Convex Duality in Martingale Optimal Transport

Author(s):
Kiiski, Matti

Publication Date:
2019

Permanent Link:
https://doi.org/10.3929/ethz-b-000360489

Rights / License:
In Copyright - Non-Commercial Use Permitted
CONVEX DUALITY IN MARTINGALE OPTIMAL TRANSPORT

A thesis submitted to attain the degree of
DOCTOR OF SCIENCES of ETH ZURICH
(Dr. Sc. ETH Zurich)

presented by
MATTI KIISKI

M.Sc., B.Sc.,
Aalto University

born on 28.4.1989

Citizen of
Finland

accepted on the recommendation of

Prof. Dr. Halil Mete Soner, examiner
Prof. Dr. Ari-Pekka Perkkio, co-examiner
Prof. Dr. Xiaolu Tan, co-examiner

2019
Abstract

This thesis is divided in two parts that share the common theme of studying convex duality in martingale optimal transportation. In the first part (Part II) we study functional analytic compactness of martingale measures on the Skorokhod space in a locally convex dual pairing of the Riesz representation theorem. The purpose of the first part is to establish a technical framework for studying the duality in martingale optimal transportation. The martingale optimal transport duality is the subject of the second part (Part III) of this thesis. We carry out systematic study of the duality using the methods of convex analysis.

The introductory section (Part I) provides a short introduction to the martingale optimal transport problem. The classical Monge-Kantorovich problem serves as a reference point. We discuss the existing results and alternative formulations. In particular, we highlight the role of compactness of the families of transportation plans.

Part II is a self-contained study of compactness of semimartingale measures on the Skorokhod space. More precisely, we characterize and study the properties of the strongest path space topology for which the families of probability measures satisfying Stricker’s uniform tightness criteria are tight, and consequently, define compact sets of probability measures in the pairing of the Riesz representation theorem. In particular, we prove the compactness of martingale transportation plans for this topology. We apply this compactness result to martingale optimal transport duality in Part III.

The content of Part III is a topological and a measurable duality result for martingale optimal transport. The convex dual for the martingale optimal transport problem is a superhedging problem. It is easy to define so-called super-hedging functionals whose bi-conjugate functional coincides with the value function for martingale transport problem. The difficulty is to close the dual gap to this value. We start with a topological duality result in the topological structure given in the previous part. Further, the topological structure, introduced in Part II, is chosen so that the topological duality can be extended to a measurable duality result using the classical results from convex analysis and function lattice theory. Both duality results require an appropriate extension of the set of pathwise hedges beyond the classical pathwise integrals.

In Appendix, we have collected the required preliminary results from ordinal arithmetic, functional analysis, convex analysis, topology and measure theory.
Zusammenfassung


Im Anhang behandeln wir die erforderliche vorläufige Ergebnisse aus der ordinalen Arithmetik, der Funktionsanalyse, der Konvexeanalyse, der Topologie und der Messtheorie und die technische Terminologie.
Acknowledgements

I thank my supervisor Prof. Soner and my co-authors Prof. Cheridito and DJ Prömel without whom this project would have not started, progressed or ended. I am happy that Prof. Perkkiö and Prof. Tan agreed to be co-examiners of this thesis. I am grateful for the financial support from the Swiss National Foundation grant SNF 200020-172815.

Shoutouts to my friends and relatives, especially those who visited me here in Zürich! I dedicate this work to my parents.
3.4.1 Simple Integrands ................................................. 45
3.4.2 Sub-linear Pathwise Integrals ................................. 46
3.4.3 European Options ................................................ 46
3.5 Price Bounds of Derivative Contracts ......................... 47
3.5.1 The No Model-Free Arbitrage Condition .................... 47
3.5.2 The Model-Free Price Bounds ................................. 47
3.6 Super-replication Functionals ..................................... 49
3.6.1 First Super-replication Functional ........................... 49
3.6.2 Upper Semicontinuous Payoffs ................................. 50
3.6.3 Measurable Payoffs .............................................. 50
3.6.4 Uniformly Continuous Payoffs and Simple Strategies .... 52
3.7 Examples and Counterexamples ................................. 55
3.7.1 Upper Semicontinuous Payoffs ................................. 55
3.7.2 Example: Simple Strategies Are Not Sufficient ............ 56
3.7.3 Example: Dynamic Strategies Are Not Sufficient .......... 57
3.8 Extensions ........................................................... 58
3.9 Properties of the Superhedging Functional $\Phi_d$ ............ 61
3.9.1 Sub-differential at the Origin .................................. 61
3.9.2 Continuity of the Superhedging Functional ................ 63
3.9.3 Compactness of the Sub-differential $\partial \Phi_d(0)$ ........... 66
3.10 Properties of the Superhedging Functional .................... 67

4 Appendix ............................................................... 69
4.0.1 Convex duality in topological vector spaces ............... 69
4.0.2 Regularity and compactness .................................... 72
4.0.3 Topologies on the Skorokhod space ......................... 73
4.0.4 Strict Topology $\beta_0$ ......................................... 77
4.0.5 Functional Choquet’s Capacity ............................... 78
4.0.6 Transfinite Schema .............................................. 79
Chapter 1

Introduction

1.0.1 The Monge problem

The classical optimal transport problem was introduced by Monge [monge1784] in 1784. We present the problem in its general form as in [bogachevot]. Assume two probability spaces \((X, A, \mu)\) and \((Y, B, \nu)\) and a non-negative measurable function \(h\) on \(X \times Y\), called a cost function. The requirement is to find a measurable mapping \(T : X \to Y\) that takes the measure \(\mu\) into the measure \(\nu\) and minimizes the integral

\[
M(\mu, \nu, T) = M_h(\mu, \nu, T) := \int_X h(x, T(x))\mu(dx)
\]

over all such measurable functions. The requirement that \(\mu\) is taken to \(\nu\) means that the measure \(\nu\) is equal to the pushforward measure \(\mu \circ T^{-1}\) of \(\mu\) under the mapping \(T\). A mapping \(T\) that gives the minimum (if it exists) is said to be 'optimal transportation' of \(\mu\) to \(\nu\). Already, in a very simple case the solution to the Monge problem fails to exist.

Example 1.0.1. Let \(X = Y = [-1, 1]\) be endowed with the usual Borel \(\sigma\)-algebra, and let \(\mu = \delta_0\), \(\nu = 2^{-1}(\delta_{-1} + \delta_1)\), and \(h(x, y) = (x - y)^2\). The there exists no \(T\) transporting \(\mu\) into \(\nu\).

1.0.2 The Kantorovich problem

Kantorovich [kantorovich1942] formulated a problem that is very close to Monge’s problem, nowadays known as the Monge-Kantorovich problem. In this problem, instead of searching for a function \(T\), it is proposed to find a probability measure \(\sigma\) on the product space \((X \times Y, A \otimes B)\) that has its marginal projections \(\mu\) and \(\nu\) on \(X\) and \(Y\), respectively, and that minimizes the integral

\[
K(\mu, \nu, \sigma) = K_h(\mu, \nu, \sigma) := \int_{X \times Y} h(x, y)d\sigma(x, y)
\]

over all probability measures \(\sigma\) on the product space \(X \times Y\) that projections \(\mu\) and \(\nu\) on \(X\) and \(Y\), respectively.

Example 1.0.2. A half-sum of the Dirac measures at the points \((0, -1)\) and \((0, 1)\) is the solution to the Monge-Kantorovich problem given in Example 1.0.1.

In contrast to the complicated non-linear Monge problem, the Kantorovich variant is linear and admits a solution whenever \(h\) is an lsc function on a completely regular space \(X \times Y\) bounded from below.
Given two measures $\mu$ and $\nu$ on completely regular spaces $X$ and $Y$, there is a compact set $K \subset X \times Y$ such that $\sigma((X \times Y) \setminus K) < \varepsilon$, as e.g. one can choose $K = K_1 \times K_2$, where the compact sets $K_1$ and $K_2$ are such that $\mu(X \setminus K_1) + \nu(Y \setminus K_2) < \varepsilon$. So, the family of all of measures satisfying these marginals is tight i.e. $\beta_0$-equicontinuous. Then $\Pi(\mu, \nu)$ weakly closed gives $\Pi(\mu, \nu)$ is weakly compact. In [kantorovich1948] Kantorovich considered only compact subsets of $\mathbb{R}^n$. For compact spaces, the family $\Pi(\mu, \nu)$ is immediately compact without any tightness argument, $\beta_0$ coincides with $\|\cdot\|_\infty$. The $\beta_0$-topology provides an appropriate functional analytic formalism for the notion of tightness. The basic facts about compactness and $\beta_0$-topology are provided in Appendix.

The functional $K(\mu, \nu, \cdot)$ is linear, so, for the existence of a minimum on $\Pi(\mu, \nu)$ it suffices to have the functional is lower semicontinuous on $\Pi(\mu, \nu)$ in the weak topology.

For instance, the lower semicontinuity and a lower bound for $h$ guarantee the lower semicontinuity of the functional $K(\mu, \nu, \cdot)$. In [kantorovich1948] Kantorovich stated that the Monge problem is a special case of his problem and wrote that "from a solution of the latter one can easily get a solution of the former". This is not true, but the opposite statement is true: if the Monge problem has a solution $T$, then the measure $\sigma$ that presents the first marginal $\mu$ concentrated on the graph of $T$ is a solution to the Kantorovich problem.

Let us denote the optimal cost for the Monge-problem

$$M(\mu, \nu) = M_h(\mu, \nu) := \inf \{ M_h(\mu, \nu, T) : T \in T(\mu, \nu) \}$$

and for the Kantorovich problem

$$K(\mu, \nu) = K_h(\mu, \nu) := \inf \{ K_h(\mu, \nu, \sigma) : \sigma \in \Pi(\mu, \nu) \}.$$ 

We have that

$$K(\mu, \nu) \leq M(\mu, \nu).$$

1.0.3 The dual problem of optimal transport

We shall now turn to the dual problem of the Monge-Kantorovich problem that goes back to the original Kantorovich’s paper [kantorovich1942]. Assume two probability measures $\mu$ and $\nu$ on $(X, \mathcal{A})$ and $(Y, \mathcal{B})$, respectively. Let $\mathcal{H}_h$ denote the set of pairs of measurable functions $(\varphi, \psi)$, $\varphi : (X, \mathcal{A}) \to \mathbb{R}$ and $\psi : (Y, \mathcal{B}) \to \mathbb{R}$, such that

$$\varphi(x) + \psi(y) \leq h(x, y), \quad x \in X, \ y \in Y.$$ 

The Kantorovich dual problem consists of finding the minimum of the expression

$$J(\mu, \nu) = J_h(\mu, \nu) := \sup \left\{ \int_X \varphi d\mu + \int_Y \psi d\nu, \ (\varphi, \psi) \in \mathcal{H}_h \right\}.$$ 

For any measurable cost function $h$, one has the estimate

$$J(\mu, \nu) \leq K(\mu, \nu) \leq M(\mu, \nu).$$ 

The duality theorem, called the Monge-Kantorovich duality, assert sufficient conditions for that

$$J(\mu, \nu) = K(\mu, \nu).$$ 

The first duality result goes back to the original Kantorovich’s paper [kantorovich1942], for $X$ and $Y$ compact subsets of an Euclidean space. Since then, various generalizations
have been provided. Our investigations in martingale optimal duality borrow heavily ideas from the classical optimal transport, in particular, from those of Kellerer [kellerer1984]. We present the main results involving the duality \( J(\mu, \nu) = K(\mu, \nu) \).

Kellerer [kellerer1984] was the first one who proved a duality result for measurable payoffs. We adapt the argument of Kellerer to extend the martingale optimal transport duality from upper semicontinuous case to the measurable case. In fact, Kellerer\’s approach based Choquet\’s capacitability can be used to extend the duality beyond measurable payoffs. The required results of capacity on function lattices are provided in Appendix.

A particularly short proof for the duality in the case of Radon measures on completely regular spaces and non-negative lower semicontinuous functions \( h \) was provided by Edwards [edwards]. The proof is based on the idea of combining the general form Riesz representation theorem with the Hahn-Banach theorem via the Fenchel-Moreau biconjugation theorem. Our approach to the martingale optimal transport duality for upper semicontinuous payoffs is very similar to that of Edwards [edwards].

Another set of necessary conditions for the duality were provided quite recently by Beiglböck and Schachermayer [bs]. In the paper [bs] a following example is constructed.

**Example 1.0.3.** Assume \( X = Y = [0, 1] \), that \( \mu = \nu \) is Lebesgue measure. Let the cost function \( h \) be \( h(x, y) = \infty \) if \( x < y \), \( h(x, x) = 1 \), and \( h(x, y) = 0 \) if \( x > y \). Then \( K(\mu, \nu) = 1 \), but \( J(\mu, \nu) = 0 \).

A similar counterexample was constructed for martingale optimal transport in [nutz] and is a guiding example also in the continuous-time framework.

The most general result on the Monge-Kantorovich duality is of Ramachandran and Rüschendorf [ruschendorf] that states that the duality holds if at least one of the measures \( \mu \) and \( \nu \) is perfect (e.g. all Radon measures are perfect). As stated in [ruschendorf], the counterexamples given in [ramachandran1996] and [ramachandran2000] for the fact that the perfectness is also a necessary property are wrong. The necessity appears to be an open question.

In the case of completely regular spaces \( X \) and \( Y \) there exists another formulation of the Kantorovich dual problem. Namely, instead of considering the class of \( H_h \) of pairs of measurable functions \( \varphi \) and \( \psi \) one considers a narrower class \( G_h \) of pairs bounded continuous functions \( \varphi \) and \( \psi \) satisfying the inequality \( \varphi(x) + \psi(y) \leq h(x, y) \) for all \( X \in X \) and \( y \in Y \). By taking the supremum over the integral over all functions in \( G_h \) one obtains the quantity \( I_h(\mu, \nu) = I(\mu, \nu) \). It is clear that \( J(\mu, \nu) \leq I(\mu, \nu) \).

In general the inequality is strict. In fact, it was shown in [levin] that the equality \( J(\mu, \nu) = I(\mu, \nu) \) is equivalent to that cost function \( h \) is lower semicontinuous. Our dual formulation for martingale optimal transport is comparable to the dual Kantorovich formulation \( I(\mu, \nu) \) when the payoff function \( h \) is upper semicontinuous while for measurable payoff functions \( h \) we accordingly allow measurable dual functions as in \( J(\mu, \nu) \).

Convex analysis is a central tool for studying the measure transport duality. The argument from convex analysis appear in the standard reference texts of [svetlozar] and [villani2003] as well as in the more recent works [apteemu] and [ekren]. In [apteemu], the Monge-Kantorovich problem and its non-linear generalizations were studied in the general duality framework of Rockafellar. In [ekren], a general approach for constrained optimal transportation was developed. We also rely in convex analysis in our approach. The principal facts about convex duality on a locally convex space are provided in Appendix.
1.0.4 Martingale optimal transport

Let us now describe the martingale optimal transport problem. The problem was introduced in [beiglböck] by Beiglböck, Henry-Labordère and Penkner and their duality result was extended for measurable functions in [nutz] by Beiglböck, Touzi and Nutz. Assume two probability spaces \((X, \mathcal{A}, \mu)\) and \((Y, \mathcal{B}, \nu)\) in convex order \(\mu \leq_{cx} \nu\) and a measurable bounded function \(h\) on \(X \times Y\), often called a payoff function. In the martingale optimal transport problem, in contrast to the Monge-Kantorovich problem of looking all probability measures on the product space \((X \times Y, \mathcal{A} \otimes \mathcal{B})\) that project \(\mu\) on \(X\) and \(\nu\) on \(Y\), one restricts oneself to the family of martingale measures. A probability measure \(\rho\) on \((X \times Y, \mathcal{A} \otimes \mathcal{B})\) is a called martingale measure, if the coordinate process \((x, y)\) is a martingale with respect to its canonical filtrations i.e.

\[
\int_{X \times Y} f(x - y) \, d\rho = 0 \quad \forall f \in \mathcal{B}(\mathcal{X}).
\]

The martingale optimal transport problem is to find a martingale measure \(\rho\) on \(X \times Y\) such that its projections on \(X\) and \(Y\) are \(\mu\) and \(\nu\), and, as in the Monge-Kantorovich problem, it maximizes the integral

\[
\int_{X \times Y} h(x, y) \, d\rho(x, y).
\]

There is a more complicated variant of the problem. As pointed out in [bogachevot], already Monge considered the measure transportation as a process in time and space, he considered the trajectories of transportation. Indeed, one can imagine a transportation from \(X = \mathbb{R}^{d}\) to \(Y = \mathbb{R}^{d}\) as a process in \(\mathbb{R}^{[0,1]}\) that presents the movement of a particle located at \(x\) with a constant speed \(T(x) = x\) along straight lines i.e. given a transport \(T\) one considers the family \(T = (T_{t})_{t \in [0,1]}\) of functions

\[
T_{t} : x \mapsto (1 - t)x + tT(x), \quad x \in \mathbb{R}^{d}
\]

while each \(T(x) : t \mapsto T_{t}(x)\) presents the trajectory of displacement at \(x\) with \(T_{0}(x) = x\) and \(T_{1} = T\). Given \(\mu\), this induces a family of measures \(\mu_{t} = \mu_{0} \circ T_{t}^{-1}, t \in [0,1]\), and \(\nu = \mu \circ T_{1}^{-1}\). Similarly, by denoting the projection of the trajectory at time \(t\) by \(\pi_{t}\), for a dynamic transportation plan \(\rho\), one can consider interpolating family of measures \(\rho_{t} = \rho \circ \pi_{t}^{-1}, t \in [0,1]\), with \(\rho_{0} = \mu\) and \(\rho_{1} = \nu\).

The martingale optimal transport problem that takes into account the trajectory of transportation was introduced on the classical Wiener space \(\mathcal{C}([0,1]; \mathbb{R}^{n})\) in by Dolinsky and Soner in [mete] and extended to the Skorokhod space \(\mathcal{D}([0,1]; \mathbb{R}^{n})\) by them in [mete2]. The problem has been actively studied since its introduction. A comprehensive list of references is provided in Section 3.

In the continuous-time formulation, we are given two probability measures \(\mu\) on \((X, \mathcal{A})\) and \(\nu\) on \((Y, \mathcal{B})\) that are in convex order \(\mu \leq_{cx} \nu\) and a payoff function \(h\) depending on the whole \(x \in \Omega\) connecting \(X\) to \(Y\). The spaces \(X\) and \(Y\) present the initial and the value of trajectory of transportation. A martingale transportation is a martingale measure \(\rho\) for the canonical process \(x\) of \(\Omega\) such that \(x_{0} \sim \mu\) and \(x_{1} \sim \nu\). Denoting by \((\mathcal{F}_{t})_{t \in [0,1]}\) the canonical filtration of the path space \(\Omega\), the martingale constraint reads as

\[
\int_{\Omega} f(x_{s})(x_{t} - x_{s}) \, d\rho = 0 \quad \forall f \in \mathcal{B}(\Omega, \mathcal{F}_{t}).
\]
The optimal martingale transportation problem is to maximize the integral
\[ D(\mu, \nu, \rho) = D_h(\mu, \nu, \rho) := \int_\Omega h(x) d\rho(x) \]
over all martingale measures \( \rho \) in the class \( Q(\mu, \nu) \) of martingale measures on the canonical space \( \Omega \) that has projections \( \mu \) and \( \nu \) at times \( t = 0 \) and \( t = 1 \), respectively. We denote the convex dual, the value of the support functional of the family \( Q(\mu, \nu) \) at \( h \), by
\[ D(\mu, \nu) = D_h(\mu, \nu) := \sup \{ D_h(\mu, \nu, \rho) : \rho \in Q(\mu, \nu) \} . \]

An important feature related to this quantity is the compactness of the family \( Q(\mu, \nu) \), observed by Guo, Tan and Touzi in [touzi]. In Part II, we study systematically this compactness in the context of the Riesz representation theorem. Compactness is almost necessary for extending the duality beyond (uniformly) continuous payoffs.

1.0.5 The Dual problem of martingale optimal transport

We denote by \( H_h \) the set of triples \((\varphi, \psi, \alpha)\) of bounded measurable functions \( \varphi, \alpha : (X, \mathcal{A}) \to \mathbb{R}, \psi : (Y, \mathcal{B}) \to \mathbb{R} \) that satisfy the inequality
\[ h(x, y) \leq \varphi(x) + \psi(y) + \alpha(x)(y - x), \quad x \in X, \ y \in Y. \]

For this Beiglböck, Henry-Labordère and Penkner considered the so-called superhedging problem of finding the minimum of the integral
\[ \int_X \varphi d\mu + \int_Y \psi d\nu. \]
In the continuous-time formulation to take into account the trajectory of transportation one replaces \( \alpha(x)(y - x) \) with a measurable function \( (\alpha \circ X)_1 \) on \( \Omega \)
\[ (\alpha \circ X)_1 := \int_{[0,1]} \alpha dX. \]

We denote quantity
\[ S(\mu, \nu) = S_h(\mu, \nu) := \inf \left\{ \int_X \varphi d\mu + \int_Y \psi d\nu, \ (\varphi, \psi, \alpha) \in H_h \right\} . \]

This problem, inspired by robust mathematical finance, is central in our circle of considerations. The choice of the triplet \((\varphi, \psi, \alpha)\) is a delicate question. Obviously, if \( (\alpha \circ X)_1 \) admits a lower bound, then one has the estimate
\[ \int_{X \times Y} h d\rho \leq \int_X \varphi d\mu + \int_Y \psi d\nu \]
for all \((\varphi, \psi, \alpha) \in H_h \) and \( \rho \in Q(\mu, \nu) \) i.e. one has that
\[ D(\mu, \nu) \leq S(\mu, \nu) . \]

In Part III, we study sufficient conditions for the equality
\[ D(\mu, \nu) = S(\mu, \nu) . \]
Part I

Radon Measure Theory for Semimartingales
Chapter 2

Radon Measure Theory for Semimartingales

2.1 Background

The Riesz representation theorem states that the operation of integration defines a one-to-one correspondence between the continuous linear functionals on continuous bounded functions and the Radon measures on a topological space. On the Skorokhod space, it provides a locally convex way of constructing all càdlàg stochastic processes on the canonical space as tight probability measures. On conceptual level, any criteria that characterizes a certain object should give rise to some kind of compactness when applied uniformly to a family of objects. We relate Stricker’s uniform tightness condition of semimartingales to the weak $\ast$ compactness in the pairing of the Riesz representation theorem.

Weak topologies on the Skorokhod space and weak convergence of stochastic processes has been earlier studied in the works of Meyer and Zheng [meyerzheng], Zheng [zheng], Stricker [stricker], Jakubowski, Mémé and Pages [jakubowski6], Kurtz [kurtz], and Jakubowski [jakubowski], [jakubowski4]. These topologies are rich in terms of convergent subsequences and have become a central tool for studying e.g. weak convergence of financial markets [prigent2003], time series in econometrics [chan2010], stochastic optimal control [kurtz2001], [bahlali2011], [tan2013] and pathwise superhedging [touzi]. We refer to [jakubowski4] for a comprehensive list of applications.

Despite the wide range of applications, compactness arising from these topologies has been lacking a solid functional analytic characterization and so far their usage has been restricted in sequential methods. Our aim is to provide such a characterization, unify and elaborate the existing results and thus enhance the applicability of weak topologies of the Skorokhod space in functional analysis, e.g., on problems arising in statistics, economics and finance.

The first objective of the paper is to prove the weak $\ast$ compactness result of semimartingales on the canonical space under general topological assumptions. The assumptions allow to study the weak convergence of stochastic processes as a weak $\ast$ topology in the pairing of the Riesz representation theorem. Thus, our approach is fully consistent with the duality theory of linear topological spaces. This is in contrast to the earlier works [meyerzheng] and [stricker], where sequential relative compactness results were established for the weak convergence of the Meyer-Zheng pseudo-path topology, and [jakubowski], where the weak converge of the $S$-topology was studied in the topology
induced by the subsequent Skorokhod’s representation theorem. Our main contribution is to unify these previous results and provide an easy method for constructing weak* compact sets of semimartingales on the canonical space. We also give examples of such sets and show that the examples are consistent with earlier results for Banach spaces of stochastic processes defined over a common probability space.

The second objective of the paper is to characterize the strongest topology on the Skorokhod space for which the weak* compactness result is true. A natural candidate is Jakubowski’s S-topology, due to its tightness criteria. However, it is an open problem whether the S-topology is sufficiently regular. We address the problem of regularity by introducing a new weak topology on the Skorokhod space that has the same continuous functions as the S-topology, suitable compact sets and additionally satisfies a strong separation axiom. The topological space is perfectly normal (Tb) in comparison to the Hausdorff property (T2) that has been verified for the S-topology. The topology is obtained from Jakubowski’s S-topology as a result of a standard regularization method that appears already in the classical works of Alexandroff [alexandroff] and Knowles [knowles]. Our contribution is to carefully show that the important properties of the S-topology are preserved in the regularization.

The rest of the paper is organized as follows.

In Section 2.2, we give rigorous definitions of semimartingale measures on the canonical space and related facts. We also provide a brief introduction to the aforementioned Riesz representation theorem that is the basis of our approach. The main results and examples are given in Section 2.3. Section 2.4 gathers the auxiliary results needed in the previous section. In Section 2.5, we characterize the strongest topology on the Skorokhod space for which the results of the two previous sections are true. Some definitions and technical results are omitted in the main part of the article and are gathered in Appendix ??.

### Conventions and notations

Throughout, the parenthesis ( ) should be understood so that they can be left out. The comparatives ‘weaker’ and ‘stronger’ should be understood in the wide sense ‘weaker or equally strong’ and ‘stronger or equally strong’, respectively. We say that two topologies are ‘comparable’, if one is stronger than another.

We fix the following notations.

- **N** denotes the family of natural numbers and \( \mathbb{N}_0 := \{0\} \cup \mathbb{N} \).
- \( \mathbb{R}_{(+)} \) denotes the family of (non-negative) real numbers and \( \mathbb{R} := \mathbb{R} \cup \{ \pm \infty \} \).
- \( \mathbb{Q} \) denotes the family of rational numbers.
- \( \mathbb{D}, \mathbb{D}(I) \) and \( \mathbb{D}(I; \mathbb{R}^d) \) denote the \( \mathbb{R}^d \)-valued càdlàg functions on \( I \).
- \( \mathbb{C}_0(I) \) denotes the family of (bounded) continuous functions.
- \( \mathbb{B}_0(I) \) denotes the family of (bounded) Borel functions.
- \( \mathbb{B}_0 \) denotes the family of bounded Borel functions vanishing at infinity.
- \( \mathcal{V} \) denotes the family of functions of finite variation.
- \( |x| := |x_1| + |x_2| + \cdots + |x_d|, \ x = (x_1, x_2, \ldots, x_d) \in \mathbb{R}^d \).
- \( x \vee y := \max\{x, y\}, \ x \wedge y := \min\{x, y\}, \ x^+ := x \vee 0, \ x^- := -(x \wedge 0), \ x, y \in \mathbb{R} \).
- \( x^+ = x_1^+ + x_2^+ + \cdots + x_d^+ \), \( x^- = x_1^- + x_2^- + \cdots + x_d^- \), \( x = (x_1, x_2, \ldots, x_d) \in \mathbb{R}^d \).
- \( ||x||_{\infty} := |x_1| \vee |x_2| \vee \cdots \vee |x_d|, \ x = (x_1, x_2, \ldots, x_d) \in \mathbb{R}^d \).
- \( ||X||_{\infty} := \sup_{t \in I} ||X(t)||_{\infty}, \) for \( X : I \rightarrow \mathbb{R}^d \).
- \( \pi \) denotes a finite partition of \( I \subset \mathbb{R}_+ \).
- \( ||X||_{\mathcal{V}} := \sum_{i=1}^{\infty} \sup_{\pi} \{|X^i(0)| + \sum_{k=1}^{\infty} |X^i(t_k) - X^i(t_{k-1})|\}, \) for \( X : I \rightarrow \mathbb{R}^d \).
2.1. BACKGROUND

\[ E_Q[f] := \int f \, dQ \] denotes the (Radon) integral.

\[
\|f\|_{L^p(Q)} := \left( E_Q[|f|^p] \right)^{1/p}, \quad p \geq 1,
\]
denotes the \(L^p\)-norm.

\[
\|X\|_{L^p,\infty(Q)} := \max \{ \|X\|_{L^p(Q)}, \|X\|_{\infty(Q)} \}, \text{ for } X : I \to \mathbb{R}^d.
\]

\( X = M + A \) denotes a canonical semimartingale decomposition.

\[
\|X\|_{\mathcal{H}^p(Q)} := \|\|M\|_\infty + \|A\|_\mathcal{V}\|_{L^p(Q)}, \quad p \geq 1,
\]
is the \((\mathcal{H}^p)\)-norm.

\( \mathcal{H} \) denotes the family of elementary predictable processes bounded by 1.

\((H \circ X) := H_0X_0 + \int H \, dX\) denotes the stochastic integral.

\[
\|X\|_{\mathcal{E}^p(Q)} := \sup_{H \in \mathcal{H}(Q)} \| (H \circ X) \|_{L^p(Q)}, \quad p \geq 1.
\]

\( \mathcal{M}_\tau \) denotes the family of \(\tau\)-additive Borel measures of finite variation.

\( \mathcal{M}_\sigma \) denotes the family of \(\sigma\)-additive Borel measures of finite variation.

If \( \mathcal{M}_1 = \mathcal{M}_\tau = \mathcal{M}_\sigma \), then we denote the whole three classes by \( \mathcal{M} \).

\( \mathcal{M}_+ \) denotes the family of non-negative elements of \( \mathcal{M} \).

\( \mathcal{P} \) denotes the family of all probability measures.

\( \mu_n \to \mu^* \) denotes the weak* convergence on \( \mathcal{M} \).

\( [A]_{\text{seq}} := \{ \mu \in \mathcal{M} : \exists (\mu_n)_{n \in \mathbb{N}} \subset A \text{ s.t. } \mu_n \to \mu \} \) is the sequential closure.

\( \mathcal{K} \), \( \mathcal{B} \) and \( \mathcal{B}_0 \) denote the family of compact, Borel and Baire sets, respectively.

\( \hat{\mathcal{F}} \) denotes the universal completion of a \(\sigma\)-algebra \( \mathcal{F} \).

\( \beta_0 \) denotes the topology generated by the family of seminorms \( \| \cdot \|_\infty, \, f \in \mathbb{B}_0 \).

\( N^x_{a,b} \) denotes the number of upcrossings of an interval \([a, b]\) w.r.t. \( \pi \).

\( N^+_{a,b} := \sup_{\pi} N^x_{a,b} \) denotes the number of upcrossings of an interval \([a, b] \).

\([\omega]^t \) denotes the restriction of \( \omega \) on \([0, t]\).

Terminology

We provide some frequently used terminology. Standard literature references for general topology and topological measure theory are [engelking] and [bogachev].

All topological spaces considered are Hausdorff \((T_2)\) and a Hausdorff space \(X\) is called:

Regular \((T_3)\), if for every point \( x \in X \) and every closed set \( Z \) in \( X \) not containing \( x \), there exists disjoint open sets \( U \) and \( V \) such that \( x \in U \) and \( Z \subset V \).

Completely regular \((T_{3\frac{1}{2}})\), if for every point \( x \in X \) and every closed set \( Z \) in \( X \) not containing \( x \), there exists a continuous function \( f : X \to [0, 1] \) such that \( f(x) = 1 \) and \( f(z) = 0 \) for all \( z \in Z \).

Perfectly normal \((T_6)\), if every closed set \( Z \subset X \) has the form \( Z = f^{-1}(0) \) for some continuous function \( f \) on \( X \).

Paracompact, if every open cover of \( X \) has an open refinement that is locally finite.

\(k\)-space, if the set \( A \subset X \) is closed in \( X \) provided that the intersection of \( A \) with any compact subspace \( Z \) of the space \( X \) is closed in \( Z \).

Sequential space, if every sequentially closed set is closed.

Fréchet-Urysohn space, if every subspace is a sequential space.

Polish space, if the space is homeomorphic to a complete separable metric space.

Lusin space, if the space is the image of a complete separable metric space under a continuous one-to-one mapping.

Souslin space, if the space is the image of a complete separable metric space under a continuous mapping.

Radon space, if every Borel measure on the space is a Radon measure.
Remark 2.1.1. For closed subspaces, all these properties are hereditary, meaning that, if the space has the property, then a closed subspace endowed with the relative topology has the property as well. So, all discussion on these properties generalizes as such for relative topologies on closed sets.

2.2 Semimartingales as linear functionals

In this preliminary section, we define the canonical space for semimartingales, and related measures and continuous linear functionals.

2.2.1 Canonical space

We fix $I$ to denote a usual time index set of a stochastic process, i.e., $I := [0, T]$ for $0 < T < \infty$ or $I := [0, \infty]$. The Skorokhod space $\mathbb{D}(I; \mathbb{R}^d)$, $d \in \mathbb{N}$, with the domain $I$ consists of $\mathbb{R}^d$-valued càdlàg functions $\omega$ on $I$ that admit a limit $\omega(t^-)$ from left, for every $t > 0$, and are continuous from right, $\omega(t) = \omega(t^+)$, for every $t < T$. The space $\mathbb{D}([0, \infty]; \mathbb{R}^d)$ is regarded as a product space $\mathbb{D}([0, \infty]; \mathbb{R}^d) \times \mathbb{R}^d$; see Appendix 4.0.3.

We write $\omega = (\omega^1, \ldots, \omega^d)$, for $\omega \in \mathbb{D}(I; \mathbb{R}^d)$, if $\omega(t) = (\omega^1(t), \ldots, \omega^d(t))$, for every $t \in I$. We denote by $X$ the canonical process of $\mathbb{D}(I; \mathbb{R}^d)$, i.e., $X_t(\omega) = \omega(t)$, for all $(t, \omega) \in I \times\mathbb{D}(I; \mathbb{R}^d)$. We write $X^i$ for each coordinate processes of the canonical process $X$, for $i \leq d$.

We endow the Skorokhod space $\mathbb{D}(I; \mathbb{R}^d)$, $d \in \mathbb{N}$, with the right-continuous version $\mathcal{F}_t := \bigcap_{\varepsilon > 0} \mathcal{F}_{t+\varepsilon}$ of the raw, i.e., unaugmented, canonical filtration $\mathcal{F}_t := \sigma(X_s : s \leq t)$ generated by the canonical process $X$ of $\mathbb{D}(I; \mathbb{R}^d)$. The right-continuous version of the raw canonical filtration is needed in the proof of Proposition 2.4.12. Alternatively, we could use the universal completion of the raw canonical filtration; see Proposition 2.4.1 (b).

A stochastic process is understood as a probability measure on the filtered canonical space $\left(\mathbb{D}(I; \mathbb{R}^d), \mathcal{F}_T, (\mathcal{F}_t)_{t \in I}\right)$, where $\mathcal{F}_T := \sup_{t \in I} \mathcal{F}_t = \sup_{t \in I} \mathcal{F}_t$ and $T = \sup_{t \in I} t \in [0, \infty]$. The family of all probability measures (processes) on $\left(\mathbb{D}(I; \mathbb{R}^d), \mathcal{F}_T\right)$ is denoted by $\mathbb{P}(\mathbb{D}(I; \mathbb{R}^d), \mathcal{F}_T)$ and two elements of $\mathbb{P}(\mathbb{D}(I; \mathbb{R}^d), \mathcal{F}_T)$ are identified as usual, i.e., $P = Q$, if $P(F) = Q(F)$, for all $F \in \mathcal{F}_T$; cf. Subsection 2.4.1.

2.2.2 Semimartingales on the canonical space

We recall some basic concepts from semimartingale theory in the present setting. We adapt the terminology of [mete2] and [touzi] and call a probability measure $Q$ on $\left(\mathbb{D}(I; \mathbb{R}^d), \mathcal{F}_T\right)$ a martingale measure, if the canonical process $X$ is a martingale on $\left(\mathbb{D}(I; \mathbb{R}^d), \mathcal{F}_T, (\mathcal{F}_t)_{t \in I}, Q\right)$. (Special) semimartingale and supermartingale measures are defined similarly. On the probability space $\left(\mathbb{D}(I; \mathbb{R}^d), \mathcal{F}_T, (\mathcal{F}_t)_{t \in I}, Q\right)$, for a fixed probability measure $Q$, let $\mathcal{H}(Q)$ denote the family of elementary predictable integrands, i.e., the family of adapted càglàd processes of the form

$$H^i = H^i_0 1\{0\} + \sum_{k=1}^n H^i_{t^i_{k-1}} 1\{t^i_{k-1}, t^i_k]\), \quad i \leq d,$$

(2.2.1)

where $n \in \mathbb{N}$, $0 = t^i_0 \leq t^i_1 \leq \cdots \leq t^i_n$ in $I$ and each $H^i_{t^i_k}$ is $\mathcal{F}_{t^i_k}$-measurable random variable in $L^\infty(Q)$ satisfying $|H^i_{t^i_k}| \leq 1$. For a family $Q \subset \mathcal{P}\left(\mathbb{D}(I; \mathbb{R}^d), \mathcal{F}_T\right)$, consider the following
condition
\[
\limsup_{c \to \infty} \sup_{Q \in \mathcal{Q}} \sup_{H \in \mathcal{H}(Q)} Q\left( |(H \circ X)_t| > c \right) = 0, \quad \forall t \in I,
\] (UT)
where
\[(H \circ X)_t = \sum_{i=1}^{d} (H^i \circ X^i)_t, \quad t \in I.\]

The condition (UT) was introduced by Stricker in [stricker]. By the classical result of Bichteler, Dellacherie and Mokobodzki, a single probability measure on \(\mathbb{D}(I; \mathbb{R}^d), \mathcal{F}_T\) is an \((\mathcal{F}_t)_{t \in I}\)-semimartingale measure if and only if it satisfies the condition (UT). The family of process (2.2.1) generates the predictable \(\sigma\)-algebra and the condition (UT) is sometimes called the \textit{predictable uniform tightness condition} (P-UT); see e.g. [hewangyan][Thm. 3.21]. Remark that, for semimartingale measures, no integrability condition is imposed on \(X_0\), i.e., the localization of the local martingale in a canonical semimartingale decomposition is understood in the sense of [hewangyan][Def. 7.1]; cf. [jacobodshiryaev][Rem. 6.3]. We say that a semimartingale measure \(Q\) is of \textit{class \(\mathcal{H}^p\)} if, on \(\mathbb{D}(I; \mathbb{R}^d), (\mathcal{F}_t)_{t \in I}, \mathcal{F}_T, Q\), the canonical process \(X\) decomposes to a (local) martingale \(M\) and a (predictable) finite variation process \(A\), \(A_0 = 0\), such that
\[
X = M + A \quad \text{and} \quad \|X\|_{\mathcal{H}^p(Q)} = \|M\|_{\infty} + \|A\|_{\mathcal{V}^p(Q)} < \infty.
\] (2.2.2)

Every semimartingale of class \(\mathcal{H}^p\), for some \(p \geq 1\), is a special semimartingale. To obtain compact statements for quasi- and supermartingales, we introduce two conditions. The first condition is
\[
\sup_{Q \in \mathcal{Q}} \sup_{t \in I} \left( E_Q[|X_t|] + \sup_{H \in \mathcal{H}(Q)} E_Q[(H \circ X)_t] \right) < \infty.
\] (UB)

The second condition is the same condition, but the \(L^1\)-boundedness is strengthened to the uniform integrability of the negative parts, for every \(t \in I\), i.e.,
\[
\text{Q satisfies (UB) and} \quad \lim_{c \to \infty} \sup_{Q \in \mathcal{Q}} E_Q[X_t^{-1}_{\{X_t^- > c\}}] = 0, \quad \forall t \in I.
\] (UI)

The uniform integrability in (UI) yields the convergence of the first moments that preserves the supermartingale property; see Proposition 2.4.11. If we insist that \(t_n^i = t\) in (2.2.1), then the second supremum in (UB) is attained, by choosing
\[
H_k^i = \text{sign}(E_Q[X_{t_k^i}^i - X_{t_{k-1}^i}^i \mid \mathcal{F}_{t_{k-1}^i}]), \quad 1 \leq k \leq n, \quad i \leq d,
\]
for which the value of the integral is equal to the \((\mathcal{F}_t)_{t \in I}\)-conditional variation of \(X^i\) on \([0, t],\)
\[
\text{Var}^Q_t(X^i) := \sup E_Q \left[ |X^i(0)| + \sum_{k=1}^{n} |E_Q[X_{t_k^i}^i - X_{t_{k-1}^i}^i \mid \mathcal{F}_{t_{k-1}^i}]| \right], \quad i \leq d,
\] (2.2.3)
where the supremum is taken over all partitions \(0 \leq t_0^i \leq t_1^i \leq \cdots \leq t_n^i = t, \ n \in \mathbb{N}; \) see e.g. [dellacheriemeyer][B, Appendix II]. A probability measure on \(\mathbb{D}(I; \mathbb{R}^d), \mathcal{F}_T\) is a \textit{quasimartingale measure} if and only if it satisfies the condition (UB); see e.g.
where the right-hand side is possibly infinite.

We have

\[(UI) \implies (UB) \implies (UT).\]  

(2.2.4)

The first implication is obvious. The second implication follows from Lemma 2.2.1.

**Lemma 2.2.1.** There exists a constant \(b > 0\) such that, for any \(Q \in \mathcal{P} (\mathbb{D}(I), \mathcal{F}_T)\), \(H \in \mathcal{H}(Q)\) and \(c > 0\), we have

\[
Q(|(H \circ X)_t| > c) \leq \frac{b}{c} \left( E_Q[|X_t|] + \sup_{H \in \mathcal{H}(Q)} E_Q[(H \circ X)_t] \right), \quad t \in I,
\]

(2.2.5)

where the right-hand side is possibly infinite.

The inequality (2.2.5) is well-known, but we provide the proof for the convenience of the reader.

**Proof.** The inequality (2.2.5) is a generalizations of Burkholder’s inequality, which states that there exists a uniform constant \(a > 0\) such that, for any \(H \in \mathcal{H}(Q)\), \(Q\)-martingale \(M\) and \(c > 0\), we have

\[
Q(|H \circ M|_t > c) \leq \frac{a}{c} E_Q[|M_t|], \quad t \in I;
\]

(2.2.6)

see e.g. [meyer] or [bichteler] for a proof of (2.2.6). For a fixed \(Q \in \mathcal{P} (\mathbb{D}(I; \mathbb{R}^d), \mathcal{F}_T)\) and \(H \in \mathcal{H}(Q)\), by [dellacheriemeyer][B, Appendix 2.3], we have

\[
\sup_{H \in \mathcal{H}(Q)} E_Q[(H \circ X)_t] = \sum_{i=1}^{d} \text{Var}_t^Q(X^i), \quad t \in I.
\]

(2.2.7)

Let us fix \(i \leq d\) and assume that \(E_Q[|X^i_t|] + \text{Var}_t^Q(X^i)\) is finite, otherwise, the result is trivial. Let \(t_0 < t_1 < \cdots < t_n = t\) and \(H = (H^i)_{i=1}^{d}\) be an element of \(\mathcal{H}(Q)\), i.e.,

\[
H^i = \sum_{k=1}^{n} H^i_{t,k} 1_{[t_{k-1}, t_k]}, \quad i \leq d,
\]

where each \(|H^i_{t,k}| \leq 1\) is \(\mathcal{F}_{t_k}\)-measurable; cf. (2.2.1). Consider the Doob decomposition

\[
X^i_{t,k} = M^i_{t,k} + A^i_{t,k}, \quad k = 1, 2, \ldots, n,
\]

where \(A^i_{t,k} = \sum_{j=1}^{k} E_Q[X^i_{t_j} - X^i_{t_{j-1}} | \mathcal{F}_{t_{j-1}}]\) and \(M^i_{t,k}\) is a \(Q\)-martingale on \(\{t^0, t^1, \ldots, t^n\}\).

We have

\[
Q(|(H^i \circ A^i)_t| > c) \leq \frac{1}{c} E_Q[|(H^i \circ A^i)_t|] \leq \frac{1}{c} \text{Var}_t^Q(X^i), \quad i \leq d.
\]

(2.2.8)

Similarly, for \(M^i\), we have

\[
E_Q[|M^i_t|] \leq E_Q[|X^i_t| + |A^i_t|] \leq E_Q[|X^i_t|] + \text{Var}_t^Q(X^i).
\]
Hence, by (2.2.6), we have

\[ Q((H^i \circ M^i)_t | > c) \leq \frac{a}{c} \left( E_Q[|X^i_t|] + \text{Var}_t^Q(X^i) \right), \quad i \leq d. \]  

(2.2.9)

Combining (2.2.7), (2.2.8) and (2.2.9), for \( H \in \mathcal{H}(Q), M = (M^i)_{i=1}^d \) and \( A = (A^i)_{i=1}^d \), we get

\[ Q((H \circ X)_t | > c) \leq Q((H \circ M)_t + (H \circ A)_t | > c) \]

\[ \leq \sum_{i=1}^d Q \left( (H^i \circ M^i)_t | > \frac{c}{2d} \right) + \sum_{i=1}^d Q \left( (H^i \circ A^i)_t | > \frac{c}{2d} \right) \]

\[ \leq \frac{2ad}{c} \sum_{i=1}^d \left( E_Q[|X^i_t|] + \text{Var}_t^Q(X^i) \right) + \frac{2d}{c} \sum_{i=1}^d \text{Var}_t^Q(X^i) \]

\[ \leq \frac{b}{c} \left( E_Q[|X_t|] + \sup_{H \in \mathcal{H}(Q)} E_Q[(H \circ X)_t] \right), \quad t \in I, \ c > 0, \]

where \( b := 2(a + 1)d \). The proof is completely similar for the filtration \((F^0_t)_{t \in I}\).

\[ \Box \]

2.2.3 The Riesz representation on the canonical space

The Riesz representation theorem for the laws of \( \mathbb{D} \)-valued random variables (stochastic processes) requires topological assumptions on the canonical space. We assume that the Skorokhod space \( \mathbb{D} \) is a completely regular Radon space on which the canonical \( \sigma \)-algebra coincides with the Borel \( \sigma \)-algebra. Under the assumption, the following Riesz representation theorem is true.

The Riesz representation theorem

The family \( \mathcal{M}(\mathbb{D}) \) of Radon measures (of finite total variation) on \( \mathbb{D} \) is isomorphic to the family of \( \beta_0 \)-continuous linear functionals on \( \mathcal{C}_b(\mathbb{D}) \) via the bilinear form

\[ u_\mu(f) := \int f \, d\mu, \quad f \in \mathcal{C}_b(\mathbb{D}), \ \mu \in \mathcal{M}(\mathbb{D}). \]  

(2.2.10)

In particular, we have \( \mathcal{P}(\mathbb{D}) \subset \mathcal{M}(\mathbb{D}) \) and, for the elements of \( \mathcal{P}(\mathbb{D}) \), we write

\[ E_Q[f] := \int f \, dQ, \quad f \in \mathcal{C}_b(\mathbb{D}), \ Q \in \mathcal{P}(\mathbb{D}). \]

The locally convex topology \( \beta_0 \), called the strict topology, is generated by the family of seminorms

\[ p_g(f) := \|fg\|_\infty, \quad f \in \mathcal{C}_b(\mathbb{D}), \ g \in \mathcal{B}_0(\mathbb{D}), \]

where

\[ \mathcal{B}_0(\mathbb{D}) := \{ f \in \mathcal{B}_b(\mathbb{D}) : \forall \varepsilon > 0 \ \exists K^\varepsilon \in \mathcal{K}(\mathbb{D}) \ \text{s.t.} \ |f(x)| < \varepsilon \ \forall x \notin K^\varepsilon \}. \]

The collection of finite intersections of the sets

\[ V_{g,\varepsilon} := \{ f \in \mathcal{C}_b(\mathbb{D}) : p_g(f) < \varepsilon \}, \ g \in \mathcal{B}_0(\mathbb{D}), \ \varepsilon > 0, \]
forms a local basis at the origin for the topology $\beta_0$. A subset $M$ of $\mathcal{M}(\mathbb{D})$ is $\beta_0$-
emph{equicontinuous} if and only if the family of linear functionals

$$u_\mu(f) = \int fd\mu, \; \mu \in \mathcal{M}(\mathbb{D}),$$

is equicontinuous in the $\beta_0$-topology on $\mathbb{C}_b(\mathbb{D})$, i.e., if, for every $\varepsilon > 0$, there exists a $\beta_0$-neighbourhood $V$ in $\mathbb{C}_b(\mathbb{D})$ such that $|u_\mu(f)| < \varepsilon$, for all $(f, \mu) \in V \times M$. The topology induced on $\mathcal{M}(\mathbb{D})$ under the pairing (2.2.10) is \emph{weak} topology. We refer to Section 2.4 and [sentilles] for details.

**Background**

The classical Riesz representation theorem is stated as a Banach space result for bounded continuous functions vanishing at infinity on a locally compact space. The strict topology $\beta_0$, introduced in [buck] to locally compact spaces, gives up the Banach space structure, but allows to relax the assumption that the bounded continuous functions are vanishing at infinity. Further observations in the 70's by Giles [giles] and Hoffman-Jorgensen [hoffmanjorgensen] lead to a generalization of the Riesz representation theorem for completely regular spaces; locally compact spaces are completely regular. A streamlined proof for the Riesz representation theorem (2.2.10) can be found, e.g., in the book of Jarchow [jarchow]. The proof relies on the fact that on a completely regular space every continuous function admits a unique continuous extension to the Stone-Čech compactification of the space. The fact that the underlying topological space is completely regular ($T_{3\frac{1}{2}}$) is also necessary for the Riesz representation theorem in the sense that the separation axiom cannot be relaxed to a weaker one as there exists examples of regular ($T_3$) spaces on which every continuous function is a constant and on such space the Riesz representation theorem cannot be true; see [herrlich]. However, in our setting it suffices to assume that the space is regular; see Subsection 2.4.1.

**2.3 Main results and examples**

A stochastic process is regarded as a probability measure on the canonical space and the family of all probability measures (processes) on the canonical space is endowed with the weak* topology of the Riesz representation theorem 2.2.3. We impose the following assumption on the canonical space.

**Assumption 2.3.1.** The Skorokhod space is endowed with a regular topology that is weaker than Jakubowski’s $S$-topology but stronger than the Meyer-Zheng topology (MZ).

The $S^*$-topology, introduced in Section 2.5, meets the previous requirements and is arguably the strongest topology on the Skorokhod space for which the results are true; see Theorem 2.5.5.

**2.3.1 Main results**

The following Theorem 2.3.2 is our main result. The statement regarding sequential compactness in Theorem 2.3.2 refines the results of [meyerzheng], [stricker] and [jakubowski] for semimartingale measures, i.e., for semimartingales on the canonical space. The statement about (non-sequential) compactness is, to the best of our knowledge, a new result.
2.3. MAIN RESULTS AND EXAMPLES

**Theorem 2.3.2.** Let \( S \) be a family of semimartingale measures satisfying the condition (UT). Under Assumption 2.3.1, the set \([S]_{\text{seq}}\) is a (sequentially) weak* compact set of semimartingale measures.

**Proof.** The condition (UT) is stronger than the condition (US*); cf. (2.4.17). So, by Proposition 2.4.5, the family \( S \) is \( \beta_0 \)-equicontinuous. Thus, by Corollary 2.4.6, the closure of \( S \) is compact and sequentially compact in the weak* topology; see Proposition 2.4.3. By Corollary 2.4.4, the closure of \( S \) coincides with the sequential closure of \( S \). It remains to show that, every element in the sequential closure \([S]_{\text{seq}}\) is a semimartingale measure. This is done in Proposition 2.4.9.

**Corollary 2.3.3.** Let \( Q \) be a family of quasimartingale measures satisfying the condition (UB). Under Assumption 2.3.1, the set \([Q]_{\text{seq}}\) is a (sequentially) weak* compact set of quasimartingale measures.

**Proof.** The condition (UB) is stronger than the condition (UT); see (2.2.4). Thus, we may invoke Theorem 2.3.2 and it suffices to show that every element in \([Q]_{\text{seq}}\) is a quasimartingale measure. This is done in Proposition 2.4.10.

**Corollary 2.3.4.** Let \( M \) be a set of supermartingale measures satisfying the condition (UI). Under Assumption 2.3.1, the set \([M]_{\text{seq}}\) is a (sequentially) weak* compact set of supermartingale measures.

**Proof.** The condition (UI) is stronger than the condition (UB); see (2.2.4). Thus, we may invoke Corollary 2.3.3 and it suffices to show that every element in \([M]_{\text{seq}}\) is a supermartingale measure. This is done in Proposition 2.4.11.

2.3.2 Examples

The following Example 2.3.5, essentially [touzi][Lemma 3.7], was our original motivation to study weak* compactness in the present setting.

**Example 2.3.5.** Let \( M^u \) be the family of uniformly integrable (\( L^1 \)-bounded) martingale measures and let \( P \) be a (sequentially) weak* compact subset of \( P(\mathbb{R}^d) \). Then the set
\[
M^u_P = \{ Q \in M^u : Q \circ \pi^{-1} \in P \}
\]
is (sequentially) weak* compact.

**Proof.** We adapt the proof of [touzi][Lemma 3.7]. For \( a > 0 \), we have
\[
E_Q[|X_t|\{|X_t|\geq a\}] \leq 2E_Q[|X_t| - a/2] + \leq 2E_Q[(X_T - a/2)^+]
\]
uniformly over \((t, Q) \in I \times M^u_P\), and
\[
E_Q[(H \cdot X)_t] = 0,
\]
for every (elementary) predictable \(|H| \leq 1\), for every \( t \in I \), for every \( Q \in M^u \). Thus, by Prokhorov’s theorem, see e.g. [bogachev][Theorem 8.6.2.], the family \( M^u_P \) satisfies the condition (UI); cf. [jacodshiryaev][IX, Lemma 1.11]. By Example 2.5.12 (b), the evaluation mapping is (sequentially) continuous at the terminal time, so, we have \( M^u_P = [M^u_P]_{\text{seq}} \). A measure \( Q \) is a martingale measure for \( X \) on \( \mathbb{D}(I; \mathbb{R}^d) \) if and only if \( Q \) is a supermartingale measure for \( X^i \) and \(-X^i\), for every \( i \leq d \), so, by Corollary 2.3.4, the set \( M^u_P \) is (sequentially) weak* compact.
Example 2.3.6. Let $\mathcal{M}^p$ denote the family of $L^p$-bounded martingale measures. Then the sets
$$\mathcal{M}_r^p := \{Q \in \mathcal{M}^p : \|X\|_{L^p,\infty}(Q) \leq r\}, \quad r \in \mathbb{R}_+,$$
are (sequentially) weak* compact, for $1 < p < \infty$.

Proof. An increasing continuous function $y \mapsto y^p$ composed with a lower semicontinuous function $y = \|\omega\|_{\infty}$ is lower semicontinuous, see Lemma 4.0.23, and non-negative, so, by [bogachev][Pro. 8.9.8.], the functional $\|X\|_{L^p,\infty}(Q)$ is lower semicontinuous in the weak* topology. Thus, the set $\mathcal{M}_r^p$ is (sequentially) weak* closed, for $r > 0$ and $p > 1$. The $L^p$-boundedness, for $p > 1$, implies that the set $\mathcal{M}_r^p$ satisfies the condition (UI), for $r > 0$ and $p > 1$, so, the set $\mathcal{M}_r^p$ is (sequentially) weak* compact, for $r > 0$ and $p > 1$; cf. Example 2.3.5. \[\square\]

Assume that a probability measure $Q$ is fixed and $p > 1$. Then the Hardy space of $L^p(Q)$-bounded (equivalence classes of indistinguishable) martingales $\mathcal{H}^p(Q) := H^p(\mathbb{D}(I;\mathbb{R}), \mathcal{F}_T, (\mathcal{F}_r)_{r \in I}, Q)$ can be identified with the Lebesgue measure space $L^p(\mathbb{D}(I;\mathbb{R}), \mathcal{F}_T, Q)$. The space $L^p(Q)$ is a reflexive Banach space and the (sequential) weak* compactness of the sets (2.3.1) follows from the Banach-Alaoglu theorem in conjunction with the Eberlein-Šmulian theorem; see [stricker]. The Dunford-Pettis theorem states that a uniformly integrable subset of a non-reflexive Banach space $L^1(Q)$ is relatively sequentially weakly compact, but the random variables of $L^1(Q)$ are not in one-to-one correspondence neither with the family of $L^{1,\infty}(Q)$-bounded, nor $L^1(Q)$-bounded martingales, for $I = [0, \infty[$.

Example 2.3.7. Let $\mathcal{H}^p$ denote the family of $\mathcal{H}^p$-semimartingale measures. Then the sets
$$\mathcal{S}_r^p := \{Q \in \mathcal{H}^p : \|X\|_{L^p,\infty}(Q) + \|X\|_{\mathcal{E}^p}(Q) \leq r\}, \quad r \in \mathbb{R}_+,$$
are (sequentially) weak* compact, for $1 \leq p < \infty$.

Proof. The sets $\mathcal{S}_r^p$, $r > 0$, $p \geq 1$, satisfy the condition (UB), so, by Corollary 2.3.3, the sets $[\mathcal{S}_r^p]_{seq}$ are (sequentially) weak* compact sets of quasimartingales. Moreover, for any sequence $(Q_n)_{n \in \mathbb{N}}$ in $\mathcal{S}_r^p$ converging in the weak* topology to some $Q$, we have
$$\|X\|_{L^p,\infty}(Q) \leq \lim\inf_{n \to \infty} \|X\|_{L^p,\infty}(Q_n) < \infty,$$
(2.3.3)
for $p \geq 1$; cf. Example 2.3.6. Thus, we have $[\mathcal{S}_r^p]_{seq} \subset \mathcal{H}^1; \text{ cf. [dellacheriemeyer][B.VII.98].}$ To show that $\mathcal{S}_r^p = [\mathcal{S}_r^p]_{seq}$ and $[\mathcal{S}_r^p]_{seq} \subset \mathcal{H}^p$, we introduce an auxiliary class $A$ of smooth elementary integrands of the form
$$A^i = \sum_{j=1}^{k} A^i_{t^i_{j-1}} \varphi_{t^i_{j-1},t^i_{j}}, \quad i \leq d,$$
(2.3.4)
where $k \in \mathbb{N}$, $0 = t^0_0 \leq t^1_1 \leq \cdots \leq t^i_k$ in $I$ and each $A^i_{t^i_{j}}$ is continuous $\mathcal{F}^i_{t^i_{j}}$-measurable function satisfying $|A^i_{t^i_{j}}| \leq 1$ and each $\varphi_{t^i_{j-1},t^i_{j}}$ is a smooth function on $I$ vanishing outside $|t^i_{j-1},t^i_{j} + \varepsilon^i_j|$, for some $\varepsilon^i_j \in (t^i_{j},t^i_{j+1}]$, and satisfies $|\varphi_{t^i_{j-1},t^i_{j}}| \leq 1$; we allow $0 \neq \varphi_{t^i_{k-1},t^i_{k}}(t^i_k) \in [-1,1]$ for $t^i_k = T$; cf. (4.0.4). Now, let $Q \in [\mathcal{S}_r^p]_{seq}$ and assume that we are given an elementary predictable integrand $H = (H^i)_{i=1}^d$, i.e., an element of $\mathcal{H}(Q)$, see (2.2.1), such that each $H^i_{t^i_j}$ in (2.2.1) is $\mathcal{F}^i_{t^i_j}$-measurable. By [dellacheriemeyer][A.IV.69 (c)], the
domain of \( F_{i,j}^o \)-measurable functions is homeomorphic to a closed subset of \( \mathbb{D}(I; \mathbb{R}^d) \); cf. Corollary 4.0.22. Thus, by Lusin’s theorem in conjunction with Tietze’s extension theorem, for every \( F_{i,j}^o \)-measurable \( |H_{i,j}^t| \leq 1 \), there exists a sequence of continuous \( F_{i,j}^o \)-measurable functions \((A_{i,j}^n)_{n \in \mathbb{N}}, \ N = \{A_{i,j}^n \}_i \leq 1 \), such that \( A_{i,j}^n \to H_{i,j}^t \) \( Q \)-a.s., as \( n \to \infty \); see e.g. [feldman] and [engelking][2.1.8.]. Moreover, càglàd step functions can be approximated from right with smooth functions (and vice versa), so, there exists a sequence \((A^n)_{n \in \mathbb{N}}, A^n = (A^n_i)_{i=1}^d \), of elements of \( \mathcal{A} \) such that \( \|A^n\|_V \leq \|H\|_V \) \( Q \)-a.s., for all \( n \in \mathbb{N} \), and \((A^n \circ X)_T \to (H \circ X)_T \) \( Q \)-a.s., as \( n \to \infty \). Integrating by parts, we get

\[
|A^n \circ X|_T \leq (|X_T A^n_T| + \|X\|_\infty \|A^n\|_V) \leq c\|X\|_\infty, \ A_0^n = 0, \ n \in \mathbb{N}, \tag{2.3.5}
\]

where \( c := 2\|H\|_V < \infty \) \( Q \)-a.s. and \( \|X\|_\infty \in L^p(Q) \), by (2.3.3). Thus, by the dominated convergence, for any \( Q \in [S^p_{\text{seq}}] \), we have

\[
\lim_{n \to \infty} \|A^n \circ X\|_{L^p(Q)} = \|(H \circ X)_T\|_{L^p(Q)}.
\]

The elements of \( \mathcal{A} \) can similarly be approximated with the elements of \( \mathcal{H}(Q) \). Moreover, due to the uniform bound (2.3.5), for any sequence of integrands bounded in total variation, by the right-continuity \( X = (X^1, X^2, \ldots, X^d) \), the \( F_{i,j}^o \)-measurability of the random variables \( H_{i,j}^t \) can be relaxed to \( F_{i,j}^o \)-measurability, and further, to \( F_{i,j}^o \)-measurability; cf. (2.4.16). Thus, for any \( Q \in [S^p_{\text{seq}}] \), we have

\[
\|X\|_{E^p(Q)} = \|X\|_{A^p(Q)} := \sup_{A \in \mathcal{A}} \|(A \circ X)_T\|_{L^p(Q)}.
\]

Now, since each \( A^i \) of \( A = (A^i)_{i=1}^d \) in \( \mathcal{A} \) is continuously differentiable in \( t \), for every \( \omega \in \mathbb{D}(I; \mathbb{R}^d) \), the function \(|(A \circ X)_T|_p \) is continuous, see (4.0.4), and non-negative, so, by Proposition 2.4.3 and [bogachev][Proposition 8.9.8.], the functional \( \|(A \circ X)_T\|_{L^p(Q)} \) is (sequentially) weak* lower semicontinuous on \( S^p_{\text{seq}} \), for every \( A \in \mathcal{A} \). Consequently, the functional \( \|X\|_{A^p(Q)} \) is (sequentially) weak* lower semicontinuous on \( S^p_{\text{seq}} \), which in conjunction with the weak* lower semicontinuity (2.3.3) of the functional \( \|X\|_{L^p,\infty(Q)} \), yields the (sequentially) weak* closedness of the sets \( S^p_{\text{seq}} \) in \( \mathcal{H}^p \). Indeed, for any \( r > 0 \) and \( p \geq 1 \), for any sequence \((Q_n)_{n \in \mathbb{N}} \) in \( S^p_{\text{seq}} \), converging in the weak* topology to some \( Q \), we have

\[
\|X\|_{L^p,\infty(Q)} + \|X\|_{E^p(Q)} = \|X\|_{L^p,\infty(Q)} + \|X\|_{A^p(Q)} \leq \liminf_{n \to \infty} \left( \|X\|_{L^p,\infty(Q_n)} + \|X\|_{A^p(Q_n)} \right) = \liminf_{n \to \infty} \left( \|X\|_{L^p,\infty(Q_n)} + \|X\|_{E^p(Q_n)} \right) \leq r,
\]

i.e., \( Q \in S^p_{\text{seq}} \). Thus, we have \( S^p_{\text{seq}} = [S^p_{\text{seq}}] \subset \mathcal{H}^p \), i.e., the sets \( S^p_{\text{seq}} \) are weak* compact, for \( r > 0 \) and \( p \geq 1 \). Every element in \( S^p_{\text{seq}} \) is indeed an \( \mathcal{H}^p \)-semimartingale measure; cf. (2.3.6)-(2.3.7).

The pseudonorm in Example 2.3.7, given by the sum of the \( L^p,\infty \)-norm and the Emery pseudonorm

\[
\|X\|_{E^p(Q)} := \sup_{H \in \mathcal{H}(Q)} \|(H \circ X)_T\|_{L^p(Q)}, \ p \geq 1, \tag{2.3.6}
\]
is equivalent to the (maximal) $\mathcal{H}^p$-norm
\[
\|X\|_{\mathcal{H}^p(Q)} := \|M\|_\infty + \|A\|_{L^p(Q)}, \quad p \geq 1,
\]
where $X = M + A$, $A_0 = 0$, denotes the canonical semimartingale decomposition of $X$ under $Q$; see [dellacheriemeyer][B.VII.104] and [dellacheriemeyer][B, p.305]. Assume that a probability measure $Q$ is fixed and $p > 1$. Then the Hardy space of $\mathcal{H}^p(Q)$-bounded (equivalence classes of indistinguishable) semimartingales $\mathcal{H}^p(Q) := \mathcal{H}^p(D(I; \mathbb{R}), \mathcal{F}_t, (\mathcal{F}_t)_{t \in I}, Q)$ is a Banach space; see [hewangyan][p.292]. In particular, for martingales, the norm $\|X\|_{\mathcal{H}^p(Q)}$ is equivalent to the norm $\|X\|_{L^p, \infty(Q)}$, and, as mentioned in the context of Example 2.3.6, there is an analogous Banach pairing $\mathcal{H}^p(Q)' = (L^p(Q))' = L^q(Q) = \mathcal{H}^q(Q)$, for $p, q > 1$, $1/p + 1/q = 1$; see [dellacheriemeyer][B, p.253] and [hewangyan][p.281].

2.4 Auxiliary results for weak* topology

In Subsection 2.4.1, we establish three basic results for the weak* topology that we used in the proof of Theorem 2.3.2. The required stability and tightness results for the weak* topology are covered in Subsection 2.4.2.

2.4.1 Weak* topology

The results of this subsection are established under the assumption that the space $\mathbb{D}$ is a regular Souslin space satisfying Property 2.5.6. Under the assumption, we obtain a stronger separation axiom than the required $T_{31/2}$; cf. Section 2.2.3. Indeed, combining the fact that the topological space is regular ($T_3$) with the fact that the space is a Souslin space, it follows, from a result of Fernique [fernique][Proposition I.6.1], that the space $\mathbb{D}$ is perfectly normal ($T_6$).

The families $\mathcal{M}_t(\mathbb{D})$, $\mathcal{M}_\tau(\mathbb{D})$ and $\mathcal{M}_\sigma(\mathbb{D})$ are defined in Section 2.1.

**Proposition 2.4.1.** The following characterize the dual space in the pairing (2.2.10).

(a) We have that
\[
\mathcal{M}_t(\mathbb{D}) = \mathcal{M}_\tau(\mathbb{D}) = \mathcal{M}_\sigma(\mathbb{D}).
\]

(b) The dual of $\mathcal{C}_b(\mathbb{D})$ can be identified with the class of measures (2.4.1) on (the universal completion of) the canonical $\sigma$-algebra under the bilinear form (2.2.10).

**Proof.** (a) Every Lusin space is a Radon space; see e.g. [schwartz][p.122]. Thus, we have $\mathcal{M}_\sigma(\mathbb{D}) \subset \mathcal{M}_t(\mathbb{D})$. The equality (2.4.1) follows from the fact that the inclusions $\mathcal{M}_t(\mathbb{D}) \subset \mathcal{M}_\tau(\mathbb{D})$ and $\mathcal{M}_\tau(\mathbb{D}) \subset \mathcal{M}_\sigma(\mathbb{D})$ are true for an arbitrary topological space; see [bogachev][Proposition 7.2.2.].

(b) For every $\mu \in \mathcal{P}(\mathcal{B}(\mathbb{D}))$, there exists a unique $\tilde{\mu} \in \mathcal{P}(\tilde{\mathcal{B}}(\mathbb{D}))$ such that
\[
\int f d\mu = \int f d\tilde{\mu}, \quad \forall f \in \mathcal{C}_b(\mathbb{D}).
\]

Since any measure of finite variation is a linear combination of two probability measures, it suffices to observe that the mapping $\mu \mapsto \tilde{\mu}$ is a bijection; see e.g. [dellacheriemeyer][A, 32 (c) (i)]. The statement then follows from (a) in conjunction with [jarchow][Theorem 7.6.3.].
A probability measure $Q$ on $(\Omega, \mathcal{F})$ is called perfect, if, for every real-valued $\mathcal{F}$-measurable function $f$ on $\Omega$, the set $f(\Omega)$ is measurable with respect to $Q \circ f^{-1}$.

**Corollary 2.4.2.** Every probability measure $Q \in \mathcal{P}(\mathbb{D})$ is perfect.

**Proof.** This follows from Proposition 2.4.1 (a); see [bogachev][Theorem 7.5.10 (i)]. □

We use the equality (2.4.1) without mentioning it when we apply the results from the book of Bogachev [bogachev].

**An analogue of the Eberlein-Šmulian theorem**

A Hausdorff topological space $X$ is called angelic if every set $S \subset X$ with the property that every infinite sequence of its elements has a limit point in $X$, possesses also the following properties: $S$ is relatively compact and each point in the closure of $S$ is the limit of some sequence in $S$. In angelic spaces, the properties of compactness and sequential compactness coincide. In addition, the closure of a relatively compact set is exhausted by the limits of sequences of points in this set. In particular, every metric space is angelic. It is also that if a regular space $X$ can be continuously injected into an angelic space $Y$, then $X$ is also angelic; see e.g. [bogachev][p. 218]. It is also known that, for a completely regular space $X$, the weak$^*$ topology on $\mathcal{M}_+(\mathbb{D})$ is metrizable if and only if the underlying topology on $X$ is metrizable; see e.g. [bogachev][p. 213].

Combining the previous facts, we obtain the following analogue of the Eberlein-Šmulian theorem on the positive orthant of the dual.

**Proposition 2.4.3.** For $M \subset \mathcal{M}_+(\mathbb{D})$, the following are equivalent:

(i) Every infinite sequence in $M$ has a weak$^*$ convergent subsequence in $\mathcal{M}_+(\mathbb{D})$,

(ii) The weak$^*$ closure of $M$ is weak$^*$ compact in $\mathcal{M}(\mathbb{D})$.

Moreover, under these conditions, the weak$^*$ closure of $M$ is metrizable.

**Proof.** The underlying topological space is a completely regular Souslin space, so, it admits a continuous injective mapping to a metric space. Thus, by [bogachev][Theorem 8.10.4.], the weak$^*$ closure of a subset of $\mathcal{M}_+(\mathbb{D})$ satisfying (i) or (ii) is a compact metrizable subspace of $\mathcal{M}(\mathbb{D})$, so, (i) and (ii) are equivalent for $M$. □

It is immediate from Proposition 2.4.3 that weak$^*$ compactness and sequential weak$^*$ compactness are equivalent for the subsets of $\mathcal{M}_+(\mathbb{D})$. In fact, a stronger statement is true.

**Corollary 2.4.4.** Assume that $M$ is a subset of $\mathcal{M}_+(\mathbb{D})$ that satisfies the equivalent conditions of Proposition 2.4.3. Then the sequential weak$^*$ closure of $M$ in $\mathcal{M}_+(\mathbb{D})$, i.e., the set

$$[M]_{\text{seq}} = \{ \mu \in \mathcal{M}_+(\mathbb{D}) : \exists (\mu_n)_{n \in \mathbb{N}} \subset M \text{ s.t. } \mu_n \to_{w^*} \mu \},$$

is weak$^*$ closed.

**Proof.** By Proposition 2.4.3, the closure of $M$ endowed with the relative topology is a first countable space. In particular, the space is a Fréchet-Urysohn space. By [engelking][Theorem 1.6.14.], the sequential closure $[M]_{\text{seq}}$ coincides with the closure of $M$. □

Various criteria that guarantee tightness and stability of a family of processes are not preserved in the weak$^*$ convergence, so, the previous results are crucial for constructing weak$^*$ compact sets of stochastic processes.
Prokhorov’s theorem
A measure $\mu \in \mathcal{M}(\mathbb{D})$ is called tight, if there exists an exhausting net of compact sets $(K^\varepsilon)_{\varepsilon > 0}$ for $\mu$, i.e., $|\mu|(\mathbb{D} \setminus K^\varepsilon) < \varepsilon$, for every $\varepsilon > 0$, where $|\mu|$ is the total variation of $\mu$. A subset $M$ of $\mathcal{M}(\mathbb{D})$ is called uniformly tight, if there exists a net of compact sets $(K^\varepsilon)_{\varepsilon > 0}$ which is uniformly exhausting for $M$, i.e., $\sup_{\mu \in M} |\mu|(\mathbb{D} \setminus K^\varepsilon) < \varepsilon$, for every $\varepsilon > 0$.

**Proposition 2.4.5.** A subset $M$ of $\mathcal{M}(\mathbb{D})$ is $\beta_0$-equicontinuous if and only if it is bounded in total variation and uniformly tight, and we have

(a) If $M$ is a $\beta_0$-equicontinuous subset of $\mathcal{M}(\mathbb{D})$, then $M$ is relatively (sequentially) compact in the weak* topology.

Moreover, we have the following useful convergence criteria.

(b) If $(\mu_n)_{n \in \mathbb{N}}$ is an uniformly tight sequence in $\mathcal{M}(\mathbb{D})$ converging in the weak* topology to $\mu \in \mathcal{M}(\mathbb{D})$, then, for any $f \in C(\mathbb{D})$ satisfying

$$\lim_{c \to \infty} \sup_n \int |f|1_{\{|f| \geq c\}} d\mu_n = 0,$$

one has $\int f d\mu_n \to \int f d\mu$.

**Proof.** The underlying topological space is completely regular, and the characterization follows directly from [sentilles][Theorem 5.1]. The compact subsets of a completely regular Souslin space are metrizable, cf. Proposition 2.4.3 and [bogachev][Lemma 8.9.2.], which, in conjunction with the fact the space is completely regular, verifies the assumptions for the sequential and non-sequential Prokhorov’s theorem (a); see [bogachev][Theorem 8.6.7.]. The convergence criteria (b) is similarly a direct consequence of the fact that the underlying space is completely regular; see [bogachevgaussian][Lemma 3.8.7.].

The characterization of the relative compactness in terms the $\beta_0$-equicontinuity thus yields a criteria for compactness of closures (of convex (circled) hulls).

**Corollary 2.4.6.** The closed convex circled hull of a $\beta_0$-equicontinuous subset of $\mathcal{M}(\mathbb{D})$ is $\beta_0$-equicontinuous and (sequentially) weak* compact. In particular, the closure and the closed convex hull of a $\beta_0$-equicontinuous set are (sequentially) weak* compact.

**Proof.** The weak* compactness of the closed convex circled hull of an equicontinuous set follows from [kelleyamioka][18.5]. Closure and closed convex hull are closed subsets of closed convex circled hull, from which the second statement follows. The $\beta_0$-equicontinuous sets are bounded in total variation, in particular, from below, so, the sequential statements are true, by Proposition 2.4.3.

**Skorokhod’s representation theorem**

By Jakubowski’s property, Property 2.5.6, we have the following variant of Skorokhod’s representation theorem.

**Proposition 2.4.7.** Let $(Q_n)_{n \in \mathbb{N}}$ be a sequence converging in the weak* topology to $Q$ in $\mathcal{P}(\mathbb{D})$. Then there exists a subsequence $(Q_{n_k})_{k \in \mathbb{N}}$, a probability space $(\Omega, \mathcal{F}, P)$ and $\mathbb{D}$-valued random variables $(Y_k)_{k \in \mathbb{N}}$ and $Y$ on $(\Omega, \mathcal{F}, P)$ such that $Q_{n_k} = P \circ (Y_k)^{-1}$, $k \in \mathbb{N}$, $Q = P \circ Y^{-1}$ and

$$f(Y_k(\omega)) \to f(Y(\omega)), \, \forall \omega \in \Omega, \, \forall f \in C_b(\mathbb{D}).$$

(2.4.2)
2.4. AUXILIARY RESULTS FOR WEAK* TOPOLOGY

Proof. The Euclidean space endowed with its usual inner product is a Hilbert space, so, by [jakubowski2][Theorem 1], the existence of an a.s. convergent subsequence follows from Property 2.5.6. The convergence in [jakubowski2][Theorem 1] is the pointwise convergence in the topology of the underlying space, which, for a sequence in a completely regular space, is equivalent to the convergence (2.4.2); cf. (2.5.3). Moreover, modifying the \( \mathbb{D} \)-valued random variables \( Y_k \) and \( Y \), given by [jakubowski2][Theorem 1], on a set of measure zero does not affect on their weak* convergence, so, their almost sure convergence can be strengthened to the pointwise convergence. \( \Box \)

In particular, by Proposition 2.4.7, every element of \( \mathcal{P}(\mathbb{D}) \) can be regarded as a law of some \( \mathbb{D} \)-valued random variable. Complementarily, any such random variable induces a probability measure on \( \mathbb{D} \).

2.4.2 Stability and tightness

In this subsection, we cover the required stability and tightness results. In particular, we provide the required multi-dimensional infinite horizon extensions of the stability results of , for the right-continuous version of the raw canonical filtration.

Stability

Under Assumption 2.3.1, it suffices to establish the required stability results for the Meyer-Zheng topology; see Appendix 4.0.3. The required stability results are classical and thoroughly studied in [meyerzheng], [jakubowski6] and [jakubowski] for scalar-valued processes. We demonstrate that, after some slight modifications, they are true in the present setting. We utilize the following multi-dimensional extension of [meyerzheng][Theorem 5], provided by Jakubowski’s subsequential Skorokhod’s representation theorem.

Lemma 2.4.8. If \( (Q_n)_{n \in \mathbb{N}} \) is a sequence converging in the weak* topology to \( Q \) in \( \mathcal{P}(\mathbb{D}(I; \mathbb{R}^d)) \), then there exists a subsequence \( (Q_{n_k})_{k \in \mathbb{N}} \) and a set \( L \subset I \) of full Lebesgue measure such that \( T \in L \), if \( I = [0, T] \), and

\[
Q_n \circ \pi_F^{-1} \rightarrow_{w^*} Q \circ \pi_F^{-1}, \quad \text{as } n \rightarrow \infty,
\]

(2.4.3)

for every finite subset \( F \) of \( L \). In particular, there exists a (countable) dense set \( D \subset I \) such that \( T \in D \), if \( I = [0, T] \), and (2.4.3) is true, for every finite subset \( F \) of \( D \).

Proof. By Proposition 2.4.7, we find a subsequence \( (Q_{n_k}) \), \( k \in \mathbb{N} \), and \( \mathbb{D} \)-valued random variables \( (Y_k)_{k \in \mathbb{N}} \) and \( Y \) on some \( (\Omega, \mathcal{F}, P) \) such that \( Q_{n_k} = P \circ Y_k^{-1} \), \( k \in \mathbb{N} \), \( Q = P \circ Y^{-1} \) and \( Y_k(\omega) \rightarrow_{MZ} Y(\omega) \), for every \( \omega \in \Omega \), as \( k \rightarrow \infty \); since the topology \( MZ \) is metrizable, (2.4.2) is equivalent to \( \rightarrow_{MZ} \). By Lemma 4.0.21, there exists a subsequence \( Y_{m_k}(\omega) \), \( m \in \mathbb{N} \), and a set \( L \) of full Lebesgue such that \( T \in L \), if \( I = [0, T] \), and \( Y_{m_k}(t, \omega) \rightarrow Y_t(\omega) \), for every \( (t, \omega) \in L \times \Omega \), as \( m \rightarrow \infty \). Hence, the finite dimensional distributions of the process \( (Y_{m_k,t})_{t \in L} \) converge to those of \( (Y_t)_{t \in L} \). The complement of the set \( L \) is a \( \lambda \)-null set. Thus, the set \( L \) contains a (countable) dense set \( D \) such that \( T \in D \), if \( I = [0, T] \); cf. Definition 4.0.19. \( \Box \)

In Proposition 2.4.9, we show that the required part of [jakubowski6][Theorem 2.1], which is an extension of [stricker2][Theorem 2] for a right-continuous canonical filtration, is true on a multi-dimensional Skorokhod space.
Proposition 2.4.9. Let \((Q_n)_{n \in \mathbb{N}}\) be a sequence of semimartingale measures satisfying the condition (UT) and converging in the weak* topology to \(Q\). Then the weak* limit \(Q\) is a semimartingale measure.

Proof. The proof is essentially a combination of [jakubowski6][Lemma 1.1 and 1.3] By Lemma 2.4.8, there exists a subsequence \((Q_{n_k})_{k \in \mathbb{N}}\) and a countable dense set \(D \subset I\) such that \(T \in D\), if \(I = [0, T]\), and \((Q_{n_k})_{k \in \mathbb{N}}\) converges to \(Q\) in finite dimensional distributions on the set \(D\). For every finite collection \(t_1 < \cdots < t_j \in D\), let \(\mathcal{A}_{t_1, \ldots, t_j}\) denote the family of continuity sets of the marginal law of \(Q\) on \(t_1 < \cdots < t_j\), i.e., \(\mathcal{A}_{t_1, \ldots, t_j}\) consists of Borel sets \(B \in \otimes_{i \leq j} \mathcal{B}(\mathbb{R}^d)\) for which \(Q \circ \pi_{t_1}^{-1} \circ \cdots \circ \pi_{t_j}^{-1}\) (\(\partial B\)) = 0, where \(\partial B\) denotes the (Euclidean) topological boundary of \(B\) on \(\mathbb{R}^{d \times j}\). Following [jakubowski6], we introduce an auxiliary class \(\mathcal{J}(D)\) of integrands, determined by the weak* limit \(Q\) and the dense set \(D\), that take the form

\[
\mathcal{J} = \sum_{i} \sum_{k=1}^{d} J_{i,k}^{-1} t_{i-1}^{-1} t_{k}^{-1}, \quad n(i) \in \mathbb{N}, \ i \leq d, \tag{2.4.4}
\]

where every \(t_{i}^{-1} t_{k}^{-1} < \cdots < t_{n}^{-1} t_{j}^{-1}\) is a finite collection of elements of \(D\) and every \(J_{i,k}^{-1}\) is a finite linear combination of indicator functions of the continuity sets of the marginal law of \(Q\) on \(s_1 < \cdots < s_j \leq t_{k}^{-1} s_k\), for every \(s_k \in D\), embedded on \(\mathbb{D}(I; \mathbb{R}^d)\) and bounded by 1 in absolute value, i.e., \(|J_{i,k}^{-1}| \leq 1\) and each \(J_{i,k}^{-1}\) is of the form

\[
J_{i,k}^{-1} = \sum_{l=1}^{p} \alpha^l_{i} \pi_{s_1}^{-1} \cdots \pi_{s_j}^{-1} A_l, \quad s_j \leq t_{k-1}^{-1}, \quad \alpha^l \in \mathbb{R}, \quad \forall_l \in A_{s_1, \ldots, s_j},
\]

for some elements \(s_1 < \cdots < s_j\) of \(D\) and \(j, p\) finite, for every \(i \leq d, k \leq n\). Now, since \((Q_{n_k})_{k \in \mathbb{N}}\) is converging to \(Q\) in finite dimensional distributions on the set \(D\), by the vectorial Portmanteau’s lemma, see e.g. [vaart][Lemma 2.2], we have

\[
Q(|(J \circ X)| > c) = \liminf_{k \to \infty} Q_{n_k} \circ \pi^{-1}_{D}(|(J \circ X)| > c)
\leq \liminf_{k \to \infty} Q_{n_k} \circ \pi^{-1}_{D}(|(J \circ X)| > c)
\leq \liminf_{k \to \infty} Q_{n_k} (|(J \circ X)| > c), \quad c > 0, \quad t \in I, \tag{2.4.4}
\]

where \(J \in \mathcal{J}(D) \subset H(Q_{n_k}), \) for all \(k \in \mathbb{N}\). Due the condition (UT), for every \(t \in I,\) the term on the last line, in (2.4.4), tends to zero, uniformly over \(\mathcal{J}(D)\), as \(c \to \infty\), i.e., the family \(\{(J \circ X)_t : J \in \mathcal{J}(D)\}\) is \(Q\)-tight, i.e. bounded in probability \(Q\), for every \(t \in I\). The topology of the convergence in probability is metrizable, so, for every \(t \in I,\) the sets, that are contained in the (sequential) closure of \(\{(J \circ X)_t : J \in \mathcal{J}(D)\}\), are bounded, which, in particular, entails that the set \(\{H \circ X)_t : H \in H(Q)\}\) is bounded in probability \(Q\). Indeed, we will show this, by adapting a sequence of approximation arguments from [jakubowski6]. First, since \(D\) is dense in \(I\) and contains \(T,\) for every \(t_0 < t_1 < \cdots < t_n\) in \(I,\) \(n \in \mathbb{N},\) there exists \(t_{i}^{-1} t_{k}^{-1} \leq \cdots \leq t_{i}^{-1} t_{n}^{-1}\) in \(D, k \leq \mathbb{N},\) such that \(t_j < t_{k}^{-1} t_{j}^{-1}\), for every \(j \leq n,\) for every \(k \geq 1,\) and \(t_j \downarrow t_{j}^{-1}\), for every \(j \leq n,\) as \(k \to \infty;\) we allow \(t_{k}^{-1} = T,\) if \(t_j = T.\) Since \(d\) and \(n\) are finite, by the right-continuity of \(X = (X^1, \ldots, X^d),\) we have

\[
X_{i,k} \to X_{i,j}, \quad \text{uniformly over } i = 1, \ldots, d \text{ and } j = 1, \ldots, n, \quad \text{as } k \to \infty. \tag{2.4.5}
\]

Secondly, for every \(i \leq d,\) for every \(j \leq n,\) for every \(t_j < T,\) any \(\mathcal{F}_{t}\)-measurable \(|H_{i,j}^{t} | \leq 1\) is \(\mathcal{F}_{t}^{o} \)-measurable, for all \(k \geq 1,\) and can therefore be expressed as an uniform limit of
2.4. AUXILIARY RESULTS FOR WEAK* TOPOLOGY

simple $\mathcal{F}^o_{t_j} \sigma$-measurable functions bounded by 1 in absolute value, for every $k \geq 1$, i.e., for every $i \leq d$, for every $j \leq n$, for every $k \geq 1$, there exists functions $|S^i_{t_j}| \leq 1$, $l \in \mathbb{N}$, such that each $S^i_{t_j}$ is of the form

$$ S^i_{t_j} = \sum_{h=1}^{q(l)} \beta^{i,j,k}_{h} 1_{F^{h,i}_{i,j,k}}, \quad \beta^{i,j,k}_{h} \in \mathbb{R}, \quad F^{h,i}_{i,j,k} \in \mathcal{F}^o_{t_j}, \quad 1 \leq q(l) < \infty, \quad (2.4.6) $$

and we have

$$ \|H^n_{t_j} - S^i_{t_j}\| \rightarrow 0 \text{ as } l \rightarrow \infty. \quad (2.4.7) $$

Further, since each $\mathcal{A}_{t_1,\ldots,t_j}$ is an algebra generating $\otimes_{i \leq j} \mathcal{B}(\mathbb{R}^d)$ on $\mathbb{R}^{d \times j}$, and the finite unions of the cylindrical sets $\pi_{t_1,\ldots,t_j}^{-1}(\otimes_{i \leq j} \mathcal{B}(\mathbb{R}^d))$ form an algebra generating the canonical $\sigma$-algebra on $\mathbb{D}(I; \mathbb{R}^d)$, for every $0 < t \in I$, the family

$$ \mathcal{A}^o_{t_j} = \left\{ \bigcup_{k=1}^{n} \pi_{t_1,\ldots,t_j(k)}^{-1} (A_k) : A_k \in \mathcal{A}_{t_1,\ldots,t_j(k)}, \quad t_1 < \cdots < t_j(k) < t, \quad j(k), n \in \mathbb{N} \right\} $$

is an algebra generating $\mathcal{F}^o_{t_j} = \sigma(X_u : u < t)$ on $\mathbb{D}(I; \mathbb{R}^d)$; cf. Corollary 4.0.22. Thus, for every $F^{h,i}_{i,j,k} \in \mathcal{F}^o_{t_j}$ in (2.4.6), there exists a sequence $(A_{h,i,j,k}^{l,m})_{m \in \mathbb{N}}$ in $\mathcal{A}^o_{t_j}$ such that

$$ \lim_{m \rightarrow \infty} 1_{A^{h,i,j,k}_{l,m}} \overset{Q}{\rightarrow} 1_{F^{h,i}_{i,j,k}}, \quad \text{as } m \rightarrow \infty. \quad (2.4.8) $$

Finally, by combining the approximations (2.4.5) and (3.6.15), and in (3.6.15), invoking the approximation (2.4.8) in the sums (2.4.6), we conclude that, for every $H \in \mathcal{H}(Q)$, there exists a sequence $(J_n)_{n \in \mathbb{N}}$, $n = n(k,l,m)$, of elements in $\mathcal{J}(Q)$ such that, for every $t \in I$, we have

$$ (J_n \circ X)_t \rightarrow Q \sum_{i=1}^{d} (H^i \circ X^i)_t = (H \circ X)_t, \quad \text{as } k \land l \land m \rightarrow \infty. $$

Thus, the family of simple integrals $\{(H \circ X)_t : H \in \mathcal{H}(Q)\}$ is contained in the closure of $\{(J \circ X)_t : J \in \mathcal{J}(D)\}$, for every $t \in I$, and, by (2.4.4), the weak* limit $Q$ is an $(\mathcal{F}_t)_{t \in I}$-quasimartingale measure - and, consequently, an $(\mathcal{F}^o_t)_{t \in I}$-quasimartingale measure; see e.g. [protter][Theorem II.4].

The following Proposition 2.4.10 is essentially [meyerzheng][Theorem 4].

**Proposition 2.4.10.** Let $(Q_n)_{n \in \mathbb{N}}$ be a sequence of quasimartingale measures satisfying the condition (UB) and converging in the weak* topology to $Q$. Then the weak* limit $Q$ is a quasimartingale measure.

**Proof.** Let $i \leq d$ be fixed. We adapt the proof of [meyerzheng][Theorem 4] and show that the coordinate process $X^i$ is a quasimartingale under $Q$ on $\mathbb{D}(I; \mathbb{R}^d)$. Using the convention

$$ \frac{1}{\varepsilon} \int_0^\varepsilon |X^i_{T+u}| \, du := |X^i_T|, \quad \text{if } I = [0,T], \quad (2.4.9) $$


we have

\[ E_Q \left[ \frac{1}{\varepsilon} \int_0^\varepsilon |X_{t+u}|^2 du \right] \leq \liminf_{n \to \infty} E_{Q_n} \left[ \frac{1}{\varepsilon} \int_0^\varepsilon |X_{t+u}^i|^2 du \right] \leq b_i^i, \]

for every \( t \in I \), for every \( \varepsilon > 0 \), where \( b_i^i := \liminf_{n \to \infty} \sup_{t \in I} E_{Q_n} ||X_t^i|| < \infty \), by the condition (UB). Thus, by Fatou’s lemma, we get

\[ E_Q[|X_t^i|] \leq \liminf_{\varepsilon \to 0} E_Q \left[ \frac{1}{\varepsilon} \int_0^\varepsilon |X_{s+u}^i| du \right] < \infty, \quad (2.4.10) \]

for every \( t \in I \). The truncated coordinate process

\[ X^{i,c} := (-c) \lor (X^i \land c) = (-c) \lor (X^i - (X^i - c)^+), \quad c > 0, \]

is a difference of two convex 1-Lipschitz functions of \( X^i \), so, we have

\[ \text{Var}_t^Q(X^{i,c}) \leq 4 \text{Var}_t^Q(X^i), \quad n \in \mathbb{N}, \quad t \in I; \quad (2.4.11) \]

see e.g. [Stricker4]. Let \( 0 = t_0 < t_1 < \cdots < t_k = t \) and \( |f_j| \leq 1, \quad j < k, \) be continuous \( \mathcal{F}_{t_j}^o \)-measurable functions. By (2.4.11), we have

\[ E_{Q_n} \left[ \sum_{j=1}^k f_{j-1}(X)(X_{u+t_j}^{i,c} - X_{u+t_{j-1}}^{i,c}) \right] \leq 4 \text{Var}_t^Q(X^i), \quad n \in \mathbb{N}, \]

so, by Fubini’s theorem, we get

\[ E_{Q_n} \left[ \frac{1}{\varepsilon} \int_0^\varepsilon \left( \sum_{j=1}^k f_{j-1}(X)(X_{u+t_j}^{i,c} - X_{u+t_{j-1}}^{i,c}) \right) du \right] \leq 4 \text{Var}_t^Q(X^i), \quad n \in \mathbb{N}, \quad \varepsilon > 0; \]

cf. (2.4.9). The function \( F(X) := \frac{1}{2} \int_0^\varepsilon \left( \sum_{j=1}^k f_{j-1}(X)(X_{u+t_j}^{i,c} - X_{u+t_{j-1}}^{i,c}) \right) du \) is lower semicontinuous and bounded from below, see (4.0.4) and Lemma 4.0.23, so, we have

\[ E_Q \left[ \frac{1}{\varepsilon} \int_0^\varepsilon \left( \sum_{j=1}^k f_{j-1}(X)(X_{u+t_j}^{i,c} - X_{u+t_{j-1}}^{i,c}) \right) du \right] \leq 4v^i, \quad \varepsilon > 0, \quad (2.4.13) \]

where \( v^i := \liminf_{n \to \infty} \sup_{t \in I} \text{Var}_t^Q(X^i) < \infty \), by the assumption (UB). Due to (2.4.10), letting \( \varepsilon \to 0 \) and then \( c \to \infty \) in (2.4.13), by the right-continuity and the monotone convergence, respectively, we get

\[ E_Q \left[ \sum_{j=1}^k f_{j-1}(X)(X_{t_j}^i - X_{t_{j-1}}^i) \right] \leq 4v^i, \quad (2.4.14) \]

for all \( \mathcal{F}_{t_j}^o \)-measurable continuous functions \( |f_j| \leq 1, \quad j < k \). Furthermore, by choosing \( f_j(X) = f(X_{t_j+u}) \) preceding (2.4.12), for a continuous function \( |f| \leq 1 \) on \( \mathbb{R}^d \), we conclude that the inequality (2.4.14) is true for a family of continuous functions that, for every \( j < k \), generates the \( \sigma \)-algebra \( \mathcal{F}_{t_j}^o \); cf. Corollary 4.0.22. Thus, by the standard \( L^1 \)-approximation via Lusin’s theorem and Tietze’s extension theorem, for any \( \mathcal{F}_{t_j}^o \)-measurable \( |H_{t_j}| \leq 1 \) in \( L^\infty(Q) \), for every \( j < k \), there exists a sequence \( (f_j^m)_{m \in \mathbb{N}}, |f_j^m| \leq 1 \), of functions satisfying (2.4.14) such that \( f_j^m \to H_{t_j} \) in \( L^1(Q) \), as \( n \to \infty \); see e.g. [feldman] and
2.4. AUXILIARY RESULTS FOR WEAK* TOPOLOGY

[engelking][2.1.8]. Thus, the inequality (2.4.14) is true for all $F_0^o$, measurable functions $|H_{t_j}| \leq 1$ in $L^\infty(Q)$, so, the process $X$ is an $(F_\tau^o)_{\tau \in I}$-quasimartingale on $(\mathbb{D}(I), F_T, Q)$, see [dellacheriemeyer][B. App. II (3.5)], which, by Rao's decomposition theorem, is a necessary and sufficient condition for the process $X$ to be decomposable to a difference $X = Y - Z$ of two càdlàg $(F_\tau^o)_{\tau \in I}$-supermartingales $Y$ and $Z$ on $(\mathbb{D}(I), F_T, Q)$; see [hewangyan][Theorem 8.13]. On the other hand, by Föllmer's lemma, $Y$ and $Z$ are $(F_\tau)_{\tau \in I}$-supermartingales, see [hewangyan][Theorem 2.46], so, by Rao's decomposition theorem, the process $X$ is an $(F_\tau)_{\tau \in I}$-quasimartingale on $(\mathbb{D}(I), F_T, Q)$.

The following Proposition 2.4.11 is essentially [meyerzheng][Theorem 11].

Proposition 2.4.11. Let $(Q_n)_{n \in \mathbb{N}}$ be a sequence of supermartingale measures satisfying the condition (UI) and converging in the weak* topology to $Q$. Then the weak* limit $Q$ is a supermartingale measure.

Proof. We adopt the proof of [meyerzheng][Theorem 11] and show that each coordinate process $X^i$, $i \leq d$, is a supermartingale under $Q$ on $(\mathbb{D}(I; \mathbb{R}^d))$. We have $E_Q[|X_t|] < \infty$, for every $t \in I; \text{ cf. } (2.4.10)$. Moreover, by Lemma 2.4.8, there exists a subsequence $(Q_{n_k})_{k \in \mathbb{N}}$ and a countable dense set $D \subset I$ such that $T \subset D$, if $I = [0, T]$, and $(Q_{n_k})_{k \in \mathbb{N}}$ converges to $Q$ in finite dimensional distributions on the set $D$. Let $X^{i,c}$ denote the coordinate process $X^i$ truncated from above at $c > 0$, i.e.,

$$X^{i,c} := X^i \land c, \quad c > 0.$$

By the condition (UI) and the fact that each $Q_{n_k}$ is a supermartingale measure for $X^{i,c}$, for every $c > 0$, we have

$$E_Q[f(X)(X^{i,c}_t - X^{i,c}_s)] \leq \text{lim inf } \kappa \rightarrow \infty E_Q_{n_k}[f(X)(X^{i,c}_t - X^{i,c}_s)] \leq 0, \quad s < t, \quad s, t \in D,$$

where

$$f(X) := f_1(X_{t_1})f_2(X_{t_2})\cdots f_n(X_{t_n}), \quad t_j \in D, \quad f_j \in \mathbb{C}_b(\mathbb{R}^d), \quad j \leq n;$$

see e.g. [vaart][Theorem 2.20]. Consequently, by Corollary 4.0.22, we have

$$E_Q[1_F(X)(X^{i,c}_t - X^{i,c}_s)] \leq 0, \quad c > 0,$$

for every $s < t$ in $D$ and $F \in F^o_\tau$. Letting $c \rightarrow \infty$ in (2.4.15), by the monotone convergence, we get the same inequality for the coordinate process $X^i$. By Föllmer's lemma, the inequality extends immediately to the whole $I$, and further, for $F \in F^\circ_\tau$, indeed, we have

$$E_Q[1_F(X)(X^i_t - X^i_s)] = \lim_{n \rightarrow \infty} E_Q[1_F(X)(X^i_t - X^i_{s+1/n})] \leq 0,$$

for every $F \in F^\circ_\tau$; cf. [hewangyan][Theorem 2.44].

Tightness

We say that a family $Q$ of probability measures on $(\mathbb{D}(I; \mathbb{R}^d), F_T)$ satisfies Jakubowski's uniform tightness criteria, if we have

$$\lim_{c \rightarrow \infty} \sup_{Q \in \mathcal{Q}} Q(||X^{i,t}||_\infty > c) = 0 \quad \text{and} \quad \lim_{c \rightarrow \infty} \sup_{Q \in \mathcal{Q}} Q(N^{a,b}(X^{i,t}) > c) = 0, \quad \forall a < b,$$

(US)
for every finite \( t \in I \), for every \( i \leq d \), where \( X^{i,t} \) denotes the coordinate process \( X^i \) restricted on \([0,t]\); cf. Corollary 2.5.11. It was shown in [jakubowski] that a family of probability measures on \((\mathbb{D}([0,T];\mathbb{R}),\mathcal{F}_T), T < \infty\), satisfies the condition (US) if and only if it is uniformly \( S \)-tight. In particular, we have the following hierarchy, cf. (2.2.4),

\[
\text{(UT)} \implies \text{(US)} \implies \text{(US*)},
\]

(2.4.17)

where \( \text{(US*)} \) stands for the uniform tightness in the \( S^* \)-topology; see Section 2.5.

The second implication in (2.4.17) is immediate from the definition of the \( S^* \)-topology; see Proposition 3.3.1 (i). The first implication in (2.4.17) follows from Proposition 2.4.12, that is essentially the result of Stricker [stricker][Theorem 2], which states that a sequence satisfying the condition (UT) admits a convergent subsequence and the limit law is a law of a semimartingale. Analogous results were obtained for the \( S \)-topology by Jakubowski in [jakubowski][Theorem 4.1]; see also [jakubowski][Proposition 3.1].

**Proposition 2.4.12.** A family of semimartingale measures satisfying the condition (UT) satisfies the condition (US).

**Proof.** Let \( X^{i,t} \) denote the coordinate processes \( X^i \) restricted on \([0,t]\), for \( i \leq d \) and \( t < \infty \). Following [stricker][Theorem 2], we define a family of stopping times

\[
\tau^{i,c} = \inf\{s \in I : |X^{i,t}_s| > c\}, \ i \leq d, \ c > 0,
\]

and each \( \tau^{i,c} \) is approximated from right with the sequence of the stopping times

\[
\tau^{i,c}_n = \min\{m/n : m \in \mathbb{N}, \ \tau^{i,c} \leq m/n\}, \ n \in \mathbb{N}.
\]

(2.4.18)

Since we are assuming a right-continuous filtration \((\mathcal{F}_t)_{t \in I}\) and \( X^i \) is right-continuous, the hitting times \( \tau^{i,c} \), and consequently, their approximations \( \tau^{i,c}_n \) are indeed stopping times. Moreover, since each \( \tau^{i,c}_n \) takes only finitely many values on \([0,t]\), every process \(|H^n| \leq 1\) of the form

\[
H^n = 1_{[0,\tau^{i,c}_n \wedge t]}, \ i \leq d, \ c > 0, \ n \in \mathbb{N}, \ t \in I,
\]

(2.4.19)

is an elementary predictable integrand; see (2.2.1). Now, due to the right-continuity of \( X^i \), by the bounded convergence, for every \( Q \in \mathcal{Q} \), we have

\[
Q \left( \|X^{i,t}\|_\infty > c \right) = Q \left( (|H^n \circ X^i|)_t > c \ \forall n \in \mathbb{N} \right),
\]

(2.4.20)

for every \( t \in I \), for every \( c > 0 \). By the condition (UT), the left-hand side in (2.4.20) tends to 0, uniformly over \( Q \in \mathcal{Q} \), for every \( i \leq d \), for every \( t \in I \), as \( c \to \infty \). Similarly, for \( a < b \), we define, recursively, for \( k \in \mathbb{N}_0 \), the stopping times

\[
\sigma^{i,a}_k = \inf\{s > \tau^{i,b}_k : |X^{i,t}_s| < a\}, \ \tau^{i,b}_k = \inf\{s > \sigma^{i,a}_k : |X^{i,t}_s| > b\}, \ \sigma^{i,a}_0 = \tau^{i,b}_0 = 0,
\]

and the respective decreasing sequences \((\sigma^{i,a}_k)_{k \in \mathbb{N}}\) and \((\tau^{i,b}_k)_{k \in \mathbb{N}}\) of approximative stopping times, taking only finitely many values on finite intervals; cf. (2.4.18). The processes \(|H^{m,n}| \leq 1\), \( m, n \in \mathbb{N} \), of the form

\[
H^{m,n} = \sum_{k=1}^m 1_{[\sigma^{i,a}_k \wedge t, \tau^{i,b}_k \wedge t]},
\]

for every finite \( t \in I \), for every \( i \leq d \), where \( X^{i,t} \) denotes the coordinate process \( X^i \) restricted on \([0,t]\); cf. Corollary 2.5.11. It was shown in [jakubowski] that a family of probability measures on \((\mathbb{D}([0,T];\mathbb{R}),\mathcal{F}_T), T < \infty\), satisfies the condition (US) if and only if it is uniformly \( S \)-tight. In particular, we have the following hierarchy, cf. (2.2.4),

\[
\text{(UT)} \implies \text{(US)} \implies \text{(US*)},
\]

(2.4.17)
2.5. WEAK S TOPOLOGY

are finite linear combinations of processes of the form (2.4.19), so, each process $|H^{m,n}| \leq 1$ is an elementary predictable integrand. Moreover, we have

$$Q \left( N^{a,b}(X^i) > c \right) \leq Q \left( \lim_{m \to \infty} |(H^{m,n} \circ X^i)_t| > a^+ + c(b-a) \ \forall n \in \mathbb{N} \right). \quad (2.4.21)$$

By the condition (UT), for every $a < b$, the right-hand side of (2.4.21) tends to zero, uniformly over $Q \in Q$, as $c \to 0$; cf. (2.4.4)-(2.4.5). Thus, by Corollary 2.5.11, the family $Q$ satisfies the condition (US).

\[ \square \]

2.5 Weak S topology

We introduce the notion of weak S-topology and study its properties and relation to other topologies on the Skorokhod space.

2.5.1 Definition

A possibility of defining a completely regular (non-sequential) S-topology is discussed already in [jakubowski]; see the page 18 therein. We describe a general method for regularizing any given topology. Our approach is inspired by [alexandroff]. Let $X = (X, \tau)$ be an arbitrary topological space and $V$ a base for the Euclidean topology on $\mathbb{R}$, then the family

$$\{ f^{-1}(V) : f \in C_b(X), \ V \in \mathcal{V} \} \quad (2.5.1)$$

is a subbase for a topology on $X$; see e.g. [jerison][3.4]. The topology is generated by the family of pseudometrics

$$\{ \rho_{f_1,f_2,\ldots,f_k} : f_1, f_2, \ldots, f_k \in C_b(X) \}, \quad (2.5.2)$$

where

$$\rho_{f_1,f_2,\ldots,f_k}(x,y) := \max \{|f_1(x) - f_1(y)|, |f_2(x) - f_2(y)|, \ldots, |f_k(x) - f_k(y)|\}$$

for $x, y \in X$, and thus, the convergence of a net $(x_\alpha)$ to $x$ in this topology is equivalent to that

$$f(x_\alpha) \to f(x), \ \forall f \in C_b(X); \quad (2.5.3)$$

see e.g. [engelking][Example 8.1.19]. We remark that by replacing $C_b(X)$ with $C(X)$ in (2.5.1), (2.5.2) and (2.5.3) one obtains an equivalent characterization. Any of these characterizations is necessary and sufficient criterion for a topological space to be completely regular ($T_{3\frac{1}{2}}$); see e.g. [jerison][3.4].

Definition 2.5.1. We will denote by $S^*$ the topology generated on the Skorokhod space by the family (2.5.1) of $S$-continuous functions, and call it weak S-topology.

Remark 2.5.2. The convergence in the weak* topology on the $\beta_0$-dual of $C_b$, cf. (2.2.10), traditionally called the 'weak convergence' for sequences of probability measures, is equivalent to the convergence (2.5.3), if the measures are Dirac measures; see [bogachev][Lemma 8.9.2.].

Remark 2.5.3. It should be emphasized that, if one could show that the $S$-topology is regular (or linear), then the $S$- and the weak $S$-topology would coincide. It was communicated to the author by Professor Jakubowski that the regularity of $S$-topology remains as an open question.
2.5.2 Relation to other topologies

The definitions of Jakubowski’s $S$-topology, the Meyer-Zheng topology ($MZ$) and Skorokhod’s $J^1$-topology are given in Appendix 4.0.3.

Proposition 2.5.4. We have the following hierarchy.

$$ MZ \subset S^* \subset S \subset J^1. $$

Proof. The functions in (4.0.4) and (4.0.5), that generate the topology $MZ$, are $S^*$-continuous; see Example 2.5.12. Moreover, the topology $MZ$ is metrizable, in particular, sequential and completely regular. Thus, by Example 2.5.12, the first inclusion $MZ \subset S^*$ is true; cf. (2.5.3). The second inclusion $S^* \subseteq S$ is obvious from (2.5.1). The final inclusion $S \subset J^1$ is proved, for a finite compact interval, already in [jakubowski], and extends immediately for the infinite interval due to (4.0.8); cf. (4.0.3).

The Skorokhod space endowed with the $J^1$-topology is a Polish space, so, the space is a Lusin space for any topology that is weaker than the $J^1$-topology. The following Theorem 2.5.5 states that the $S^*$-topology, which is the strongest (completely) regular topology that is weaker than the $S$-topology, is the strongest (completely) regular Souslin topology on the Skorokhod space for which the sets (2.5.7) are compact, and consequently, Jakubowski’s uniform tightness criteria (US) is a sufficient tightness criteria; cf. Subsection 2.4.2.

Theorem 2.5.5. Let $T$ be a completely regular Souslin topology on the Skorokhod space, comparable to $S$, and $K(T) = K(S)$. Then

$$ T \subset S. $$

Proof. Assume that $S \subset T$ and let $T_s$ denote the sequential topology generated by $T$. Since the compact sets of a completely regular Souslin space are metrizable, we have $K(T) \subset K(T_s)$; see e.g. [bogachev][p. 218]. Consequently, we have

$$ K(S) = K(T) = K(T_s), \quad (2.5.4) $$

where

$$ S \subset T \subset T_s, \quad (2.5.5) $$

and $S$ and $T_s$ are sequential; see Appendix 4.0.3. By [engelking][Theorem 3.3.20.], the Skorokhod space is a (Hausdorff) $k$-space for $S$ and $T_s$, so, by (2.5.4) and (2.5.5), we have $S = T$. \qed

2.5.3 Properties

Consider the following property of a topological space $X = (X, \tau)$, extensively studied in [jakubowski5].

Property 2.5.6. There exists a countable family of real-valued $\tau$-continuous functions $f_k, k \in \mathbb{N}$, such that, for all $x, y \in X$, we have

$$ f_k(x) = f_k(y), \ \forall k \in \mathbb{N} \implies x = y. $$
Jakubowski’s fundamental observation was that Property 2.5.6 yields a subsequential Skorokhod representation theorem; cf. Subsection 2.4.1. In fact, all key properties of the $S$-topology follow immediately from Property 2.5.6 and the property is preserved in the regularization (2.5.1).

**Proposition 2.5.7.** The $S$-topology has the following properties:

(a) $S$ is Hausdorff,

(b) Each $K \in \mathcal{K}(S)$ is metrizable,

(c) A set is compact if and only if it is sequentially compact,

(d) The Borel $\sigma$-algebra $\mathcal{B}(S)$ and the canonical $\sigma$-algebra coincide,

(e) The Skorokhod space endowed with $S$ is a Lusin space.

The $S^*$-topology has the properties (a)-(e) and additionally:

(f) The Skorokhod space endowed with $S^*$ is perfectly normal and paracompact,

(g) The Borel $\sigma$-algebra $\mathcal{B}(S^*)$ and the Baire $\sigma$-algebra $\mathcal{B}_{a}(S^*)$ coincide,

(h) $C(S) = C(S^*)$,

(i) $\mathcal{K}(S) = \mathcal{K}(S^*)$.

**Proof.** The properties (a), (b) and (c) follow immediately from the fact that the (weak) $S$-topology satisfies Property 2.5.6; see [jakubowski5][pages 10-11]. Indeed, the mappings

$$\omega \mapsto \frac{1}{r} \int_{q}^{q+r} \omega^i(t)dt, \text{ and } \omega \mapsto \omega^i(T), \text{ for } I = [0,T],$$

where $q$ and $q + r$ run over the rationals in $I$ and $i$ over the spatial dimensions $1, \ldots, d$, constitute a countable family of continuous functions that separates the Skorokhod space; cf. Example 2.5.12.

(d) We prove the claim for $I = [0,T]$. The proof is completely similar for $I = [0,\infty[$. Fix a coordinate $i \leq d$. For all $0 \leq t < T$, we have

$$\omega^i(t) = \lim_{\delta \to 0} \frac{1}{\delta} \int_{t}^{t+\delta} \omega^i(u)du,$$

i.e., the mapping $\omega \mapsto \omega^i(t)$ is a limit of elements of $C(S)$, for every $t$ in $I$, while for $t = T$, the mapping $\omega \mapsto \omega^i(t)$ is an element of $C(S)$. Consequently, we have $\sigma(X_u : u \in I) \subset \mathcal{B}(S^*)$ and since $S^*$ is weaker than $S$, we have $\mathcal{B}(S^*) \subset \mathcal{B}(S)$. On the other hand, by Proposition 2.5.4, $S$ is weaker than $J^1$, so, we have $\mathcal{B}(S) \subset \mathcal{B}(J^1)$, where $\mathcal{B}(J^1) = \sigma(X_u : u \in I)$. Thus, $\mathcal{B}(S^*) = \mathcal{B}(S) = \sigma(X_u : u \in I)$. By Proposition 2.5.4, we have $S^* \subset S \subset J^1$ and the Skorokhod space endowed with $J^1$ is a Polish space, so, the Skorokhod space endowed with $S$ or $S^*$ is a Lusin space. Thus, we have (e).

The Skorokhod space endowed with $S^*$ is a (completely) regular Souslin space. By the result of Fernique, every regular Souslin space is perfectly normal and paracompact; see [fernique][Proposition I.6.1]. Thus, we have (f). Now, by (f), the Skorokhod space endowed with $S^*$ is perfectly normal, and consequently, by [bogachev][Proposition 6.3.4],
we have $\mathcal{B}(S^*) = \mathcal{B}_a(S^*)$, i.e., (g) is true. The claim (h) follows directly from Definition 2.5.1.

To prove (i), we first observe that $\mathcal{K}(S) \subset \mathcal{K}(S^*)$, by Definition 2.5.1. To prove the converse inclusion we use the topology $\Sigma$; see Appendix 4.0.3. By [jakubowski4][Remark 3.6], we have $\Sigma \subset S$, so, we have $C(\Sigma) \subset C(S) = C(S^*)$. Thus, we have that $\Sigma \subset S^*$ since the topology $\Sigma$ is completely regular. Indeed, topological vector spaces are completely regular; see e.g. [bogachevtvs][Theorem 1.6.5]. Consequently, by [jakubowski4][Remark 3.8], we get $\mathcal{K}(S^*) \subset \mathcal{K}(\Sigma) = \mathcal{K}(S)$. Thus, we have shown $\mathcal{K}(S) = \mathcal{K}(S^*)$. □

Remark 2.5.8. A countable product of regular Souslin spaces is a regular Souslin space. Thus, by the result of Fernique [fernique][Proposition I.6.1], the previous properties (after the obvious modifications) are inherited for a countable product topology; cf. Subsection 2.4.1.

2.5.4 Compact sets and continuous functions

In this subsection, we recall the compactness and continuity criteria for the $S$-topology from [jakubowski] and [jakubowski4].

Compactness criteria

The necessity and sufficiency of the condition (2.5.7) for the relative (sequential) compactness in the $S$-topology was proved in [jakubowski], for $I = [0, T]$, $T < \infty$, and the multi-dimensional infinite horizon extension was provided in [jakubowski4].

Proposition 2.5.9. A subset $K$ of $\mathbb{D}([0, T]; \mathbb{R})$, $T < \infty$, is relatively sequentially $S$-compact if and only if the following conditions are satisfied:

$$\left\{ \begin{array}{l}
\sup_{\omega \in K} \|\omega\|_\infty < \infty, \\
\sup_{\omega \in K} N^{a,b}(\omega) < \infty, \ \forall a < b, \ a, b \in \mathbb{R}.
\end{array} \right. \tag{2.5.7}$$

Proposition 2.5.10. A subset $K$ of $\mathbb{D}([0, \infty]; \mathbb{R})$ is relatively sequentially $S$-compact if and only if the set $K$ restricted on $[0, t]$ satisfies the conditions (2.5.7) for every $0 < t < \infty$.

We make the following observations.

1. For any two real numbers $a < b$ one can find rationals $r < q$ such that $a < r < q < b$, so, it is sufficient to let $a < b$ range rationals in Proposition 2.5.9.

2. The mappings of $\omega$ in Proposition 2.5.9 are (sequentially) lower semicontinuous in the $S^*$-topology, so, their lower level sets are (sequentially) closed in the $S^*$-topology; cf. Example 2.5.13.

3. A Cartesian product set in a multi-dimensional Skorokhod space is relatively sequentially $S$-compact if and only if each set in the product is relatively sequentially $S$-compact; cf. Definition 4.0.12.

4. $S$-compact set are $S^*$-compact; cf. Proposition 3.3.1 (i).

Combining the previous facts we obtain the following compactness criteria.
Corollary 2.5.11. Let $K = K^1 \times \cdots \times K^d$ be a Cartesian product set on the Skorokhod space $D(I; \mathbb{R}^d)$ endowed with $S$ or $S^*$. Then the set $K$ is compact, if, for each $i \leq d$, there exists a (non-decreasing) function $C^i_{q,r} : I \rightarrow \mathbb{R}_+$, for all $q < r$ in $\mathbb{Q}$, and a (non-decreasing) function $M^i : I \rightarrow \mathbb{R}_+$ such that

$$K^i := \bigcap_{q < r} \{ \omega^i : N^q_{q,r} \left( \lfloor \omega^i \rfloor t \right) \leq C^i_{q,r}(t) \text{ and } \| \lfloor \omega^i \rfloor t \|_\infty \leq M^i(t) \ \forall t < \infty \},$$

where the intersection is taken over all rationals $q < r$ in $\mathbb{Q}$ and $\lfloor \omega^i \rfloor t$ denotes the restriction of $\omega^i$ on $[0,t]$.

Remark that, for $I = [0,T], T < \infty$, it suffices to consider constant $C^i_{q,r}$ and $M^i$ in Corollary 2.5.11.

Examples of (semi-)continuous functions

By Proposition 3.3.1 (h), $S^*$-continuous functions are precisely the $S$-continuous ones. In particular, we would like to emphasize that the evaluation mapping at $t$ is not continuous for any $t < T$; see [jakubowski][p.11].

Example 2.5.12. (a) The following mappings are $S^*$-continuous on $D(I; \mathbb{R}^d)$

$$\omega \mapsto \int_I G(t, \omega^i(t))d\mu(t), \ i \leq d,$$

whenever $G$ is measurable as a mapping of $(t, x)$, continuous as a mapping of $x$, for every $t \in I$, and such that

$$\sup_{|t| \leq c} \sup_{|x| \leq c} |G(t,x)| < \infty, \ \forall c > 0, \quad (2.5.8)$$

and $\mu$ is an atomless measure on $I$; see [jakubowski][Corollary 2.11].

(b) The mapping

$$\omega \mapsto \omega(T)$$

is $S^*$-continuous on $D([0,T]; \mathbb{R}^d)$; see [jakubowski][Remark 2.4].

By Proposition 2.5.4, the $S^*$-topology is stronger than the Meyer-Zheng topology ($MZ$), so, the uniform norm and the number of upcrossings of an interval $[a,b]$ are (sequentially) lower semicontinuous; see Lemma 4.0.23.

Example 2.5.13. The mappings

$$\omega \mapsto \| \omega \|_\infty \text{ and } \omega \mapsto N^{a,b}(\omega^i), \ a < b, \ i \leq d,$$

are (sequentially) lower semicontinuous in the $S^*$-topology on $D(I; \mathbb{R}^d)$. 
Part II

Martingale Optimal Transport
Duality
Chapter 3

Martingale Optimal Transport Duality

Chapter 3 is based on the joint work [martdual].

3.1 The superhedging principle

As an example, consider the following Asian type derivative contract with payoff

$$\xi_a(X) = \frac{1}{T} \int_0^T X_t \, dt,$$

where $X = (X_t)_{t \in [0,T]}$ is a path modelling the discounted prices of underlying assets that can consist of common stocks or liquid options. The fair initial price for $\xi_a$ is $X_0$ because the payoff $\xi_a(X)$ can be replicated with initial capital $X_0$. Indeed, one can re-express the payoff as

$$\xi_a(X) = \frac{1}{T} \int_0^T X_t \, dt = X_T - \int_0^T \frac{t}{T} \, dX_t.$$

The integral on the right-hand side is well-defined pathwise as a Lebesgue-Stieltjes integral and has zero initial cost and holding one share of the underlying yields $X_T$ at maturity. Therefore, the contract $\xi_a$ can be priced without any probabilistic assumptions on the behaviour of $X$. Next we consider the following option with a given strike price $K$,

$$\xi_c(X) = \left( \frac{1}{T} \int_0^T X_t \, dt - K \right)^+.$$

It is now no longer possible to deduce the price for $\xi_c$ by direct replication as done for $\xi_a$. However, the price of $\xi_c$ should be less than the initial capital requirement of any portfolio that produces a value greater than $\xi_c(X)$ for every possible realization of the underlying asset $X$. Indeed, we can consider all financial instruments that are liquidly priced in the market. Then, it should not be possible to make sure profits by creating portfolios of long and short positions in these options and the underlying assets. A basic problem is then to express these price bounds in terms of market parameters. In the pathwise formulation, the price bound for a payoff $\xi$ is (related to) the maximal expected value over all martingale laws $Q$ for $X$ that is also consistent with the prices observed in the market, i.e.,

$$\mathcal{E}^{\star}_Q(\xi) = \sup_{Q \in \mathcal{Q}^{\star}} E_Q[\xi],$$

37
where $\mathcal{Q}^*$ denotes the set of all such measures. In particular, since the (discounted) price process $X$ has to be a martingale, all elements of $\mathcal{Q}^*$ must be martingale measures and satisfy additional linear constraints.

A ‘derivative’ means that its payoff is derived from the existing assets on the market and fully determined by them. It is then a natural question to ask if the price for a derivative contract can be derived from the prices of the existing assets too - without imposing any assumption on the behavior of the existing assets, and thus affecting the payoff. This is indeed possible using the superhedging principle. We show that the highest price admits a sharp expression in terms of the existing assets on the market expressed by the superhedging functionals.

Price bounds by superhedging is a classical method in mathematical finance. The model-free version of the problem in discrete time is first introduced in [beiglbock]. The pathwise superhedging duality in continuous time is proved in [mete] with continuous paths and extended to the Skorokhod space in [mete2]. Since then both in discrete and continuous time many related problems have been studied. In [touzi], stability and asymptotics of the bid-ask spread are studied under topological assumptions (similar to the first part of this article) and the authors also provided a duality result for a general measurable claim using the abstract dynamic programming principle. In [obloj] and [david2], the authors use the Wiener space endowed with the uniform norm topology and study the connection to the Knightian uncertainty and the classical diffusion models. [mete3] considers a generalization of the problem that takes into account the proportional transaction costs.

Previous results in the Skorokhod space, [mete2, mete3, touzi] assume that the payoff is uniformly continuous in the Skorokhod metric. Whereas the assumption of uniform continuity is a natural first step from discrete to continuous time, relaxing this assumption is crucial (necessary) to our approach. Indeed, uniform continuity (uniformity) is not a topological property, i.e., it is a property that is not preserved in a homeomorphism. Without the assumption of uniform continuity, the approximation methods of [mete2, touzi] are no longer at our disposal. Instead, our approach is based on the Riesz representation theorem and associated convex conjugate duality. Indeed, the pricing-hedging duality is arguably a convex duality result, closely related to model-free arbitrage. We refer to [matteo], [cheridito2] and [ekren] for a clear exposition of this statement in discrete time. In this paper, adapting this view, we show that the pathwise Doob’s up-crossing inequality translates to a (lower semi-)continuity of the superreplication functional in an appropriately chosen topology. Moreover, this topology admits the family of countable additive measures as its topological dual, which, by the Fenchel-Moreau theorem, is the necessary and sufficient condition for the duality gap to be closed. Our goal is thus to close this duality gap.

To achieve this, we endow the Skorokhod space with a weak (though, arguably the strongest possible) topology and derive the dual representation for the minimal superhedging price for continuous payoffs. The novelty of our approach is to provide a duality result for an entire class of continuous payoffs: the topological structure is chosen so that the generalization to upper semicontinuous payoffs is immediate and the further extension to measurable payoffs possible. The first extension to upper semicontinuous functions is a classical minimax argument that appears already e.g. in the seminal paper of Beiglböck et al. [beiglbock]. The second step to measurable claims is the classical Choquet capacity argument of Kellerer [kellerer].

In the main part of the article, we work with a single terminal marginal. However, due
to our approach finitely many marginal constraints are easily incorporated in the model via simple measurable lifting argument. We also show that there is market calibration determined by these marginals is stable (lsc) in the sense of [touzi] from which we obtain the recover the dual attainment results of [cheridito].

The guiding idea is that more general duality requires more hedging positions. Indeed, even for an ill-behaving continuous payoff, the family of simple strategies, or more generally the family of finite variation strategies, is not sufficiently large to close the dual gap. The counter-example, Example 3.7.1, provides an upper semicontinuous claim for which there is duality gap with simple integrands and in Proposition 3.8.1 we show that duality holds for upper semicontinuous payoffs if and only if it only holds for continuous ones. Theorem 3.6.2 shows that the duality is achieved for (upper semi-)continuous functions by the use of the sub-linear integral of Vovk [vovk]. This integral, defined in the Subsection 3.4.2, allows to define gain processes for dynamic strategies beyond the class of simple integrands. However, Example 3.7.2 shows that this extension alone is not sufficient to achieve the duality for all measurable payoffs and further enlargement of the hedging set is needed. This is done by taking an appropriate closure in Subsection 3.6.3 and the duality is proved in Theorem 3.6.4. Therefore, the enlargements of the hedging set arise naturally.

This article is organized as follows. We start by listing the notation used in the paper in Section 3.2. After outlining the market structure in Section 3.3, the general hedging class and the integration via Vovk are defined in Section 3.4. The main pricing-hedging duality results are stated in Section 3.6. The duality for uniformly continuous claims is discussed in Section 3.6.4. Examples and counterexamples are presented in Section 3.7 and further extensions of the duality results are provided in Section 3.8. Section 3.9 and Section 3.10 are devoted for the proofs, in particular, showing the required properties of the superhedging functional. In the Appendix we recall some known auxiliary results.

### 3.2 Notation

We start by listing some frequently used notation.

- $\mathbb{N} = \{1, 2, \ldots \}$ denotes the family of natural numbers and $\mathbb{N}_0 := \mathbb{N} \cup \{0\}$.
- $\mathbb{R}$ (resp. $\mathbb{R}_+$) denotes the family of (non-negative) real numbers.
- $\mathbb{R} := \mathbb{R} \cup \{\pm \infty\}$ denotes the family of extended real numbers.
- $\mathbb{C}([0, T]; \mathbb{R}^d)$ denotes the $\mathbb{R}^d$-valued, continuous paths.
- $\mathbb{D}([0, T]; \mathbb{R}_+^d)$ denotes the $\mathbb{R}_+^d$-valued, càdlàg paths.

Let $\Omega$ be a topological space. All functions are assumed extended real valued.

- $\mathcal{B} := \mathcal{B}(\Omega; \mathbb{R})$ denotes the family of Borel functions.
- $\mathcal{C} := \mathcal{C}(\Omega; \mathbb{R})$ denotes the family of continuous functions.
- $\mathcal{U} := \mathcal{U}(\Omega; \mathbb{R})$ denotes the family of upper semicontinuous functions.

Below, $S$ is a place holder for $\mathbb{C}([0, T]; \mathbb{R}^d)$, $\mathbb{D}([0, T]; \mathbb{R}_+^d)$, $\mathcal{B}$, $\mathcal{C}$, or $\mathcal{U}$.

- $S^u$ denotes the family of uniformly continuous functions of $S$.
- $S_b$ denotes the family of bounded functions of $S$.
- $S_0$ is the family of bounded functions of $S$ vanishing at infinity, i.e.,

$$S_0 := \{f \in S_b : \forall \varepsilon > 0 \exists \ a \text{ compact set } K \ s.t. \ |f(x)| \leq \varepsilon \ \forall \ x \notin K\}.$$

- $S_p := \left\{ \xi \in S : \sup_{\omega \in \Omega} \frac{|\xi(\omega)|}{1+\|\omega\|_\infty} < \infty \right\}$ for $p \geq 1$.\n
\( S_{< p} := \bigcup_{1 \leq q < p} S_q \) with the convention \( S_{< 1} := S_0 \) for \( p = 1 \).
\( | \cdot | \) denotes the Euclidean norm.
\( \| \cdot \|_\infty \) denotes the supremum norm.
\( \mathcal{M} \) is the set of all Radon measures of finite total variation.
\( \mathcal{M}_p := \{ \mu \in \mathcal{M} : \int |f| \, d\mu < \infty \forall f \in \mathbb{B}_p \} \).
\( \mathcal{P} \) denotes the family of probability measures on \( \Omega \).
\( \mathcal{Q}(\mu) := \mathcal{Q} \cap \mathcal{P}(\mu) \).
\( \beta_0 \) is the topology generated by the semi-norms \( \| \cdot \|_\infty, \eta \in \mathbb{B}_0 \).
\( \langle f, \mu \rangle := E^\mu[f] := \int f \, d\mu \) denotes the Radon integral setting
\[ \int f \, d\mu := \infty \quad \text{if} \quad \int f^+ \, d\mu = \int f^- \, d\mu = \infty. \]
For a dual pair of vector spaces \( X, Y \) with pairing \( \langle \cdot , \cdot \rangle \), we recall
\( \mathcal{E}_C(x) := \sup_{y \in C} \langle x, y \rangle, \quad x \in X; \)
\( \partial \Phi_d(x) := \{ y \in Y : \Phi_d(x) + \langle \bar{x} - x, y \rangle \leq \varphi(\bar{x}) \forall \bar{x} \in X \}, \quad x \in X. \)

All measures in this manuscript are Radon.

For two real numbers \( x, y \in \mathbb{R} \), we set
\( x \wedge y := \min\{x, y\}, \)
\( x \vee y := \max\{x, y\}. \)

### 3.3 Structure

In this section, we set the structure used in the paper.

#### 3.3.1 The Skorokhod Space and Jakubowski’s \( S \)-topology

Let \( \mathbb{D}([0, T]; \mathbb{R}^d_+) \) be the set of all non-negative càdlàg paths. Recall that a function \( \omega \) on \([0, T]\) is called càdlàg, if it is right-continuous and has only discontinuities of the first kind. The latter is equivalent to the condition
\[ N^{a,b}(\omega) < \infty \quad \text{for all} \quad a < b \quad \text{and} \quad \|\omega\|_\infty < \infty, \] (3.3.1)
where \( a \) and \( b \) range over rational (or real) numbers, \( \omega \) denotes the restriction of \( \omega \) on \([0, t]\), and \( N^{a,b}(\omega) \) denotes the number of (strict) upcrossings of the interval \([a, b]\) by \( \omega \) on \([0, T]\), i.e., we write \( N^{a,b}(\omega) \geq n \), if one can find \( 0 \leq t_1 \leq \cdots \leq t_{2n} \leq T \) such that \( \omega(t_{2k-1}) < a \) and \( \omega(t_{2k}) > b \), for \( 1 \leq k \leq n \).
3.3. STRUCTURE

The Topology $S$ and $S^*$

The topology $S^*$ is given in Part II. Here we state the key properties that are proved in [jakubowski] and [semimart]; see Section 5 in [semimart].

Proposition 3.3.1. The Skorokhod space endowed with $S^*$ is a perfectly normal Hausdorff space. In particular, it is completely regular; cf. 4.0.4. Moreover, the following statements are true:

(b) The universal completion of the Borel $\sigma$-algebra coincides with the universal completion of the canonical $\sigma$-algebra.

(a) Every measure on the universal completion is a unique extension of a Radon measure on the Borel $\sigma$-algebra.

(c) The sets of martingales measures $Q(\mu)$ are (sequentially) weak$^*$ compact.

The Skorokhod space endowed with the topology $S$ is a Hausdorff space and the statements (a) and (b) are true.

The $S$-topology was introduced by Adam Jakubowski [jakubowski] in the Skorokhod space and in the context of martingale transport context, the topology was first used in [touzi].

A subset $K$ of $D([0,T];\mathbb{R}_d^+)$ is (relatively) $S$-compact if and only if

$$\sup_{\omega \in K} \sup_{i \leq d} N^{a,b}(\omega^i) < \infty \text{ for all } a < b \text{ and } \sup_{\omega \in K} \sup_{i \leq d} \|\omega^i\|_{\infty} < \infty. \quad (3.3.2)$$

We endow $D([0,T];\mathbb{R}_d^+)$ with the topology $S^*$ generated by the family of $S$-continuous functions that is arguably the strongest topology suitable to our approach. In particular, the compactness criteria (3.3.2) remains sufficient. We refer to [semimart] for details. The necessary facts regarding the $S$- and $S^*$-topology are covered in Appendix 3.3.1.

Path Space

We impose the following standing assumption.

The path space $\Omega$ is a closed subset of $D([0,T];\mathbb{R}_d^+)$. \hspace{1cm} (3.3.3)

In particular, the usual canonical spaces satisfy the previous assumption.

Example 3.3.2. The usual discrete and continuous time-models are included.

(a) $\Omega = D([0,T];\mathbb{R}_d^+), \; 0 < T < \infty$.

(b) $\Omega = \{\omega \in D([0,T];\mathbb{R}_d^+) : \omega(t) = \omega(k), \; k \leq t < k+1, \; 0 \leq k \leq T\}, \; T \in \mathbb{N}$.

Remark 3.3.3. The choice of $\Omega$ is not a modeling choice, but determined by the domain (input) of a derivative contract in question. In particular, the choice of $\Omega$ is not a prediction - or a foretelling of any other sort.

Remark 3.3.4. We assume a finite time-horizon, but we want to emphasize that due to the Vovk integral, introduced in Section 3.4, the extension to the infinite horizon case $T = \infty$ are immediate. One simply, invokes the recent infinite extesion of the $S$-topology [jakubowski4] as was done in [semimart]. However, analogously to the classical case, in the infinite discrete time, the Vovk integral is necessary.
For \( t \in [0, T] \), let \( X = (X_t) \) be the canonical process of \( \Omega \), i.e., \( X_t(\omega) = \omega(t) \). In particular, each component of \( X \) is assumed non-negative.

**Remark 3.3.5.** It suffices to assume that each component of \( X \) is bounded from below by a deterministic function.

The process \( X \) models the prices of the underlying (stocks or liquid options). The interest rate is assumed to be deterministic and can therefore be omitted.

We impose our second standing assumption.

The initial value \( X_0 \) is normalized to a unit vector \( 1 \). (3.3.4)

We endow \( \Omega \) with the relative topology; cf. [semimart][Remark 1.1].

**Remark 3.3.6.** Combining the previous assumption (3.3.4) with our earlier assumption (3.3.3) it follows that the mapping \( X_0(\omega) = \omega(0) \) is continuous on \( \Omega \).

We endow \( \Omega \) with the filtration \( \mathbb{F} = (\mathcal{F}_t)_{t \in [0, T]} \) generated by the canonical process \( X \) of \( \Omega \). We assume that the canonical space satisfies the usual conditions, i.e., we assume the universal completion of the right-continuous version of the raw canonical filtration; cf. [semimart][Proposition 4.1 (b)].

### 3.3.2 Càdlàg Peacocks

To avoid discontinuous (measurable) static positions we define a set \( \mathbb{T} \) of **continuity points** of \( \Omega \), i.e., the times \( t \in [0, T] \) for which the evaluation mapping \( X_t(\omega) = \omega(t) \) is continuous. The initial time \( 0 \) and the terminal time \( T \) are always continuity points i.e. elements of \( \mathbb{T} \); see Remark 3.3.6. See also Remark 3.6.5.

Suppose that probability measures \( \mu = (\mu_t)_{t \in \mathbb{T}} \) on \( \mathbb{R}^d_+ \) having finite first moments are given. A probability measure \( P \in \mathcal{P} \) is said to satisfy the marginals \( \mu \), if for each \( t \in \mathbb{T}, P \circ \pi_t^{-1} = \mu_t \). \( \mathcal{P}(\mu) \) is the set of all probability measures satisfying the marginals. A probability measure \( Q \in \mathcal{P} \) is called a **martingale measure** if the canonical process is a martingale. The family of all martingale measures is denoted by \( \mathcal{Q} \). Finally, we set

\[
\mathcal{Q}(\mu) := \mathcal{Q} \cap \mathcal{P}(\mu).
\] (3.3.5)

The set \( \mathcal{Q}(\mu) \) can be written as

\[
\mathcal{Q}(\mu) = \left\{ Q \in \mathcal{P} : X \text{ is a } (Q, \mathbb{F})\text{-martingale and } Q \circ \pi_t^{-1} = \mu_t, \ t \in \mathbb{T} \right\}.
\]

We say that \( \mu \) is a **peacock** if it is increasing in convex order i.e. if

\[
\int_{\mathbb{R}^d} \psi(x) \, d\mu_t(x)
\]

is increasing in time \( t \) for every convex function \( \psi : \mathbb{R}^d \to \mathbb{R} \). A peacock \( \mu \) is said to be càdlàg if it is right-continuous and admits left-limits in the weak topology; see [touzi][Definition 2.1].

**Lemma 3.3.7.** We have \( \mathcal{Q}(\mu) \neq \emptyset \) if and only if \( \mu = (\mu_t)_{t \in \mathbb{T}} \) is a càdlàg peacock.
Proof. The necessity follows by Jensen’s inequality and the sufficiency is a direct consequence of the multidimensional extension of Kellerer’s theorem by Hirsch & Roynette [Hirsch2013][Theorem 3.2]. Indeed, a right-hand limit of continuity points is a continuity point. Therefore, the set $\mathcal{T}$ is closed in the right-limit topology. Thus, $\mu = (\mu_t)_{t \in \mathcal{T}}$ can be extended to a càdlàg peacock on $\mu = (\mu_t)_{t \in [0,T]}$; see [touzi][Remark 2.2. (ii)].

Remark 3.3.8. The previous lemma requires the usual condition that the filtration is right-continuous. Indeed, the final step (5) of the proof of [Hirsch2013][Theorem 3.2] on page [Hirsch2013] is Föllmer’s lemma; see e.g. [hewangyan][Theorem 4.22].

Lemma 3.3.9. (a) There exist a peacock $\mu$ with a marginal at $t \notin \mathcal{T}$ such that $\mathcal{Q}(\mu) = \emptyset$.

(b) There exists a càdlàg peacock with a marginal at $t \notin \mathcal{T}$ such that $\mathcal{Q}(\mu) \neq \partial\Phi_d(0;\mu)$.

Proof. For (a), choose a non-càdlàg peacock; e.g. a càglàd peacock $\mu_s = \delta_{\{1\}}$, for $0 \leq s \leq t$, and $\mu_s = \frac{1}{2}\delta_{\{0\}} + \frac{1}{2}\delta_{\{2\}}$, for $t < s \leq T$, for some $0 < t < T$, will do. Naturally, a càdlàg $X$ of $\mathbb{D}([0,T];\mathbb{R}^+)$ cannot satisfy these marginals. Therefore, we have $\mathcal{Q}(\mu) = \emptyset$.

For (b), choose an appropriate non-closed $\mathcal{P}(\mu)$ in (3.3.5); e.g. $\mu_t = \delta_{\{1\}}$ and $\mu_T = \frac{1}{2}\delta_{\{0\}} + \frac{1}{2}\delta_{\{2\}}$ for $X$ of $\mathbb{D}([0,T];\mathbb{R}^+)$ will do. We find $(Q^n)_{n \in \mathbb{N}}$ in $\mathcal{Q}(\mu)$ converging in the weak$^*$ topology to Consequently, $\mathcal{Q}(\mu)$ is not weak$^*$ closed. Therefore, we have $\mathcal{Q}(\mu) \neq \partial\Phi_d(0;\mu)$.

Remark 3.3.10. Càdlàg peacock were considered in [touzi]; see Subsection 2.1 and Example 3.8 therein.

The Growth Assumption

Throughout the paper we assume that

We are given a càdlàg peacock $\mu = (\mu_t)_{t \in \mathcal{T}}$.

The important structural characteristic of the peacock $\mu = (\mu_t)_{t \in \mathcal{T}}$ is that, for some $p \geq 1$, it satisfies

$$c_T := \int_{\mathbb{R}_+^d} |x|^p \, d\mu_t(x) < \infty.$$  \hspace{1cm} (3.3.6)

The letter $p$ is fixed to denote the largest exponent $p \geq 1$ for which (3.3.6) is true.

Remark 3.3.11. The classical theorems by Strassen and Keller state that the assumption (3.3.6) is necessary, i.e., one cannot relax the assumption to $p < 1$.

Remark that since the power function with exponent $p \geq 1$ is convex, the assumption (3.3.6) implies that

$$E_Q [|X_t|^p] = \int_{\mathbb{R}_+^d} |x|^p \, d\mu_t(x) \leq c_T < \infty, \quad \forall t \in [0,T],$$

and, the assumption (3.3.4) implies that

$$\int_{\mathbb{R}_+^d} x \, d\mu_t(x) = \int_{\mathbb{R}_+^d} x \, d\mu_T(x) = 1, \quad \forall t \in [0,T].$$  \hspace{1cm} (3.3.7)
The expected value (3.3.7) is an abstraction of the notion of the center of mass. Likewise, there exists a close connection between the moment assumption (3.3.6) that describes the spread of the probability mass around the expected value (3.3.7) and the pathwise growth assumption on the payoff. This is a natural trade off. For a heavy-tailed distribution one must impose a stronger growth assumption on the payoff, and vice versa.

The statement of Doob’s maximal inequality in martingale pricing is that a payoff that exceeds the growth rate admits no finite price.

Throughout the paper, we will utilize the convention that

\[ S_{<p} := S_{b} \text{ for } p = 1 \text{ for a function space } S. \]  

(3.3.8)

The duality of Theorem 3.6.2 and Theorem 3.6.4 is true regardless of the value of the exponent \( p \geq 1 \) in (3.3.6) - in particular, for \( p = 1 \) under relaxed assumptions. Indeed, the convention (3.3.8) also enters in the assumptions on the hedging instruments that will be introduced in Section 3.4. This is the content of the subsequent remarks.

**Remark 3.3.12.** The growth conditions in the admissibility criteria of dynamic strategies (3.4.1) and static positions (3.4.4) are necessary for (super-)hedging polynomially growing claims. However, for bounded payoffs it is sufficient to consider only dynamic strategies that admit a deterministic lower bound and consequently the assumption 3.3.6 on the terminal marginal is not needed. Likewise, for bounded payoffs, it is sufficient to consider bounded static positions. For uniformly continuous payoffs on the other hand, it indeed is sufficient to consider only admissible simple strategies - in particular, in finite discrete time the admissibility criteria can be removed altogether. These are discussed in detail in Proposition 3.6.6 and Remark 3.3.13.

**Remark 3.3.13.** The extension of Proposition 3.8.3 is the only part of the paper, where we utilize the growth assumption (3.3.6) for \( p > 1 \). It is clear from the proofs of Proposition 3.8.1 and Proposition 3.9.2 that, for bounded claims, it is sufficient to consider only dynamic strategies and static positions bounded from below whence neither the growth assumptions (3.4.1) and (3.4.4) nor the integrability assumption (3.3.6) are needed.

**Remark 3.3.14.** Likewise, it is apparent from the proofs of Theorem 3.6.4 and Proposition 3.10.1 that, for bounded claims, the closure in (3.6.6) can be relaxed to be taken over sequences that admit a deterministic lower bound.

**Marginals in Practice**

Let us consider a one-dimensional case. Postulate the existence of call and put function

\[ C(T, K) = E[(X_T - K)^+] \text{ and } P(T, K) = E[(K - X_T)^+]. \]

The cumulative distribution of the terminal marginal is given as

\[ F_{X_T}^Q(K) = \lim_{\varepsilon \to 0} \frac{1}{\varepsilon} (C(T, K) - C(T, K + \varepsilon)) , \quad K \in \mathbb{R}_+. \]

In particular, if the terminal marginal is absolutely continuous, and call and put function are twice continuously differentiable, the density is determined by the Breeden-Litzenberger formula(s) as

\[ f_{X_T}(K) = \frac{\partial^2 C}{\partial K^2}(T, K) = \frac{\partial^2 P}{\partial K^2}(T, K), \]
3.4. PATHWISE INTEGRALS AND HEDGES

whence the price for any $G \in C^2([0, \infty])$ is given by the Carr-Madan formula as

$$E[G(X_T)] = G(X_0) + \int_0^{X_0} \frac{\partial^2 G}{\partial K^2}(K)P(T, K)dK + \int_{X_0}^{\infty} \frac{\partial^2 G}{\partial K^2}(K)C(T, K)dK,$$

where $X_0 = E[X_T]$. For a $C \in C^{1,2}([0, T] \times [0, \infty])$ it can be shown that the law of $X$ is obtained as a solution to the stochastic differential equation

$$dX_t = \sigma(t, X_t)dB_t,$$

for a Brownian motion $B$ and with a local volatility $\sigma$ given by the forward equation

$$\frac{\partial C}{\partial t} C(t, x) = \frac{1}{2} \sigma(t, x)^2 \frac{\partial^2 C}{\partial t^2} (t, x);$$

see e.g. [lowther].

3.4 Pathwise Integrals and Hedges

Since no probabilistic structure on $\Omega$ is given, stochastic integration with respect to $X$ is not immediate. Appropriate pathwise generalization of the Itô’s integral plays a central role in the main duality results.

3.4.1 Simple Integrands

Following the standard terminology, a portfolio process $H = (\tau_k, h_k)_{k \in \mathbb{N}}$ is called a simple integrand (or simple strategy) if the followings are satisfied:

- $0 \leq \tau_1 \leq \tau_2 \leq \ldots$ are $\mathbb{F}$-stopping times;
- for each $\omega \in \Omega$, there is an integer $k(\omega)$ so that $\tau_{k(\omega)}(\omega) = T$;
- for each $k$, $h_k : \Omega \to \mathbb{R}^d$ is a bounded $\mathcal{F}_{\tau_{k+1}}$-measurable function.

For a simple integrand $H = (\tau_k, h_k)_{k \in \mathbb{N}}$, we define the simple gain process by

$$(H \bullet X)_t(\omega) := \lim_{n \to \infty} (H \bullet X)^n_t(\omega),$$

for $(t, \omega) \in [0, T] \times \Omega$, where

$$(H \bullet X)^n_t(\omega) := \sum_{k=1}^{n} h_k(\omega) \cdot (X_{\tau_{k+1} \wedge t}(\omega) - X_{\tau_k \wedge t}(\omega)).$$

A simple integrand $H$ is called an admissible simple strategy if

$$(H \bullet X)^n_t(\omega) \geq -\lambda(\omega) \quad \forall (t, \omega, n) \in [0, T] \times \Omega \times \mathbb{N}$$

for some $\lambda \in \mathbb{B}$ whose positive part $\lambda^+ := \lambda \lor 0$ satisfies the growth condition

$$\sup_{\omega \in \Omega} \frac{\lambda^+(\omega)}{1 + \|\omega\|^p} < \infty \quad (3.4.1)$$

for $p \geq 1$ determined by (3.3.6). The class of all admissible simple strategies is denoted by $\mathcal{H}_s$. The set $\mathcal{H}_s$ also depends on $p$ but this implicit dependence is omitted in our notation as $p$ is fixed throughout the paper.

Remark 3.4.1. We remark that if $k(\omega)$ is finite uniformly over $\Omega$, the lower bound is unnecessary and the present definition of an admissible simple integrand coincides with the classical definition of a simple strategy.
3.4.2 Sub-linear Pathwise Integrals

A *dynamic strategy* is a countable collection $H = (H^k)_{k \in \mathbb{N}}$ with $H^k \in \mathcal{H}_s$. Following [vovk, perkowski, david, david2], we define the corresponding integral by

$$(H \cdot X)_t(\omega) := \lim_{k \to \infty} (H^k \cdot X)_t(\omega), \quad (t, \omega) \in [0, T] \times \Omega,$$

(3.4.2)

which possibly takes the value infinite.

We call $H = (H^k)_{k \in \mathbb{N}}$ an *admissible dynamic strategy* if

$$(H^k \cdot X)_t(\omega) \geq -\lambda(\omega) \quad \forall (t, \omega, k) \in [0, T] \times \Omega \times \mathbb{N}$$

for some $\lambda \in \mathbb{B}$ with $\lambda^+ \in \mathbb{B}_p^+$. We write $\mathcal{H}_d$ for the set of all admissible dynamic strategies, again suppressing the dependence on $p$.

**Comments**

The process $(H \cdot X)$ is not an integral in the classical sense - it is neither linear nor finite. Indeed, it is possible to make infinite profits starting from an arbitrary small initial capital and trading thereafter using a dynamic strategy, i.e., an element of $\mathcal{H}_d$, so, the definition allows the existence of arbitrage - a fact manifested by the divergence of the sum in its definition; we refer to Vovk [vovk, vovk2] and Perkowski & Prömel [david, david2] for details on these facts. In well established literature on stochastic integration, e.g. in the book of Protter [protter], this approach to pathwise integration is referred as the *naïve stochastic integration*.

In comparison, the classical case [DS] assumes that we are given a probability measure $\mathbb{P}$ such that the process $X$ is a semimartingale under $\mathbb{P}$, after which the superhedging is understood as an event of probability one, and thus the stochastic integration with predictable integrands $\mathbb{P}$-almost surely is justified. However, it is important to understand that the stochastic integral is not finite or linear in a pathwise manner; the stochastic integral is linear (and finite) only in terms of $\mathbb{P}$-almost sure equivalence classes, the evaluation at a given path is a priori meaningless. In particular, this means that all critic towards the sublinear pathwise integral can be guided towards the classical definition of the stochastic integral, too.

3.4.3 European Options

In addition to dynamic trading, we assume that all European options with maturing at times $t \in \mathbb{T}$ are available for hedging. Indeed, let $G \colon \Omega \to \mathbb{R}$ be given by

$$G(\omega) = G(X(\omega)) = g(X_t(\omega))$$

(3.4.3)

where $t \in \mathbb{T}$ and $g \colon \mathbb{R}_d^+ \to \mathbb{R}$ is continuous and satisfies the growth condition

$$\sup_{x \in \mathbb{R}_d^+} \frac{|g(x)|}{1 + |x|^p} < \infty$$

(3.4.4)

for $p \geq 1$ determined by (3.3.6). The set $\mathcal{G} = \mathcal{G}(\mu)$ is the set finite linear combinations of all such functions satisfying the additional linear constraint,

$$\mu(G) := \int_{\mathbb{R}_d^+} g(x) \, d\mu(x) = 0.$$

Note that $\mathcal{G}$ is a vector subspace of $\mathbb{C}_p(\Omega)$. 
3.5 Price Bounds of Derivative Contracts

We state the two fundamental theorems of derivative pricing without model assumptions. The first fundamental theorem states the necessary and sufficient conditions for the existence of an arbitrage-free price. The second fundamental theorem states the sharp bound of the arbitrage-free price.

3.5.1 The No Model-Free Arbitrage Condition

A model-free arbitrage opportunity is a semi-static portfolio such that
\[
\mu(G) + (H \cdot X)_T + G(X) \geq r, \quad \text{for all } \omega \in \Omega, \quad \text{for some } r > \mu(G).
\]

We say that a semi-static market \((\Omega, \mu)\) admits "No Model-Free Arbitrage", if there does not exist any such portfolio, i.e., if, for all \(H \in \mathcal{H}_d(\mu)\) and \(G \in \mathcal{G}(\mu)\), one has
\[
\mu(G) + (H \cdot X)_T + G(X) \geq r \quad \forall \omega \in \Omega \implies \mu(G) \geq r, \quad \forall r > 0. \tag{NMFA}
\]

The condition (NMFA) is equivalent to that \(Q(\mu) \neq \emptyset\).

A derivative \(\xi\) is introduced to a market \((\Omega, \mu)\) for a price \(\xi_0\). Analogously to the condition (NMFA), the price \(\xi_0\) must satisfy
\[
\xi_0 \leq \mathcal{E}_{Q(\mu)}(\xi) := \sup_{Q \in Q(\mu)} E_Q[\xi]. \tag{3.5.1}
\]

There exists a (finite) price for any derivative contract on the market if and only if the market satisfies one of the equivalent conditions of the Theorem 3.5.1.

**Theorem 3.5.1.** The market \((\Omega, \mu)\) is free of model-free arbitrage if and only if there exist a martingale measure on \(\Omega\) satisfying the marginals \(\mu\).

3.5.2 The Model-Free Price Bounds

The sharp price bound (3.5.1) is given by Theorem 3.5.2.

**Theorem 3.5.2.** The model-free price bound for a derivative contract \(\xi\) is
(a) \(\mathcal{E}_{Q(\mu)}(\xi) = \Phi_s(\xi; \mu)\) on \(\mathbb{U}_{\geq P}(\Omega)\),
(b) \(\mathcal{E}_{Q(\mu)}(\xi) = \Phi_d(\xi; \mu)\) on \(\mathbb{U}_{< P}(\Omega)\),
(c) \(\mathcal{E}_{Q(\mu)}(\xi) = \Phi_f(\xi; \mu)\) on \(\mathbb{B}_{< P}(\Omega)\).

The model-free superhedging principle is explained in the introduction. The functionals \(\Phi_s, \Phi_d\) and \(\Phi_f\) are defined later in the following section. The proof of our main theorem has been divided in the several following sections.
CHAPTER 3. MARTINGALE OPTIMAL TRANSPORT DUALITY

The Classical and the Existing Model-Free Arbitrage Notions: An Overview

Unable to find an appropriate notion for a (strong) model-free arbitrage in continuous time, that we are trying to model, we introduced (NMFA) above. Thus, we find it appropriate to try to argue that it is also an appropriate one. Let us begin by looking at the classical case.

In the classical continuous time, one has

\[ \text{(NFLVR)} \iff \text{(NA)} + \text{(NUPBR)}, \]

which in the classical finite discrete time reduces to

\[ \text{(NFLVR)} \iff \text{(NA)}; \]

see e.g. the excellent paper by Kabanov [kabanov]. Indeed, in the infinite horizon one still needs both of the conditions as shown by Schachermayer in [infhor].

The (NMFA) in finite discrete time. The notion of (NA) in the model-free setting was introduced in finite discrete time by the Vienna group in [schachermayer]. One obtains their finite discrete time notion (NA) from (NMFA) by replacing the family of dynamic strategies \( \mathcal{H}_d \) with the family of simple strategies \( \mathcal{H}_s \) in the definition of (NMFA). Indeed, they are sufficient in finite discrete-time; see Corollary 3.6.7. Thus, in finite discrete-time, our (NMFA) is consistent with the earlier model-free (NA) of [schachermayer] and we have

\[ \text{(NMFA)} \iff \text{(NA)}, \]

for the model-free notion (NA) of [schachermayer].

In [matteo], two different model-free arbitrage notions were also introduced. The notion of a uniform arbitrage that is equivalent to (NA) and the notion of strong arbitrage that is equivalent to (NA) without a lower bound - the two are known to be equivalent in the classical finite discrete time.

The (NMFA) in continuous (and infinite discrete) time. The notion of a weak (pathwise) arbitrage opportunity in the model-free setting, the one allowed in the definition of sub-linear pathwise integral, was studied in detail by Perkowski & Prömel in [perkowski]. The authors call such a weak pathwise arbitrage opportunity a model-free arbitrage opportunity, but acknowledge that their notion of a model-free arbitrage is completely analogous to the classical case of unbounded profit with bounded risk in a semimartingale model. Indeed, the necessary and sufficient condition for the non-existence of model-free arbitrage opportunity, in the sense of [perkowski], is that there does not exist unbounded profit with bounded risk on the canonical space endowed with Vovk’s outer measure; thus, let us write (NUPBR) for the non-existence of a model-free arbitrage in the sense of [perkowski]. Consequently, such opportunities exists only in Vovk’s null sets. In particular, the paths that are too rough or too smooth to be paths of a semimartingale are contained in the Vovk null sets, as shown in the seminal paper by Vovk [vovk]. So, by the result of [perkowski], the paths that violate the model-free (NUPBR) do not impose additional super-hedging cost for a functional allowing sub-linear pathwise trading, and, on the other hand, by the result of Vovk, these paths do not affect the price of a payoff for a sub-linear expectation defined over all martingale measures satisfying the given marginals. Thus, in continuous (infinite discrete) time, the classical notion of (NUBPR) is built in the definition of sub-linear pathwise integral and we have

\[ \text{(NMFA)} \iff \text{(NA)} + \text{(NUPBR)}, \]
3.6. **SUPER-REPLICATION FUNCTIONALS**

for the model-free notions (NA) of [schachermayer] and (NUPBR) of [perkowski].

It is noteworthy that in the (NMFA) formulation one completely avoids any abstract
norm closure of superhedgeable positions in contrast to the classical (NFLVR). In the
pathwise setting, an abstract (Fatou) closure plays a role only at a later stage of obtaining
a sharp bound for the price of a general measurable payoff.

### 3.6 Super-replication Functionals

In this section, we state the main duality results for upper semicontinuous and measurable
claims.

#### 3.6.1 First Super-replication Functional

We are now ready to define the first superreplication functional. The second one,
appropriate for general, measurable claims is given in 3.6.7 below.

We define a subset of $\mathcal{F}_T$-measurable random variables achievable by trading with
zero initial cost by

$$
\mathcal{K} = \mathcal{K}(\mu) := \{ \zeta = G + (H \bullet X)_T : G \in \mathcal{G}(\mu_T), \ H \in \mathcal{H}_d \}.
$$

(3.6.1)

The first superreplication functional for an $\mathcal{F}_T$-measurable claim $\xi$ is defined by

$$
\Phi_d(\xi; \mu) = \inf \{ x \in \mathbb{R} : \exists \zeta \in \mathcal{K}(\mu) \text{ s.t. } x + \zeta(\omega) \geq \xi(\omega), \ \forall \omega \in \Omega \}.
$$

(3.6.2)

We use the convention that $\Phi_d(\xi; \mu) = +\infty$ whenever $\xi$ is not superreplicable, or
equivalently, when the above set is empty. It is clear that the functional $\Phi_d$ is increasing,
translation invariant and sublinear.

The first result follows almost directly from known results. We recall that by the
Strassen’s theorem [strassen][Theorem 8], $Q(\mu)$ is non-empty.

**Proposition 3.6.1.** Under the assumption (3.3.6), $\Phi_d$ is increasing, translation invariant
and sublinear. Moreover, one has

$$
\Phi_d(0; \mu) = 0.
$$

**Proof.** The positive monotonicity, translation invariance and sublinearity follows directly
from the definition of $\Phi_d$. It is clear that $\Phi_d(\xi; \mu) \leq \|\xi\|_{\infty}$ for every bounded $\xi \in \mathbb{B}$.

Let $Q \in Q(\mu)$. Then, it is clear that $E_Q[G] = E_Q[(H \bullet X)_T] = 0$ for all $G \in \mathcal{G}$ and
$H \in \mathcal{H}_d$. Also by Fatou’s lemma, this extends for every dynamic strategy $H \in \mathcal{H}_d$ as
well. Therefore, for every $\mathcal{F}_T$-measurable $\xi$,

$$
\Phi_d(\xi; \mu) \geq E_Q(\mu)(\xi) = \sup_{Q \in Q(\mu)} E_Q[\xi].
$$

(3.6.3)

Notice that if $\xi$ is not superreplicable, then $\Phi_d(\xi)$ is infinity and the above inequality
holds trivially.

Since $Q(\mu)$ is non-empty, 3.6.3 implies that

$$
0 = E_Q(\mu)(0) \leq \Phi_d(0; \mu) \leq \|0\|_{\infty} = 0.
$$

$\square$
3.6.2 Upper Semicontinuous Payoffs

In this subsection, we consider the claims $\xi$ that are upper semicontinuous and satisfy the growth condition
\[
\sup_{\omega \in \Omega} \frac{|\xi(\omega)|}{1 + \|\omega\|_\infty^q} < \infty
\] (3.6.4)
for some $1 \leq q < p$ with the convention that they are bounded for $p = 1$. This set of claims is denoted by $U_{<p}$, and some examples of common derivative contracts in $U_{<p}$ are provided in the next section.

Our first result is the duality on the set $U_{<p}$.

**Theorem 3.6.2.** For all $\xi \in U_{<p}(\Omega)$, one has
\[
\Phi_d(\xi; \mu) = E_{Q(\mu)}(\xi),
\] (3.6.5)
for $E_{Q(\mu)}(\xi) := \sup_{Q \in Q(\mu)} E_Q[\xi]$.

In view Proposition 3.8.1, it suffices to prove 3.6.5 only for $\xi \in C_b$. The following proof uses several key technical results: Propositions 3.8.1, 3.8.3, 3.9.1, 3.9.2 and Lemma 3.8.2.

**Proof.** We prove the duality on $C_b$. First note that $\Phi_d$ is a proper sub-linear functional on $C_b$. Proposition 3.9.1 proves that $\partial \Phi_d(0; \mu) = Q(\mu)$. Also, the lower semicontinuity of $\Phi_d$ on $C_b$ is established in Proposition 3.9.2. Then, the classical Fenchel-Moreau theorem, see for instance [zalinescu][Theorem 2.3.3.], proves 3.6.5 on $C_b$. \qed

**Remark 3.6.3.** The value $E_{Q(\mu)}(\xi)$ is attained for some $Q \in Q(\mu)$, for every $\xi \in U(\Omega)$ bounded from above; see Corollary 3.9.7.

3.6.3 Measurable Payoffs

In general, the payoff of a derivative contract is a function of $X$. Thus, it is natural to assume that the function is $F_T$-measurable.

As discussed earlier, there is a duality gap for continuous claims when superreplication is defined through simple integrands $H_s$ and it is closed by considering the larger class of dynamic strategies $H_d$. Example 3.7.2 shows that there is also a duality gap for general measurable claims if the hedging class $K$ defined by $H_d$ is not further extended. Therefore, one needs to enlarge the hedging set by taking an appropriate closure. Alternatively, one could achieve duality by keeping the superreplication functional unchanged and enlarge the dual measures by allowing finitely additive martingale laws. This approach is studied in [ekren] and not pursued in this paper.

We enlarge the hedging set $K$ by taking its Fatou closure. Vovk [vovk3] introduced this approach to establish a superhedging duality for lower semicontinuous payoffs in the classical Wiener space - without static positions or marginal distributions. Furthermore, the seminal result of Delbaen & Schachermayer [DS] boils down to showing that the set of (the equivalence classes) of stochastic integrals with admissible predictable integrands admits the closure property defined below. Hence, in the classical case of [DS] the Fatou closure below is given already by the family of predictable integrands.

**Fatou closure.** Recall the set $K$ as defined in 3.6.1. Let $\overline{K}$ be the smallest set of measurable functions on $\Omega$ containing $K$ and having the following property:
\[
\text{if } (\zeta^n)_{n \in \mathbb{N}} \subset \overline{K} \text{ and } \exists \lambda \in \mathbb{B} \text{ s.t. } \lambda^+ \in \mathbb{B}_p \text{ and } \zeta^n \geq -\lambda, \text{ then } \liminf_{n \to \infty} \zeta^n \in \overline{K},
\] (3.6.6)
where \( p \geq 1 \) is determined by (3.3.6).

The Fatou closure is motivated by the sequential order lower semicontinuity of the functional that is a necessary condition for the conjugate duality to hold as shown already in [biagini2010]. However, the closure cannot be achieved via countable sequential procedures and we settle to characterize the elements of \( \overline{K} \) by transfinite recursion. A first proof of a general theorem justifying the definition by transfinite recursion was provided by John von Neumann in [Neumann1928]. Transfinite recursion was introduced to mathematical finance by Vladimir Vovk in [vovk3]. More explicit characterization up to now is obtained only for the one-dimensional martingale optimal transport problem in [nutz]. We leave it as an interesting open problem to achieve a similar characterization in the present or more general setting.

We now define the second superreplication functional enlarging \( K \) with \( K \),

\[
\Phi_f(\eta; \mu) := \inf \{ x \in \mathbb{R} : \exists \zeta \in \overline{K} \text{ s.t. } x + \zeta(\omega) \geq \eta(\omega), \quad \forall \omega \in \Omega \}
\]  

(3.6.7)

with the convention that \( \Phi_f(\eta; \mu) = +\infty \) whenever \( \eta \) is not superreplicable. As before, we sometimes write \( K = K(\mu) \) to emphasize the dependence on the marginals.

We now define the second superreplication functional enlarging \( K \) with \( K \),

\[
\Phi_f(\eta; \mu) := \inf \{ x \in \mathbb{R} : \exists \zeta \in \overline{K} \text{ s.t. } x + \zeta(\omega) \geq \eta(\omega), \quad \forall \omega \in \Omega \}
\]  

(3.6.7)

with the convention that \( \Phi_f(\eta; \mu) = +\infty \) whenever \( \eta \) is not superreplicable. As before, we sometimes write \( K = K(\mu) \) to emphasize the dependence on the marginals.

We continue by stating the extension of Theorem 3.6.2 to \( F_T \)-measurable claims. This result is based on the previous duality and Choquet’s capacibility theorem which is recalled in the Appendix.

Let us write \( \mathbb{B}_{<p} \) for the set of all measurable claims satisfying the growth assumption 3.6.4.

**Theorem 3.6.4.** For every \( \eta \in \mathbb{B}_{<p}(\Omega) \), one has

\[
\Phi_f(\eta; \mu) = \mathcal{E}_{Q(\mu)}(\eta),
\]  

(3.6.8)

for \( \mathcal{E}_{Q(\mu)}(\eta) := \sup_{Q \in Q(\mu)} E_Q[\eta] \).

In view of Proposition 3.8.1, it suffices to prove 3.6.8 for bounded, measurable claims. As in the proof of Theorem 3.6.2, we prove this result using several key technical results, Propositions 3.10.1, 3.3.1 and 4.0.31.

**Proof.** We first observe that

\[
\Phi_d(\xi; \mu) \geq \Phi_f(\xi; \mu) \geq \mathcal{E}_{\partial \Phi_f(0)}(\xi) = \mathcal{E}_{Q(\mu)}(\xi),
\]

for every \( \xi \in \mathbb{B}_b \). Indeed, the first inequality follows from \( K \subset \overline{K} \), the second one from the Young-Fenchel inequality, see for instance [zalinescu][Theorem 2.3.1. (iv)], and the last equality follows from Proposition 3.10.1 (b). Hence, by Theorem 3.6.2,

\[
\Phi_d(\xi; \mu) = \Phi_f(\xi; \mu) = \mathcal{E}_{Q(\mu)}(\xi), \quad \forall \xi \in \mathbb{B}_b.
\]

Moreover, Proposition 4.0.31 proves that for every \( (\xi^n)_{n \in \mathbb{N}} \) and \( \xi \in \mathbb{B}_b \),

\[
\xi^n \downarrow \xi \quad \Rightarrow \quad \mathcal{E}_{Q(\mu)}(\xi^n) \downarrow \mathcal{E}_{Q(\mu)}(\xi).
\]

The sequential continuity of the functional \( \Phi_f \) from below on \( \mathbb{B}_b \) is proved in Proposition 3.10.1 (c), i.e.,

\[
\eta^n \uparrow \eta \quad \Rightarrow \quad \Phi_f(\eta^n; \mu) \uparrow \Phi_f(\eta; \mu)
\]

\[\text{1} \]We are very grateful for Glenn Shafer and Vladimir Vovk for sharing an early pre-print of their book [Shafer/Vovk:2019] with us.
CHAPTER 3. MARTINGALE OPTIMAL TRANSPORT DUALITY

for every \((\eta^n)_{n \in \mathbb{N}}\) and \(\eta\) in \(\mathbb{B}_b\).

Let \(U_0^+\) be the set of all non-negative functions in \(U_0\). Then, \(U_0^+\) is closed with respect to finite suprema and countable infima. Moreover, by Proposition 3.3.1, the underlying topological space is perfectly normal. Hence, the assumptions of Corollary 4.0.30 are satisfied and we conclude that \(\Phi_f\) is a \(U_0^+\)-capacity on \(\mathbb{B}_b^+\). Consequently, the duality 3.6.8 holds for bounded measurable claims \(\eta \in \mathbb{B}_b^+\).

Let \(U_{b}^+\) be the set of all non-negative functions in \(U_b\). Then, \(U_{b}^+\) is closed with respect to finite suprema and countable infima. Moreover, by Proposition 3.3.1, the underlying topological space is perfectly normal. Hence, the assumptions of Corollary 4.0.30 are satisfied and we conclude that \(\Phi_f\) is a \(U_{b}^+\)-capacity on \(\mathbb{B}_b\).

3.6.4 Uniformly Continuous Payoffs and Simple Strategies

This section gathers some complementary facts regarding our first duality that will establish a link to older results.

Uniformity

In order to study uniform continuity, we endow \(\Omega\) with the uniformity generated by the metric

\[
D(\omega, \overline{\omega}) := J^1(\omega, \overline{\omega}) + L^1(\omega, \overline{\omega}),
\]

where \(J^1\) denotes the usual Skorokhod metric (3.6.11) and

\[
L^1(\omega, \overline{\omega}) := \int_0^T |\omega(t) - \overline{\omega}(t)| dt.
\]

The metric \(d\) was introduced in \cite{mete2}. In particular, we have

\[
J^1(\omega, \overline{\omega}) \leq D(\omega, \overline{\omega}) \leq (1 + T)\|\omega - \overline{\omega}\|_\infty.
\]

We want to emphasize that uniformity and therefore uniform continuity is not a topological concept; see e.g. \cite{engelking}[Chapter 8] for elaboration. In particular, it is not sufficient to prove the duality for a family of uniformly continuous payoffs and proceed from there to measurable payoffs, as we did. Moreover, we want to emphasize that the family \(U_{u}^+\) of uniformly continuous upper semicontinuous is non-redundant since the uniformity is assumed in a stronger topology than the semicontinuity.

Skorokhod’s metric

The Skorokhod’s \(J^1\)-metric on \(\mathbb{D}([0, T]; \mathbb{R}^d)\) is

\[
J^1(\omega, \tilde{\omega}) := \inf_{\lambda \in \Lambda} \left\{ \sup_{s < t} \left| \log \frac{t - s}{t - \lambda s} \right| \vee \|\omega - \tilde{\omega} \circ \lambda\|_\infty \right\},
\]

where \(\Lambda\) denotes the class of strictly increasing, continuous mappings of \([0, T]\) onto itself and \(i\) is the identity map on \([0, T]\). The modulus of continuity is

\[
m_\delta(\omega) := \inf_i \sup_{s, t \in [t_{i-1}, t_i]} \|\omega(t) - \omega(s)\|,
\]

where the infimum is taken over all finite meshes with \(t_i - t_{i-1} > \delta\). A subset \(K\) of \(\mathbb{D}([0, T]; \mathbb{R}^d)\) is (relatively) \(J^1\)-compact if and only if it is equicontinuous and bounded, i.e.,

\[
\lim_{\delta \to 0} \sup_{\omega \in K} m_\delta(\omega) = 0 \quad \text{and} \quad \sup_{\omega \in K} \|\omega\|_\infty < \infty.
\]
3.6. SUPER-REPLICATION FUNCTIONALS

Remark 3.6.5. The mapping \( X_t(\omega) = \omega(t) \) is continuous on \( \Omega \), if \( X \) of \( \Omega \) is \( J^1 \)-equicontinuous and bounded (e.g. constant) on some \([t, t + \varepsilon]\); cf. (3.6.12). Indeed, one can find a \( J^1 \)-convergent subsequence from the set of paths of \( X \), from which the continuity of the evaluation at time \( t \) follows. Moreover, stretching each path an \( \varepsilon \) amount of time at the marginals defines a uniformly continuous isomorphism between the spaces and consequently the duality result of [mete2] is at our disposal; the restriction to an \( S^* \) closed subset \( \Omega \) of \(([0, T]; \mathbb{R}^d)\) is straightforward.

Thus, we recover (essentially) the structure of [mete2], where the dynamic trading of finite variation is sufficient, for uniformly continuous payoffs.

The Dolinsky-Soner Super-replication Functional

The super-replication functional of [mete2] is

\[
\Phi_{bv}(\xi; \mu) := \inf \{ x : \exists \zeta \in K_{bv}(\mu) \text{ s.t. } x + \zeta \geq \xi \},
\]

where

\[
K_{bv}(\mu) := \{ G(X) + (H \cdot X)_T : G \in \mathcal{G}(\mu) \text{ and } H \in \mathcal{H}_{bv} \},
\]

where \( \mathcal{H}_{bv} \) denotes the space of admissible, càdlàg functions \( H : [0, T] \times \Omega \to \mathbb{R}^d \) of finite variation and

\[
(H \cdot X)_T := H_TX_T - H_0X_0 + (H \cdot X)_T,
\]

where the last integral is a Lebesgue-Stiefel integral integral; cf. (3.4.1). The same super-replication functional was considered in [touzi]. For more details we refer to [mete2][Section 2.3] and [touzi][Section 2.2].

Simple Strategies

Recall the class of simple strategies and its role in the definition of the sublinear pathwise integral à la Vovk, Perkowski & Prömel; Subsection 3.4.1-3.4.2. We show that their sublinear extension of the pathwise integral does not lower the superhedging cost under the assumption that the payoff is uniformly continuous. The underlying uniformity of the space is the basis for the seminal time-space discretization solution to the pathwise superhedging duality problem of Dolinsky & Soner that approximates the pathwise problem, uniformly, with its probabilistic counterpart - in the probabilistic case, the superhedging duality is a consequence of the optional decomposition theorem.

We wish to embed the aforementioned result for uniformly continuous payoffs in the present framework and thus let us define an auxiliary superreplication functional

\[
\Phi_s(\xi; \mu) := \inf \{ x : \exists \zeta \in K_s \text{ s.t. } x + \zeta \geq \xi \},
\]

where

\[
K_s = K_s(\mu) := \{ G(X) + (H \cdot X)_T : G \in \mathcal{G}(\mu) \text{ and } H \in \mathcal{H}_s \}.\]

(3.6.15)

References are provided after the proposition at the end of this subsection. Recall the uniformity from Subsection 3.6.4.

Proposition 3.6.6. For all uniformly \( \xi \in \mathcal{U}_{<\mu}(\Omega) \), one has

\[
\Phi_s(\xi; \mu) = \mathcal{E}_{\mathcal{Q}(\mu)}(\xi),
\]

for \( \mathcal{E}_{\mathcal{Q}(\mu)}(\xi) := \sup_{\mathcal{Q} \in \mathcal{Q}(\mu)} E_{\mathcal{Q}}[\xi] \).
Proof. We prove the statement in steps by first invoking the marginal approximation result of [touzi] to reduce the problem to the finite marginal case of [mete2], where we can apply a simple approximation argument on the canonical space to show that the simple strategies are sufficient, and extend thereafter using the growth condition (3.3.6).

Step 1. Let us begin by assuming that $\xi \in \mathcal{U}_b^b(\Omega)$. Recall that the time-index set $[0, T]$ of $\Omega$ is endowed with the right-limit topology and the set $T$ of continuity points of $\Omega$ is a closed subset of $[0, T]$ containing the terminal time $T$. Thus, there exists an increasing sequence $(T_n)_{n \in \mathbb{N}} \subset T$ such that $T \in T_n$ for all $n \in \mathbb{N}$ and the set $\cup_{n \in \mathbb{N}} T_n$ is dense in $T$. Set $\mu_n := (\mu_t)_{t \in T_n}$, for all $n \in \mathbb{N}$, for a given $\mu = (\mu_t)_{t \in \mathbb{N}}$. Each $\mu^n$ is a peacock and the payoff $\xi$ is uniformly continuous in the metric (3.6.11), so, the necessary assumptions of [mete2] are satisfied. Thus, by [mete2][Theorem 5.2] in conjunction with [mete2][Remark 3.6.5], we have

$$\Phi_{bv}(\xi; \mu^n) = \mathcal{E}_{\mathbb{Q}}(\mu^n)(\xi), \text{ for all } n \in \mathbb{N};$$

(3.6.17)

cf. Remark 3.6.5.

Step 2. Since $S^* \subset S$ and $\xi \in \mathcal{U}_b^b(\Omega) \subset \mathcal{U}_b(\Omega)$ is bounded from above, the assumptions of [touzi][Theorem 3.16] are satisfied. From (3.6.17), by (i) in conjunction with (ii) of [touzi][Theorem 3.16], we obtain the duality

$$\Phi_{bv}(\xi; \mu) = \mathcal{E}_{\mathbb{Q}}(\mu)(\xi) \text{ for all } \xi \in \mathcal{U}_b^b(\Omega).$$

(3.6.18)

Step 3. Let $\xi \in \mathcal{U}_b^b(\Omega)$, $\varepsilon > 0$ and, for every $\omega \in \Omega$, define $\nu_\varepsilon(\omega)$ as

\[
\nu_\varepsilon(\omega)(t) := \omega(\tau_k(\omega)) - \omega(\tau_{k+1}(\omega)), \; t \in [0, T], \; k \in \mathbb{N}_0,
\]

with respect to the sequence $(\tau_k^\varepsilon)_{k \in \mathbb{N}_0}$ of $\mathbb{F}$-stopping times

\[
\begin{cases}
\tau_0^\varepsilon(\omega) := 0, \\
\tau_k^\varepsilon(\omega) := \inf\{t > \tau_{k-1}^\varepsilon(\omega) : |\omega(t) - \omega(\tau_{k-1}^\varepsilon(\omega))| > \varepsilon\}, \; k \in \mathbb{N}.
\end{cases}
\]

For every $G \in \mathcal{G}(\mu)$, by (3.6.10), for $\delta := (1 + T)^{-1}\varepsilon$, we have

$$\xi(\nu_\varepsilon(\omega)) - G(\nu_\varepsilon(\omega)) \geq \xi(\omega) - G(\omega) - \delta, \; \forall \omega \in \Omega.$$

In particular, for every $H \in \mathcal{H}_{bv}$ and $G \in \mathcal{G}(\mu)$ superhedging $\xi$, for every $\omega \in \Omega$, we have

$$H_0 + (H \bullet X)_T(\nu_\varepsilon(\omega)) \geq \xi(\nu_\varepsilon(\omega)) - G(\nu_\varepsilon(\omega)) \geq \xi(\omega) - G(\omega) - \delta.$$

The fact $H \in \mathcal{H}_{bv}$ is of finite variation is crucial for what follows. Let us denote

$$\Delta X_t(\omega) := \omega(t) - \omega(t^-), \forall (\omega, t) \in \Omega \times [0, T].$$

For the solution $H^\varepsilon$ of the following pathwise difference equation

$$H_0^\varepsilon = H_0 \text{ and } \Delta(H^\varepsilon \bullet X)_t(\omega) = \Delta(H \bullet X)_t(\nu_\varepsilon(\omega)), \; \forall (t, \omega) \in [0, T] \times \Omega,$$

we have

$$H_0^\varepsilon + (H^\varepsilon \bullet X)_T(\omega) \geq \xi(\omega) - G(\omega) - \delta, \; \forall \omega \in \Omega.$$

Since a càdlàg path has only finitely many jumps greater than a fixed size, $\varepsilon > 0$, and $H$ was assumed admissible and is of finite variation, the process $H^\varepsilon$ is an element of $\mathcal{H}_s$. Moreover, since $\varepsilon > 0$, and therefore $\delta > 0$, above was arbitrary and $H_0^\varepsilon = H_0$, we have

$$\Phi_s(\xi; \mu) = \Phi_{bv}(\xi; \mu) \text{ for all } \xi \in \mathcal{U}_b^b(\Omega).$$

(3.6.19)
Step 4. Finally, by Proposition 3.8.3 in conjunction with Remark 3.8.4, the duality (3.6.18)-(3.6.19) extends to
\[ \Phi_s(\xi; \mu) = \mathcal{E}_{\mathcal{Q}(\mu)}(\xi) \quad \text{for all} \quad \xi \in U_p^\mu(\Omega), \]
where the exponent \( p \geq 1 \) is determined by (3.3.6).

Corollary 3.6.7. In (finite) discrete time, for all \( \xi \in U_p(\Omega) \), one has
\[ \mathcal{E}_{\mathcal{Q}(\mu)}(\xi) = \Phi_s(\xi; \mu), \quad (3.6.20) \]
for \( \mathcal{E}_{\mathcal{Q}(\mu)}(\xi) := \sup_{Q \in \mathcal{Q}(\mu)} E_Q[\xi] \).

Proof. In finite discrete time all dynamic strategies are simple. Likewise, continuous functions are uniformly continuous on bounded sets; cf. Lemma 3.9.3.

Background
In the Skorokhod space, the time-space discretization method was first deployed in \([\text{mete2}]\), later refined in \([\text{touzi}]\) - whose result we could have as well relied on above. Prior to the aforementioned Skorokhod space results, the time-space discretization approach was originally developed in the classical Wiener space in \([\text{mete}]\), recently refined in \([\text{obloj}]\). The classical work of Föllmer & Kramkov on the optional decomposition theorem \([\text{follmer1997}]\) can be seen as a predecessor of the continuous-time pathwise superhedging duality results.

The problem was introduced in discrete time in the seminal paper of Beiglöck, Henry-Labordère & Penkner \([\text{beiglbock}]\). The notion of model-free arbitrage was introduced in \([\text{schachermayer}]\). The discrete-time model-free arbitrage notions were further developed in \([\text{matteo}]\). These papers concentrated on topological (upper semicontinuous) payoffs. The duality was extended for measurable payoffs in \([\text{nutz}]\), in the one-dimensional single-step case.

3.7 Examples and Counterexamples
In this section, we provide examples of upper semicontinuous claims. We also give counter-examples showing the necessity of the enlargements from simple integrands to dynamic strategies and also to the Fatou closure.

3.7.1 Upper Semicontinuous Payoffs
The following payoffs satisfy the assumptions of Theorem 3.6.2 on \( \mathbb{D}([0,T]; \mathbb{R}_+) \).

Some examples of continuous derivatives are Asian put and call option with fixed strike
\[ \left( \frac{1}{T} \int_0^T X_t \, dt - K \right)^+ \quad \text{and} \quad \left( K - \frac{1}{T} \int_0^T X_t \, dt \right)^+ \]
and Asian put and call option with floating strike
\[ \left( \frac{1}{T} \int_0^T X_t \, dt - X_T \right)^+ \quad \text{and} \quad \left( X_T - \frac{1}{T} \int_0^T X_t \, dt \right)^+. \]
Some examples of upper semicontinuous derivatives are lookback put and call option
with fixed strike
\[
\left( K - \max_{0 \leq t \leq T} X_t \right)^+ \quad \text{and} \quad \left( \min_{0 \leq t \leq T} X_t - K \right)^+
\]
and digital options
\[
1_{\{\max_{0 \leq t \leq T} X_t \leq K\}} \quad \text{and} \quad 1_{\{\min_{0 \leq t \leq T} X_t \geq K\}}.
\]

The previous Asian and lookback options are uniformly continuous; digital options are not.

### 3.7.2 Example: Simple Strategies Are Not Sufficient

The following example is a negative result for Lemma 3.9.5 that yields the continuity of \( \Phi_d \). Lemma 3.9.5 is not true if the functional \( \Phi_d \) is replaced with the functional \( \Phi_s \). We assume \( \Omega \) bounded that allows us to omit the marginal distributions and static positions altogether.

Recall that \( \Phi_s \) and \( K_s \) are defined in Section 3.4.1. In fact, Example 3.7.1 remains true if one replaces \( \Phi_s \) with \( \Phi_{bv} \) as defined in Subsection 3.6.4.

**Example 3.7.1.** Let \( d = 1, T = 1, \Omega := D([0, 1]; [0, 2]) \) bounded and \( G = \emptyset \). Given any compact set \( K \) in \( \Omega \), there exists a set \( A \) in \( \Omega \setminus K \) such that
\[
\Phi_s(1_A) = 1 > 0 = \sup_{Q \in \mathcal{Q}} Q(A).
\]

**Proof.** Our example is related to the pathwise \( p \)-variation given by
\[
v^p(\omega) := \sup_{\pi} \sum_{k=1}^{n} |\omega(\tau_k) - \omega(\tau_{k-1})|^p, \omega \in D([0, 1]; \mathbb{R}_+), \quad p > 2,
\]
where \( \pi \) ranges over all finite partitions \( 0 = \tau_0 < \tau_1 < \cdots < \tau_n = 1 \) of \([0, 1]\).

Let \( f(x) := x^{1/p}, x \geq 0, t_n := 1/n, s_n := \frac{1}{2}(t_{n+1} + t_n), n \in \mathbb{N} \). Set
\[
\bar{\omega}_p(t) := \begin{cases} 
  k_n(t - t_{n+1}) & \text{for } t_{n+1} < t \leq s_n, \\
  -k_n(t - s_n) + f(s_n) & \text{for } s_n < t \leq t_n,
\end{cases} \quad t \in [0, 1],
\]
where \( k_n \) is the slope of the line connecting \((t_n, 0)\) to \((s_n, f(s_n))\). It is clear that \( \bar{\omega}_p \in C([0, 1]; \mathbb{R}_+) \subset D([0, 1]; \mathbb{R}_+) \).

For any \( s \in [0, 1] \) and \( \omega \in C([0, 1]; \mathbb{R}_+) \), let \( \omega^s \) be the stopped path, i.e.,
\[
\omega^s(t) := \omega(t \wedge s), \quad t \in [0, 1].
\]

We have
\[
v^p(\bar{\omega}_p^t) \geq \sum_{k=n}^{\infty} (f(s_k) - f(t_k))^p = \infty,
\]
for every \( n \in \mathbb{N} \), for every \( p > 2 \). Finally, consider the collection of stopped paths started at a point \( a \in \mathbb{R}_+ \), and set
\[
A^{a,p} := \{ a + \bar{\omega}_p^t_n : n \in \mathbb{N}, \ a \in \mathbb{R}_+ \} \quad (3.7.1)
\]
and

\[ A := \bigcup_{a \geq 0, p > 2} A^{a,p}. \]

A set \( K \) is compact if and only if the set \( K \) is closed and \( \sup_{\omega \in K} |\omega|_{\infty} \leq C_{\infty} \) for some finite constants \( C_{a,b}, 0 \leq a < b \leq 2, \) and \( C_{\infty} \). It is apparent that no matter how these constant are chosen, for any \( a \geq 0 \), by choosing \( p = p(a) > 2 \) sufficiently large, we find a set \( A := A^{a,p(a)} \subset \Omega \setminus K \). We show that \( \Phi_{a}(1) = 1 \). Indeed, suppose that \( c \in \mathbb{R} \) and \( H \in H_{a} \) satisfy

\[ c + (H \cdot X)_{t_n} = c + (H \cdot X)_{t_n}(\omega) \geq 1_{A^{t_n}} = 1 \]

for every \( n \in \mathbb{N} \). By integrating by parts,

\[ 1 - c \leq (H \cdot X)_{t_n}(\omega) \]

\[ = (X_{t_n}(\omega) - 1)H_{t_n}(\omega) - \int_{0}^{t_n} (X_s - 1) dH_s(\omega) \]

\[ \leq |X_{t_n}(\omega) - 1| |H_{t_n}(\omega)| + \|((\omega) - 1)1_{[0,t_n]}\| \int_{0}^{t_n} |dH| (\omega). \]

Since \( H \) is left-continuous and of finite variation, it must be bounded in a neighborhood of the origin. Therefore, for all sufficiently large \( n \),

\[ 1 \leq c + |X_{t_n}(\omega) - 1| \cdot C + \|((\omega) - 1)1_{[0,t_n]}\| \cdot C \]

for some constant \( C \) depending on the variation of \( H \). Since \( \omega \) is right-continuous at 0, as \( n \) tends to infinity,

\[ |X_{t_n}(\omega) - 1| \to 0, \quad \text{and} \quad \|((\omega) - 1)1_{[0,t_n]}\| \to 0. \]

We now let \( n \) tend to infinity in 3.7.6 to conclude that \( c \geq 1 \). Hence, \( \Phi_{a}(\xi) \geq 1 \). Since the constant 1 superreplicates \( \xi \), this proves our claim.

On the other hand, by \([\text{Lepingle1976}]\) [Theorem 1 (a)], for every \( Q \in Q(\mu) \)

\[ E_{Q}[\xi] = Q(A) \leq Q \left( \{v^3 = \infty \} \right) + Q(\{1\}) = 0. \]

So, there is a duality gap.

\[ \Box \]

3.7.3 Example: Dynamic Strategies Are Not Sufficient

The following example motivates the use of the closure \( \overline{K} \) to establish the duality for measurable payoffs in Theorem 3.6.4. In particular, the example below shows that \( \overline{K} \) is strictly larger than \( K \) and thus \( \Phi_{d} \) (as defined by 3.6.2) fails to have the sequential continuity from below.
Example 3.7.2. There exists a peacock $\mu$ and a non-negative $\eta$ in $\mathbb{B}_b$ such that
\[
\Phi_d(\eta; \mu) = 1 > 0 = \mathcal{E}_Q(\eta) = \Phi_f(\eta; \mu).
\]
Proof. Let $d = 1$ and $\mu_0 = \mu_T$ be the Dirac measure at $1$. Consider the following lower semicontinuous payoff
\[
\eta(\omega) := 1_{\{X_T \neq X_0\}}(\omega).
\]
Let $\omega^*(t) := 1$ for all $t \in [0, T]$. Then, the only measure in $Q(\mu)$ is the Dirac measure at $\omega^*$. Hence, we get
\[
\Phi_f(\eta; \mu) = \mathcal{E}_Q(\mu)(\eta) = 0
\]
where we used Theorem 3.6.4 for the first equality.

However, one can show that $\Phi_d(\eta; \mu) = 1$. Indeed, $\Phi_d(\eta; \mu) \leq 1$ follows immediate from $\eta \leq 1$, that is, it remains to show that $\Phi_d(\eta; \mu) < 1$ is not possible. For is purpose, let us assume there exist $0 \leq c < 1$, $g \in G(\mu_T)$ and a dynamic strategy $H = (H^k)_{k \in \mathbb{N}} \subset \mathcal{H}_s$ such that
\[
c + g(X_T(\omega)) + (H \cdot X)_T(\omega) \geq \eta(\omega), \quad \omega \in \Omega.
\]
Note, since $g \in G(\mu)$, we have
\[
0 = \int_{\mathbb{R}_+} g \, d\mu_T = g(1)
\]
and thus for all $\delta \in (0, 1 - c)$ there exists an $\varepsilon > 0$ such that
\[
g(x) \leq \delta \quad \forall x \in \mathbb{R}_+ \text{ with } |x - 1| \leq \varepsilon
\]
as $g$ is continuous by assumption. Hence, for all $\omega \in \Omega$ satisfying $\omega(0) \neq \omega(T)$ and $|\omega(T) - 1| \leq \varepsilon$ we observe that
\[
c + \delta + \liminf_{k \to \infty} (H^k \cdot X)_T(\omega) \geq 1
\]
or equivalently
\[
\liminf_{k \to \infty} (H^k \cdot X)_T(\omega) \geq 1 - c - \delta > 0. \tag{3.7.8}
\]
However, (3.7.8) cannot be true for all $\omega \in \Omega$ satisfying $0 \neq |\omega(T) - \omega(0)| \leq \varepsilon$. Consider the paths $\tilde{\omega}$ such that $\tilde{\omega}(t) = \tilde{\omega}(0)$ for all $t \in [0, T)$ and $\tilde{\omega}(T) \neq \tilde{\omega}(0)$. Any simple strategy $H^k$ which satisfies (3.7.8) in case of an upwards jump at time $T$ cannot satisfy (3.7.8) in case of a downwards jump at time $T$, and vice versa. Therefore, $\Phi_d(\eta; \mu) < 1$ is not possible.

Remark 3.7.3. In the context of martingale optimal transport on the real line it was recently shown that a (similar) appropriate closure of dual elements is necessary in order to obtain duality results for measurable functions, see [nutz][Example 8.1].

3.8 Extensions

We first extend the duality from $\mathbb{C}_b$ to $\mathbb{U}_b$ by a minimax argument and further to $\mathbb{U}_{<p}$ via Lemma 3.8.2.

Proposition 3.8.1. The duality 3.6.5 holds on $\mathbb{C}_b$ if and only if it holds on $\mathbb{U}_b$. 
3.8. EXTENSIONS

Proof. Assume that 3.6.5 holds on \( C_b \) and let \( \eta \in \mathbb{U}_b \). Since the inequality \( \Phi_d \geq \mathcal{E}_{\mathbb{Q}(\mu)} \) holds directly on \( \mathbb{B}_b \), we need to show that

\[
\mathcal{E}_{\mathbb{Q}(\mu)}(\eta) \geq \Phi_d(\eta).
\]

By [dellacheriemeyer][Theorem 49 (c)] we have

\[
E_Q[\eta] = \inf_{\eta \leq \xi \in C_b} E_Q[\xi]
\]

for every \( \eta \in \mathbb{U}_b \) and Radon measure \( Q \in \mathcal{M} \). Clearly, \( \{ \xi \in C_b : \eta \leq \xi \} \) is a convex subset of \( C_b \) and the mapping \( \langle \xi, Q \rangle := E_Q[\xi] \) is bicontinuous bilinear on \( C_b \times \mathcal{M} \). Moreover, by Proposition 3.9.6 below, \( \mathcal{Q}(\mu) \) is a convex, weak* compact subset of \( \mathcal{M} \). Hence, the assumption for a standard minimax argument are satisfied; see e.g. [strasser][Theorem 45.8]. Since \( \Phi_d \) is increasing,

\[
\mathcal{E}_{\mathbb{Q}(\mu)}(\eta) = \sup_{Q \in \mathcal{Q}(\mu)} \inf_{\eta \leq \xi \in C_b} E_Q[\xi] = \inf_{\eta \leq \xi \in C_b} \sup_{Q \in \mathcal{Q}(\mu)} E_Q[\xi] = \inf_{\eta \leq \xi \in C_b} \Phi_d(\xi) \geq \Phi_d(\eta).
\]

So, the duality holds on \( \mathbb{U}_b \).

The finite \( p \)th moment of the terminal marginal \( \mu_T \) allows us to extend the duality from bounded functions to polynomially growing ones. To achieve this we use the pathwise superhedging argument of [mete2] based on the pathwise Doob’s \( L^p \)-maximal inequality of [schachermayer], which we now recall. Let \( p \geq 1 \) and \( x_0, \ldots, x_T \) be non-negative real numbers. Then,

\[
\max_{0 \leq k \leq T} x_k^p \leq \sum_{n=0}^{T-1} h \left( \max_{0 \leq k \leq n} x_k \right) (x_{k+1} - x_k) - \frac{p}{p-1} x_0^p + \left( \frac{p}{p-1} \right)^p x_T^p,
\]  

(3.8.1)

where

\[
h(x) := -\frac{p^2}{p-1} x^{p-1}.
\]  

(3.8.2)

**Lemma 3.8.2.** The power option \( \xi(X) := \|X\|_p^p \) with exponent \( p \geq 1 \) as in (3.3.6) is superreplicable, i.e., \( \Phi_d(\|X\|_p^p) < \infty \). In particular, every \( \xi \in \mathbb{B}_p \) is superreplicable.

**Proof.** Let \( \varepsilon > 0 \) and define a sequence \( (\tau_n^\varepsilon)_{n \in \mathbb{N}} \) of \( \mathbb{F} \)-stopping times by

\[
\tau_0^\varepsilon(\omega) := 0,
\]

(3.8.3)

\[
\tau_n^\varepsilon(\omega) := \inf \{ t > \tau_{n-1}^\varepsilon(\omega) : \|x(t) - \omega(\tau_{n-1}^\varepsilon(\omega))\| \geq \varepsilon \} \land T, \quad n \in \mathbb{N}.
\]

Let \( h \) be given by (3.8.2) and set

\[
h_0^\varepsilon(\omega) := 0,
\]

(3.8.3)

\[
h_n^\varepsilon(\omega) := \left( h \left( \max_{0 \leq k \leq n} \omega_k^1 \left( \tau_k^\varepsilon(\omega) \right) \right), \ldots, h \left( \max_{0 \leq k \leq n} \omega_k^d \left( \tau_k^\varepsilon(\omega) \right) \right) \right).
\]

(3.8.4)

These stopping times and the above functions, define an admissible simple integrand \( H^\varepsilon \in \mathcal{H}_\varepsilon \).

Next, consider the static option given by

\[
g(x) := \left( \frac{p}{p-1} \right)^d \sum_{i=1}^{d} x^p - \frac{pd}{p-1}, \quad \forall x \in \mathbb{R}_+^d.
\]
Then, \( g(X_T) - g_0 \in \mathcal{G} \) where
\[
g_0 := \int_{\mathbb{R}^d} g(x) \, d\mu_T(x) < \infty.
\]

By (3.8.1) and the constructions of the stopping times,
\[
g(X_T) + (H^\varepsilon \cdot X)_T \geq \max_{0 \leq k} X^p_k \geq \left( \|X\|_\infty - \varepsilon \right)^p, \quad \forall X \in \Omega.
\]

Moreover, for sufficiently small \( \varepsilon > 0 \),
\[
\left( (y - \varepsilon)^+ \right)^p \geq \frac{y^p}{2^p} - 1, \quad \forall y \geq 0.
\]

Therefore,
\[
\Phi_d(\|X\|_\infty^p) \leq 2^p(1 + g_0) < \infty.
\]
The monotonicity of \( \Phi_d \) implies that every element of \( \mathbb{B}_p \) is also superreplicable. \( \square \)

The following proposition holds in particular when \( \mathcal{V} \) is \( \mathbb{U} \) or \( \mathcal{V} \) is \( \mathbb{B} \), whence by the duality we refer to 3.6.5 or 3.6.8, respectively. Recall the convention \( \mathcal{V}_{<p} = \mathcal{V}_b \) for the case \( p = 1 \).

**Proposition 3.8.3.** Assume that \( \mathcal{V} \) is closed with respect to finite minima and maxima. If the duality holds on \( \mathcal{V}_b \), then the duality holds on \( \mathcal{V}_{<p} \).

**Proof.** We first record an elementary inequality. For \( 1 \leq q < p \) and \( y \geq K \geq 1 \),
\[
(1 + y^q) \leq \left( 1 + \frac{y^q}{K^q} \right)^{\frac{q-1}{q}} \leq \frac{(1 + y^q)^{\frac{p}{q}}}{K^{p-q}} \leq \frac{2^p}{K^{p-q}} (1 + y^p).
\]

Consider a claim \( \xi \in \mathcal{V}_q \) which is bounded from below. Set
\[
M := \sup_{\omega \in \Omega} \frac{\xi(\omega)}{1 + \|\omega\|_\infty^q} < \infty.
\]

For \( K \geq 1 \), we use the above inequality and the definition of \( M \) to arrive at
\[
\xi(X) \leq M(1 + K^q) \wedge \xi(X) + M(1 + \|X\|_\infty^q) 1_{\|X\|_\infty \geq K}
\leq M(1 + K^q) \wedge \xi(X) + \frac{M 2^p}{K^{p-q}} \|X\|_\infty^p.
\]

Since \( \xi \) is bounded from below, \( M(1 + K^q) \wedge \xi \) is bounded. We now use the sub-linearity, monotonicity of \( \Phi_d \) and the duality on \( \mathcal{V}_b \) to obtain
\[
\Phi_d(\xi) \leq \Phi_d(M(1 + K^p) \wedge \xi) + \frac{M 2^p}{K^{p-1}} \Phi_d(\|X\|_\infty^p)
\leq \mathcal{E}_{Q|\mu}(M(1 + K^q) \wedge \xi) + \frac{M 2^p}{K^{p-1}} \Phi_d(\|X\|_\infty^p)
\leq \mathcal{E}_{Q|\mu}(\xi) + \frac{M 2^p}{K^{p-1}} \Phi_d(\|X\|_\infty^p).
\]
3.9. PROPERTIES OF THE SUPERHEDGING FUNCTIONAL $\Phi_d$

We let $K$ tend to infinity. The result is $\Phi_d(\xi) \leq E_{Q(\mu)}(\xi)$ for every $\xi \in \mathbb{V}_{<p}$ which is bounded below. Since the opposite inequality is immediate, the duality is proved for claims which are bounded from below. To remove this restriction, we deploy the same argument. Indeed, for $K \geq 1$,

$$\xi(X) \geq -M(1 + K^p) \vee \xi(X) - M(1 + \|X\|_\infty^q) 1_{\{|\omega|\|_\infty \geq K\}}$$

$$\geq -M(1 + K^p) \vee \xi(X) - \frac{M 2^\frac{q}{p}}{K^{p-q}} \|X\|_\infty^q.$$ We rewrite the above inequality as

$$\xi_{M,K}(X) := -M(1 + K^p) \vee \xi(X) + \frac{M 2^\frac{q}{p}}{K^{p-q}} \|X\|_\infty^q.$$

Note that the claim $\xi \leq \xi_{M,K} \in \mathbb{V}_q$ and is bounded from below. Hence, we can use the duality that is just proved. The result is the following,

$$\Phi_d(\xi) \leq \Phi_d(\xi_{M,K}) = E_{Q(\mu)}(\xi_{M,K}) \leq E_{Q(\mu)}(\xi) + \frac{M 2^\frac{q}{p}}{K^{p-q}} E_{Q(\mu)}(\|X\|_\infty^q).$$

We again let $K$ to infinity. The result is

$$\Phi_d(\xi) \leq E_{Q(\mu)}(\xi), \quad \forall \xi \in \mathbb{V}_q.$$ Since the opposite inequality is immediate this completes the proof. \qed

**Remark 3.8.4.** Since the integrand (3.8.3) is simple, it is apparent that one can replace $\Phi_d$ with $\Phi_s$, $\Phi_b$, or $\Phi_f$ above; see (3.6.14), (3.6.13) and (3.6.7), respectively.

**Remark 3.8.5.** Proposition 3.8.3 is a moderate extension of the corresponding result of [mete2] and thus extends their duality result therein. Indeed, the authors [mete2] only considered the case of sub-linear growth ($q = 1$), but the sub-polynomial growth of order $q$ determined by the condition $1 \leq q < p$, where $p$ is given by (3.3.6), is sufficient.

The previous proof also shows the following technical result.

**Lemma 3.8.6.** For $Q \in \mathcal{P}(\mu)$ and $\xi \in \mathbb{B}_{<p}$, one has

$$E_Q[\xi] = \lim_{K \to \infty} E_Q[(\xi \land K) \lor (-K)]$$

and

$$\Phi_d(\xi) = \lim_{K \to \infty} \Phi_d((\xi \land K) \lor (-K)).$$

3.9 Properties of the Superhedging Functional $\Phi_d$

This section proves the important technical properties of $\Phi_d$ used in the proof of Theorem 3.6.2. In view of the previous section, we only consider bounded claims.

3.9.1 Sub-differential at the Origin

Recall that the sub-differential at $\xi \in \mathbb{C}_b$ with respect to the dual pairing $\langle \mathbb{C}_b, \mathcal{M} \rangle$ is given by

$$\partial \Phi_d(\xi) := \{ Q \in \mathcal{M} : \Phi_d(\xi) + \langle \eta - \xi, Q \rangle \leq \Phi_d(\eta) \quad \forall \eta \in \mathbb{C}_b \}, \quad \xi \in \text{dom } \Phi_d,$$

and $\partial \Phi_d(\xi)$ is set to be the empty set when $\xi$ is not in the (effective) domain. The main goal of this subsection is to identify the sub-differential at the origin with the family of martingale measures satisfying the given marginals

$$Q(\mu) := \{ Q \in \mathcal{P} : X \text{ is a } (Q, F)\text{-martingale } \& \ Q \circ \pi^{-1} = \mu_t \ \forall t \in \mathbb{T} \}.$$
We first claim that
\[ \partial \Phi_d(0; \mu) = \mathcal{Q}(\mu). \]

**Proof.** We prove the statement in several steps. First observe that under the hypothesis, \( \Phi_d(c) = c \) for every constant \( c \) and \( \Phi_d \) is finite valued. In particular, domain of \( \Phi_d \) is the whole space \( \mathbb{C}_b \) and by definition the sub-differential at the origin is given by
\[ \partial \Phi_d(0) = \{ Q \in \mathcal{M} : E_Q[\eta] \leq \Phi_d(\eta) \ \forall \eta \in \mathbb{C}_b \}, \]
where as before \( E_Q[\eta] := \langle \eta, Q \rangle \). Moreover, by Lemma 3.8.6 the above holds for every \( \eta \in \mathbb{C}_{<p}, \) i.e.,
\[ \partial \Phi_d(0) = \{ Q \in \mathcal{M} : E_Q[\eta] \leq \Phi_d(\eta) \ \forall \eta \in \mathbb{C}_{<p} \}. \]

**Step 1.** We first claim that
\[ \mathcal{Q}(\mu) \subset \partial \Phi_d(0; \mu). \]

Fix \( Q \in \mathcal{Q}(\mu) \). For any \( t \leq T \), set \( \nu_t := Q \circ \pi_t^{-1} \). In view of (3.3.6), \( \nu_t \) has finite \( p \)-moment. Then, one directly verifies that \( E_Q[G] = E_Q[(H \bullet X)_T] = 0 \) for every \( G \in \mathcal{G} \) and a simple integrand \( H \in \mathcal{H}_s \). Let \( H = (H^n) \) be an admissible dynamic strategy. The admissibility implies that there is uniform lower bound which is \( p \) integrable. Since each \( H^n \in \mathcal{H}_s \) and Fatou’s Lemma,
\[ E_Q[(H \bullet X)_T] = E_Q[\lim_n (H^n \bullet X)_T] = \lim_n E_Q[(H^n \bullet X)_T] = 0. \]

Suppose that \( c + G(X) + (H \bullet X)_T \geq \eta \) for some \( c \in \mathbb{R}, \eta \in \mathbb{B}_p \) and admissible dynamic strategy \( H \). Then,
\[ c = E_Q[c + G(X) + (H \bullet X)_T] \geq E_Q[\eta]. \]

Hence, \( E_Q[\eta] \leq \Phi_d(\eta) \) and consequently \( Q \in \partial \Phi_d(0; \mu) \).

**Step 2.** In the opposite direction, we start by showing that \( \partial \Phi_d(0; \mu) \subset \mathcal{P} \).

We fix \( Q \in \partial \Phi_d(0; \mu) \). We need to show that it is positive and \( \mathcal{Q}(1) = 1 \). Indeed, We choose \( \eta \equiv c \) in the definition of the super-differential to conclude that any \( Q \in \partial \Phi_d(0; \mu) \) satisfies, \( E_Q[c] \leq \Phi_d(c) = c \). Hence, \( Q(c) = c \) for any constant. Now, suppose \( \eta \leq 0 \). Since \( \Phi_d \) is increasing, \( \Phi_d(\eta) \leq 0 \). Hence, \( E_Q[\eta] \leq \Phi_d(\eta) \leq 0 \) for every \( \eta \leq 0 \) and the every element of the sub-differential is a probability measure.

**Step 3.** In this step, we show that
\[ \partial \Phi_d(0; \mu) \subset \mathcal{Q}. \]

We fix \( Q \in \partial \Phi_d(0; \mu) \). We know that \( Q \in \mathcal{P} \) and need to show that
\[ E_Q[\gamma(X) \cdot (X_s - X_t)] = 0, \]
for every \( 0 \leq u < s \leq T \) and \( \mathcal{F}_t \)-measurable, bounded vector valued function \( \gamma \). By standard density arguments we may assume that \( \gamma \) is continuous; see e.g. [semimart][Proposition 4.9. and Proposition 4.10]. Also, it suffices to consider the case \( s = T \). Indeed, since \( X \) has right limits at every point,
\[ E_Q[\gamma(X) \cdot (X_T - X_u)] = \lim_{M \to \infty} E_Q[I(u, M)_T], \]
3.9. PROPERTIES OF THE SUPERHEDGING FUNCTIONAL \( \Phi_D \)

where for \( M \geq 1/(T - u) \),

\[
I(u, M)_t := \gamma(X) \cdot M \int_u^{u+1/M} (X_{T_\tau} - X_{T_\tau}) \, d\tau.
\]

The function \( I(u, M)_T = I(u, M)_T(\omega) \) is \( S^* \)-continuous and it is growing linearly, cf. [semimart][Example 5.11]. Hence, \( I(u, M)_T \in \mathbb{C}_{<p} \) and therefore, the sub-differential inequality holds, i.e.,

\[
E_Q[I(u, M)_T] \leq \Phi_d(I(u, M)_T).
\]

For each \( M \) there is a sequence \( (H^k_{u, M})_{k \in \mathbb{N}} \subset \mathcal{H} \) satisfying

\[
(H^k_{u, M} \cdot X)_t = \gamma(X) \cdot \sum_{i=1}^{k} M \int_u^{u+1/M} (X_{T_\tau} - X_{(u+1/M)_\tau}) \, d\tau,
\]

\( t \in [0, T] \).

The integral for the dynamic strategy \( H_{u, M} := (H^k_{u, M}) \) is by definition given by

\[
(H_{u, M} \cdot X)_t = \lim_{k \to \infty} (H^k_{u, M} \cdot X)_t.
\]

Since \( X \) is càdlàg, the above limit is equal to \( I(u, M)_T \). Hence, \( I(u, M)_T \in \mathcal{K} \) and consequently, \( \Phi_d(I(u, M)_T) = 0 \). We combine the above to arrive at

\[
E_Q[\gamma(X) \cdot (X_T - X_u)] = \lim_{M \to \infty} E_Q[I(u, M)_T] \leq 0.
\]

Since above holds for all \( \gamma \) regardless its sign, we conclude that the above expectation is in fact zero for all \( \gamma \).

**Step 4.** In this final step, we show that \( Q \circ \pi_t^{-1} = \mu_t \), for every \( t \in \mathbb{T} \). Indeed, for any \( g \in C(\mathbb{R}_+^d) \) and \( t \in \mathbb{T} \), set

\[
\xi_{g,t}(\omega) := g(\omega_t) - \int g \, d\mu_t.
\]

Then, \( \xi_{g,t} \in \mathcal{G} \). Hence, for any \( Q \in \partial \Phi_d(0; \mu) \),

\[
E_Q[\xi_{g,t}] \leq \Phi_d(\xi_{g,t}; \mu) = 0.
\]

We write the above inequality as follows

\[
\int g \, d[Q \circ \pi_t^{-1}] = E_Q[\xi_{g,t}] + \int g \, d\mu_t \leq \int g \, d\mu_t.
\]

Since above holds for every \( g \in C(\mathbb{R}_+^d) \) and \( t \in \mathbb{T} \) was arbitrary, we conclude that \( Q \circ \pi_t^{-1} = \mu_t \), for every \( t \in \mathbb{T} \).

3.9.2 Continuity of the Superhedging Functional

The goal of this subsection is to prove the \( \beta_0 \)-continuity of the superreplication functional. While the continuity of \( \Phi_d \) with respect to the uniform norm is immediate, the continuity in the \( \beta_0 \)-topology is one of the key steps in establishing the duality for continuous claims. Indeed, the topological dual of \( \mathcal{C}_b \) with the \( \beta_0 \)-topology is the set of countably additive measures where as with the uniform norm the dual is larger; see Appendix and reference therein for elaboration. Since we want to establish duality with countably additive measures, it is essential that continuity is proved in the \( \beta_0 \)-topology.
Proposition 3.9.2. The superreplication functional $\Phi_d$ is $\beta_0$-continuous on $C_b$.

Before the proof of this proposition we prove several lemmata. First result enables us to essentially bound the underlying asset process.

Lemma 3.9.3. For all $K > 0$,

$$\Phi_d(1_{\{\|X\|_\infty > K\}}) \leq \frac{1}{K} \sum_{i=1}^d \int x^i d\mu_T.$$  \hspace{1cm} (3.9.1)

Proof. We use the pathwise Doob’s maximal $L^1$-inequality of [hobson] and its connection to the superhedging price [touzi]. Indeed, as a special case of [hobson][Lemma 2.3], for any $K > 0$, we have

$$1_{\{\|X_i\|_\infty > K\}} \leq \frac{X^i_T}{K} + 1_{\{\|X_i\|_\infty > K\}} - \frac{X^i_{\tau_1}}{K},$$

for $i = 1, \ldots, d$. The second term can be superhedged with a zero initial capital and the admissible simple strategy given by

$$h^i_1(\omega) = -\frac{1}{K}, \quad \tau_1(\omega) := \inf\{t : \omega^i(t) \in [K, \infty) \wedge T \} \quad \text{and} \quad \tau_2(\omega) = T.$$ 

By right continuity, on the set $\{\|X_i\|_\infty > K\}$ we have $X^i_{\tau_1} \geq K$. Consequently,

$$(H \cdot X)_T = -\frac{1}{K} (X^i_T - X^i_{\tau_1}) \geq \frac{1}{K} (K - X^i_T) 1_{\{\|X_i\|_\infty > K\}}.$$ 

Hence,

$$\Phi_d \left(1_{\{\|X_i\|_\infty > K\}} \frac{K - X^i_T}{K}\right) \leq 0.$$ 

Moreover,

$$\Phi_d(X^i_T) = \int x^i d\mu_T.$$ 

These imply the estimate 3.9.1. \hfill \Box

Next we prove a pathwise Doob’s up-crossing inequality which is essentially due to [vovk]; see Lemma 1 and Remark 2 there in.

Lemma 3.9.4. Let $i \in \{1, \ldots, d\}$, $0 < a < b$ be given. Then, there exists an admissible simple strategy $H^{a,b} \in H_s$ satisfying $a + (H^{a,b} \cdot X) \geq 0$ and

$$a + (H^{a,b} \cdot X)_T(\omega) \geq (b - a) N^{a,b}(\omega^i), \quad \forall \omega \in \Omega.$$ 

In particular, for any $0 < a < b$ and $C > 0$,

$$\Phi_d(1_{\{N^{a,b}(\omega^i) > C\}}) \leq \frac{a}{C(b-a)}.$$  \hspace{1cm} (3.9.2)

Proof. For $k \in \mathbb{N}$, set $I_k := [0, a]$ when $k$ is odd and $I_k := [b, \infty)$ when $k$ is even. Set $\tau_0 = 0$. For $k \geq 1$, define a sequence of stopping times recursively by

$$\tau_k(\omega) := \inf\{t \geq \tau_{k-1} : \omega^i(t) \in I_k \} \wedge T$$
Also let \(d\) with the same argument, the consecutive hitting times \(k\) with the usual convention that the infimum over an empty set is infinity. Next let \(h_k := 1\) when \(k\) is odd and set \(h_k := 0\) for \(k\) even. Let \(n(\omega)\) denote the largest integer for which \(\tau_n < T\). Then, when \(n(\omega)\) is even, 
\[
a + (H^{a,b} \cdot X)_T(\omega) \geq (b - a)N^{a,b}(\omega)
\]
and for odd \(n(\omega)\),
\[
(H^{a,b} \cdot X)_T(\omega) \geq \delta(b - a)N^{a,b}(\omega) - a.
\]
Hence, in both cases \(H^{a,b}\) satisfies the claimed inequalities.

Moreover, since \(X\) is càdlàg, and \(I_1\) is closed, \(\tau_1\) is a \(\mathbb{F}\)-stopping time. Recursively, with the same argument, the consecutive hitting times \(\tau_k, k \geq 2\), are also \(\mathbb{F}\)-stopping times. Hence, \(H^{a,b}\) is an admissible, simple strategy. Then, the admissible simple strategy \((1/C(b - a))H^{a,b}\) with initial capital \(a/C(b - a)\) superhedges \(1_{\{N^{a,b}(\omega) > C\}}\). This proves 3.9.2.

We now use the above two results to prove an essential localization result.

**Lemma 3.9.5.** There exist an increasing sequence of compact sets \((K^n)_{n \in \mathbb{N}}\) such that \(\Phi_d(1_{\Omega \setminus K^n})\) decreases to zero as \(n\) approaches to infinity.

**Proof.** For a positive integer let \(B^n := \{\omega : \|\omega\|_\infty \leq n\}\). Let \(D(n)\) be a dense countable subset of \((0, n]\) and \(C_i^{a,b} > 0\) be a sequence of real numbers. Then, the set

\[
K := B^n \cap \bigcap_{i=1}^{d} \bigcap_{a < b \in D(n)} \{\omega : N^{a,b}(\omega) \leq C_i^{a,b} < \infty\}
\]

is an \(S^*\)-compact subset of \(\Omega\); cf. [seminart][Corollary 5.11.].

Since \(\Omega \setminus K \subset B^n \setminus K \cap \Omega \setminus B^n\), the sub-linearity of \(\Phi_d\) implies that

\[
\Phi_d(1_{\Omega \setminus K}) \leq \Phi_d\left(1_{B^n \setminus K}\right) + \Phi_d\left(1_{\Omega \setminus B^n}\right).
\]

Moreover, by Lemma 3.9.3, \(\Phi_d\left(1_{\Omega \setminus B^n}\right)\) tends to zero as \(n\) goes to infinity. We continue by choosing the sequence \(C_i^{a,b}\) so that the second term also converges to zero as \(n\) gets larger. Indeed, let \(\{(a_n^k, b_n^k) : k = 1, 2, \ldots\}\) be a denumerization of the countable set \(\{(a, b) \in D(n) \times D(n) : a < b\}\). Set

\[
C_i^{a_n^k, b_n^k} := 2^{n+k} \frac{d}{b_n^k - a_n^k} =: C_n^k, \quad k = 1, 2, \ldots
\]

Also let \(K_n\) be the compact set given by 3.9.3 with the sequence \(\{C_n^k\}_{k=1,2,\ldots}\). By 3.9.2,

\[
\Phi_d\left(1_{\{\omega \in B_n : N^{a_n^k, b_n^k}(\omega) > C_n^k\}}\right) \leq \frac{a_n^k}{C_n^k(b_n^k - a_n^k)} \leq \frac{a_n^k}{d} \frac{n}{2^{k+n}} \leq \frac{1}{d} 2^{k+n}.
\]

Moreover, by Lemma 3.9.4 there is \(H_n^k \in \mathcal{H}_s\) such that

\[
\frac{1}{d} 2^{k+n} + (H_n^k \cdot X)_T \geq 1_{\{\omega \in B_n : N^{a_n^k, b_n^k}(\omega) > C_n^k\}},
\]

and

\[
\frac{1}{d} 2^{k+n} + (H_n^k \cdot X)_t \geq 0, \quad \forall t \in [0, T].
\]
For $M > 0$ set,
\[
H_n^M := \sum_{i=1}^d \sum_{k=1}^M H_{ik}^n.
\]
Then,
\[
(H_n^M \cdot X)_t \geq -\sum_{i=1}^d \sum_{k=1}^M \frac{1}{2^{k+n}} \geq -2^{-n}, \quad \forall M > 0,
\]
and
\[
2^{-n} + \lim_{M \to \infty} (H_n^M \cdot X)_T \geq \sum_{i=1}^d \sum_{k=1}^\infty 1_{\{\omega \in B_n : N_{ik}^n \cap C_k^n(\omega) > C_k^n\}} \geq 1_{B^n \setminus K_n}.
\]
Hence, for each $n$, the dynamic strategy $(H_{ik}^n)$ is admissible and therefore, $\Phi_d(1_{B^n \setminus K_n}) \leq 2^{-n}$.

Proof of Proposition 3.9.2. By Lemma 3.9.5, there exists a sequence of compact sets $(K_j)_{j \in \mathbb{N}_0}$ with $K_0 = \emptyset$ such that
\[
\Phi(1_{K_j \setminus K_{j-1}}) \leq 4^{-j}
\]
for each $j \in \mathbb{N}$. Define
\[
\eta^n := \sum_{j=1}^n 2^{-j} 1_{K_j \setminus K_{j-1}}
\]
for $n \in \mathbb{N}$ and $\eta := \lim_{n \to \infty} \eta^n$. Since $\Phi$ is monotone and sublinear, for $\xi \in C_b(\Omega)$, we get
\[
|\Phi(\xi 1_{K^n})| = \left| \Phi \left( \left| \frac{\xi}{\eta^n} 1_{\eta^n} \right| 1_{K^n} \right) \right| \leq \|\xi\eta^n\|_\infty \Phi \left( \frac{1}{\eta^n} 1_{K^n} \right)
\]
\[
\leq \|\xi\eta^n\|_\infty \sum_{j=1}^n 2^j \Phi(1_{K_j \setminus K_{j-1}}) \leq \|\xi\eta^n\|_\infty (1 - 2^{-n}).
\]
Consequently,
\[
|\Phi(\xi)| \leq \limsup_{n \to \infty} \|\Phi(\xi 1_{K^n})\| + \|\xi\|_\infty \Phi(1_{\Omega \setminus K^n}) \leq \|\xi\|_\infty,
\]
where $\eta \in B_0$, so, the functional $\Phi$ is $\beta_0$-continuous on $C_b(\Omega)$.

3.9.3 Compactness of the Sub-differential $\partial \Phi_d(0)$

In this subsection, we combine the results of the previous two subsections and establish a connection to the compactness of martingale measures, studied in [seminart] via Prohorov’s theorem.

Proposition 3.9.6. The family $Q(\mu)$ is weak$^*$ compact in the pairing $(C_b, M)$.

Proof. By the Banach-Alaoglu-Bourbaki theorem, the sub-differential $\partial \Phi(0)$ is weak$^*$ compact, if $\Phi$ is continuous at 0 and 0 $\in$ dom $\Phi$. Thus, the weak$^*$ compactness of $Q(\mu)$ follows from the sub-differential characterization $Q(\mu) = \partial \Phi_d(0)$ in conjunction with the $\beta_0$-continuity of $\Phi_d$, i.e., from Proposition 3.9.1 and Proposition 3.9.2, respectively.

As a curiosity we obtain the following as a corollary; cf. [cheridito][Pro. 1.1].
Corollary 3.9.7. For every \( \xi \in \mathbb{U}(\Omega) \) bounded from above, we have
\[
\mathcal{E}_{Q(\mu)}(\xi) = \max_{Q \in Q(\mu)} E_Q[\xi].
\]

Proof. An upper semicontinuous functional attains its maximum on a compact domain. The functional \( Q \mapsto E_Q[\xi] \) is upper semicontinuous, for every \( \xi \in \mathbb{U}(\Omega) \) bounded from above; see e.g. [bogachev][Proposition 8.9.8.]. \( \square \)

3.10 Properties of the Superhedging Functional

We start with the basic properties of the functional \( \Phi_f \).

Proposition 3.10.1. Let \( \Phi_f \) be given by (3.6.7). Then, the following are true:

(a) \( \Phi_f \) is increasing, translation invariant and sublinear.

(b) \( \partial \Phi_f(0; \mu) = Q(\mu) \).

(c) \( \Phi_f \) is sequentially continuous from below on \( \mathbb{B}_p \).

Proof. We use transfinite recursion and induction in the proofs. In the Appendix below, we provide a brief introduction and the necessary references. To simplify the notation, for \( x \in \mathbb{R} \) set
\[
K_x := \{ x + \zeta : \zeta \in K \}.
\]

Then, it is clear that the \( K_0 = K \) and \( K_x = x + K_0 \).

Proof of (a): The set \( K_x \) can be characterized by transfinite recursion over countable ordinals. Indeed, we let \( K_0^x := K_x \). For every countable ordinal \( \alpha > 0 \), define \( K_\alpha^x \) recursively by
\[
(\zeta^n)_{n \in \mathbb{N}} \subset \bigcup_{\beta < \alpha} K_\beta^x \quad \text{and} \quad \exists \lambda \in \mathbb{B}_p^+ \text{ s.t. } \zeta^n \geq -\lambda \implies \liminf_{n \to \infty} \zeta^n \in K_\alpha^x.
\]
The set \( K_x \) is given as \( \bigcup_\alpha K_\alpha^x \) where the union is taken over all countable ordinals. The functional \( \Phi_f \) is increasing by the definition.

It is also clear that \( K_x = x + K_0 \). This implies that \( \Phi_f \) is translation invariant. Also, it directly follows that
\[
K_{\alpha x} = \alpha K_x \quad \text{and} \quad K_{x+y} = K_x + K_y
\]
for every \( \alpha > 0 \) and \( x, y \in \mathbb{R} \). These imply that \( \Phi_f \) is sublinear.

Proof of (b): We first prove that the elements of \( K_x \) are measurable. Indeed, \( \liminf_{n \to \infty} \zeta^n \in \mathbb{B} \) whenever \( (\zeta^n)_{n \in \mathbb{N}} \subset \mathbb{B} \). Since \( K \subset \mathbb{B} \), it follows from transfinite induction that \( K \subset \mathbb{B} \).

Similarly, for \( Q \in Q(\mu) \), we have
\[
E_Q[\zeta] \leq 0
\]
for every \( \zeta \in K_0 \). Now consider a sequence \( (\zeta^n)_{n \in \mathbb{N}} \subset \mathbb{B} \) for which \( E_Q[\zeta^n] \leq 0 \) and \( \zeta^n \geq -\lambda \) for some \( \lambda \in \mathbb{B} \) with \( \lambda^+ \in \mathbb{B}_p \). Then, by Fatou’s lemma,
\[
E_Q[\liminf_{n \to \infty} \zeta^n] \leq \liminf_{n \to \infty} E_Q[\zeta^n] \leq 0.
\]
Consequently, a transfinite induction argument implies that
\[
E_Q[\zeta] \leq 0, \quad \forall \zeta \in K_0, \quad Q \in Q(\mu).
\]
Therefore, $E_Q[\eta] \leq \Phi_d(E_Q[\eta]) \leq \Phi_d(\eta)$ for every $\eta \in \mathbb{B}_p$. Hence, $Q(\mu) \subset \partial \Phi_f(0; \mu)$.

To prove the opposite implication, we observe that $\Phi_f(\eta; \mu) \leq \Phi_d(\eta; \mu)$ for every $\eta \in \mathbb{B}_b$ and $\mathbb{C}_p \subset \mathbb{B}_p$. Thus, it follows directly from the definition of sub-differential and Proposition 3.9.1 that we have

$$
\partial \Phi_f(0; \mu) \subset \partial \Phi_d(0; \mu) = Q(\mu).
$$

**Proof of (c):** Let $(\eta^n)_{n \in \mathbb{N}}$ and $\eta$ be in $\mathbb{B}_p$. Suppose that $\eta^n \uparrow \eta$. Since $(\eta^n)_{n \in \mathbb{N}} \subset \mathbb{B}$ is uniformly bounded from below by $\eta_1$ whose positive part $(\eta_1^+)$ is in $\mathbb{B}_p$, each $\eta^k$ is superreplicable. Fix $\varepsilon > 0$. Then, for all sufficiently large $k$,

$$
\Phi_f(\eta^k; \mu) \leq x^\varepsilon := \liminf_{n \to \infty} \Phi_f(\eta^n; \mu) - \varepsilon.
$$

Hence, there are $\zeta^k \in \overline{K}_{x^\varepsilon}$ superreplicating $\eta^k$. Set

$$
\zeta := \liminf_{k \to \infty} \zeta^k.
$$

Since $\zeta^k \geq \eta^k$ and $\eta^k \uparrow \eta$, we conclude that $\zeta \geq \eta$. Moreover, by the definition of closure $\zeta \in \overline{K}_{x^\varepsilon}$. Therefore, $\Phi_f(\eta; \mu) \leq x^\varepsilon$. Since $\varepsilon > 0$ is arbitrary, we conclude that

$$
\Phi_f(\eta; \mu) \leq \liminf_{n \to \infty} \Phi_f(\eta^n; \mu).
$$

The opposite inequality follows directly from the monotonicity of $\Phi_f$. \qed
Chapter 4

Appendix

The appendix collects auxiliary results and technical definitions used in the main part.

4.0.1 Convex duality in topological vector spaces

We cover the required preliminary results from convex analysis and topology. We assume the usual axioms of set theory including the axiom of choice. The proofs are omitted; consult e.g. [bogachevtvs], [zalinescu].

A topological vector space $X$ over the scalar field $\mathbb{R}$ is a vector space on which the addition of vectors and the multiplication of vectors by scalars are continuous operations on the respective product spaces i.e. the topology is compatible with vector space structure. A topological vector space is locally convex if the topology possesses a base of convex neighborhoods of zero. Any such topology can be defined by a family of seminorms and any family of seminorms defines such topology.

Two vector spaces $X$ and $Y$ form a dual pair $\langle X, Y \rangle$ if there exists a bilinear functional $\langle x, y \rangle$ on $X \times Y$ such that the space $Y$ separates the points of $X$, $x \neq 0$ implies $\langle x, y \rangle \neq 0$ for some $y \in Y$ for every $x \in X$, and the space $X$ separates the points of $Y$, similarly.

For a topological vector space $X$, the symbol $X'$ denotes the topological dual to $X$ that is the vector space of all continuous linear functionals on $X$. If $X$ is a locally convex space and $Y = X'$, then the topology $\sigma(X, X')$ generated on $X$ by $X'$ is called the weak topology in $X$; note that $(X, \sigma(X, X')) = X'$. The topology $\sigma(X, X')$ is defined similarly and usually called the weak* topology. We mention that the topological dual $X'$ is (often) much smaller space than the algebraic dual to $X$ that consists of all linear functionals on $X$, denoted by $X^*$. For a dual pair $\langle X, Y \rangle$, among the locally convex topologies on $X$ with a given topological dual $Y$ there exists the strongest one, called the Mackey topology and denoted by $\tau(X, Y)$.

A Hausdorff locally convex space $X$ is in natural duality with its topological dual space $Y = X'$ under the canonical bilinear form $\langle x, y \rangle = y(x)$, $x \in X$, $y \in Y$.

Example 4.0.1. On a completely regular topological space, $\langle C_b, \mathcal{M}_r \rangle$ is a dual pair under the operation of integration and

\[ \sigma(C_b, \mathcal{M}_r) \subset \beta_0 \subset \tau(C_b, \mathcal{M}_r). \]

Various necessary conditions for the equivalence $\beta_0 = \tau(C_b, \mathcal{M}_r)$ are provided in [sentilles].

Let $X$ be a topological vector space, $H$ be a closed hyperplane of $X$ and $A$ and $B$ be subsets of $X$. We say that $H$ separates $A$ and $B$ if $A$ and $B$ are contained in the different
CHAPTER 4. APPENDIX

closed halfspaces of $H$. If $A$ and $B$ are contained in the different open halfspaces defined $H$, we say that $H$ strictly separates $A$ and $B$.

The following assertion is the geometric version of the Hahn-Banach theorem, Eidelheit's separating hyperplane theorem. It follows from the axiom of choice, but is not true if the axiom of choice is removed.

**Theorem 4.0.2.** Let $A$ and $B$ be two nonempty convex subsets of the topological vector space $X$. If $\text{int } A \neq \emptyset$ and $A \cap \text{int } A = \emptyset$ then there exists non-zero $y \in X'$ and $c \in \mathbb{R}$ such that

$$\forall a \in A \\forall b \in B : \langle a, y \rangle \leq c \leq \langle b, y \rangle,$$

or equivalently, $\sup y(A) \leq \inf y(B)$.

One can further state the conditions for strict separation.

**Theorem 4.0.3.** Let $A$ and $K$ be nonempty disjoint convex subsets of a real locally convex space $X$ such that $A$ is closed and $K$ is compact. Then, there exists non-zero $y \in X'$ and $c_1, c_2 \in \mathbb{R}$ such that

$$\forall a \in A \\forall b \in B : \langle a, y \rangle \leq c_1 < c_2 \leq \langle b, y \rangle,$$

or equivalently, $\sup y(A) < \inf y(B)$.

Let $(X,Y)$ be a dual pair of two locally convex spaces and $f$ be an extended real valued functional on $X$. Set

$$f^*(y) := \sup \{ \langle x, y \rangle - f(x) \mid x \in X \}, \ y \in Y.$$ 

The functional $f^*$ is the *Fenchel conjugate* of $f$. The conjugate of a functional $h$ on $Y$ is defined similarly:

$$h^*(x) := \sup \{ \langle x, y \rangle - h(x) \mid y \in Y \}, \ x \in X.$$ 

Invoking the strict separating hyperplane theorem for a closed convex proper epigraph of functional and an a point on the graph of its bi-conjugate we arrive at the Fenchel-Moreau theorem. The conjugation defines an isomorphism between proper, lower semicontinuous, convex functionals on $X$ and $Y$. This result is the basis of our study of martingale optimal transport duality.

**Theorem 4.0.4.** Assume that $f$ is a proper, lower semicontinuous, convex functional on $E$. Then

$$f^{**} := (f^*)^* = f.$$ 

Consider a functional $f$ on $X$ and let $\bar{x} \in X$ such that $f(\bar{x}) \in \mathbb{R}$. An element $y \in Y$ is a *subgradient* of $f$ at $\bar{x} \in X$ if

$$\forall x \in X : \langle x - \bar{x}, y \rangle \leq f(x) - f(\bar{x}). \quad (4.0.1)$$

The set $\partial f(\bar{x})$ of all the subgradients of $f$ at $\bar{x}$ is called the *subdifferential* of $f$ at $\bar{x}$. If this set is non-empty, we say that $f$ is *subdifferentiable* at $\bar{x}$.

For $A \subset Y$, the indicator functional of $A$ on $Y$ is $\delta_A(y) = 0$, for $y \in A$, and $\delta_A(y) = \infty$, otherwise. The support functional of $A$ on $X$ is $\sigma_A(x) = \sup_{y \in A} \langle x, y \rangle$, for $x \in X$.

If the functional is sublinear i.e. its epigraph is a convex cone, then the bi-conjugate takes a particularly convenient form.
Corollary 4.0.5. Let $f$ be sublinear. Then

$$f^* = \delta_{\partial f(0)} \quad \text{and} \quad f^{**} = \sigma_{\partial f(0)},$$

where $\partial f(0) \neq \emptyset \iff f$ is lower semicontinuous at 0.

The subdifferential is always weak$^*$ closed convex (possibly empty) set. The subdifferential inequality (4.0.1) should be though as a generalization of the polar inequality (4.0.2).

The (absolute) polar of a set $A \subset X$ in $Y$ is the set

$$A^\circ = \{ y \in Y : |\langle x, y \rangle| \leq 1 \ \forall x \in A \}.$$  \hfill (4.0.2)

The following important property of polars is known as the Banach-Alaoglu-Bourbaki theorem. It is the main source of compact sets in locally convex spaces.

Theorem 4.0.6. Let $X$ be a locally convex space. Then the polar of any neighborhood of zero in $X$ is weak$^*$ compact.

Let $X$ be a locally convex space. A subset $B$ in $X$ is called equicontinuous if, for every $\varepsilon > 0$, there exists a neighborhood of zero $V \subset X$ such that $|u(x)| < \varepsilon$ for all $x \in V$, $u \in B$. The equicontinuous sets in the topological dual $X'$ are precisely the subsets of the polars of neighborhoods of zero in $X$. Thus,

Corollary 4.0.7. Let $X$ be a locally convex space. Then every equicontinuous set of $X'$ is weak$^*$ compact.

Example 4.0.8. On a completely regular topological space, a subset $B \subset M$ is $\beta_0$-equicontinuous if and only if the $B$ is bounded in total variation and tight.

We obtain the following as a corollary of the Banach-Alaoglu-Bourbaki theorem.

Theorem 4.0.9. Let $f$ be a proper, lower semicontinuous, convex functional on $X$. If $f$ is continuous at $\varpi \in \text{dom } f$, then $\partial f(\varpi)$ is nonempty weak$^*$ compact.

The previous theorem suggests a method for constructing weak$^*$ compact sets. The weak$^*$ compactness is central for extending the martingale optimal transport duality.

Finally, we recall the classical minimax theorem that is another ingredient of our extensions. For this, consider two nonempty sets $A$ and $B$ and a function $f$ on $A \times B$. It is apparent that

$$\sup_{x \in A} \inf_{y \in B} f(x, y) \leq \inf_{y \in B} \sup_{x \in A} f(x, y).$$

Results establishing the equality of two quantities are known as minimax theorems.

Theorem 4.0.10. Let $X$ and $Y$ be locally convex spaces, $A \subset X$ be a nonempty convex set and $B \subset Y$ be a nonempty convex compact set. Let also $f$ be a function on $A \times B$ such that $f(\cdot, y)$ is concave and upper semicontinuous for every $y \in B$ and $f(x, \cdot)$ is convex and lower semicontinuous for every $x \in A$. Then

$$\sup_{x \in A} \inf_{y \in B} f(x, y) = \inf_{y \in B} \sup_{x \in A} f(x, y).$$

Due to compactness, the infima are in fact minima.
4.0.2 Regularity and compactness

Separation axioms

A topological space $X$ is called a Hausdorff (separated) space or a $T_2$-space if, for every two distinct points $a \neq b$ in $X$, there exists disjoint open sets $A, B$, $A \cap B = \emptyset$, such that $a \in A$, $b \in B$. We consider only Hausdorff topological spaces. In a Hausdorff space, singletons are closed. A topological space $X$ is called completely regular space if for every closed set $F \subset X$ and every point $x \not\in F$, there exists a continuous function $f : X \to [0,1]$ such that $f(x) = 0$ and $f = 1$ on $F$. A Hausdorff completely regular space is called a Tychonoff space or a $T_{3\frac{1}{2}}$-space. In particular, every topological vector space is completely regular. A Hausdorff space is called normal or $T_4$ space if for every two disjoint closed sets $A$ and $B$ in $T$ there exists a pair of disjoint open sets $U$ and $V$ such that $A \subset U$, $B \subset V$. A Hausdorff space is called perfectly normal or $T_6$ space, if every closed set has the form $f^{-1}(0)$ for some continuous function $f$. Equivalently, a Hausdorff space is perfectly normal if the space is normal and every closed set is a $G_\delta$ set i.e. every closed set can be expresses as a countable intersection of open sets.

We denote by $\mathcal{B}(X)$ the Borel $\sigma$-algebra on $X$, that is, the smallest $\sigma$-algebra containing all open sets. If $X$ is perfectly normal, then every Borel set can be constructed from open sets by transfinite recursion of the operations of in union and intersection over countable ordinals. The Baire $\sigma$-algebra $\mathcal{Ba}(X)$ the $\sigma$-algebra is generated by the sets $\{ f > 0 \}$, where $f \in C_b(X)$. It is clear that $\mathcal{Ba}(X) \subset \mathcal{B}(X)$. For all perfectly normal spaces this inclusion is an equality, but in the general case it is strict (even if the space is compact).

Compactness and related notions

A cover of a set is any collection of sets whose union contains this set. A subset of a topological space $X$ is called compact if every cover of this set by open sets contains a finite subcover. A subset of a topological space is called relatively compact if its closure is compact, or equivalently, the set is a subset of a compact set.

A partial order is a reflexive, antisymmetric and transitive binary relation and a partially ordered set $I$ is called directed, if for every two elements $s, t \in I$ there exists an element $u$ such that $s \leq u$ and $t \leq u$. A net in $X$ is a family of element of $X$ indexed over a directed partially ordered set.

In particular, the set of neighborhoods of a given point in a topological space $X$ partially ordered by the inverse inclusion form a direct set whereas the set of open sets partially ordered by the inverse inclusion is not directed.

We may thus re-express compactness in terms of nets. A subset of a topological space is compact if and only if every net of its elements contains a subnet converging in it.

Let us introduce some useful concepts related to compactness.

A set is called sequentially compact if every infinite sequence of its elements contains a subsequence converging to an element in the set. A set is called relatively sequentially compact if every infinite sequence of its elements contains a subsequence converging to some element in the space.

It is known that for subsets of metric spaces the compactness and sequential compactness are equivalent and the same is true for the corresponding relative notions. In general case, one does not imply another.
Example 4.0.11. Let $X$ be the product of the continuum of copies of the real line, realized as the space of all real functions on the interval $[0, 1]$, endowed with the topology of pointwise convergence. The subset $K = \{ f \in X : \sup_{t \in [0, 1]} |f(t)| \leq 1 \}$ is compact, but not sequentially compact, since the sequence $f_n(t) = \sin(nt)$ in $K$, does not contain a convergent subsequence. Further, the subset $K_0$ of $K$ consisting of all function in $K$ not vanishing on at most many points on $[0, 1]$ is sequentially compact, but not compact.

The following is a key concept in our construction of compact sets.

A Hausdorff topological space $X$ is said to be angelic if every set $S \subset X$ with the property that every infinite sequence of its elements has a limit point in $X$, has also the following two properties: $S$ is relatively compact and each point in the closure of $S$ is the limit of some sequence in $S$. In angelic spaces, the properties of compactness and sequential compactness coincide. In addition, the closure of a relatively compact set is exhausted by the limits of sequences of points in this set. For example, every metric space is angelic. Moreover, if a regular space $X$ can be continuously injected into an angelic space $Y$, then $X$ is also angelic.

4.0.3 Topologies on the Skorokhod space

We recall the definitions of Jakubowski’s $S$-topology, the Meyer-Zheng pseudo-path topology and Skorokhod’s $J^1$-metric topology. We define each topology separately on $\mathbb{D}([0, T]; \mathbb{R}^d)$, for $T < \infty$, and $\mathbb{D}([0, \infty]; \mathbb{R}^d)$. In particular, we use a formal definition of the Meyer-Zheng pseudo-path topology (MZ) that takes into account the fluctuations of the terminal value in the case of a finite time-horizon. The space $\mathbb{D}([0, \infty]; \mathbb{R}^d)$ is regarded as a product space $\mathbb{D}([0, \infty]; \mathbb{R}^d) = \mathbb{D}([0, \infty]; \mathbb{R}^d) \times \mathbb{R}^d$, where the space $\mathbb{R}^d$ is endowed with the Euclidean topology.

The $S$-topology

Jakubowski’s $S$-topology, introduced in [jakubowski], is a sequential topology. The following definition of the $S$-convergence on $\mathbb{D}([0, T]; \mathbb{R})$ is taken from [jakubowski]; the multi-dimensional version can be found in [jakubowski4].

Definition 4.0.12. On $\mathbb{D}([0, T]; \mathbb{R}^d)$, we write $\omega_n \rightarrow_S \omega_0$, if, for every $i \leq d$, for every $\varepsilon > 0$, one can find $(\nu_n^{i, \varepsilon})_{n \in \mathbb{N}_0} \subset \mathbb{V}([0, T])$ such that

$$
\|\omega_n^{i, \varepsilon} - \nu_n^{i, \varepsilon}\|_{\infty} \leq \varepsilon, \forall n \in \mathbb{N}_0, \quad \nu_n^{i, \varepsilon} \rightarrow_{w^*} \nu_0^{i, \varepsilon}, \quad \text{as } n \rightarrow \infty,
$$

where the convergence $\rightarrow_{w^*}$ is in the weak* topology on $\mathbb{V}([0, T])$, which can be identified with the Banach dual of $\mathbb{C}([0, T])$, under the uniform norm.

The following definition of the $S$-convergence on $\mathbb{D}([0, \infty]; \mathbb{R}^d)$ is taken from [jakubowski4].

Definition 4.0.13. On $\mathbb{D}([0, \infty]; \mathbb{R}^d)$ we write $\omega_n \rightarrow_S \omega_0$, if, for every $i \leq d$, one can find a sequence of positive real numbers $(T^r)_{r \in \mathbb{N}}$, increasing to $\infty$, such that

$$
[\omega_n^{i}]^{T^r} \rightarrow_S [\omega_0^{i}]^{T^r}, \quad \text{for every } r \in \mathbb{N},
$$

where $[\omega^{i}]^{T^r}$ denotes the restriction of a path $\omega^{i} \in \mathbb{D}([0, \infty]; \mathbb{R})$ on $\mathbb{D}([0, T^r]; \mathbb{R})$.

A topological convergence is obtained by requiring that every subsequence admits a further $S$-convergent subsequence; see [jakubowski4][Theorem 6.3]. The following definition for the $S$-topology on the Skorokhod space $\mathbb{D}([0, T]; \mathbb{R}^d)$ and $\mathbb{D}([0, \infty]; \mathbb{R}^d)$ are taken from [jakubowski] and [jakubowski4], respectively.
**Definition 4.0.14.** The $S$-topology is the topology generated on the Skorokhod space by the subsequential $S$-convergence.

The Skorokhod space endowed with the $S$-topology is a Hausdorff ($T_2$) space and a stronger separation axiom is an open problem. A weak separation axiom is a well-known issue for topologies defined via the subsequential convergence (KVPK recipe); see [jakubowski4] for elaboration. The difficulties encountered in establishing the regularity of the $S$-topology are explained in [jakubowski][Rem. 3.12].

The $\Sigma$-topology

Jakubowski’s $\Sigma$-topology was introduced in [jakubowski4]. The Skorokhod space endowed with $\Sigma$ is a locally convex vector space. Following [jakubowski4], we start by defining an auxiliary mode of convergence $\rightarrow_\tau$ on the space of continuous functions of finite variation $A([0,T];\mathbb{R}) := C([0,T];\mathbb{R}) \cap \mathbb{V}([0,T];\mathbb{R})$. Namely, for $(A_n)_{n \in \mathbb{N}_0} \subset A([0,T];\mathbb{R})$, we will write $A_n \rightarrow_\tau A_0$, if

$$\|A_n - A_0\|_\infty \rightarrow 0, \text{ as } n \rightarrow \infty,$$

and

$$\sup_{n \in \mathbb{N}_0} \|A_n\|_\mathbb{V} < \infty,$$

where $\| \cdot \|_\mathbb{V}$ denotes the total variation norm.

**Definition 4.0.15.** The topology $\Sigma$ on $\mathbb{D}([0,T];\mathbb{R}^d)$ is the topology generated by the seminorms

$$\rho^i_A = \sup_{A \in \mathcal{A}} \left| \int_{[0,T]} \omega^i(u) dA(u) \right|, \ i \leq d,$$

where $\mathcal{A}$ ranges over relatively $\tau$-compact subsets of $A([0,T];\mathbb{R})$.

**Definition 4.0.16.** The topology $\Sigma$ on $\mathbb{D}([0,T];\mathbb{R}^d)$ is the topology generated by the seminorm $\rho^i_T(\omega) = |\omega^i(T)|, \ i \leq d$, and the seminorms

$$\rho^i_A = \sup_{A \in \mathcal{A}} \left| \int_{[0,T]} \omega^i(u) dA(u) \right|, \ i \leq d,$$

where $\mathcal{A}$ ranges over relatively $\tau$-compact subsets of $A([0,T];\mathbb{R})$.

The topology $\Sigma$ was defined on the Skorokhod space $\mathbb{D}([0,T];\mathbb{R})$ for $T = 1$ in [jakubowski4]. The following properties were shown to be true for $\Sigma$ on $\mathbb{D}([0,T];\mathbb{R})$.

**Proposition 4.0.17.** The topology $\Sigma$ has the following properties:

(i) The Skorokhod space endowed with $\Sigma$ is a locally convex vector space.

(ii) The topology $\Sigma$ is weaker than the topology $S$.

(iii) A set is $\Sigma$-compact if and only if it is $S$-compact.

**Remark 4.0.18.** It was communicated to the author by Professor Jakubowski that the properties of Proposition 4.0.17 are true for the infinite horizon extension of the topology $\Sigma$. 
The Meyer-Zheng topology

The Meyer-Zheng topology, introduced in [meyerzheng], is a relative topology, on the image measures on graphs \((t, \omega(t)) \in [0, \infty)\) of trajectories \((\omega(t))_{t \in [0, \infty)}\) under the measure \(\lambda(dt) := e^{-t} dt\) (called pseudo-paths), induced by the weak topology on probability laws on compactified space \([0, \infty) \times \mathbb{R}\). We refer to the Meyer-Zheng topology formally \((MZ)\) as the topology on the Škorokhod space \(\mathbb{R}(I; \mathbb{R}^d)\) generated by the coordinatewise convergence in measure; see (4.0.4). The following definition is adapted from [meyerzheng][Lemma 1], which states that, on \(\mathbb{D}([0, \infty]; \mathbb{R})\), the convergence in measure (4.0.4) is indeed equivalent to the convergence in the pseudo-path topology.

**Definition 4.0.19.** The topology \(MZ\) on \(\mathbb{D}(I; \mathbb{R}^d)\), where \(I = [0, \infty]\), is the topology generated by the convergence:

\[
\int_I f(t, \omega_n^i(t)) \lambda(dt) \rightarrow \int_I f(t, \omega^i(t)) \lambda(dt), \quad \forall f \in C_b(I \times \mathbb{R}), \quad \forall i \leq d,
\]

where \(\lambda(dt) := e^{-t} dt\).

On \(\mathbb{D}([0, T]; \mathbb{R}^d)\), we additionally require the convergence of the terminal value (4.0.5). Without this addition, the topology is not a Hausdorff topology on \(\mathbb{D}([0, T]; \mathbb{R}^d)\).

**Definition 4.0.20.** The topology \(MZ\) on \(\mathbb{D}(I; \mathbb{R}^d)\), where \(I = [0, T]\), is the topology generated by the convergence (4.0.4) in conjunction with the convergence:

\[
\omega_n(T) \rightarrow \omega(T).
\]

The key lemma, [meyerzheng][Lemma 1], extends to \(I = [0, T]\), for \(T\) finite, and \(d > 1\) via a simple iterative argument; cf. Subsection 2.4.2.

**Lemma 4.0.21.** Let \((\omega_n)_{n \in \mathbb{N}}\) and \(\omega\) be paths in \(\mathbb{D}(I; \mathbb{R}^d)\) such that \(\omega_n \rightarrow_{MZ} \omega\). Then \(\omega_n^i \rightarrow \omega^i, \) for every \(i \leq d\). Moreover, there exists a subsequence \((\omega_{n_k})\) and a set \(L \subset I\) of full Lebesgue measure such that \(T \in L\), if \(I = [0, T]\), and \(\omega_{n_k}^i(t) \rightarrow \omega^i(t), \) for every \(i \leq d\), for every \(t \in L\). In particular, there exists a (countable) dense set \(D \subset I\) such that \(T \in D\), if \(I = [0, T]\), and \(\omega_{n_k}^i(t) \rightarrow \omega^i(t), \) for every \(i \leq d\), for every \(t \in D\).

**Proof.** Let \(\omega_n \rightarrow_{MZ} \omega\). By the definition (4.0.4), we have

\[
\int_I f(t, \omega_n^i(t)) \lambda(dt) \rightarrow \int_I f(t, \omega^i(t)) \lambda(dt), \quad \forall f \in C_b(I \times \mathbb{R}^d), \quad \forall i \leq d,
\]

where the measure \(\lambda(dt) = e^{-t} dt\) is equivalent to the Lebesgue measure. Taking \(f(t, x) := \alpha(t) \arctan(x), \alpha \in C_b(I)\), we deduce that \(u_n^i := \arctan(\omega_n^i)\) converges weakly to \(u^i := \arctan(\omega^i)\) in \(L^2(\lambda)\), for every \(i \leq d\). Further, taking \(f(t, x) := \alpha(t) \arctan^2(x), \alpha \in C_b(I)\), we deduce that \(u_n^i\) converges strongly to \(u^i\) in \(L^2(\lambda)\), and consequently, \(\omega_{n_k}^i\) converges in measure \(\lambda\) to \(\omega^i\) in \(I\), i.e., \(\omega_{n_k}^i \Rightarrow \omega^i\), for every \(i \leq d\). Thus, for \(i = 1\), there exists a subsequence \((\omega_{n_l})_{l \in \mathbb{N}} = (\omega_{n_1}^1, \ldots, \omega_{n_l}^d)_{l \in \mathbb{N}}\) of \((\omega_n)_{n \in \mathbb{N}}\) such that

\[
\omega_{n_l}^1(t) \rightarrow \omega^1(t),
\]

for every \(t\) in some set \(L_1\) of full Lebesgue measure. By the bounded convergence, we have

\[
\int_I f(t, \omega_{n_l}^1(t)) \lambda(dt) \rightarrow \int_I f(t, \omega^1(t)) \lambda(dt), \quad \forall f \in C_b(I \times \mathbb{R}^d), \quad \forall i \leq d.
\]
Now, by replacing \( i = 1 \) with \( i = 2 \) and \((\omega_n)_{n \in \mathbb{N}}\) with \((\omega_n)_{n \in \mathbb{N}}\) preceding (4.0.6), we obtain a set \( L_2 \) of full Lebesgue measure and a further subsequence \((\omega_{n_{lm}})_{m \in \mathbb{N}} = (\omega_{n_{1m}}, \ldots, \omega_{n_{dm}})_{m \in \mathbb{N}}\) of \((\omega_n)_{n \in \mathbb{N}}\) such that \( \omega_{n_{lm}}^1 (t) \to \omega^1(t) \), for every \( t \in L_2 \). We have

\[
\omega_{n_{lm}}^1 (t) \to \omega^1(t) \quad \text{and} \quad \omega_{n_{lm}}^2 (t) \to \omega^2(t),
\]

for every \( t \in L_1 \cap L_2 \), where the set \( L_1 \cap L_2 \) is of full Lebesgue measure. By repeating the argument \( d - 2 \) more times, we obtain a set \( L := L_1 \cap L_2 \cap \cdots \cap L_d \) and a subsequence \((\omega_{n_k})_{k \in \mathbb{N}} = (\omega_{n_{k1}}, \ldots, \omega_{n_{kd}})_{k \in \mathbb{N}}\) such that

\[
\omega_{n_k}^i (t) \to \omega^i(t), \quad \forall i \leq d,
\]

for every \( t \in L \), where the set \( L \) is of full Lebesgue measure. Moreover, by (4.0.5), for \( I = [0,T] \), we have \( \omega_{n_k}(T) \to \omega(T) \), so, the set \( L \) can be chosen to contain the terminal time \( T \). The complement of \( L \) is a \( \lambda \)-null set, so, the set \( L \) contains a (countable) dense set \( D \) such that \( T \in D \), if \( I = [0,T] \).

**Corollary 4.0.22.** We have \( \mathcal{F}_T^\sigma := \sigma(\omega(s) : s \in [0,t]) = \sigma(\omega(s) : s \in D \cap [0,t]) \), for any countable dense subset \( D \) of \([0,t]\), for every \( t \leq T \). Moreover, we have

\[
\mathcal{F}_T^\sigma = \sigma(\mathcal{G}_T^\sigma, \omega(t)), \quad t \leq T,
\]

where \( \mathcal{G}_T^\sigma \) denotes the \( \sigma \)-algebra generated by the family of \( \mathcal{F}_T^\sigma \)-measurable \( MZ \)-continuous functions.

**Proof.** Let \( \mathcal{G}_T^\sigma \) denote the \( \sigma \)-algebra generated by the family of \( \mathcal{F}_T^\sigma \)-measurable \( MZ \)-continuous functions. We have \( \mathcal{G}_T^\sigma \subset \mathcal{F}_T^\sigma \) and \( \mathcal{F}_T^\sigma \subset \mathcal{G}_T^\sigma \) from Lemma 4.0.21. Moreover, we have

\[
\omega^i(t) = \lim_{\varepsilon \to 0} \frac{1}{\varepsilon} \int_0^\varepsilon \omega^i(t+u)du, \quad i \leq d, \quad \varepsilon < T - t,
\]

where each \( \omega \mapsto \frac{1}{\varepsilon} \int_0^\varepsilon \omega^i(t+u)du \) is \( MZ \)-continuous.

**Lemma 4.0.23.** The mappings

\[
\omega \mapsto \|\omega\|_\infty \quad \text{and} \quad \omega \mapsto N^{a,b}(\omega^i), \quad a < b, \quad i \leq d,
\]

are (sequentially) lower \( MZ \)-semicontinuous.

**Proof.** The proof is adapted from [meyerzheng]. Let \( i \leq d \) and \( \omega_n^i \to MZ \omega^i \) with \( \sup_n \|\omega_n^i\| \leq c \). If \( \|\omega^i\|_\infty > c \), then there exists \( s < t \) such that \( \omega^i(u) > c \), for all \( u \in [s,t] \), or we have \( \omega^i(T) > c \), either way, there exists an \( MZ \)-continuous function \( F \) for which \( \lim_{n \to \infty} F(\omega_n^i) < F(\omega^i) \), cf. (4.0.4) and (4.0.5), so, this is a contradiction. Thus, the mapping \( \omega \mapsto \|\omega\|_\infty := \|\omega^1\|_\infty \vee \cdots \vee \|\omega^d\|_\infty \) is lower \( MZ \)-semicontinuous. Similarly, one can show that the sets of the form \( \{ \omega : \exists u \in [s,t] \text{ s.t. } \omega^i(u) > b \} \) and \( \{ \omega : \exists u \in [s,t] \text{ s.t. } \omega^i(u) < a \} \), \( s < t, \ a < b \) are open in the \( MZ \)-topology, from which the lower \( MZ \)-semicontinuity of the mappings \( N^{a,b}, \ a < b \), follows. Indeed, let \( a < b \) be fixed and consider a finite partition \( \pi := \{ t_0 < t_1 < \cdots < t_n \} \) of \([0,t_n]\). We write \( N^{a,b}(\omega^i) \geq k \), if one can find

\[
l_1 < m_1 \leq l_2 < m_2 \leq \cdots < l_k < m_k \leq n
\]
such that, for all $j < k$, $\omega^i(s) < a$, for some $s \in [t_{i-1}, t_i]$, and $\omega^i(t) > b$, for some $t \in [t_{mj-1}, t_{mj}]$ (or, for $t = T$, if $j = k$ and $m_k = n$). The partition $\pi$ is finite, so, the sets

$$\{ \omega : N_{\pi}^a,b(\omega^i) \geq k \} = \{ \omega : N_{\pi}^a,b(\omega^i) > k - 1 \}, \quad k \in \mathbb{N},$$

are open in the $MZ$-topology. Consequently, the mapping $\omega \mapsto N_{\pi}^a,b(\omega^i)$ and the mapping $N_{\pi}^a,b(\omega^i) := \sup_{\pi} N_{\pi}^a,b(\omega^i)$ are lower $MZ$-semicontinuous, for every $i \leq d$.

We refer the reader to [dellacheriemeyer][A, IV] and [meyerzheng] for details on pseudo-paths and the Meyer-Zheng topology, respectively.

The Skorokhod’s $J^1$-topology

**Definition 4.0.24.** The Skorokhod’s $J^1$-topology on $\mathbb{D}([0, T]; \mathbb{R}^d)$ is the topology generated by the (complete) metric

$$J_{\pi}^1(\omega, \tilde{\omega}) := \inf_{\lambda \in \Lambda} \left\{ \sup_{s < t} \left| \log \lambda \frac{t - s}{t - s} \right| \vee \| \omega - \tilde{\omega} \circ \lambda \|_{\infty} \right\}, \quad (4.0.7)$$

where $\Lambda$ denotes the class of strictly increasing, continuous mappings of $[0, T]$ onto itself and $i$ is the identity map on $[0, T]$.

**Definition 4.0.25.** The Skorokhod’s $J^1$-topology on $\mathbb{D}([0, \infty]; \mathbb{R}^d)$ is the topology generated by the (complete) metric

$$J^1(\omega, \tilde{\omega}) := \sum_{r=1}^{\infty} 2^{-r} \left( 1 \wedge J_{\pi}^1([\omega]^r, [\tilde{\omega}]^r) \right), \quad (4.0.8)$$

where $[\omega]^r$ indicates the restriction of $\omega$ on $[0, r]$.

We refer the reader to [billingsley][Section 12, 16] for details on the Skorokhod $J^1$-metric on $\mathbb{D}([0, T]; \mathbb{R}^d)$ and $\mathbb{D}([0, \infty]; \mathbb{R}^d)$, respectively.

### 4.0.4 Strict Topology $\beta_0$

Given a topological space $X$, the family of (signed) regular measures of finite total variation on the (universal completion of the) Borel $\sigma$-algebra $\mathcal{B}(X)$ is denoted by $\mathcal{M}(X)$ and the family of bounded continuous functions from $X$ into $\mathbb{R}$ is denoted by $C_b(X)$.

We sketch the proof of the general form of the Riesz representation theorem; see [jarchow] for details. The proof is based on the classical Riesz representation theorem of compact spaces.

**Theorem 4.0.26.** Let $X$ be a compact space and $C(X)$ endowed with the uniform norm. If $u$ is a continuous linear form on $C(X)$, then there exists a unique regular Borel measure $\mu$ on $X$ such that

$$u(f) = \int_X f \, d\mu, \quad \forall f \in C(X).$$

Moreover, the map $u \mapsto \mu$ establishes an isometric isomorphism of $[C(X), \| \cdot \|_\infty]'$ onto $\mathcal{M}(X)$. Here, $\mathcal{M}(X)$ is endowed with the total variation norm.
On a completely regular space $X$, every continuous function $f \in \mathcal{C}_b(X)$ admits a unique continuous extension to the Stone–Čech compactification $\beta X$ of $X$, and the mapping $R: \mathcal{C}(\beta X) \to \mathcal{C}_b(X) : g \mapsto g|_X$ obtained this way is an isometric isomorphism for the corresponding uniform norms $\| \cdot \|_\infty$. In particular, the corresponding duals can be identified.

**Corollary 4.0.27.** Let $X$ be a completely regular Hausdorff space. For every $u \in [\mathcal{C}_b(X), \| \cdot \|_\infty]'$ there exists a unique $\mu \in \mathcal{M}(\beta X)$ such that $u(f) = \int_{\beta X} f d\tilde{\mu}$, $\forall f \in \mathcal{C}_b(X)$.

Let us introduce the locally convex topology $\beta_0$ generated by the family of seminorms

$$p_g(f) := \|fg\|_\infty, \; f \in \mathcal{C}_b(\mathbb{D}), \; g \in \mathbb{B}_0(X),$$

where $\mathbb{B}_0(X)$ denotes bounded Borel functions vanishing at infinity.

$$\mathbb{B}_0(X) := \{ f \in \mathcal{B}_b(X) : \forall \varepsilon > 0 \exists K^\varepsilon \in \mathcal{K}(\mathbb{D}) \text{ s.t. } |f(x)| < \varepsilon \; \forall x \notin K^\varepsilon \}.$$ 

Further analysis of Radon measures on completely regular space $X$ leads to the following result.

**Corollary 4.0.28.** Let $X$ be a completely regular Hausdorff space, and let $\beta_0$ be the strict topology on $\mathcal{C}_b(X)$. For every $\beta_0$-continuous linear form $u$ on $\mathcal{C}_b(X)$, there is a unique measure $\mu \in \mathcal{M}(X)$ such that $u(f) = \int_X f d\mu$, $\forall f \in \mathcal{C}_b(X)$. The map $[\mathcal{C}_b(X), \beta_0]' \to \mathcal{M}(X) : u \mapsto \mu$ obtained in this way is an isomorphism.

Details for the proof can be found e.g. in the book of Jarchow [jarchow]. Remark that the proof relies on the fact that a completely regular space admits the Stone–Čech compactification so it requires the axiom of choice.

### 4.0.5 Functional Choquet’s Capacity

The extension of the duality from upper semicontinuous to measurable payoffs in Theorem 3.6.4 is based on Choquet’s capacitation theorem for functional capacities. We briefly recall a result from [kellerer]; see also [choquet1959] or [dellacherie1972].

**Proposition 4.0.29.** Let $\mathcal{F} = \mathcal{F}(\mathcal{X})$ be the set of all real-valued functions on a Hausdorff space $\mathcal{X}$, let $\mathcal{G}$ be a subset of $\mathcal{F}$ that is closed w.r.t. finite suprema and countable infima, and let $I: \mathcal{F} \to \mathbb{R}$ be a $\mathcal{G}$-capacity on $\mathcal{F}$, i.e., $I$ is increasing on $\mathcal{F}$, $I$ is sequentially continuous from below on $\mathcal{F}$, and $I$ is sequentially continuous from above on $\mathcal{G}$, then every $\mathcal{G}$-Suslin function $f$ is $I$-capacitable, i.e.,

$$I(f) = \sup \{ I(g) : g \in \mathcal{G}, g \leq f \}.$$ 

**Corollary 4.0.30.** If $\mathcal{X}$ is perfectly normal Hausdorff space, then every Borel function $f \in \mathcal{B}_b$ (resp. $f \in \mathcal{B}_{<p}$) is $I$-capacitable for a $\mathcal{U}_b$-capacity $I$ (resp. $\mathcal{U}_{<p}$-capacity $I$).

**Proof.** The space $\mathcal{U}_b$ is closed with respect to finite suprema and countable infima. The family of $\mathcal{U}_b$-Suslin functions contains the indicators of closed sets $\mathcal{C}$. Since every Baire set is $\mathcal{C}$-Suslin, see e.g. [fremlin][Proposition 241L], the family of $\mathcal{U}_b$-Suslin functions contains the indicators of Baire sets. It follows by a monotone class argument that every Baire function is $I$-capacitable. Since, on a perfectly normal space, the families of Borel and Baire functions coincide, see e.g. [bogachev][Proposition 6.3.4.], we conclude that every Borel function is $I$-capacitable for a $\mathcal{U}_b$-capacity $I$. \qed
Moreover, we used the following continuity property of functional $E_Q$.

**Proposition 4.0.31.** Let $X$ be topological space and $Q$ be a sequentially weak* compact subset of $P = P(X)$. Then the support functional $E_Q$ is sequentially continuous from above on $U_b$.

**Proof.** For $(\xi^n)_{n \in \mathbb{N}}$ and $\xi \in U_b$ such that $\xi^n \downarrow \xi$, extracting a subsequence if necessary, we will find a sequence $(Q^k, \xi^{nk})_{k \in \mathbb{N}}$ in $Q \times U_b$ such that $(Q^k)_{k \in \mathbb{N}}$ is converging sequentially in the weak* topology to $Q$ and

$$E_{Q^k}[\xi^{nk}] \downarrow \inf_{n \geq 1} E_Q(\xi^n).$$

By the fact that each $\xi^{nk}$ is upper semicontinuous, we get

$$E_Q[\xi^{nk}] \geq \limsup_{k \to \infty} E_{Q^k}[\xi^{nk}]$$

for every $k_0 \geq 1$, and further

$$\limsup_{k \to \infty} E_{Q^k}[\xi^{nk}] \geq \lim_{k \to \infty} E_{Q^k}[\xi^{nk}] = \inf_{n \geq 1} E_Q(\xi^n).$$

Hence, one gets

$$E_Q(\xi) \geq E_Q[\xi] \geq \inf_{n \geq 1} E_Q(\xi^n)$$

and consequently $E_Q(\xi^n) \downarrow E_Q(\xi)$ as claimed.

Proposition 4.0.31 is essentially [elkarouitan][Proposition 2.2] for non-Polish spaces but assuming the set of measures to be sequentially compact.

### 4.0.6 Transfinite Schema

Transfinite induction is an extension of induction over finite ordinals (natural numbers $\mathbb{N}_0$) up to the first infinite ordinal $\omega$ to up to an arbitrary given ordinal. We deploy transfinite induction up to the first uncountable ordinal $\omega_1$. The ordinals are well-ordered by inclusion. A comprehensive treatment of transfinite induction and recursion over ordinals is provided e.g. in [suppes][Section 7.1]. We provide the proofs for the sake of completeness.

**Principle of Transfinite Induction.** Let $P(\alpha)$ be a statement for the ordinal $\alpha \leq \omega_1$. Assume that

(i) $P(0)$ is true for the ordinal 0,

(ii) for any successor ordinal $\alpha + 1$ of $\alpha$, $P(\alpha + 1)$ follows from $P(\beta)$ for all $\beta < \alpha$,

(iii) for the limit ordinal $\omega_1$, $P(\omega_1)$ follows from $P(\alpha)$ for all $\alpha < \omega_1$.

Then $P(\alpha)$ is true for all $0 \leq \alpha \leq \omega_1$.

**Proof.** Given the hypothesis, assume there is an ordinal $\alpha$ such that $P(\alpha)$ is false. Let $L(\alpha) = \{\beta : \beta \leq \alpha \text{ and } P(\beta) \text{ is false}\}$. Let $\beta^*$ be the first element of $L(\alpha)$, well-ordered by the inclusion. By the hypothesis, for every $\gamma < \beta^*$, we have $P(\gamma)$. But then, by the inductive hypothesis, that is (i) of the theorem, $P(\beta^*)$, a contradiction. \qed
Definition by Transfinite Recursion. Let $H$ be any function. For any ordinal $\alpha \leq \omega_1$, there exists a unique $F$ such that

(i) $F$ is a function on $\alpha$,

(ii) for every $\beta < \alpha$

$$F(\beta) = H(F|\beta).$$

Proof. Suppose that for some $\beta$ there is $F$ and $G$ such that (i) $F$ and $G$ are functions on $\beta$, and (ii) for every $\gamma$, if $\gamma < \beta$, then $F(\gamma) = H(F|\gamma)$ and $G(\gamma) = H(G|\gamma)$. Let us show

$$\text{graph}(F) \subset \text{graph}(G) \text{ or } \text{graph}(G) \subset \text{graph}(F). \quad (4.0.9)$$

Then $\text{dom}(F) \cap \text{dom}(G)$ is $\text{dom}(F)$ or $\text{dom}(F)$. Now assume that there is an ordinal $\gamma$ in $\text{dom}(F) \cap \text{dom}(G)$ such that $F(\gamma) \neq G(\gamma)$. It follows, by the transfinite induction, that there is the smallest ordinal $\gamma^*$ such that $F(\gamma^*) \neq G(\gamma^*)$. We have $F|\gamma^* = G|\gamma^*$, and consequently, $H(F|\gamma^*) = H(G|\gamma^*)$, but this implies that $F(\gamma^*) = G(\gamma^*)$, a contradiction, so, we conclude that $(4.0.9)$ is true. Now, we set

$$\text{graph}(F) = \bigcup \text{graph}(G),$$

where the union is taken over all $G$ satisfying (i) and (ii). By $(4.0.9)$, the $F$ is a function. Moreover, if $\beta$ is in $\text{dom}(F)$, then for some $G$, $\beta$ is in $\text{dom}(G)$, and consequently $G(\beta) = H(G|\beta)$, so, $F(\beta) = H(F|\beta)$. It remains show that $\alpha$ is the domain of $F$. Assume for a contradiction that it is not, and let $\beta^*$ be the least ordinal less than $\alpha$ that does not belong to the domain of $F$. Then there is $G$ whose domain is $\beta^*$, and by the definition of $F$, this is a contradiction. Thus, $\alpha$ is in $\text{dom}(F)$. It similarly follows that $F$ is unique. \qed