J^{(RT)}$ (for an attainable claim) in the form given in Definition 6.6. Again, we will only consider the discounted case, i.e. the case $A = 0$.

Definition 6.6 *(running and terminal cost hedging problem)*
Let the assumptions on the financial market coefficients hold with $A = 0$. Assume $M \in L^\infty_{\mathcal{F}_T}(\mathbb{R})$ and $Q \in L^\infty_{\mathcal{F}}(0,T;\mathbb{R})$ with $M \gg 0$ and $Q \gg 0$. Fix some $q > 1$. By the running and terminal cost hedging problem we mean the stochastic control problem

$$J^{(RT)}(u) := \frac{1}{q}E[\int_0^T Q(s)|x(s)|^q ds \; + \; M\,|x(T)|^q] = \min_{u\in\mathcal{V}}! \tag{6.19}$$

where x is the solution of

$$dx(t) \;=\; B(s)u(s)ds + \sum_{i=1}^d D^i(s)u(s)dw^i(s), \tag{6.20}$$

$$x(0) \;=\; x_0, \tag{6.21}$$

with $D^i := (\sigma_{1,i}, \dots, \sigma_{m,i})$ and $x_0 \in \mathbb{R}$, $x_0 \neq 0$.

Like the terminal cost hedging problem, the above problem has already been studied for $q = 2$, although much less intensively than the mean variance hedging problem. See for example [KZ:RBBS] and [KT:MLQC].
These two problems can be treated directly via the BSRDE approach, as they obviously satisfy Assumption A1 or A3. Theorem 5.11 and Corollary 5.12 show that the corresponding Riccati equation is uniquely solvable. Theorem 5.6 and Corollary 5.7 indicate the optimal control and the optimal cost.
For both "abstract" problems (6.15) and (6.19), the optimal state is seen to be stochastic exponential, thus has always the sign of x_0. Reconsider the initial problem of hedging an attainable claim ξ that can be represented as $\xi = x^{\xi_0,u_\xi}(T)$, $P-a.s.$, and assume that $\xi \geq 0$, $P-a.s.$. Let $e_0 \in [0,\xi_0]$ be the initial endowment of the agent, and let $\hat{x}$ be the optimal state for the hedging problem with performance criterion as given in (6.14) or (6.18). As the "abstract" control problems (6.15) and (6.19) with initial value $x_0 = \xi_0 - e_0 \geq 0$ have a non-negative optimal state, we get that $0 \leq \hat{x}(T) \leq \xi$, $P-a.s.$, and $0 \leq \hat{x} \leq x^{\xi_0,u_\xi}$, $Leb \otimes P-a.s.$.
In the next section we will introduce a duality relation, and it turns out that the dual problems arising from this relation can be (in most cases) treated with BSRDE theory, too. It is interesting to observe that if the primal problem e.g. satisfies Assumption A1, then the dual problem is of type A4 or some mixture of Assumptions A3 and A4.

6.2 A duality

Looking at the definitions in (6.15) and (6.19) of $J^{(T)}$ and $J^{(RT)}$, one may perceive that these convex functionals belong to a particularly convenient class of functionals. Due to the uniform positivity of M respectively Q and M they can be seen as defining norms (up to the exponent $\frac{1}{q}$) on $L^q_{\mathcal{F}_T}(\mathbb{R})$ respectively $L^q_{\mathcal{F}}(0,T;\mathbb{R}) \times L^q_{\mathcal{F}_T}(\mathbb{R})$ that are applied to $x(T)$ respectively $(x, x(T))$, and that are equivalent to the usual L^q-norms. We take this observation as a starting point for a duality approach based on the following theorem. We take the statement from [L:O], Thm.1 in § 5.8, where also its proof can be found.

6.2.1 The framework

Let $\mathcal{Z}$ be a real normed linear space with norm $\|\cdot\|_{\mathcal{Z}}$ and let $\mathcal{Z}^*$ be its dual with norm $\|\cdot\|_{\mathcal{Z}^*}$. Denote the duality product of these spaces by $< \cdot,\cdot >_{\mathcal{Z}}$. For a subspace $\mathcal{M} \subset \mathcal{Z}$ its orthogonal complement $\mathcal{M}^\perp \subset \mathcal{Z}^*$ is defined by $\mathcal{M}^\perp := \{m^* \in \mathcal{Z}^* : < m, m^* >_{\mathcal{Z}} = 0 \text{ for all } m \in \mathcal{M}\}$.

Theorem 6.7 *Let $\mathcal{Z}$ and $\mathcal{M}$ be as above. Then for every $z \in \mathcal{Z}$ we have*

$$\inf_{m\in\mathcal{M}} \|z - m\|_{\mathcal{Z}} = \max_{\substack{m^*\in\mathcal{M}^\perp \\ \|m^*\|_{\mathcal{Z}^*}\leq 1}} < z, m^* >_{\mathcal{Z}} . \tag{6.22}$$

The maximum on the right is achieved for some $m_0^* \in \mathcal{M}^*$. If the infimum on the left is *achieved for some* $m_0 \in \mathcal{M}$*, then* $z - m_0$ *and* m_0^* *are* aligned *, i.e.* $< z - m_0, m_0^* >_{\mathcal{Z}} = \|z - m_0\|_{\mathcal{Z}} \; \|m_0^*\|_{\mathcal{Z}^*}$. ■

We will address the maximization problem on the right hand side of (6.22) as *dual problem.* It will turn out that, in the case of terminal cost, the dual problem essentially leads to the dual problem investigated in [GLP:MVH].
In the next subsection we will do some preparation so that we can apply this theorem to the problems (6.15) and (6.19). Let us sketch how the theorem fits in the situation of these problems. Choose some $q > 1$. Let M respectively Q and M be as in the Definitions 6.5 respectively 6.6. Define the normed linear spaces $\mathcal{Z}_T$ and $\mathcal{Z}_{RT}$ by

$$\mathcal{Z}_T := (L^q_{\mathcal{F}_T}(\mathbb{R}), \|\cdot\|_{\mathcal{Z}_T}) \text{ with } \|Z\|_{\mathcal{Z}_T} := \left(\frac{1}{q} E[M\,|Z|^q]\right)^{\frac{1}{q}}, \tag{6.23}$$

and

$$\begin{aligned} \mathcal{Z}_{RT} &:= (L^q_{\mathcal{F}}(0,T;\mathbb{R}) \times L^q_{\mathcal{F}_T}(\mathbb{R}), \|\cdot\|_{\mathcal{Z}_{RT}}) \text{ with} \\ \|(z,Z)\|_{\mathcal{Z}_{RT}} &:= \left(\frac{1}{q} E[\int_0^T Q(s)|z(s)|^q ds \; + \; M\,|Z|^q]\right)^{\frac{1}{q}}. \end{aligned} \tag{6.24}$$

(The use of the letters z and Z is *not* meant to allude to the solution of some FBSDE). Further, let H_T and H_R be solution operators for the SDE (6.16) and (6.17) - actually integral operators - with initial value 0, i.e.

$$H_T : \mathcal{V} = H_q(0,T;\mathbb{R}) \longrightarrow L^q_{\mathcal{F}_T}(\mathbb{R}) \tag{6.25}$$

$$u \mapsto \int_0^T B'(s)u(s)ds + \int_0^T u'(s)\sigma(s)dw(s), \tag{6.26}$$

respectively

$$H_R : \mathcal{V} \longrightarrow L^q_{\mathcal{F}}(0,T;\mathbb{R}) \tag{6.27}$$

$$u \mapsto \int_0^{\cdot} B'(s)u(s)ds + \int_0^{\cdot} u'(s)\sigma(s)dw(s). \tag{6.28}$$

Of course we have $H_T u = (H_R u)(T)$. Note that the solution x of the state equation (6.16), (6.17) for the pair $(x_0, u) \in \mathbb{R} \times \mathcal{V}$ can be written as

$$x = x_0 + H_R u.$$

Finally, define the normed linear subspaces $\mathcal{M}_T$ and $\mathcal{M}_{RT}$ of $\mathcal{Z}_T$ and $\mathcal{Z}_{RT}$ by

$$\mathcal{M}_T := \{H_T u : u \in \mathcal{V}\} \subset \mathcal{Z}_T, \tag{6.29}$$

$$\mathcal{M}_{RT} := \{(H_R u, H_T u) : u \in \mathcal{V}\} \subset \mathcal{Z}_{RT}. \tag{6.30}$$

With these notations, problem (6.15) can be equivalently formulated as

$$\|x_0 + Z\|_{\mathcal{Z}_T} = \|-x_0 - Z\|_{\mathcal{Z}_T} = \min_{Z\in\mathcal{M}_T}!, \tag{6.31}$$

and for (6.19) we get the equivalent problem

$$\|(x_0 + z, x_0 + Z)\|_{\mathcal{Z}_{RT}} = \|(-x_0 - z, -x_0 - Z)\|_{\mathcal{Z}_{RT}} = \min_{(z,Z)\in\mathcal{M}_{RT}}!. \tag{6.32}$$

We will call these problems *primal problems*. In a precise meaning, (6.31), (6.32) aren't really identical to (6.15), (6.19), since the cost functional of the latter does not have an exponent $\frac{1}{q}$. But the optimal controls of course coincide, and the optimal values can be easily calculated form one another. It is essentially just a rephrasing if we shift from (6.15), (6.19) to (6.31), (6.32), in contrast to the shift from a *dynamic* to a *static* problem as described e.g. in [Ph:DLPH].

Looking at the duality statement in Theorem 6.7, this rephrasing indicates what we want to plug in the left hand side of (6.22). For evaluating the right hand side we must find out what $\mathcal{M}_T^\perp$ respectively $\mathcal{M}_{RT}^\perp$ is. This will be done in the next subsection. Note that the duality in Theorem 6.7 is as a special case of a more general theory of the minimization of convex functionals. See, for example, [ET:CA], Chapter 3, and [L:UDP], §§ 1.5.3 , 1.5.4.

6.2.2 The adjoint operators and orthogonal complements

As $\mathcal{M}_T$ and $\mathcal{M}_{RT}$ are images of the operators H_T and H_{RT} we may determine the orthogonal complement of these spaces as the kernel of the adjoint operators. Thus we are lead to look for the adjoint operators. We start by noting what the dual spaces of $\mathcal{Z}_T$ and $\mathcal{Z}_{RT}$ are.

Remark 6.8 *Consider the spaces $\mathcal{Z}_T$ and $\mathcal{Z}_{RT}$. Set $\tilde{M} := (\frac{1}{q}M)^{-\frac{1}{q-1}}$ and $\tilde{Q} := (\frac{1}{q}Q)^{-\frac{1}{q-1}}$. Consider the usual dual pairings $< \cdot,\cdot >_{\mathcal{Z}_T}$: $L^q_{\mathcal{F}_T}(\mathbb{R}) \times L^{q'}_{\mathcal{F}_T}(\mathbb{R}) \longrightarrow \mathbb{R}$, $(Z, Z^*) \mapsto E[ZZ^*]$ and $< \cdot,\cdot >_{\mathcal{Z}_{RT}}$: $(L^q_{\mathcal{F}}(0,T;\mathbb{R})\times L^q_{\mathcal{F}_T}(\mathbb{R}))\times(L^{q'}_{\mathcal{F}}(0,T;\mathbb{R})\times L^{q'}_{\mathcal{F}_T}(\mathbb{R})) \longrightarrow \mathbb{R}$, $((z,Z),(z^*,Z^*)) \mapsto E[\int_0^T z(s)z^*(s)ds + ZZ^*]$. Based on these pairings we have $\mathcal{Z}_T^* = (L^{q'}_{\mathcal{F}_T}(\mathbb{R}), \|\cdot\|_{\mathcal{Z}_T^*})$ with*

$$\|Z^*\|_{\mathcal{Z}_T^*} = \left(E[\tilde{M}\,|Z^*|^{q'}]\right)^{\frac{1}{q'}}$$

for $Z^* \in L^{q'}_{\mathcal{F}_T}(\mathbb{R})$, and $\mathcal{Z}_{RT}^* = (L^{q'}_{\mathcal{F}}(0,T;\mathbb{R}) \times L^{q'}_{\mathcal{F}_T}(\mathbb{R}), \|\cdot\|_{\mathcal{Z}_{RT}^*})$ with

$$\|(z^*, Z^*)\|_{\mathcal{Z}_{RT}^*} = \left(E[\int_0^T \tilde{Q}(s)|z^*(s)|^{q'}ds \,+\, \tilde{M}\,|Z^*|^{q'}]\right)^{\frac{1}{q'}}$$

for $(z^*, Z^*) \in L^{q'}_{\mathcal{F}}(0,T;\mathbb{R}) \times L^{q'}_{\mathcal{F}_T}(\mathbb{R})$.

Proof: From the uniform boundedness and positivity of M and Q it is clear that $\tilde{M}$ and $\tilde{Q}$ are well defined and that the norms $\|\cdot\|_{\mathcal{Z}_T}$ respectively $\|\cdot\|_{\mathcal{Z}_{RT}}$ are mutually equivalent to the norm $Z \mapsto |Z|_q$ respectively $(z,Z) \mapsto (\|z\|^q_{L^q_{\mathcal{F}}} + |Z|^q_q)^{\frac{1}{q}}$ on $L^q_{\mathcal{F}_T}(\mathbb{R})$ respectively $L^q_{\mathcal{F}}(0,T;\mathbb{R}) \times L^q_{\mathcal{F}_T}(\mathbb{R})$. The dual of $L^q_{\mathcal{F}_T}(\mathbb{R})$ respectively $L^q_{\mathcal{F}}(0,T;\mathbb{R}) \times L^q_{\mathcal{F}_T}(\mathbb{R})$ with respect to these latter norms are known to be $L^{q'}_{\mathcal{F}_T}(\mathbb{R})$ respectively $L^{q'}_{\mathcal{F}}(0,T;\mathbb{R})$ with the usual dual pairings that we also use. Hence, every continuous linear functional on $\mathcal{Z}_T$ respectively $\mathcal{Z}_{RT}$ can be represented with some

$$Z^* \in L^{q'}_{\mathcal{F}_T}(\mathbb{R}) \text{ respectively } (z^*, Z^*) \in L^{q'}_{\mathcal{F}}(0,T;\mathbb{R}) \times L^{q'}_{\mathcal{F}_T}(\mathbb{R})$$

by

$$Z \mapsto < Z, Z^* >_{\mathcal{Z}_T} \text{ respectively } (z,Z) \mapsto < (z,Z),(z^*,Z^*) >_{\mathcal{Z}_{RT}} .$$

Let us calculate the dual norm of $\mathcal{Z}_T$. By definition, we have

$$\|Z^*\|_{\mathcal{Z}^*_T} := \sup_{\|Z\|_{\mathcal{Z}_T}=1} < Z, Z^* >_{\mathcal{Z}_T} .$$

If $\|Z\|_{\mathcal{Z}_T} = 1$ we have for every Z^* by Hölder's inequality

$$\begin{aligned}
< Z, Z^* >_{\mathcal{Z}_T} &= E[ZZ^*] \\
&= E[(\frac{1}{q}M)^{\frac{1}{q}} Z\, (\frac{1}{q}M)^{-\frac{1}{q}} Z^*] \\
&\leq \underbrace{\left(E[\frac{1}{q}M|Z|^q]\right)^{\frac{1}{q}}}_{=1} \left(E[(\frac{1}{q}M)^{-\frac{q'}{q}} |Z^*|^{q'}]\right)^{\frac{1}{q'}} \\
&= \left(E[(\frac{1}{q}M)^{-\frac{1}{q-1}} |Z^*|^{q'}]\right)^{\frac{1}{q'}} ,
\end{aligned}$$

hence, $\|Z^*\|_{\mathcal{Z}^*_T} \leq \left(E[(\frac{1}{q}M)^{-\frac{1}{q-1}}|Z^*|^{q'}]\right)^{\frac{1}{q'}}$. Now, given a fixed $Z^* \neq 0$, set $M_0 := \frac{1}{q}M$ and $Z_0 := M_0^{-\frac{1}{q}} \operatorname{sign}(M_0^{-\frac{1}{q}} Z^*)|M_0^{-\frac{1}{q}} Z^*|^{q'-1} \in L^q_{\mathcal{F}_T}(\mathbb{R})$. Using the criterion for equality in Hölders's inequality (see [DS:LO], Ex. III.9.42) in the second line, we get

$$\begin{aligned}
E[Z_0 Z^*] &= E[\operatorname{sign}(M_0^{-\frac{1}{q}} Z^*) \left|M_0^{-\frac{1}{q}} Z^*\right|^{q'-1} M_0^{-\frac{1}{q}} Z^*] \\
&= \left(E[M_0^{-\left(\frac{q'-1}{q}\right)q} |Z^*|^{(q'-1)q}]\right)^{\frac{1}{q}} \left(E[M_0^{-\frac{q'}{q}} |Z^*|^{q'}]\right)^{\frac{1}{q'}} \\
&= \left(E[\left|M_0^{-\frac{1}{q}} Z^*\right|^{(q'-1)q}]\right)^{\frac{1}{q}} \left(E[M_0^{-\frac{1}{q-1}} |Z^*|^{q'}]\right)^{\frac{1}{q'}} \\
&= \|Z_0\|_{\mathcal{Z}_T} \left(E[\tilde{M}\, |Z^*|^{q'}]\right)^{\frac{1}{q'}} .
\end{aligned}$$

Combined with the above estimate for $\|Z^*\|_{\mathcal{Z}_T}$, this shows that $\|Z^*\|_{\mathcal{Z}_T^*} = \left(E[\tilde{M}\,|Z^*|^{q'}]\right)^{\frac{1}{q'}}$.
Let us turn to the dual norm of $\|\cdot\|_{\mathcal{Z}_{RT}}$. Using the same arguments, one can show that if we equip the linear space $L^q_{\mathcal{F}}(0,T;\mathbb{R})$ with the norm $z \mapsto \left(\frac{1}{q}E[\int_0^T Q|z|^q ds]\right)^{\frac{1}{q}}$, the dual space is $L^{q'}_{\mathcal{F}}(0,T;\mathbb{R})$ with norm $z^* \mapsto \left(E[\int_0^T \tilde{Q}|z^*|^{q'} ds]\right)^{\frac{1}{q'}}$ (given the usual pairing). It then follows that the dual norm on the product space $L^{q'}_{\mathcal{F}_T}(\mathbb{R}) \times L^{q'}_{\mathcal{F}}(0,T;\mathbb{R})$ is given by $\|\cdot\|_{\mathcal{Z}^*_{RT}}$, see [A:LF], Chap. 4, Ex. 4.4 and 4.5. This finishes the proof of the remark. ■

Let us turn to the adjoint operators. Up to a technical detail, we draw the following result from [YZ:SC], Prop. 4.1 in § 6.4.

Lemma 6.9 *Consider the operators H_T and H_R introduced in (6.26) and (6.28). Their adjoint operators $H_T^* : L^{q'}_{\mathcal{F}_T}(\mathbb{R}) \longrightarrow H_{q'}(0,T;\mathbb{R}^m)$ and $H_R : L^{q'}_{\mathcal{F}}(0,T;\mathbb{R}) \longrightarrow H_{q'}(0,T;\mathbb{R}^m)$ are given as follows:*
For $Z^* \in L^{q'}_{\mathcal{F}_T}(\mathbb{R})$ let (α,β) be the unique solution in $L^{q'}_{\mathcal{F}}(\Omega, C([0,T];\mathbb{R})) \times H_{q'}(0,T;\mathbb{R}^d)$ of the BSDE

$$d\alpha = \beta' dw,\ \alpha(T) = Z^*. \tag{6.33}$$

Then, $H_T^* Z^* = B'\alpha + \sigma\beta$.
For $z^* \in L^{q'}_{\mathcal{F}}(0,T;\mathbb{R})$ let $(\alpha,\beta) \in L^{q'}_{\mathcal{F}}(\Omega, C([0,T];\mathbb{R})) \times H_{q'}(0,T;\mathbb{R}^d)$ be the unique solution of the BSDE

$$d\alpha = -z^* ds + \beta' dw,\ \alpha(T) = 0. \tag{6.34}$$

Then, $H_R z^* = B'\alpha + \sigma\beta$.

Proof: Let us first note that both BSDEs mentioned in the assertion are uniquely solvable. For the first one it's just martingale representation, for the second one it follows from Theorem 5.1 in [EPQ:BSDE], though the present situation does not precisely match the hypothesis of this theorem, as it would require that $z^* \in H_{q'}(0,T;\mathbb{R})$. For $q > 2$ this does not hold true. The proof, though, shows that in our particular situation of a linear equation $z^* \in L^{q'}_{\mathcal{F}}(0,T;\mathbb{R})$ is sufficient.
The proof now consists in both cases of merely applying Itô's formula to $\alpha \cdot H_R u$ and to evaluate $E[\int_0^T d(\alpha \cdot H_R u)]$. This leads to

$$\begin{aligned} E[(H_T u) Z^*] &= E[u(H_T^* Z^*)] \text{ for all } u \in \mathcal{V},\ Z^* \in \mathcal{Z}_T^* \\ E[\int_0^T (H_R u) z^* ds] &= E[\int_0^T u(H_R z^*) ds] \text{ for all } u \in \mathcal{V},\ z^* \in L^{q'}_{\mathcal{F}}(0,T;\mathbb{R}). \end{aligned}$$

■

We now can describe $\mathcal{M}_T^\perp$ and $\mathcal{M}_{RT}^\perp$. The third part of the following lemma states a parametrization of the kernel $\ker \sigma$. It will be helpful when stating dual problems as linear isoelastic stochastic control problems.

Lemma 6.10 *Set* $\theta := \sigma'(\sigma\sigma')^{-1}B'$ *. Then we have*

1.

$$\begin{aligned}\mathcal{M}_T^{\perp} \;=\; & \Big\{\alpha_1(T) \in L^{q'}_{\mathcal{F}_T}(\mathbb{R}) : \alpha_1 \textit{ is the solution of the SDE} \\ & d\alpha_1 = \{-\theta'\alpha_1 + k'\}\,dw,\ \alpha_1(0) = \alpha_{1,0},\ \textit{for some } k \in H_{q'}(0,T;\mathbb{R}^d) \textit{ with} \\ & \sigma k = 0,\ Leb \otimes P - a.s.,\ \textit{and some } \alpha_{1,0} \in \mathbb{R}\Big\}, \qquad (6.35)\end{aligned}$$

2.

$$\begin{aligned}\mathcal{M}_{RT}^{\perp} \;=\; & \Big\{(z^*, \alpha_2(T)) \in L^{q'}_{\mathcal{F}}(0,T;\mathbb{R}) \times L^{q'}_{\mathcal{F}_T}(\mathbb{R}) : \alpha_2 \textit{ is the solution of the SDE} \\ & d\alpha_2 = -z^* ds + \{-\theta'\alpha_2 + k'\}\,dw,\ \alpha_2(0) = \alpha_{2,0},\ \textit{for some} \\ & z^* \in L^{q'}_{\mathcal{F}}(0,T;\mathbb{R}),\ \textit{some } k \in H_{q'}(0,T;\mathbb{R}^d) \\ & \textit{with } \sigma k = 0,\ Leb \otimes P - a.s.,\ \textit{and some } \alpha_{2,0} \in \mathbb{R}\Big\}. \qquad (6.36)\end{aligned}$$

3. *If* $m < d$ *choose a matrix-valued process* $\widetilde{\sigma} \in L^{\infty}_{\mathcal{F}}(0,T;\mathbb{R}^{(d-m)\times d})$ *such that*
   (i) $\widetilde{\sigma}\widetilde{\sigma}' \gg 0$, *i.e* $\widetilde{\sigma}\widetilde{\sigma}'$ *is uniformly positive;*
   (ii) the matrix $\left[\begin{smallmatrix}\sigma\\ \widetilde{\sigma}\end{smallmatrix}\right]$ *is* $Leb \otimes P - a.s.$ *regular;*
   (iii) $\sigma\widetilde{\sigma}' = 0,\ Leb \otimes P - a.s.;$
   Then,

$$\begin{aligned}&\big\{k \in H_{q'}(0,T;\mathbb{R}^d) : \sigma k = 0,\ Leb \otimes P - a.s.\big\} \\ &\quad = \big\{\widetilde{\sigma}'v \in H_{q'}(0,T;\mathbb{R}^d) : v \in H_{q'}(0,T;\mathbb{R}^{d-m})\big\}. \qquad (6.37)\end{aligned}$$

Proof:

1. With the definition of $\mathcal{M}_T^{\perp}$ and $\mathcal{M}_T$ we have

$$\begin{aligned}\mathcal{M}_T^{\perp} &= \Big\{Z^* \in L^{q'}_{\mathcal{F}_T}(\mathbb{R}) : < Z, Z^* >_{\mathcal{Z}_T} = 0 \text{ for all } Z \in \mathcal{M}_T\Big\} \\ &= \Big\{Z^* \in L^{q'}_{\mathcal{F}_T}(\mathbb{R}) : < H_T u, Z^* >_{\mathcal{Z}_T} = 0 \text{ for all } u \in \mathcal{V}\Big\} \\ &= \Big\{Z^* \in L^{q'}_{\mathcal{F}_T}(\mathbb{R}) : E[\int_0^T u'(H_T^* Z^*)ds] = 0 \text{ for all } u \in H_q(0,T;\mathbb{R}^m)\Big\} \\ &= \Big\{Z^* \in L^{q'}_{\mathcal{F}_T}(\mathbb{R}) : H_T^* Z^* = 0,\ Leb \otimes P - a.s.\Big\}.\end{aligned}$$

   Hence, $\mathcal{M}_T^*$ coincides with the kernel of H_T^* and we have to show that the set described in the assertion is a parametrization of this kernel.
   First assume that we are given an element $\alpha_1(T)$ from (6.35). Then there is an $\alpha_{1,0} \in \mathbb{R}$ and a k with $\sigma k = 0$, such that $d\alpha_1 = \{-\theta'\alpha_1 + k'\}\,dw,\ \alpha_1(0) = \alpha_{1,0}$. We have to show that $\alpha_1(T) \in \ker H_T^*$. Let (α, β) be the solution of equation (6.33) with

$Z^* := \alpha_1(T)$. From unicity of solutions it is clear that $\alpha = \alpha_1$, and $\beta = -\theta\alpha_1 + k$. Hence,

$$H_T^*(\alpha_1(T)) = B'\alpha + \sigma\beta = B'\alpha_1 - \sigma\theta\alpha_1 + \sigma k = \sigma k = 0,\ Leb \otimes P - a.s.$$

Now assume that some Z^* belongs to $\ker H_T^*$. We have to find $\alpha_{1,0}$ and k such that for the resulting α_1 we have $\alpha_1(T) = Z^*$. Let (α, β) be the solution of (6.33) for the Z^* under consideration. Set $k := \beta + \theta\alpha \in H_{q'}(0, T; \mathbb{R}^d)$ and $\alpha_{1,0} := \alpha(0)$. As $H_T^* Z^* = 0$ we have $B'\alpha + \sigma\beta = 0$, hence

$$\sigma k = \sigma(\beta + \theta\alpha) = B'\alpha + \sigma\beta = 0,$$

$Leb \otimes P - a.s.$. By definition, $\alpha_1 := \alpha$ satisfies $d\alpha_1 = \{-\theta\alpha_1 + k\}\, dw$, $\alpha_1(0) = \alpha_{1,0}$. Hence, $\mathcal{M}_T^*$ has the form as stated in the assertion.

2. As above we have

$$\begin{aligned}
\mathcal{M}_{RT}^\perp &= \Big\{(z^*, Z^*) \in L_{\mathcal{F}}^{q'}(0, T; \mathbb{R}) \times L_{\mathcal{F}_T}^{q'}(\mathbb{R}) : \\
&\quad < (z, Z), (z^*, Z^*) >_{\mathcal{Z}_{RT}} = 0 \text{ for all } (z, Z) \in \mathcal{M}_{RT}\Big\} \\
&= \{(z^*, Z^*) :< (H_R u, H_T u), (z^*, Z^*) >_{\mathcal{Z}_{RT}} = 0 \text{ for all } u \in \mathcal{V}\} \\
&= \Big\{(z^*, Z^*) : E[\int_0^T u'(H_R^* z^* + H_T^* Z^*) ds] = 0 \text{ for all } u \in H_q(0, T; \mathbb{R}^m)\Big\} \\
&= \{(z^*, Z^*) : H_R z^* + H_T^* Z^* = 0,\ Leb \otimes P - a.s.\}\,. \qquad (6.38)
\end{aligned}$$

We first want to characterize the set in (6.38) via a BSDE and claim that

$$\begin{aligned}
&\{(z^*, Z^*) : H_R z^* + H_T^* Z^* = 0\} \\
= &\{(z^*, Z^*) : \text{ the solution of } d\alpha = -z^* + \beta' dw,\ \alpha(T) = Z^*, \\
&\text{satisfies } B'\alpha + \sigma\beta = 0\}\,. \qquad (6.39)
\end{aligned}$$

To see this, assume $H_R z^* + H_T Z^* = 0$ for some (z^*, Z^*). Let (a_1, b_1) rectively (a_2, b_2) be the solution of (6.33) respectively (6.34) for this Z^* respectively z^* and set $\alpha := a_1 + a_2$ and $\beta = b_1 + b_2$. Then, (α, β) obviously satisfies the BSDE on the right hand side of (6.39) and

$$0 = H_R^* z^* + H_T^* Z^* = B' a_1 + \sigma b_1 + B a_2 + \sigma b_2 = B'\alpha + \sigma\beta,$$

what shows the inclusion "$\subset$" in (6.39). Conversely, assume that we are given a pair (z^*, Z^*) from the set on the right hand side of (6.39). Define (a_1, b_1), (a_2, b_2) just as above. Again, set $\alpha = a_1 + a_2$, $\beta = b_1 + b_2$, and $B'\alpha + \sigma\beta = 0$ now entails $B' a_1 + \sigma b_1 + B' a_2 + \sigma b_2 = H_R^* z^* + H_T^* Z^* = 0$. This proves (6.39).
Finally, it's easy to see that the set on the right hand side of (6.39) coincides with the "parametrization" of $\mathcal{M}_{RT}^*$ in (6.36): Given a triplet z^*, k and α_0 from this

parametrization with solution (α_2, β_2) of the resulting SDE, set $\alpha := \alpha_2$ and $\beta := -\theta\alpha_2 + k$. (α, β) then satisfies the BSDE on the right hand side of (6.39), and due to the definition of θ we have $B'\alpha + \sigma\beta = 0$. Conversely, if (z^*, Z^*) belongs to the right hand side of (6.39) with corresponding (α, β), then $k := \beta + \theta\alpha$ and $\alpha_{2,0} := \alpha(0)$ will yield the desired representation of Z^*.

3. At first sight it may not be completely clear that a process $\widetilde{\sigma}$ with the properties $(i) - (iii)$ exists. A precise argument is given in Appendix A.
What follows is meant to hold $Leb \otimes P - a.s.$. From (iii) we have that the image of $\widetilde{\sigma}'$ is contained in $\ker\sigma$, from (ii) we get that it actually coincides with $\ker\sigma$. Hence, for every $k \in \ker\sigma$ there is a v such that $\widetilde{\sigma}'v = k$. It is easily seen that $v = (\widetilde{\sigma}\widetilde{\sigma}')^{-1}\widetilde{\sigma}k$. Passing to the consideration of processes (i) now shows that $v \in H_{q'}(0, T; \mathbb{R}^{d-m})$ if $k \in H_{q'}(0, T; \mathbb{R}^d)$.

■

Let $\alpha_1(T)$ be as in the definition of $\mathcal{M}_T^\perp$ with $\alpha_1(0) = 1$. $dR := \alpha_1(T)dP$ then defines a signed martingale measure on Ω for $\mathcal{S}(t) := \int_0^t B'ds + \int_0^t \sigma dw$, i.e. $\alpha_1(t)\mathcal{S}(t)$ is a P-martingale, where $\alpha_1(t) = E[\alpha_1(T)|\mathcal{F}_t]$. If conversely $\alpha_1(T)$ is a q'-integrable r.v. such that $\alpha_1(t)\mathcal{S}(t)$ is a martingale (with $\alpha_1(t) = E[\alpha_1(T)|\mathcal{F}_t]$), then arguments similar to those in the proof show that $\alpha_1(T)$ belongs to $\mathcal{M}_T^\perp$. Hence, the dual space of $\mathcal{M}_T$ consists of signed martingale measures for S that are multiplied with some real. A characterization of equivalent martingale measures for S in terms of a SDE is given in [EQ:DPPC], Proposition 1.8. Martingale measures in different generality (equivalent, local, signed) play a central role in mathematical finance. See [L:UDP], Section 1.4 for a short overview.
The representation of $\mathcal{M}_T^*$ and $\mathcal{M}_{RT}^*$ in terms of SDEs will allow us to treat the dual problems as stochastic control problems that fit into the linear isoelastic framework.
So much for the objects involved in the dual problems for (6.15) and (6.19). The last lemma of this section indicates how to determine the solution of the primal problem, once the solution for dual problem is found. The result for the terminal cost problem are quite similar to those of [Ph:DLPH], Theorems 4.1 and 5.1, where L^q-hedging is considered in discrete time with cone constraints on the portfolio.

Lemma 6.11 *Consider the recast forms of the hedging problems with terminal respectively running and terminal cost (6.31), (6.32), and their dual problems. Let $\tilde{M}$, $\tilde{Q}$ be as in Remark 6.8*

1. *Let $\bar{u}$ be the solution for (6.15) and let $\bar{Z}^*$ be the solution for the dual problem to (6.31). (the existence of $\bar{Z}^*$ follows from Theorem 6.7). Let $\overline{\alpha}_{1,0}$ be the initial value in the representation of $\bar{Z}^*$ in (6.35). Then there is a $c_T > 0$ such that $-x_0 - (H_R\bar{u})(T) = c_T\,\tilde{M}\,sign(\bar{Z}^*)|\bar{Z}^*|^{q'-1}$; more precisely, $c_T = |x_0\overline{\alpha}_{1,0}|(E[\tilde{M}\,|\bar{Z}^*|^{q'}])^{-\frac{1}{q}}$. $\bar{Z}^*$ $P-a.s$ never has the sign of x_0.*

2. *Let $\bar{u}$ be the solution for (6.19) and let $(\bar{z}^*, \bar{Z}^*)$ be the solution for the dual problem to (6.32) (see Theorem 6.7). Let $\overline{\alpha}_{2,0}$ be the initial value in the representation of $\bar{Z}^*$ in*

(6.36). Then, there constants $c_R, c_T > 0$ such that. $-x_0 - H_R\bar{u} = c_R\,\tilde{Q}\,sign(\bar{z}^*)|\bar{z}^*|^{q'-1}$ and $-x_0 - (H_R\bar{u})(T) = c_T\,\tilde{M}\,sign(\bar{Z}^*)|\bar{Z}^*|^{q'-1}$. $\bar{z}^*$ and $\bar{Z}^*$ $Leb \otimes P-$ respectively $P-a.s.$ never have the sign of x_0. Let Γ be the solution of $d\Gamma = -\theta'\Gamma dw$, $\Gamma(0) = 1$, $\theta := \sigma'(\sigma\sigma')^{-1}B'$. Then c_T can be calculated from $x_0 + c_T E[\Gamma(T)\,sign(\bar{Z}^*)|\bar{Z}^*|^{q'-1}] = 0$. c_R can be calculated from

$$-x_0\overline{\alpha}_{2,0} = E[c_R^q \int_0^T \tilde{Q}|\bar{z}^*|^{q'}ds + c_T^q\,\tilde{M}|\bar{Z}^*|^{q'}]. \tag{6.40}$$

Proof: Let $\bar{x}$ be the optimal state process for the problems (6.15) or (6.19). By Remark 4.4, $\bar{x}(T)$ cannot be identical zero unless $x_0 = 0$. In particular, the optimal cost for problems (6.15) and (6.19) is strictly positive. The representation of the optimal cost in Theorem 6.7 then shows that $\bar{Z}^* \not\equiv 0$ respectively $(\bar{z}^*, \bar{Z}^*) \not\equiv (0,0)$.

1. Set $\overline{h} = H_R\bar{u}$. From Theorem 6.7 we know that $-x_0 - \overline{h}(T)$ and $\bar{Z}^*$ are aligned i.e. $< -x_0 - \overline{h}(T), \bar{Z}^* >_{\mathcal{Z}_T} = \left\|-x_0 - \overline{h}(T)\right\|_{\mathcal{Z}_T} \left\|\bar{Z}^*\right\|_{\mathcal{Z}_T^*}$, and as neither $-x_0 - \overline{h}(T)$ nor $\bar{Z}^*$ are identically zero, this expression is strictly greater than zero. Set $M_0 := \frac{1}{q}M$. Since $M_0^{\frac{1}{q}}\tilde{M}^{\frac{1}{q'}} = 1$, the alignment (used in the second line of the following equation) means that

$$\begin{aligned} E[(-x_0 - \overline{h}(T))\bar{Z}^*] &= E[M_0^{\frac{1}{q}}(-x_0 - \overline{h}(T))\,\tilde{M}^{\frac{1}{q'}}\,\bar{Z}^*] \\ &= \left\|-x_0 - \overline{h}(T)\right\|_{\mathcal{Z}_T} \left\|\bar{Z}^*\right\|_{\mathcal{Z}_T^*} \\ &= \left(E[\left|M_0^{\frac{1}{q}}(-x_0 - \overline{h}(T))\right|^q]\right)^{\frac{1}{q}} \left(E[\left|\tilde{M}^{\frac{1}{q'}}\bar{Z}^*\right|^{q'}]\right)^{\frac{1}{q'}}. \end{aligned}$$

Hölder's inequality gives that the right hand side of the first line of this is less or equal the right hand side of the third line. The criterion for equality in Hölder's inequality thus gives that there is a $c_T \in \mathbb{R}$ such that

$$M_0^{\frac{1}{q}}(-x_0 - \overline{h}(T)) = c_T\,\tilde{M}^{\frac{1}{q}}\,\mathrm{sign}(\bar{Z}^*)|\bar{Z}^*|^{q'-1}. \tag{6.41}$$

As $M_0^{-\frac{1}{q}}\tilde{M}^{\frac{1}{q}} = \tilde{M}$, this yields

$$-x_0 - \overline{h}(T) = c_T\,\tilde{M}\,\mathrm{sign}(\bar{Z}^*)|\bar{Z}^*|^{q'-1}. \tag{6.42}$$

Hence, $< -x_0 - \overline{h}(T), \bar{Z}^* >_{\mathcal{Z}_T} = c_T \left\|\bar{Z}^*\right\|_{\mathcal{Z}_T^*}^{q'} > 0$, what gives $c_T > 0$. By the representation of $\bar{Z}^*$ as in (6.35) there is an $\overline{\alpha}_{1,0} \in \mathbb{R}$ and a $k \in \ker\sigma$ such that $\bar{Z}^*(T) = \alpha_1(T)$, where α_1 is the solution of $d\alpha_1 = \{-\theta\alpha_1 + k\}\,dw$, $\alpha_1(0) = \overline{\alpha}_{1,0}$. This entails that

$$< -x_0, \bar{Z}^* >_{\mathcal{Z}_T} = -x_0\overline{\alpha}_{1,0}. \tag{6.43}$$

Noting that we have to modify the optimal values, we get from Theorem 6.7 that $(-x_0\overline{\alpha}_{1,0})^q = J^{(T)}(\bar{u})$. Now we can calculate c_T from

$$\begin{aligned}(-x_0\overline{\alpha}_{1,0})^q &= J^{(T)}(\bar{u}) \\ &= \left\| -x_0 - \overline{h}(T) \right\|^q_{\mathcal{Z}_T} \\ &= |c_T|^q \, E[M_0 \left(\tilde{M} \, |\bar{Z}^*|^{q'-1} \right)^q] \\ &= |c_T|^q \, E[\tilde{M} |\bar{Z}^*|^{q'}] = c_T^q \, \left\| \bar{Z}^* \right\|^{q'}_{\mathcal{Z}^*_T}, \end{aligned} \tag{6.44}$$

since $M_0 \tilde{M}^q = \tilde{M}$. $\bar{x}(T) = x_0 + \overline{h}(T)$ is the optimal terminal state for problem (6.15), and Lemma 3.8 thus gives that this r.v. is only of one sign - that of x_0 (not counting 0 as an extra sign). So, from (6.42) it follows that $\bar{Z}^*$ never is of the sign of x_0.

2. Again, set $\overline{h} = H_R \bar{u}$. It is easily seen that the alignment of $(-x_0 - \overline{h}, -x_0 - \overline{h}(T))$ with $(\bar{z}^*, \bar{Z}^*)$ guaranteed by Theorem 6.7 entails the alignment of $-x_0 - \overline{h}$ with $\bar{z}$ and of $-x_0 - \overline{h}(T)$ with $\bar{Z}^*$. With the same arguments as in part *1.* of this proof it follows that there are constants $c_R, c_T \in \mathbb{R}$ such that

$$-x_0 - \overline{h} = c_R \, \tilde{Q} \, \mathrm{sign}(\bar{z}^*) |\bar{z}^*|^{q'-1}, \;\; -x_0 - \overline{h}(T) = c_T \, \tilde{M} \, \mathrm{sign}(\bar{Z}^*) |\bar{Z}^*|^{q'-1}. \tag{6.45}$$

By Remark 4.4, $\bar{x}(T) = x_0 + \overline{h}(T)$ is not identical zero, and by Lemma 3.8 it is only of one sign. Let us consider the SDE (6.16) for $\bar{x}$ as an BSDE by setting $S = \bar{u}'\sigma$. Then, $\bar{x}$ satisfies

$$d\bar{x} = \theta' S' ds + S dw, \;\; \bar{x}(T) = -c_T \, \tilde{M} \, \mathrm{sign}(\bar{Z}^*) |\bar{Z}^*|^{q'-1}.$$

The solution part $\bar{x}$ of this BSDE is given by

$$\bar{x}(t) = \Gamma^{-1}(t) E[\Gamma(T) \bar{x}(T) | \mathcal{F}_t],$$

(compare [EM:BSDE], §2.2, Prop. 2.4), hence $x_0 = -c_T \, E[\Gamma(T) \, \tilde{M} \, \mathrm{sign}(\bar{Z}^*) |\bar{Z}^*|^{q'-1}]$, where the expectation does not vanish. This enables us to compute c_T. As $\bar{x}(T) \not\equiv 0$, the alignment shows that $c_T > 0$. Further, we have $\bar{z}^* \not\equiv 0$, since otherwise from (6.45) it follows that $x_0 + \overline{h} = \bar{x} = 0$, what would imply that $\bar{x}(T) \equiv 0$. So again the alignment argument shows $c_R > 0$, and as $\bar{x}$ never changes its sign (by Lemma 3.8), we see that $\bar{z}^*$ is $Leb \otimes P - a.s.$ of the opposite sign of x_0.
Now let $k \in \ker \sigma$ be the process in the representation of $\bar{Z}^*$ in (6.36). Let α_2 be the solution of the SDE $d\alpha_2 = -\bar{z}^* ds + \{-\theta\alpha_2 + k'\} \, dw$, $\alpha_2(0) = \overline{\alpha}_{2,0}$, then $\bar{Z}^* = \alpha_2(T)$. With this representation we get from Theorem 6.7

$$\begin{aligned} J^{(RT)}(\bar{u}) &= \, < (-x_0, -x_0), (\bar{z}^*, \bar{Z}^*) >_{\mathcal{Z}_{RT}} \\ &= E[\int_0^T -x_0 \, \bar{z}^* ds - x_0 \, \bar{Z}^*] \\ &= -x_0 E[\int_0^T \bar{z}^* ds + \alpha_{2,0} - \int_0^T \bar{z}^* ds + \int_0^T -\theta\alpha_2 + k dw] \\ &= -x_0 \alpha_{2,0}. \end{aligned} \tag{6.46}$$

Plugging the alignments (6.45) in the cost functional $J^{(RT)}$ now yields (6.40) (with the same calculations that lead to (6.44)). As $\bar{z}^* \not\equiv 0$, this latter relation allows us to determine c_R. This finishes the proof.

■

The relation between the solution of the primal dual problem established in the previous lemma enables us to completely determine the optimal state $\bar{x}$ of the primal problem if we have complete knowledge of $\bar{Z}^*$ respectively $(\bar{z}^*, \bar{Z}^*)$. It is now a matter of BSDE-theory (compare (3.3), (3.4)) to determine the optimal control $\bar{u}$ for the primal problem. If we construct $\bar{x}(T)$ from $\bar{Z}^*$, we can solve the BSDE

$$\begin{aligned} dR &= \theta' S' ds + S dw, \\ R(T) &= \bar{x}(T). \end{aligned}$$

R must coincide the solution of the state equation for $\bar{x}$,

$$d\bar{x} = B\bar{u} + \bar{u}'\sigma dw, \quad \bar{x}(0) = x_0,$$

since $(\bar{x}, u'\sigma)$ solves the BSRDE for (R, S). Thus, S' is in the range of σ' and $\bar{u}$ is given by $\bar{u} = (\sigma\sigma')^{-1} \sigma S'$. Note that $\sigma' (\sigma\sigma')^{-1} \sigma$ is the identity on the range of σ', so this choice of $\bar{u}$ indeed satisfies $\sigma' u = S'$.
Hence we will consider the primal problem as solved if the dual problem is.

6.3 Solutions to the problems

This section gives the solutions for the hedging problems with terminal respectively running and terminal cost. All problems can be handled with the BSRDE results that we developed. So, the introduction of dual problems may seem superfluous. But these dual problems exhibit interesting properties and may offer some advantages, depending on the point of view. E.g., in the case of running cost and complete market, the dual approach delivers directly the optimal final state for problem (6.15). In the case of terminal cost and incomplete market, the dual problem may have fewer control variables than the primal problem. When considering running and terminal cost in an complete market, we find that the control variable in the dual problem is one dimensional and does not enter the diffusion part of the state equation. Running and terminal cost in an incomplete market yield a dual problem that seems interesting on its own right and can be handled using the solution of the primal problem (admittedly, this is in general not what the introduction of a dual problem is meant for).
Lemma 6.10 enables us to handle the dual problem basically as a kind of linear isoelastic stochastic control problem. In the most easiest case - hedging of terminal cost in a complete market - this may hardly be recognized, since there will be no real control variable

and the minimization is performed solely with respect to some initial value. However, it's a nice and very basic illustration of the duality. Besides, we will see that in this particular case we are spared (if we wish) from invoking BSRDEs.

6.3.1 Terminal cost, complete market

Let us consider problem (6.15) in a complete market setting. By Remark 6.4 this means, that $d = m$ and thus σ is $Leb \otimes P - a.s.$ regular. So, the orthogonal complement of $\mathcal{M}_T$ boils down to (see Lemma 6.10) $\mathcal{M}_T^{\perp} = \{\alpha_{1,0}\,\alpha_1(T) : \alpha_{1,0} \in \mathbb{R},\, d\alpha_1 = -\theta\alpha_1 dw,\, \alpha_1(0) = 1\}$, i.e. is one-dimensional. For simplicity, we will fix within this subsection the notation $\alpha_1(T)$ for the r.v. denoted by this in the above parametrization of $\mathcal{M}_T^{\perp}$. Plugging this into (6.22) gives

$$\min_{Z \in \mathcal{M}_T} \|-x_0 - Z\|_{\mathcal{Z}_T} = \max_{\substack{\alpha_{1,0} \in \mathbb{R} \\ \|\alpha_{1,0}\,\alpha_1(T)\|_{\mathcal{Z}_T^*} \le 1}} < -x_0, \alpha_{1,0}\alpha_1(T) >_{\mathcal{Z}_T} .$$

As $< -x_0, \alpha_{1,0}\alpha_1(T) >_{\mathcal{Z}_T} = -x_0\alpha_{1,0}$ this is means

$$\min_{Z \in \mathcal{M}_T} \|-x_0 - Z\|_{\mathcal{Z}_T} = \max_{\substack{\alpha_{1,0} \in \mathbb{R} \\ \|\alpha_{1,0}\alpha_1(T)\|_{\mathcal{Z}_T^*} \le 1}} -x_0\alpha_{1,0}.$$

It is clear that the maximum on the right hand side of this last equation is achieved in some $\overline{\alpha}_{1,0}$ for which $\|\overline{\alpha}_{1,0}\alpha_1(T)\|_{\mathcal{Z}_T^*} = |\overline{\alpha}_{1,0}|\ \|\alpha_1(T)\|_{\mathcal{Z}_T^*} = 1$ and which has the opposite sign of x_0. Hence, $\overline{\alpha}_{1,0} = -\text{sign}(x_0)(\|\alpha_1(T)\|_{\mathcal{Z}_T^*})^{-1}$, i.e.

$$\overline{\alpha}_{1,0} = -\text{sign}(x_0) \left(E[\tilde{M}|\alpha_1(T)|^{q'}]\right)^{-\frac{1}{q'}} .$$

Let $\bar{x}$ be the optimal state for the primal problem (6.15). We determine $\bar{x}(T)$ by Lemma 6.11-1. In the notation of this lemma we have $\bar{Z}^* = \overline{\alpha}_{1,0}\,\alpha_1(T)$, hence

$$c_T = |x_0||\alpha_{1,0}|^{1-\frac{q'}{q}} \left(E[\tilde{M}|\alpha_1(T)|^{q'}]\right)^{-\frac{q'}{q}}$$

and

$$\bar{x}(T) = x_0\,|\overline{\alpha}_{1,0}|^{q'}\ \tilde{M}\ \alpha_1(T)^{q'-1}. \tag{6.47}$$

Let us have a look at the BSRDE that corresponds to the primal problem (6.15) in this situation, i.e. when σ is a square matrix. If we set $A = 0$ and $C = (C^1, \ldots, C^d)' = 0$ in (4.27) we get

$$\begin{aligned} dK &= \left\{ \frac{q}{2(q-1)^2} B\,(\sigma\sigma')^{-1} B'K + q'B\,(\sigma\sigma')^{-1}\sigma L + \frac{q}{2}\frac{1}{K} L'\sigma'\,(\sigma\sigma')^{-1}\sigma L \right. \\ & \left. + \frac{2-q}{2}\frac{1}{K}|L|^2 \right\} ds + L'dw. \end{aligned}$$

Let us replace $\sigma' (\sigma\sigma')^{-1} B'$ by θ. This yields

$$dK = \left\{ \frac{q}{2(q-1)^2}|\theta|^2 K + q'\theta' L + \frac{q}{2}\frac{1}{K} L'\sigma' (\sigma\sigma')^{-1} \sigma L + \frac{2-q}{2}\frac{1}{K}|L|^2 \right\} ds + L'dw. \quad (6.48)$$

If we use the regularity of σ this finally gives the BSRDE for problem (6.15)

$$dK = \left\{ \frac{q}{2(q-1)^2}|\theta|^2 K + q'\theta' L + \frac{1}{K}|L|^2 \right\} ds + L'dw, \quad (6.49)$$

$$K(T) = f(M). \quad (6.50)$$

This BSRDE is solvable by Theorem 5.5. The optimal control for the corresponding problem is given in Theorem 5.11. If we let $q = 2$ in this BSRDE, no terms vanish, and so the change in the form of the BSRDE, if we switch here from linear-quadratic to linear isoelastic problems, is not so big, especially if $M = 1$.
When we proved that the BSRDE we considered is solvable, our approach was to show that the optimal state $\bar{x}$ of the corresponding control problem does not vanish. In the specific framework of the present subsection we encountered via the duality a situation, where such knowledge about the optimal state is a priori available, by (6.47). This allows us to give an alternative proof of the solvability of (6.49), (6.50). We want to sketch it in the following remark, just for theoretical interest and because it's not much work.

Remark 6.12 *Consider problem (6.15) in a complete market setting (i.e. for $d = m$). Then the corresponding BSRDE (6.49), (6.50) has a solution (K, L).*

Proof: Let $\bar{x}$ be the optimal state process for problem (6.15) and $\bar{u}$ be the optimal control. From (6.47) we have $\bar{x}(T) = x_0 \, |\overline{\alpha}_{1,0}|^{q'} \, \tilde{M}\alpha_1(T)^{q'-1}$. Without loss of generality we can assume that $x_0 > 0$. As $\tilde{M} \gg 0$ and α_1 is a stochastic exponential we get that $\bar{x}(T) > 0$, $P-a.s.$. Consider the SDE (6.16), (6.17) as an BSDE. Set $R := \bar{x}$ and $S' := \bar{u}'\sigma$. Then,

$$dR = \theta' S ds + S' dw, \quad R(T) = \bar{x}(T).$$

The solution part R of this BSDE is given by

$$R(t) = x_0 \, |\overline{\alpha}_{1,0}|^{q'} \, \alpha_1(t)^{-1} E[\tilde{M} \, \alpha_1(T)^{q'} |\mathcal{F}_t],$$

$t \in [0, T]$. It follows, that $P - a.s.$ we have $\bar{x}(t) > 0$ for all $t \in [0, T]$. The rest of the proof now follows the line of the proof of Theorem 5.5 - 2. ■

The advantage here is that one need not invoke the a priori estimates on L as given in Theorems 4.13 or 4.16. The obvious disadvantage is that this proof of existence indicates nothing about uniqueness or about the question, if one can construct an optimal control for the underlying control problem out of (K, L).

We will now turn to the incomplete market case.

6.3.2 Terminal cost, incomplete market

Let us consider problem (6.15) in the incomplete market setting, i.e. when σ is not a square matrix. Again, the problem can be solved via BSRDE theory. The existence of a solution for the corresponding BSRDE is made sure by Theorem 5.5, and the optimal control is given in Theorem 5.11.
More precisely, let (K, L) be the unique solution of (6.48) with terminal condition $K(T) = f(M)$. Then, the optimal state process for the hedging problem (6.15) follows the SDE

$$d\bar{x} = BG(K, L)\bar{x}ds + \bar{x}G'(K, L)\sigma dw, \ \bar{x}(0) = x_0,$$

with

$$G(K, L) = -(\sigma\sigma')^{-1}\left(\frac{1}{q-1}B' + \frac{1}{K}\sigma L\right).$$

The optimal portfolio $\bar{u}$ is given by

$$\bar{u} = G(K, L)\bar{x},$$

and the minimal cost is

$$J^{(T)}(\bar{u}) = \frac{1}{q}K^{q-1}(0)|x_0|^q.$$

Though the situation is covered by our BSRDE results, we want to consider the dual problem. We will see that the dual problem leads to a control problem with a $(d-m)$-dimensional control variable. So, if one wishes, one could eventually "save" control variables by switching to the dual problem.
In the incomplete market case, $\mathcal{M}_T^{\perp}$ does not reduce to a one dimensional vector space, hence the dual problem turns out to be a bit more complicated, it is in fact stochastic control problem. The duality relation of Theorem 6.7, applied to problem (6.15) reads as

$$\min_{Z \in \mathcal{M}_T} \|-x_0 - Z\|_{\mathcal{Z}_T} = \max_{\substack{\alpha_1(T) \in \mathcal{M}_T^{\perp} \\ \|\alpha_1(T)\|_{\mathcal{Z}_T^*} \leq 1}} < -x_0, \alpha_1(T) >_{\mathcal{Z}_T} .$$

$\mathcal{M}_T^{\perp}$ is given by (6.35) and (6.37) in Lemma 6.10 as

$$\{\alpha_1(T) : \alpha_1 \text{ satisfies } d\alpha_1 = \{-\theta'\alpha_1 + v'\widetilde{\sigma}\}\, dw, \alpha_1(0) = \alpha_{1,0},$$
$$\text{for some } v \in H_{q'}(0, T; \mathbb{R}^{d-m}) \text{ and some } \alpha_{1,0} \in \mathbb{R}\},$$

for some $\widetilde{\sigma}$ that enjoys the properties stated in Part *3.* of Lemma 6.10. As the processes α_1 in this parametrization of $\mathcal{M}_T^{\perp}$ are martingales, we have

$$< -x_0, \alpha_1(T) >_{\mathcal{Z}_T} := E[-x_0\alpha_1(T)] = -x_0\alpha_1(0),$$

so that the above equality can be written as

$$\min_{Z \in \mathcal{M}_T} \|-x_0 - Z\|_{\mathcal{Z}_T} = \max_{\substack{\alpha_1(T) \in \mathcal{M}_T^{\perp} \\ \|\alpha_1(T)\|_{\mathcal{Z}_T^*} \leq 1}} -x_0\alpha_1(0). \tag{6.51}$$

Let us have a look at the right hand side of this expression. It is clear that it becomes maximal if $\alpha_1(0)$ has the opposite sign of x_0 and if $|\alpha_1(0)|$ is maximal under those values for which there is an $v \in H_{q'}(0,T;\mathbb{R}^{d-m})$ such that the resulting $\alpha_1(T)$ still satisfies $\|\alpha_1(T)\|_{\mathcal{Z}_T^*}^{q'} = E[\tilde{M}|\alpha_1(T)|^{q'}] \leq 1$. If $\overline{\alpha}_1(T)$ is optimal, we must necessarily have $\|\overline{\alpha}_1(T)\|_{\mathcal{Z}_T^*} = 1$, since otherwise $\hat{\alpha}_1(T) := \frac{1}{\|\overline{\alpha}_1(T)\|_{\mathcal{Z}_T^*}}\overline{\alpha}_1(T)$ would yield a larger value on the right hand side of (6.51).
We can determine the optimal $\alpha_1(T)$ in the following way:
Consider the stochastic control problem

$$J_*^{(T)}(v) = \frac{1}{q'}E[q'\tilde{M}|a(T)|^{q'}] = \min_{v\in H_{q'}(0,T;\mathbb{R}^{d-m})}! \tag{6.52}$$

where a is the solution of

$$da = \{-\theta a + v'\widetilde{\sigma}\}\, dw,\ a(0) = 1. \tag{6.53}$$

$\tilde{M}$ is uniformly positive, and also $\widetilde{\sigma}\widetilde{\sigma}'$ can be assumed to be uniformly positive. Hence, the above control problem (6.52) satisfies Assumption A1 or A3, depending on q. So, by Lemma 3.2, it has a unique solution. Let $\bar{a}, \bar{v}$ be the optimal state and the optimal control for this problem. In Theorem 5.5 we have seen, that $\bar{a}$ is $P-a.s.$ not zero, hence $\|\bar{a}(T)\|_{\mathcal{Z}_T^*} > 0$. Clearly, we have

$$\overline{\alpha}_1(T) := \frac{-\text{sign}(x_0)}{\|\bar{a}(T)\|_{\mathcal{Z}_T^*}}\bar{a}(T) \in \mathcal{M}_T^{\perp}$$

with $\|\overline{\alpha}_1(T)\|_{\mathcal{Z}_T^*} = 1$. $\overline{\alpha}_1(T)$ now is optimal for the right hand side of (6.51) - if there were a different minimizing $\hat{\alpha}_1$ with $|\hat{\alpha}_1(0)| > |\overline{\alpha}_1(0)|$ and $\|\hat{\alpha}_1(T)\|_{\mathcal{Z}_T^*} = 1$ one could construct out of this a control $\hat{v}$ with $J_*^{(T)}(\hat{v}) < J_*^{(T)}(\bar{v})$, a contradiction.
Thus, the dual problem reduces to the stochastic control problem (6.52), which, like the primal problem, can be solved with a stochastic Riccati equation. The BSRDE for problem (6.52) reads as

$$\begin{aligned} d\mathcal{K} &= \left\{-\frac{q'}{2}|\theta|^2\mathcal{K} + \frac{q'}{2}\theta'\widetilde{\sigma}\left(\widetilde{\sigma}\widetilde{\sigma}'\right)^{-1}\widetilde{\sigma}\theta\mathcal{K} + q'\theta'\mathcal{L} - q'\theta'\widetilde{\sigma}'\left(\widetilde{\sigma}\widetilde{\sigma}'\right)^{-1}\widetilde{\sigma}\mathcal{L}\right. \\ &\quad \left. +\frac{q'}{2}\frac{1}{\mathcal{K}}\mathcal{L}\widetilde{\sigma}'\left(\widetilde{\sigma}\widetilde{\sigma}'\right)^{-1}\widetilde{\sigma}\mathcal{L} + \frac{2-q'}{2}\frac{1}{\mathcal{K}}|\mathcal{L}|^2\right\} ds + \mathcal{L}'dw, \\ \mathcal{K}(T) &= \left(q'\tilde{M}\right)^{\frac{1}{q'-1}}. \end{aligned}$$

Note that $\widetilde{\sigma}$ is assumed to be chosen such that $\sigma\widetilde{\sigma}' = 0$. As $\theta = \sigma'\left(\sigma\sigma'\right)^{-1}B'$, this entails that $\widetilde{\sigma}\theta = 0$. Thus, the BSRDE reduces to

$$\begin{aligned} d\mathcal{K} &= \left\{-\frac{q'}{2}|\theta|^2\mathcal{K} + q'\theta'\mathcal{L} + \frac{q'}{2}\frac{1}{\mathcal{K}}\mathcal{L}'\widetilde{\sigma}'\left(\widetilde{\sigma}\widetilde{\sigma}'\right)^{-1}\widetilde{\sigma}\mathcal{L} + \frac{2-q'}{2}\frac{1}{\mathcal{K}}|\mathcal{L}|^2\right\} ds + \mathcal{L}'dw, \\ \mathcal{K}(T) &= \left(q'\tilde{M}\right)^{\frac{1}{q'-1}}. \end{aligned}$$

Since problem (6.52) satisfies Assumption A1 or A3, this BSRDE is solvable. The optimal state $\bar{a}$ and optimal control $\bar{v}$ for problem (6.52) are now as given in Theorem 5.11. The optimal cost is given by $J_*^{(T)}(\bar{v}) := \|\bar{a}(T)\|_{\mathcal{Z}_T^*}^{q'} = \frac{1}{q'}\mathcal{K}^{q'-1}(0)$. As $\overline{\alpha}_1$ is a martingale, we have

$$\overline{\alpha}_1(0) = E[\overline{\alpha}_1(T)] = \frac{-\text{sign}(x_0)}{\|\bar{a}(T)\|_{\mathcal{Z}_T^*}} E[\bar{a}(T)] = -\text{sign}(x_0)(q')^{\frac{1}{q'}}\mathcal{K}^{-\frac{1}{q}}(0).$$

This means that the optimal value on the right hand side of (6.51) is given by $|x_0|(q')^{\frac{1}{q'}}\mathcal{K}^{-\frac{1}{q}}(0)$. The optimal value for the left hand side of (6.51) can also be expressed in terms of BSRDE. If (K, L) is a solution of (6.48), (6.50), then the optimal value on the left hand side is given by $(\frac{1}{q}K^{q-1}(0)|x_0|^q)^{\frac{1}{q}}$. The equality (6.51) hence gives that

$$K(0) = q' q^{\frac{q'}{q}} \mathcal{K}^{-\frac{q'}{q}}(0) = q'\left(\frac{1}{q}\mathcal{K}\right)^{-\frac{q'}{q}}(0).$$

6.3.3 Running and terminal cost, complete market

In this section we consider problem (6.15) in a complete market setting. The direct solution via BSRDE is easily stated. The BSRDE for this problem is just that for problem (6.15) in a complete market, with an additional term $-\frac{1}{q-1}QK^{2-q}$, i.e.

$$\begin{aligned} dK &= \left\{\frac{q}{2(q-1)^2}|\theta|^2 K + q'\theta' L - \frac{1}{q-1}QK^{2-q} + \frac{1}{K}|L|^2\right\} ds + L' dw, && (6.54)\\ K(T) &= f(M). && (6.55) \end{aligned}$$

As the underlying problem satisfies Assumption A1 or A3, this BSRDE is solvable, and the optimal state and the optimal control can be constructed like in Theorem 5.11.

The dual problem, more precisely the stochastic control problem resulting from the dual problem, exhibits an interesting property: whereas in the primal problem the control enters the time integral as well as the stochastic integral, in the dual problem, the control variable is only found in the ds-integral.

Let us have a closer look at the dual problem. Recall the representation of $\mathcal{M}_{RT}^{\perp}$ from Lemma 6.10. As in (6.46) it is easily seen that for all $(z^*, \alpha_2(T)) \in \mathcal{M}_{RT}^{\perp}$ we have

$$< (-x_0, -x_0), (z^*, \alpha_2(T)) >_{\mathcal{Z}_{RT}} = -x_0\alpha_2(0).$$

Thus, the dual problem to (6.15) reads as

$$\min_{(z,Z)\in\mathcal{M}_{RT}} \|(-x_0 - z, -x_0 - Z)\|_{\mathcal{Z}_{RT}} = \max_{\substack{(z^*,\alpha_2(T))\in\mathcal{M}_{RT}^{\perp} \\ \|(z^*,\alpha_2(T))\|_{\mathcal{Z}_{RT}^*}\le 1}} -x_0\alpha_2(0). \qquad (6.56)$$

As in Subsection 6.3.2 it is also easily seen that the optimal pair $(\bar{z}^*, \overline{\alpha}_2(T))$ for the right hand side of (6.56) satisfies $\|(\bar{z}^*, \overline{\alpha}_2(T))\|_{\mathcal{Z}_{RT}^*} = 1$. In order to solve the right hand side of

(6.56) consider the stochastic control problem

$$J_*^{(RT)}(\zeta) := \|(\zeta, a(T))\|_{\mathcal{Z}_{RT}^*}^{q'} = \frac{1}{q'}E[\int_0^T \left(q'\tilde{Q}\right)|\zeta|^{q'}ds + \left(q'\tilde{M}\right)|a(T)|^{q'}] = \min_{\zeta \in L_{\mathcal{F}}^{q'}(0,T;\mathbb{R})}!, \tag{6.57}$$

where a is the solution of

$$da = -\zeta ds - \theta' a dw, \quad a(0) = 1. \tag{6.58}$$

Since $\tilde{Q}$ is uniformly positive, problem (6.57) satisfies Assumption A4 and is hence solvable via the BSRDE theory that we developed if $q' > 2$. With the same arguments following (6.53) we can see, that the optimal vector $(\bar{z}^*, \overline{\alpha}_2(Z))$ for the right hand side of the duality relation (6.56) is given by $(\bar{z}^*, \overline{\alpha}_2(Z)) = \frac{1}{\|(\bar{\zeta}, \bar{a}(T))\|_{\mathcal{Z}_{RT}^*}}(\bar{\zeta}, \bar{a}(T))$, where $(\bar{\zeta}, \bar{a})$ is the optimal control and the optimal state for problem (6.57). Note that in this dual problem there is only a one-dimensional control variable and that this variable does not enter the dw-part of the state equation. Looking at the adjoint operators in the complete market case, this is not surprising, though, very convenient.

We want to write down the BSRDE for problem (6.57). Here, the function G reads as

$$G(\mathcal{K}, \mathcal{L}) = \mathcal{K}\left(q'\tilde{Q}\right)^{\frac{1}{q'-1}}.$$

So, the BSRDE for problem (6.57) is

$$\begin{aligned} d\mathcal{K} &= \left\{-\frac{q'}{2}|\theta|^2\mathcal{K} + \left(q'\tilde{Q}\right)^{\frac{1}{q'-1}}\mathcal{K}^2 + q'\theta'\mathcal{L} + \frac{2-q'}{2}\frac{|\mathcal{L}|^2}{\mathcal{K}}\right\}ds + \mathcal{L}'dw, & (6.59)\\ \mathcal{K}(T) &= \left(q'\tilde{M}\right)^{\frac{1}{q'-1}}. & (6.60)\end{aligned}$$

So far, we can guarantee the solvability of this BSRDE only for $q' > 2$, due to the fact that we have no a-priori estimates for $\mathcal{L}$ as, for example, stated in Theorem 4.16. To be precise, if $q' < 2$ a solution is here meant to be a pair of adapted processes $(\mathcal{K}, \mathcal{L})$ such that $P - a.s.\ \mathcal{K}(t) > 0$ for all $t \in [0, T]$, $\mathcal{K}$ is essentially bounded, $\int_0^T |\mathcal{L}|^2 ds < \infty\ P - a.s.$ and $(\mathcal{K}, \mathcal{L})$ satisfies the BSRDE (6.59), (6.60).

By considering the differential of $\mathcal{K}^2$ and following the line of the proof of e.g. Theorem 4.16 it would be quite easy to see, that for a solution $(\mathcal{K}, \mathcal{L})$ of (6.59), (6.60) we have the a-priori estimate, i.e. $\mathcal{L} \in H_p(0, T; \mathbb{R}^d)$ for all $p > 1$.

Yet, in the very special situation of BSRDE (6.59), (6.60) with $q' < 2$ one may get along to proof solvability without the a priori estimate. Let us assume that all the results of Chapter 3 that concern problems that satisfy Assumption A4 also hold for problem (6.57). Let $\bar{a}, \bar{\zeta}$ be the optimal state and the oqtimal control for this problem. Then, all the calculations of Chapter 4 go through. Define now $(\mathcal{K}, \mathcal{L})$ according to (5.1) (with x, u replaced by $\bar{a}, \bar{\zeta}$ etc.). $\mathcal{K}$ is then essentially bounded and we would have $\bar{\zeta} = G(\mathcal{K}, \mathcal{L})\bar{a} = \left(q'\tilde{Q}\right)^{\frac{1}{q'-1}}\mathcal{K}\bar{a}$ on

$[0, \tau_0)$ (τ_0 being the first time that $\bar{a}$ attains 0). By Lemma 3.8, the optimal state $\bar{a}$ satisfies the SDE

$$d\bar{a} = -\mathbf{1}_{[0,\tau_0)} G(\mathcal{K}, \mathcal{L}) \bar{a} ds - \mathbf{1}_{[0,\tau_0)} \theta \bar{a} dw, \; \bar{a}(0) = 1,$$

i.e.

$$d\bar{a} = -\mathbf{1}_{[0,\tau_0)} \left(q'\tilde{Q}\right)^{\frac{1}{q'-1}} \mathcal{K} \bar{a} ds - \mathbf{1}_{[0,\tau_0)} \theta \bar{a} dw, \; \bar{a}(0) = 1.$$

As $\left(q'\tilde{Q}\right)^{\frac{1}{q'-1}} \mathcal{K}$ is essentially bounded, this shows that $\bar{a}$ does not vanish, i.e. $\tau_0 = T$, since we can apply Remark 5.4. In the general situation, at this point a a-priori estimate is needed. Further, we have $\bar{a} \in L^p_{\mathcal{F}}(\Omega, C([0,T]; \mathbb{R}))$ for every $p > 1$. Consequently, the BSRDE (6.59), (6.60) is solvable and $\bar{\zeta} = G(\mathcal{K}, \mathcal{L})\bar{a}$ satisfies the hypothesis of Lemma 5.6. One can now proceed as in Corollary 5.12 to show uniqueness of a solution of (6.59), (6.60).

In this shortly sketched approach to the solvability of (6.59), (6.60) the main point is that one need not be much concerned about $\mathcal{L}$. For example, one need not carry out the localizations of Chapter 5. This "lesser importance" of $\mathcal{L}$ technically reflects the fact that in the state equation for problem (6.57) the control variable does not enter the diffusion term.

6.3.4 Running and terminal cost, incomplete market

In this subsection we consider the problem (6.15) when the market is not supposed to be complete, i.e. when $m < d$. This problem satisfies Assumption A1 or A3, depending on q. Again, we can immediadetly apply our results to this situation. The BSRDE for the control problem is given by

$$\begin{aligned} dK &= \left\{ \frac{q}{2(q-1)^2} |\theta|^2 K - \frac{1}{q-1} Q K^{q-2} + q'\theta' L + \frac{q}{2} \frac{1}{K} L'\sigma' \left(\sigma\sigma'\right)^{-1} \sigma L \right. \\ &\quad \left. + \frac{2-q}{2} \frac{1}{K} |L|^2 \right\} ds + L' dw, & (6.61) \\ K(T) &= f(M). & (6.62) \end{aligned}$$

This BSRDE is solvable by Theorem 5.5. The optimal control, the optimal state and the optimal cost are given as in Theorem 5.11 and Corollary 5.7.

In the present situation, we cannot claim that there may be any advantage in considering the dual problem. But it will turn out that the dual problem is of a less conventional type. Our knowledge about the primal problem will entitle us to tackle the dual problem. Let us have a look at this dual problem.

As in the case of running and terminal cost in a complete market we have

$$\min_{(z,Z) \in \mathcal{M}_{RT}} \|(-x_0 - z, -x_0 - Z)\|_{\mathcal{Z}_{RT}} = \max_{\substack{(z^*, \alpha_2(T)) \in \mathcal{M}^{\perp}_{RT} \\ \|(z^*, \alpha_2(T))\|_{\mathcal{Z}^*_{RT}} \leq 1}} -x_0 \alpha_2(0), \qquad (6.63)$$

but here the parametrization of $\mathcal{M}_{RT}^{\perp}$ involves an additional control variable, see Lemma 6.10. In order to maximize the right hand side of (6.63) consider the stochastic control problem

$$\|(z^*, \alpha_2(T)\|_{\mathcal{Z}_{RT}^*} = \min_{\substack{(z^*,\alpha_2(T))\in\mathcal{M}_{RT}^{\perp} \\ \alpha_2(0)=1}} !,$$

i.e. the control problem

$$\begin{aligned} J_*^{(RT)}((\zeta, v)) &= \|(\zeta, a(T))\|_{\mathcal{Z}_{RT}^*}^{q'} = \frac{1}{q'} E[\int_0^T \left(q'\tilde{Q}\right) |\zeta|^{q'} ds + \left(q'\tilde{M}\right) |a(T)|^{q'}] \\ &= \min_{\substack{\zeta\in L_{\mathcal{F}}^{q'}(0,T;\mathbb{R}) \\ v\in H_{q'}(0,T;\mathbb{R}^{d-m})}} , \end{aligned} \tag{6.64}$$

with state equation

$$da = -\zeta ds + \{-\theta' a + v'\widetilde{\sigma}\} dw, \quad a(0) = 1. \tag{6.65}$$

Here, $\widetilde{\sigma}$ is supposed to satisfy the properties stated in Lemma 6.10-3.
Problem (6.64) in an essential way does not fit into the types of control problems we considered so far: the minimization if performed over a product space, and the components of the compound control variable (ζ, v) are treated in a very different manner. The part ζ of the control variable does not enter the dw-integral. Actually, it couldn't for $q' < 2$, since otherwise $a(T)$ need not be q'-integrable. Concerning the integrability of a, it would not matter if the control variable part v did enter the time integral of the state equation. Yet, v could not appear in the immediate running control cost for $q' > 2$, as it it is not sufficiently integrable. If considered isolated, the control variable ζ would satisfy an assumption like A4 and the control variable v would satisfy A1 or A3.
It is clear that $J_*^{(RT)}((\zeta, v)) \to \infty$ if $\|\zeta\|_{L_{\mathcal{F}}^{q'}} \to \infty$ or $\|v\|_{H_{q'}} \to \infty$, see the proof of Lemma 3.2. From Proposition 3.1 it then follows, that (6.64) is uniquely solvable. Let $(\bar{\zeta}, \bar{v})$ be the optimal control and $\bar{a}$ be the optimal state. As in the previous sections, one can see that the maximizing element $(\bar{z}^*, \overline{\alpha}_2(T))$ for the right hand side of (6.63) is $\frac{-\text{sign}(x_0)}{\|(\zeta,\bar{a}(T))\|_{\mathcal{Z}_{RT}^*}}(\bar{\zeta}, \bar{a}(T))$.
Since the primal problem has an optimal state $\bar{x}$ with $P-a.s.$ $\bar{x}(t) \neq 0$ for all t, Lemma 6.11 shows, that $P-a.s.$ we have

$$\bar{\zeta}(t) \neq 0 \text{ for all } t \in [0, T] \text{ and } \bar{a}(T) \neq 0. \tag{6.66}$$

We want to sketch how to establish a BSRDE for problem (6.64). A detailed representation of this procedure would essentially repeat the work done in Chapters 3 and 4. We will not go through the respective proofs and confine ourselves to a statement of the main points. With the same ideas leading to Lemma 3.5 it can be seen that the optimal state $\bar{a}$ and the optimal control $(\bar{\zeta}, \bar{v})$ for problem (6.64) can be characterized as part of the unique solution $(a, (\zeta, v), y, z)$ of the FBSDE

$$da = -\zeta ds + \{-\theta a + v'\widetilde{\sigma}\} dw, \tag{6.67}$$

$$dy = \theta' z ds + z' dw, \tag{6.68}$$

$$a(0) = 1, \quad y(T) = \left(q'\tilde{M}\right)\varphi_{q'}(x(T)) \tag{6.69}$$

$$-y + \left(q'\tilde{Q}\right)\varphi_{q'}(\zeta) = 0, \quad \widetilde{\sigma} z = 0. \tag{6.70}$$

To avoid confusion about the exponents in φ and f we write $\varphi_{q'}(r) = r|r|^{q'-2}$ for $r \in \mathbb{R}^m, r \neq 0$. $f_{q'}$ then denotes the inverse of $\varphi_{q'}$. Note that the auxiliary condition (6.70) decouples in two separate conditions. Besides, $\widetilde{\sigma} z = 0$ implies that there is a unique $u \in H_{q'}(0,T;\mathbb{R}^m)$ such that $\sigma' u = z$. Note the formal coincidence with the state equation of the primal problem if one plugs this relation in (6.68).
In the forthcoming we will use the notation $(a,(\zeta,v),y,z)$ for the solution of this FBSDE. In particular we will omit the bar on the optimal state and control.
Like in Lemma 3.8, one can show that $(a,(\zeta,v),y,z)$ is identical zero after the first time that a attains zero. But from (6.66) we know that $P-a.s.\ a(T) \neq 0$. This entails that

$$P-a.s. \;\; a(t) \neq 0 \text{ for all } t \in [0,T]. \tag{6.71}$$

Along the lines of Propositions 3.11, 4.1, Lemma 4.2 and Remark 4.3 it can be seen that there is a adapted, essentially bounded family of r.v. $\mathcal{K}(t)$, $t \in [0,T]$, such that $f_{q'}(y(t)) = \mathcal{K}(t)a(t)$. We have, for $t \in [0,T]$, that $\mathcal{K}(t) > 0\ P-a.s.$. Since a does not vanish, we get the representation

$$\mathcal{K}(t) = \frac{f_{q'}(y(t))}{a(t)} \text{ on } [0,T]. \tag{6.72}$$

The availability of this relation on all of $[0,T]$ is the crucial difference to the situation encountered in Chapter 4!
Set

$$\mathcal{L} := \frac{f'_{q'}(y)}{a} z + \theta\mathcal{K} - \mathcal{K}\widetilde{\sigma}\frac{1}{a}v. \tag{6.73}$$

Differentiating $\mathcal{K}(t) = \frac{f_{q'}(y(t))}{a(t)}$ yields (on [0,T])

$$\begin{aligned} & d\mathcal{K}(t) \\ = & \left\{ -|\theta|^2\mathcal{K} + 2\theta'\mathcal{L} + \mathcal{K}\frac{\zeta}{a} - \left(\widetilde{\sigma}\theta\mathcal{K} + \widetilde{\sigma}\mathcal{L}\right)'\frac{v}{a} + \frac{2-q'}{2}\frac{1}{\mathcal{K}}\left|\mathcal{L} - \theta\mathcal{K} + \mathcal{K}\widetilde{\sigma}'\frac{v}{a}\right|^2 \right\} ds \\ & +\mathcal{L}' dw. \end{aligned} \tag{6.74}$$

Let us replace $\frac{\zeta}{a}$ and $\frac{1}{a}v$. We use the auxiliary conditions (6.70) and the definition of $\mathcal{K}$ and $\mathcal{L}$ to get

$$\frac{\zeta}{a} = \frac{1}{f_{q'}\left(q'\tilde{Q}\right)}\mathcal{K}, \quad \frac{1}{a}v = -\frac{1}{\mathcal{K}}\left(\widetilde{\sigma}\widetilde{\sigma}'\right)^{-1}\widetilde{\sigma}\mathcal{L}. \tag{6.75}$$

When we plug this into (6.74) we arrive at the Riccati equation for problem (6.64) (for a slight simplification we observe that $\widetilde{\sigma}\theta = 0$):

$$\begin{aligned} d\mathcal{K}(t) &= \Big\{ -|\theta|^2\mathcal{K} + 2\theta'\mathcal{L} + \left(f_{q'}\left(q'\tilde{Q}\right)\right)^{-1}\mathcal{K}^2 + \frac{1}{\mathcal{K}}\mathcal{L}'\widetilde{\sigma}'\left(\widetilde{\sigma}\widetilde{\sigma}'\right)^{-1}\widetilde{\sigma}\mathcal{L} \\ &\quad + \frac{2-q'}{2}\frac{1}{\mathcal{K}}\left|\mathcal{L} - \theta\mathcal{K} - \widetilde{\sigma}'\left(\widetilde{\sigma}\widetilde{\sigma}'\right)^{-1}\widetilde{\sigma}\mathcal{L}\right|^2 \Big\} ds + \mathcal{L}'dw, \quad (6.76) \\ \mathcal{K}(T) &= f_{q'}(q'\tilde{M}). \quad (6.77) \end{aligned}$$

Obviously, $(\mathcal{K}, \mathcal{L})$ is a solution for this BSRDE, if we impose regularity requirements for a solution as in the case of Assumption A4.

Remarks on Chapter 6 The financial problems investigated in this chapter were our main reason to address the linear isoelastic control problems. The introduction of the dual problems can considerably simplify the problems, especially in the case of complete markets. Considering the case of terminal cost, our duality approach is of course the same as the one used for example in [GLP:MVH] in the case of mean-variance hedging. Our setting allows us to determine precisely the orthogonal complement of the attainable terminal values. This enabled us to treat the problems (6.15), (6.19) completely as minimum norm problems according to Theorem 6.7. Note that in general semimartingale settings duality approaches are widely used. But it is not always clear that the optimization of the dual problem is actually performed over a set that is precisely the dual space or a subset thereof.

The treatment of BSRDEs for a problem of the type of (6.64) seems also in the quadratic case to be new. Though we have stated nothing whether one can construct a solution of the control problem out of the solution of the BSRDE also in this case, this example illustrates that the method developed independently in [T:GLQO] and by the author (i.e. to show that the optimal state never attains zero) is quite flexible in its application.

Summary

Chapter 1: Linear isoelastic stochastic control problem $\mathcal{P}(\tau, h)$

$$J(u) = \frac{1}{q} E[\int_{\tau}^{T} Q(t)|x(t)|^q + N(t)|u(t)|^q dt + M|x(T)|^q] = \min_{u \in \mathcal{U}}!$$

where

$$\begin{aligned} dx(t) &= \{A(s)x(s) + B(s)u(s)\} \, ds + \sum_{i=1}^{d} \left\{ C^i(s)x(s) + D^i(s)u(s) \right\} dw^i(s), \\ x(\tau) &= h. \end{aligned}$$

$\mathcal{U} = H_q(\tau, T; \mathbb{R}^m)$ in case of Assumptions A1 and A3, $\mathcal{U} = L^q_{\mathcal{F}}(\tau, T; \mathbb{R}^m)$ in case of Assumptions A2 and A4.
$A, B, (C^i)_{1 \le i \le d}, (D^i)_{1 \le i \le d}, Q, N$ and M adapted/measurable and essentially bounded;

Assumption A1: $q \le 2$, $\sum_{i=1}^{d} (D^i)' D^i \gg 0$, $M \gg 0$, $Q \ge 0$, $N \ge 0$;

Assumption A2: $q \ge 2$, $N \gg 0$, $Q \ge 0$, $M \ge 0$;

Assumption A3: $q \ge 2$, $\sum_{i=1}^{d} (D^i)' D^i \gg 0$, $M \gg 0$, $N = 0$, $Q \ge 0$;

Assumption A4: $q \ge 2$, $N \gg 0$, $M > 0$, $Q \ge 0$;

Chapter 2 : $H_q(\tau, T; \mathbb{R}^m)$ is reflexive, $H_{q'}(\tau, T; \mathbb{R}^m)$ is isomorphic to $H^*_q(\tau, T; \mathbb{R}^m)$.

$$\begin{aligned} dx(t) &= \{a(s)x(s) + \alpha(s)\} \, ds + \{x(s)c(s) + \beta(s)\} \, dw(s) \\ x(\tau) &= h, \end{aligned}$$

has a unique solution $x \in L^q_{\mathcal{F}}(\Omega, C([\tau, T]; \mathbb{R}))$ for $\alpha \in H_q(0, T; \mathbb{R})$ or $\alpha \in L^q_{\mathcal{F}}(0, T; \mathbb{R})$, $\beta \in H_q(0, T; \mathbb{R}^d)$ and $h \in L^q_{\mathcal{F}_\tau}(\mathbb{R})$; x depends continuously on (h, α, β) with a boundedness constant independent of τ;

Chapter 3: Under Assumptions A1-A4 the linear isoelastic control problem is solvable; it can remain solvable if N becomes not too much negative.
(x,u) is the optimal state and the optimal control for problem $\mathcal{P}(\tau,h)$ if and only if it is part of the unique solution $(\bar{x},\bar{u},\bar{y},\bar{z}) \in L^q_{\mathcal{F}}(\Omega, C([\tau,T];\mathbb{R})) \times \mathcal{U} \times L^{q'}_{\mathcal{F}}(\Omega, C([\tau,T];\mathbb{R})) \times H_{q'}(\tau,T;\mathbb{R}^d)$ of the FBSDE with auxiliary condition

$$\begin{aligned}
dx(t) &= \{A(s)x(s)+B(s)u(s)\}\,ds + \sum_{i=1}^d \left\{C^i(s)x(s)+D^i(s)u(s)\right\} dw^i(s),\\
dy(t) &= \left\{-A(s)y(s) - \sum_{i=1}^d C^i(s)z^i(s) - Q(s)\varphi\left(x(s)\right)\right\} ds + \sum_{i=1}^d z^i(s)dw^i(s),\\
& \quad x(\tau)=h, \qquad y(T)=M\varphi(x(T))\\
& \quad B'y + \sum_{i=1}^d (D^i)'z^i + N\varphi(u) = 0, \quad Leb\otimes P - a.s..
\end{aligned}$$

$(\bar{x},\bar{u},\bar{y},\bar{z}) = \mathbf{1}_{[\tau,\tau_0]}(\bar{x},\bar{u},\bar{y},\bar{z})$, where τ_0 is the first time that $\bar{x}$ attains zero;

$h \mapsto (\bar{x}^{\tau,h},\bar{u}^{\tau,h},f(\bar{y}^{\tau,h}))$ is linear; the optimal cost is given by $\frac{1}{q}E[y(\tau)h]$;

$h \mapsto (\bar{x}^{\tau,h},\bar{u}^{\tau,h},\bar{y}^{\tau,h},\bar{z}^{\tau,h})$ is continuous; $(\bar{x}^{\tau,h},\bar{u}^{\tau,h}) = h(\bar{x}^{\tau,1},\bar{u}^{\tau,1})$;

Chapter 4: There is a uniformly bounded, adapted family of r.v. $K(t\vee\tau)$, $t\in[0,T]$, such that $f\left(y(t\vee\tau)\right) = K(t\vee\tau)x(t\vee\tau)$ for $t\in[0,T]$; $(K(t\vee\tau))_t$ is strictly positive respectively uniformly positive if Assumption A1 or A3 respectively A4 holds; K satisfies the BSRDE (4.17) on $[\tau,\tau_0)$;

$|G(B,(C^i)_i,(D^i)_i,N,K,L)| \le a + b|L|$ in case of Assumptions A1 and A3;

$|G(B,(C^i)_i,(D^i)_i,N,K,L)| \le a + b|L|^{\frac{1}{q-1}}$ in case of Assumption A4;

If (K,L) is a solution of the BSRDE, then $L\in H_p(\tau,T;\mathbb{R}^d)$ for all $p>1$;

Chapter 5: $\tau_0 = T$, $P-a.s.$ and

$$K := \frac{f(y)}{x}, \quad L^i := \frac{f'(y)z^i}{x} - C^i\frac{f(y)}{x} - D^i\frac{f(y)}{x^2}u, \quad i=1,\dots,d,$$

is the unique solution of the BSRDE;

The optimal state x and the optimal control u for problem $\mathcal{P}(\tau,h)$ are related by $x =$

$G(K,L)u$; the optimal cost is given by $\frac{1}{q}E[K(\tau)^{q-1}|h|^q]$;

Chapter 6: Introduction of a financial market model and two financial market hedging problems

$$J^{(T)}(u) := \frac{1}{q}E[M\,|x(T)|^q] = \min_{u\in\mathcal{V}}!,$$

$$J^{(RT)}(u) := \frac{1}{q}E[\int_0^T Q(s)|x(s)|^q ds \;+\; M\,|x(T)|^q] = \min_{u\in\mathcal{V}}!,$$

where

$$\begin{aligned} dx(t) &= B(s)u(s)ds + \sum_{i=1}^{d} D^i(s)u(s)dw^i(s), \\ x(0) &= x_0, \end{aligned}$$

$u \in \mathcal{V} := H_q(0,T;\mathbb{R}^m)$, $\sum_{i=1}^d (D^i)'\,D^i \gg 0$, $Q \gg 0$, $M \gg 0$; these problems can be solved with the previous results on BSRDEs;
Introduction of duality approach via a formulation as minimum-norm problem; alternative proofs of the solvability of some special BSRDEs by using the duality approach; consideration of the control problem

$$\begin{aligned} J_*^{(RT)}((\zeta,v)) &= \frac{1}{q'}E[\int_0^T \left(q'\tilde{Q}\right)|\zeta|^{q'}ds + \left(q'\tilde{M}\right)|a(T)|^{q'}] \\ &= \min_{\substack{\zeta\in L^{q'}_{\mathcal{F}}(0,T;\mathbb{R}) \\ v\in H_{q'}(0,T;\mathbb{R}^{d-m})}} !, \end{aligned}$$

where

$$da = -\zeta ds + \{-\theta' a + v'\widetilde{\sigma}\}\,dw, \;\; a(0) = 1,$$

statement of a solvable BSRDE for this problem.

Appendix A

Some measurable selection arguments

In the proofs of Lemma 3.4 and 6.10-3 we made use of some processes (m_λ and $\widetilde{\sigma}$) from which we claimed that they can be chosen measurable or adapted. For the reader's convenience we give the proofs of these statements in this Appendix. We will rely on measurable selection arguments which we take from [AF:SVA] and [CV:CAMM]. The next section sums up the results that we use.
Before this, let us introduce some notation and a standing assumption.
For a set X we denote by $\mathrm{Pow}(X)$ the power class of X. Let Y be another set. The notation $H : Y \rightsquigarrow X$ means a set-valued map $H : Y \longrightarrow \mathrm{Pow}(X)$.
In the forthcoming we will always assume that Y is a measurable space, equipped with the Sigma-algebra $\mathcal{A}$. $X = (X, d)$ is a complete, separable metric space, equipped with its Borel-Sigma-algebra.

A.1 Measurable Selections

Let $H : Y \rightsquigarrow X$ be a set-valued map. We call $h : Y \longrightarrow X$ a *measurable selection of* H if h is measurable (i.e. $\mathcal{A}$ -Borel-measurable) and if $h(y) \in H(y)$ for all $y \in Y$ (see [AF:SVA], Definition 8.1.2).
We call a set-valued map $H : Y \rightsquigarrow X$ with closed images *measurable*, if for all open sets $O \subset X$ we have

$$H^{-1}(O) := \{y \in Y : H(y) \cap O \neq \emptyset\} \in \mathcal{A},$$

see [AF:SVA], Definition 8.1.1.
If X is compact, the measurability of a set valued map $H : Y \rightsquigarrow X$ with closed images is equivalent to the following: for all closed sets $C \subset X$ we have

$$H^{-1}(C) \in \mathcal{A}, \tag{A.1}$$

see [CV:CAMM], Theorem III.2.
Measurable set valued maps H with closed images allow a measurable selection. Of course, we must exclude the case that the empty set occurs in the range of H. We cite Theorem 8.1.3 of [AF:SVA].

Theorem A.1 *Let $H : Y \rightsquigarrow X$ be a set valued map with closed, nonempty images. Assume that H is measurable. Then there exists a measurable selection h for H.*

We will apply this Theorem.

A.2 To the proof of Lemma 3.4.

Let u, v be in $H_q(\tau, T; \mathbb{R}^m)$. In the proof of Lemma 3.4 we claimed, that for every $\lambda > 0$ there is a measurable process m_λ such that $0 \leq m_\lambda \leq 1$ and

$$\frac{1}{q}\left(|u(s) + \lambda v(s)|^q - |u(s)|^q\right) = \lambda v'(s)\varphi(u(s) + m_\lambda \lambda v(s)), \tag{A.2}$$

$Leb \otimes P - a.s.$.
To proof this, we want to find a measurable function $h : \mathbb{R}^m \times \mathbb{R}^m \longrightarrow [0,1]$ such that $m_\lambda := h(u,v)$ will do. To this end, fix some $\lambda > 0$ and consider the set valued function

$$H : \mathbb{R}^m \times \mathbb{R}^m \rightsquigarrow [0,1],$$

with $H(u,v) := \left\{m \in [0,1] : \frac{1}{q}\left(|u + \lambda v|^q - |u|^q\right) = \lambda v' \varphi(u + m\lambda v)\right\}$. To apply Theorem A.1, we first have to check that $H(u,v) \neq \emptyset$ for all u, v. This is clear by the Mean Value Theorem. Next, we must convince ourselves that $H(u,v)$ is closed for all u, v. If m_n is a sequence in $H(u,v)$ with $m_n \to m$, $n \to \infty$, then continuity yields

$$\lambda v' \varphi(u + m\lambda v) = \lim_{n \to \infty} \lambda v' \varphi(u + m_n \lambda v) = \frac{1}{q}\left(|u(s) + \lambda v(s)|^q - |u(s)|^q\right),$$

hence $m \in H(u,v)$.
Finally, let us show that H is measurable. We apply criterion (A.1). Let C be a closed subset of $[0,1]$. We claim that $H^{-1}(C)$ is closed, hence a Borel-set in $\mathbb{R}^m \times \mathbb{R}^m$. So, let (u_n, v_n) be a sequence in $H^{-1}(C)$ with $(u_n, v_n) \to (u,v)$, $n \to \infty$. We have to show that $(u,v) \in H^{-1}(C)$. Let (m_n) be a sequence with $m_n \in C \cap H(u_n, v_n)$. Since C is compact, there is a subsequence, for simplicity still denoted by (m_n), that convergers to some $m \in C$. We have for all n

$$\begin{array}{ccc} \frac{1}{q}\left(|u_n + \lambda v_n|^q - |u_n|^q\right) & = & \lambda v_n' \varphi(u_n + m_n \lambda v_n) \\ \downarrow n \to \infty & & \downarrow n \to \infty \\ \frac{1}{q}\left(|u + \lambda v|^q - |u|^q\right) & = & \lambda v' \varphi(u + m\lambda v), \end{array}$$

i.e. $m \in H(u,v)$, hence $(u,v) \in H^{-1}(C)$.
By Theorem A.1, there is now a measurable selection h of H. For processes $u, v \in H_q(\tau, T; \mathbb{R}^m)$, $m_\lambda := h(u,v)$ now yields (A.2).

A.3 To the proof of Lemma 6.10-3.

Let $\sigma \in L^\infty_{\mathcal{F}}(0,T;\mathbb{R}^{m\times d})$ satisfy $\sigma\sigma' \gg 0$. This in particular implies, that σ has $Leb \otimes P - a.s.$ maximal rank. In the proof of Lemma 6.10-3 we claimed, that there is a $\widetilde{\sigma} \in L^\infty_{\mathcal{F}}(0,T;\mathbb{R}^{(d-m)\times d})$ such that $Leb \otimes P - a.s.$

$$\widetilde{\sigma}\widetilde{\sigma}' \gg 0,\ \begin{bmatrix}\sigma \\ \widetilde{\sigma}\end{bmatrix} \text{ is regular },\ \sigma\widetilde{\sigma}' = 0. \tag{A.3}$$

As in the previous Section, we want to show the existence of $\widetilde{\sigma}$ by determining a measurable $h : \mathbb{R}^{m\times d} \longrightarrow \mathbb{R}^{(d-m)\times d}$ such that one can choose $\widetilde{\sigma} = h(\sigma)$.
Consider the set valued map $H : \mathbb{R}^{m\times d} \rightsquigarrow \mathbb{R}^{(d-m)\times d}$,

$$\sigma \mapsto \left\{\widetilde{\sigma} \in \mathbb{R}^{(d-m)\times d} : \widetilde{\sigma}\widetilde{\sigma}' = \mathrm{id}_{d-m},\ \sigma\widetilde{\sigma} = 0\right\}.$$

For a given σ, $H(\sigma)$ contains all $(d-m)$-tuples that are mutually orthogonal, have norm 1, and are in the orthogonal complement of of the linear space generated by the rows of σ. If σ happens to have maximal rank this implies that the rows of σ and $\widetilde{\sigma}$ form a basis of $\mathbb{R}^d$. In particular, $H(\sigma)$ is not empty for any σ, and it's also easily seen to be closed. As $\sigma\sigma' = \mathrm{id}_{d-m}$, we may also consider H as a set valued map $H : \mathbb{R}^{m\times d} \rightsquigarrow X$ for some compact set $X \subset \mathbb{R}^{(d-m)\times d}$. Thinking of the application that we have in mind we can also restrict the domain of H to some sufficiently large, compact set $Y \subset \mathbb{R}^{m\times d}$. With arguments quite similar to those it is easy to check that $H^{-1}(C)$ is closed if $C \subset \mathcal{K}$ is closed. Hence, we can apply Theorem A.1, which yields a measurable selection h for H. If we have chosen Y big enough such that for a given process $\sigma \in L^\infty_{\mathcal{F}}(0,T;\mathbb{R}^{m\times d})$ we have $\sigma \in Y$ $Leb \otimes P - a.s.$, then we may set $\widetilde{\sigma} := h(\sigma)$ and get the desired measurability of $\widetilde{\sigma}$.

Bibliography

[A:LF] H. W. Alt (1992): Lineare Funktionalanalysis, Springer.

[AF:SVA] J.-P. Aubin, H. Frankowska (1990): Set-Valued Analysis, Birkhäuser

[B:LQOC] J.-M. Bismut (1976): Linear Quadratic Optimal Control with Random Coefficients, SIAM J. Control and Optimization 14(3), pp. 419-444

[B:CDSL] J.-M. Bismut (1978): Contrôle des systemès lineaires quadratiques: application de l'integrale stochastique; *in* : C. Dellacherie, P. A. Meyer, M. Weil (eds.): Séminaire des Probabilités XII, Lecture Notes in Mathematics, pp. 180-264, Springer Verlag,

[BR:ODE] G. Birkhoff, G.-C. Roda (1969): Ordinary Differential Equations, Blaisdell Publishing Company

[BS:PO] F. Black, M. Scholes (1973): The pricing of options and corporate liabilities, J. Polit. Economy, 3, pp. 637 - 659.

[CLZ: SLQ] S. Chen, X. Li, X. Zhou (1998): Stochastic linear regulators with with indefinite control weight costs, SIAM J. Control and Optimization 36(5), pp. 1685 -1702.

[CZ:SLQ2] S. Chen, X. Y. Zhou (2000): Stochastic linear regulators with with indefinite control weight costs. II, SIAM J. Control and Optimization 39(4), pp. 1065-1081

[CK:TSMP] A. Cadenillas, I. Karatzas (1995): The Stochastic Maximum Principle for Linear Convex Optimal Control with Random Coefficients, SIAM J. Control and Optimization 33(2), pp. 590-624.

[CV:CAMM] C. Castaing, M. Valadier (1977): Convex Analysis and Measurable Multifunctions, Springer

[CW:ISI] K. L. Chung, R. J. Williams (1983): Introduction to Stochastic Integration, Birkhäuser.

[D:VM] N. Dinculeanu (1966): Vector Measures, Deutscher Verlag der Wissenschaften.

[DM:PPB] C. Dellacherie, P.-A. Meyer (1982): Probability and Potential B, Theory of Matingales, North-Holland

[DeS:ACPM] F. Delbaen, W. Schachermayer (1996): Attainable Claims with $p'th$ moments, Ann. Inst. Henri Poincaré, 32(6), pp. 743-763.

[DS:LO] N. Dunford, J. T. Schwartz (1957): Linear Operators, Part I: General Theory, Interscience Publishers.

[ET:CA] I. Ekeland, R. Temam (1976): Convex Analysis and Variational Problems, North-Holland.

[EM:BSDE] N. El Karoui, L. Mazliak, eds. (1997): Backward Stochastic Differential Equations, Longman.

[EPQ:BSDE] N. El Karoui, S. Peng, M. C. Quenez (1997): Backward Stochastic Differential Equations in Finance, Mathematical Finance 7, pp. 1-71

[EQ:DPPC] N. El Karoui, M.-C. Quenez (1995): Dynamic programming and pricing of contingent claims in an incomplete market, SIAM J. Control and Optim. 33(1), pp. 29 -66.

[FL:EH] H. Föllmer, P. Leukert (2000): Efficient hedging: cost versus shortfall risk, Finance Stoch. 4(2), pp. 117-146.

[FL:QH] H. Föllmer, P. Leukert (1999): Quantile hedging, Finance Stoch. 3(3), pp. 251-273.

[G:EUS] L.I. Gal' chuk (1978): Existence and uniqueness of a solution fir a stochastic equation with respect to semimartingales, Theory Probab Appl. 23, pp. 751-763.

[GLP:MVH] C. Gourieroux, J. P. Laurent, H. Pham (1998): Mean-Variance Hedging and Numeraire, Mathematical Finance 8(3), pp. 179-200

[H:MT] P. Halmos (1974): Measure Theory, Springer.

[HZ:ISRE] Y. Hu, X. Y. Zhou (2003): Indefinite stochastic Riccati equations, SIAM J. Control Optim. 42(1), pp. 123-127

[K:LMF] I. Karatzas (1997): Lectures on the Mathematics of Finance, CRM Monograph Series, AMS, Rhode Island.

[KS:BM] I. Karatzas, S. E. Shreve (1991): Brownian Motion and Stochastic Calculus, Springer.

[KS:MMF] I. Karatzas, S. E. Shreve (1998): Methods of Mathematical Finance, Springer.

[K:REU] M. Kobylanski (1997): Resultats d'existence et d'unicite pour des equations differentielles stochastiques retrogrades avec des generateurs a croissance quadratique; C. R. Acad. Sci. Paris 324, Serie I, pp. 81-86

[KT:GAS] M. Kohlmann, S. Tang (2002): Global Adapted Solutions of One-Dimenional Backward Stochastic Riccati Equations, with Application to the Mean-Variance Hedging, Stochastic Processes and their Applications 97, pp. 255-288.

[KT:MBSR] M. Kohlmann, S. Tang (2003): Multidimensional Backward Stochastic Riccati Equations and Applications, SIAM J. Control and Optim. 41(6), pp. 1696-1721.

[KT:MLQ] M. Kohlmann, S. Tang (2002): Minimization of Risk, LQ Theory, and Mean Variance Hedging with a Stochastic Income Process, Preprint, University of Konstanz.

[KT:MLQC] M. Kohlmann, S. Tang (2003): Minimization of Risk and Linear Quadratic Optimal Control Theory, Siam J. Control Optim 42(3), pp. 1118-1142

[KZ:RBBS] M. Kohlmann, X. Y. Zhou (2000): Relationship between Backward Stochastic Differential Equations and stochastic controls: a linear quadratic approach, Siam J. Control Optim 38(5), pp. 1392-1407

[KwS:LO] H. Kwakernaak, R. Sivan (1972): Linear Optimal Control Systems, Wiley-Interscience

[LLV:MIT] G.S. Ladde, V. Lakshmikantham, A.S. Vatsala (1985) Monotone Iterative Techniques for Nonlinear Differential Equation, Pitman Advanced Publishing Program

[L:UDP] J. Leitner (2001): Utility Maximation, Duality, Price for Risk, Semimartingale Representations & Continuous Time CAPM, Doctoral Dissertation, Universität Konstanz

[LSM:EBS] J. P. Lepeltier, J. San Martin (1998): Existence for BSDE with superlinear-quadratic coefficient, Stochastics stochastics Rep. 63, pp.227-240

[L:OCSG] J.L. Lions (1971): Optimal Control of Systems Governed by Partial Differential Equations, Springer.

[L:O] D. Luenberger (1968): Optimization by Vector Space Methods, Wiley, New York

[Leu:AMV] P. Leukert (1999): Absicherungsstrategien zur Minimierung des Verlustrisikos, Dissertation, Humboldt-Universität zu Berlin.

[M:IP] P. Malliavin (1982): Integration et Probabilites, Analyse de Fouries et Analyse Spectrale, Masson.

[MY:FBSDE] J. Ma, J. Yong (1999): Forward-Backward stochastic Differential Equations and Their Applications, Springer.

[Ph:DLPH] H. Pham (2000): Dynamic L^p-Hedging in discrete time under cone constraints, SIAM J. Control Optim 38(3), pp. 665-682.

[PP:ASBS] E. Pardoux, S. Peng (1990): Adapted solution of a backward stochastic equation, Systems Control Lett. 14, pp. 55 -61

[P:SIDE] P. Protter (1995): Stochastic Integration and Differential Equations, Springer.

[R:FA] W. Rudin (1973): Functional Analysis, McGraw-Hill.

[S:RDE] T. T. Soong (1973): Random Differential Equations in Science and Engineering, Academic Press

[T:GLQO] S. Tang (2003): General Linear Quadratic Optimal Control Problems with Random Coefficients: Linear Stochastic Hamilton Systems and Backward Stochastic Riccati Equations, SIAM J. Control Optim 42(1), pp 53-75.

[T:MP] J. C. Taylor (1998): An Introduction to Measure and Probability, Springer-Verlag.

[WW:SI] H. v. Weizsäcker, G. Winkler (1990): Stochastic Integrals, Vieweg Verlagsgesellschaft

[W:MRE] W.M. Wonham (1968): On a Matrix Riccati equation of stochastic control, SIAM J. Control and Optimization 6, pp. 681-697.

[YZ:SC] J. Yong, X. Y. Zhou (1999): Stochastic Controls: Hamiltonian Systems and HJB-Equations, Springer-Verlag.

[Z:NFA] E. Zeidler (1984): Nonlinear Functional Analysis and its Applications , Volume III, Springer-Verlag