universität wien

DISSERTATION

Titel der Dissertation:

„Conditional Acceptability Mappings: Convex Analysis in Banach Lattices“

Verfasser:

Mag. Raimund Kovacevic

angestrebter akademischer Grad:

Doktor der Sozial- und Wirtschaftswissenschaften (Dr. rer. soc. oec.)

Wien, im Oktober 2008

Studienkennzahl lt. Studienblatt: A 084 136

Dissertationsgebiet lt. Studienblatt: Statistik

Betreuer: o. Univ.-Prof. Mag. Dr. Georg Pflug

For Alice and Juliana

Acknowledgement. I would like to thank the following persons. Without them this work would not have been possible for me:

Prof. Georg Pflug - my supervisor and first refereer - for his deep questions, for his patience and the many exciting discussions. In addition I would like to thank him for bringing me back to science.

Prof. Immanuel Bomze - my second referee - for his helpful comments.

Dr. Alice Mladenka - my wife - for tolerating my research-activities in addition to my ordinary job, but also for proofreading.

Dr. Brian Sallans for reading the raw versions of my dissertation as a native speaker.

SUMMARY. Conditional Acceptability Mappings quantify the degree of desirability of random variables modeling financial returns, accounting for available non-trivial information. They are defined as mappings from spaces $\boldsymbol{L}_p(\Omega, \mathcal{F}, \mu)$ to spaces $\boldsymbol{L}_{p'}(\Omega, \mathcal{F}_1, \mu)$, where the $\sigma-$algebra $\mathcal{F}_1 \subseteq \mathcal{F}$ describes the available information. Additionally, such mappings have to be concave, translation-equivariant and monotonically increasing.

Based on the order characteristics (in particular the order completeness) of $\boldsymbol{L}_p(\Omega, \mathcal{F}, \mu)$-spaces, superdifferentials and concave conjugates for conditional acceptability mappings are defined and analyzed. The novelty of this work is that the almost sure partial order is consequently used for this purpose, which results in simpler definitions and proofs, but also accounts for all requirements concerning continuity, integrability and measurability of the supergradients and conjugates.

Furthermore, the results about conditional mappings are used to show properties of multiperiod acceptability functionals that are based on conditional acceptability mappings, such as SEC-functionals and acceptability compositions. A chain rule for superdifferentials as well as the conjugate of multiperiod functionals and their properties are derived.

Contents

CHAPTER 1

Introduction

Most economic decisions are affected by uncertainty and risk. Historically, insurance firms were the first institutions to deal with risk in a mathematical way - mainly by pooling individual risks and referring to the laws of large numbers. Although evidently existent, risk was neglected in economic theory until the mid 20th century. At this time Von Neumann and Morgenstern [**34**] raised the idea of expected utility which was enhanced by Arrow and Pratt [**4, 26**]. On the other hand Markowitz [**15**] developed the idea of using quadratic optimization in the field of portfolio selection - using the standard deviation of returns as a risk measure and mean returns as a measure of profit. The idea of optimizing a tradeoff between these antagonistic objectives - which results in the construction of an efficient frontier - was very fruitful in the following decades, where a lot of alternative risk and acceptability measures were developed.

While at the beginning - because of limited computational capacities - these methods were more or less a matter of theory, the revolution in computer technology made it possible to cheaply use quantitative methods in practical financial management [**16**].

Generally all kinds of risks and their quantification and management, were a key issue in the rapid development of modern finance and financial mathematics. Another influence came from the literature on derivative pricing, beginning with Black and Scholes [**8**] and Merton [**17**].

While all of these ideas had their origin in finance and corporate finance, the idea of measuring and managing risk has recently entered all areas of business like supply chains, telecommunication, electricity and general energy management.

Another input into the development of risk-management came directly from practical considerations: In the last decades most countries enacted laws constraining banks and insurance companies to hold a substantial amount of risk capital. Initially - until the recent past - such prescriptions were very crude, defining minimum requirements as a certain percentage of assets or other suitable reference figures from financial statements.

In Europe these requirements accompanied the process of liberalization in the financial sector, replacing the more direct influence of the state under the old regime.

The last step in this process - for the time being - is the activities well known as “Basel II” for banks and as “Solvency II” for insurance companies: One of the so called “pillars” of this regulatory system requires - under the name of “Value at Risk” - the usage of quantiles of the loss distribution to measure the required risk capital.

Quantiles were well known for a long time in statistics. Also it should be noted that in principle Markowitz based his theory - for the special case of Gaussian distributions - on this concept. Nevertheless it is clear that the introduction of Value at Risk was a great advance for the regulatory system. Value at Risk is easy to understand and communicate and also seems easy to calculate. Although the latter point is not true for the general case, a lot of work has been done in the last years to make more and more difficult stochastic models tractable for calculation.

The main drawback of Value at Risk lies in the fact that for the general case it has a lot of problematic mathematical properties, making it difficult to control by a structured optimization procedure. This was stated by Artzner et al. in their seminal papers [**6, 5**]. Since then a lot of work has been done on risk measures that calculate required risk capital and have nice mathematical properties under a variety of theoretical and practical aspects. A recent example is [**14**].

Initially the focus was on coherent risk functionals, based on subadditivity and positive homogeneity. Föllmer and Schied [**10**] and Frittelli and Rosazza [**11**] independently weakened these axiomatic assumptions, using convexity instead. This in turn opened the possibility to use the whole arsenal of convex analysis: Based on the work of Moreau and Rockafellar ([**18**], [**27, 30**]), Ruszczyński and Shapiro [**33**], Pflug and Römisch [**22, 23**] and others were able to formulate and prove a lot of results about risk and acceptability functionals. A key role in this process - as a basic example - was played by the average (or conditional) value at risk [**29**] .

Coherent risk measures remain an important field of study, but new types of functionals for the valuation of random variables in finance also have been developed in the last years:

- acceptability functionals - which can be seen as negative coherent risk measures, but lacking positive homogeneity
- deviation risk functionals, generalizing the properties of the standard deviation in the context of convex analysis.

While there seems to be agreement in principle about how to handle one-period functionals that valuate random variables by assigning a real number, much less agreement exists regarding the valuation of (multiperiod) stochastic processes. Here, an important building block seems to be the notion of conditional acceptability or risk mappings.

Such mappings evaluate a random outcome at the end of a time period relative to the information available at the beginning of the period. Such an evaluation generally will result not in a real number, but in another random variable and is in principle a mapping between two random spaces. In addition these mappings will be defined in a similar way to acceptability or risk functionals: It is important that they have nice mathematical properties like monotonicity and concavity or convexity etc.

There is consensus in the literature on how to define such mappings, but - as we will discuss in chapter 4 - there are different views about how to define and use their conjugate mappings. In the following work we will consequently use the almost sure partial order and the associated notion of an infimum for this purpose.

This approach seems to be totally new in the context of conditional risk and acceptability mappings and therefore we will recapitulate some important concepts from probability, order theory and nonsmooth analysis in the first part, to construct a firm ground for the following study of conditional mappings in the second part. In particular we use the properties of $\boldsymbol{L}_p(\Omega, \mathcal{F}, \mu)$-spaces as Banach lattices and some helpful characteristics of the p-norms for showing that $\boldsymbol{L}_p(\Omega, \mathcal{F}, \mu)$-spaces are order complete. This fact makes the almost sure infimum a reasonable concept for defining conjugates.

Our approach combines the following benefits, while each of the previously discussed concepts has some weaknesses in these fields:

- There is a simple, clear and general concept regarding the spaces involved. In particular we do not require that all random variables are essentially bounded.
- There is a close connection between supergradients and the conjugate representation.
- The approach is based on sound mathematical principles and accounts for all requirements regarding continuity, integrability and measurability of supergradients and conjugate mappings. In particular, there are clear requirements for the dual variables involved, ensuring that the linear mappings used map into the correct random spaces.

- Most results - even for the applications in the multiperiod case - can be proved in a relatively simple manner.

As a starting point we will - mainly following **[23]** - characterize deviation risk and acceptability functionals in the following sections and compare them with coherent risk measures. This is done in an informal way, stating only the main facts necessary to understand the similarities and dissimilarities with conditional mappings in part two of this work. One main focus will be the notion of a dual or conjugate representation of such functionals.

1.1. Single Period Deviation Risk and Acceptability

While the term "risk measure" is widely used in economics and in the practical area, from now on in this work we will avoid the terms "risk measure" or "acceptability measure": In mathematics measures are a topic of measure theory and are related to sets, while risk and acceptability are related to random variables. Instead we will speak of acceptability or risk functionals.

A *probability functional* is an extended real valued function defined on some random space, or on a suitable subset of a random space. Examples are well known functionals like the expectation, the median, value at risk, average (or "conditional") value at risk, variance etc. If the value of a probability functional depends only on the distribution of the random variable under consideration, it is called *version-independent.* Such functionals could also be defined on spaces of distribution functions. In this context they are well known under the name of *statistical parameters.*

If a functional is interpreted in the sense that higher values are preferable to lower values it is called an *acceptability-type functional. Acceptability functionals* are acceptability type functionals with some additional properties:

DEFINITION 1.1.1. (acceptability) A real valued functional $\mathcal{A}$, defined on a linear space $\mathcal{Y}$ of random variables on $(\Omega, \mathcal{F}, \mathbb{P})$ is called an *acceptability functional*, if the following properties are true for all $Y \in \mathcal{Y}$:

(A1) ***Translation Equivariance***. $\mathcal{A}(Y + c) = \mathcal{A}(Y) + c$ holds a.s. for all constants c.

(A2) ***Concavity***. $\mathcal{A}\left(\lambda \cdot X + (1-\lambda) \cdot Y\right) \geq \lambda \cdot \mathcal{A}\left(X\right) + (1-\lambda) \cdot \mathcal{A}\left(Y\right)$ holds for $\lambda \in [0, 1]$.

(A3) ***Monotonicity***. $X \leq Y \; a.s. \Rightarrow \mathcal{A}(X) \leq \mathcal{A}(Y)$.

Another relevant property is positive homogeneity and strictness: An acceptability functional is called *positively homogeneous*, if it satisfies the condition $\mathcal{A}(\lambda Y) = \lambda \cdot \mathcal{A}(Y)$ for all $\lambda \geq 0$. It is called strict, if $\mathcal{A}(Y) \leq \mathbb{E}(Y)$ holds.

The second important group of probability functionals are deviation type functionals. While the prototype for acceptability type functionals could be expectation, standard deviation is the ideal model for deviation risk functionals. In principle such functionals measure the risk of a deviation from some target value.

DEFINITION 1.1.2. (deviation risk) A real valued functional $\mathcal{D}$ defined on a linear space $\mathcal{Y}$ of random variables on $(\Omega, \mathcal{F}, \mathbb{P})$ is called a *deviation risk functional*, if the following properties hold for all $Y \in \mathcal{Y}$:

(D1) ***Translation Invariance***. $\mathcal{D}(Y + c) = \mathcal{D}(Y)$ holds a.s. for all constants c.

(D2) ***Convexity***. $\mathcal{D}(\lambda \cdot X + (1 - \lambda) \cdot Y) \leq \lambda \cdot \mathcal{D}(X) + (1 - \lambda) \cdot \mathcal{D}(Y)$ holds for $\lambda \in [0, 1]$.

(D3) ***Monotonicity***. $X \leq Y \ a.s. \Rightarrow \mathbb{E}(X) - \mathcal{D}(X) \leq \mathbb{E}(Y) - \mathcal{D}(Y)$.

From this definitions it follows that $\mathcal{D}$ is a deviation risk functional if and only if the functional $\mathbb{E}(X) - \mathcal{D}(X)$ is an acceptability functional. In this work the focus will be on acceptability: Although most of the results about acceptability functionals could easily be reformulated for deviation risk functionals, we will not do so for the sake of avoiding redundancy.

Again it is possible to define some additional properties: *positive homogeneity* for deviation risk functionals means that $\mathcal{D}(\lambda Y) = \lambda \cdot \mathcal{D}(Y)$ holds for all $\lambda \geq 0$. On the other hand a deviation risk functional is called strict if it satisfies $\mathcal{D}(X) \geq 0$ for any X. Typical examples are the standard deviation or the lower partial moments.

A lot of interesting probability functionals are not continuous, but as we shall see it is crucial that they are at least semicontinuous: A convex functional is lower semicontinuous (l.s.c.) if its epigraph is a closed set. A concave functional is upper semicontinuous (u.s.c.) if its hypograph is a closed set.

As mentioned above, the first axiomatic treatment of risk functionals was Artzner et al. [**5**]. The key notion of this approach was coherence:

DEFINITION 1.1.3. (coherence) A real valued functional ρ defined on a linear space $\mathcal{Y}$ of random variables on $(\Omega, \mathcal{F}, \mathbb{P})$ is called a *coherent risk (capital) functional*, if the following properties are valid for all $Y \in \mathcal{Y}$:

(R1) ***Translation Antivariance***. $\rho(Y + c) = \rho(Y) - c$ holds for all constants c.

(R2) ***Convexity***. $\rho(\lambda \cdot X + (1 - \lambda) \cdot Y) \leq \lambda \cdot \rho(X) + (1 - \lambda) \cdot \rho(Y)$ holds for $\lambda \in [0, 1]$.

(R3) ***Pointwise Antimonotonicity***. $X \leq Y$ *a.s.* $\Rightarrow \rho(X) \geq \rho(Y)$.

(R4) ***Positive Homogeneity.***$\rho(\lambda Y) = \lambda \cdot \rho(Y)$ for $\lambda \geq 0$.

REMARK. In the original paper **[5]** subadditivity was used instead of convexity, which is equivalent due to homogeneity.

Coherent risk functionals valuate risk in terms of a risk reserve, necessary to make the distribution of a risk acceptable. This is the reason why they must be translation-antivariant. In addition it is clear that coherent risk functionals are risk functionals in the sense that lower values are preferable to higher values.

A prototype would be Value at Risk. Unfortunately it is not convex if we consider distributions other than just Gaussian.

Although the axiomatic definition of coherent risk functionals is older than the definition of acceptability functionals we prefer to use the latter:

ρ is a coherent risk functional if and only if $\mathcal{A} = -\rho$ is a positively homogeneous acceptability functional. This means that the class of coherent risk functionals can be reconstructed by a subset of the class of acceptability functionals. Coherence requires positive homogeneity but in fact there are important acceptability functionals that are not positively homogeneous.

Additionally: if we later on deal with conditional and multi-period mappings, translation-equivariance will be easier to handle than translation antivariance.

Another important group of acceptability type probability functionals is well known and widely used in economic theory under the name of *expected utility*. Such functionals are given by $\mathbb{E}(U(X))$, where U is an utility function. As a functional operating on random variables X expected utility generally is not an acceptability functional, because it lacks translation equivariance if the utility function is not linear. At least expected utility is concave, if the utility function is concave . If the utility is monotone, the expected utility is monotone with respect to first order stochastic dominance. If the utility is both monotonic and concave, the functional is monotone with respect to second order stochastic dominance[1].

[1]**[23]**, Proposition 2.59

1.2. Dual Representation of Acceptability functionals

In the following we consider acceptability functionals as defined on a linear space of random variables. If we use a $\boldsymbol{L}_p(\Omega, \mathcal{F}, \mu)$-space as domain-space, for each p the dual space can be identified with the space $\boldsymbol{L}_q(\Omega, \mathcal{F}, \mu)$, where $\frac{1}{p} + \frac{1}{q} = 1$ holds. The dual pairing between elements X of the space and elements Y of its dual is given by the expectation[2] $\mathbb{E}(X \cdot Y)$.

From the Fenchel-Moreau-Rockafellar Theorem[3] it follows that concave u.s.c. functionals can be represented in the following way:

$$\mathcal{A}(X) = \inf_{Z \in \boldsymbol{L}_q(\Omega,\mathcal{F},\mu)} \left\{ \mathbb{E}(X \cdot Z) - \mathcal{A}^*(Z) \right\}, \tag{1.2.1}$$

where $\mathcal{A}^*(Z) = \inf_{X \in \boldsymbol{L}_p(\Omega,\mathcal{F},\mu)} \left\{ \mathbb{E}(X \cdot Z) - \mathcal{A}(X) \right\}$ is the concave conjugate or Fenchel-Moreau conjugate of the functional $\mathcal{A}$. Basically this means that the functional equals its biconjugate.

An alternative representation is given by $\mathcal{A}(X) = \inf_{Z \in \boldsymbol{L}_q(\Omega,\mathcal{F},\mu)} \left\{ \mathbb{E}(X \cdot Z) - \mathcal{A}^*(Z) : Z \in Z_{\mathcal{A}} \right\}$. $Z_{\mathcal{A}}$ denotes the set of random variables from the dual space where $\mathcal{A}^*(Z)$ is finite.

The notion of conjugates has its origin in the Legendre-transform of a convex real function $f(x)$, which is a widely used concept in physics:

$$f^*(p) = \max_x \left\{ p \cdot x - f(x) \right\}$$

For differentiable f it is easy to see that x is a maximizer if $p = \frac{df(x)}{dx}$. While the normal representation of a function is given by a set of points (the graph) the conjugate represents the function by a set of tangent lines, specified by their slope and intercept.

The conjugate of a probability functional is a generalization of this idea: u.s.c. acceptability functionals can be represented by sets of tangent hyperplanes for the hypograph. These hyperplanes are defined by their intercept and their normal vectors, given by the supergradients of the functional.

The connection to the superdifferential of a functional is only one interesting feature of conjugates: It is also remarkable that most of the crucial properties of a functional can be derived from its dual representation.

For example by inserting into the conjugate representation it can easily be seen that a functional $\mathcal{A}$ is translation- equivariant if $\mathbb{E}(Z) = 1$ for all $Z \in Z_{\mathcal{A}}$ and it is monotonic if

[2] For more details see section 2.1.

[3] See [**28**], Theorem 5

all $Z \in Z_{\mathcal{A}}$ are nonnegative. Positive homogeneity follows from $\mathcal{A}^*(Z) \equiv 0$. There are also sufficient conditions for version independence and monotonicity with respect to stochastic dominance[4]. Moreover, under some additional assumptions these properties are not only sufficient but also necessary[5].

Using translation equivariance and monotonicity, any proper u.s.c. acceptability functional $\mathcal{A}$ has the representation[6]

$$\mathcal{A}(X) = \inf_{Z \in \boldsymbol{L}_q(\Omega, \mathcal{F}, \mu)} \{\mathbb{E}(X \cdot Z) - \mathcal{A}^*(Z) : \mathbb{E}(Z) = 1,\ Z \geq 0\} \tag{1.2.2}$$

In addition, if $\mathcal{A}$ can be represented by 1.2.2, then it must be an u.s.c acceptability functional.

One of the simplest examples for an acceptability functional is the average value-at-risk

$$AV@R_\alpha(X) = \frac{1}{\alpha}\int_0^\alpha G_X^{-1}(u)du,$$

where $G_X(u)$ is the c.d.f. of the random variable X.

This acceptability functional is also known under the name of conditional value-at-risk or tail value-at-risk. Its conjugate representation is given as follows[7]

Proposition 1.2.1. *(AV@R: dual properties) The dual representation of the average value-at-risk is given by* $AV@R_\alpha(X) = \inf\left\{\mathbb{E}(X \cdot Z) : \mathbb{E}(Z) = 1,\ 0 \leq Z \leq \frac{1}{\alpha}\right\}$.

In the rest of this work we will generalize the concepts of this introduction and develop the notion of conditional acceptability mappings. Those are mappings between $\boldsymbol{L}_p(\Omega, \mathcal{F}, \mu)$-spaces based on different σ -algebras, representing nontrivial information available. We will define and analyze them and also try to generalize the dual properties of probability functionals as much as possible to the case of acceptability mappings.

1.3. Outline

This dissertation is composed of five chapters and can broadly be divided into two parts: Chapter one gave a short introduction into the problem and recapitulates the most crucial results about unconditional acceptability functionals in an informal manner. The first main

[4][**23**], Theorem 2.30
[5][**23**], Theorem 2.31
[6][**23**], Corollary 2.32
[7][**23**], Theorem 2.34 (ii)

part contains two chapters that lay the mathematical fundamentals for the discussion of conditional acceptability mappings. The second part - consisting of chapters 4 and 5 - first of all analyzes conditional mappings and their properties based on the notions and results of part one. The other main theme in part two is the application of conditional acceptability mappings for constructing multi-period acceptability functionals.

Chapter 2 treats useful properties of $\boldsymbol{L}_p(\Omega, \mathcal{F}, \mu)$-spaces, the standard partial order for these spaces and the implications for defining an infimum for sets in such spaces. Such infima will be used later to define conjugate mappings.

Chapter 3 gives an overview of the main analytical results about mappings between partially ordered vector spaces. These definitions and theorems will be used as a basis for defining supergradients and conjugate mappings in part two.

In Chapter 4 concave (conditional) mappings between $\boldsymbol{L}_p(\Omega, \mathcal{F}, \mu)$-spaces as well as their supergradients and conjugate mappings are defined. Moreover we prove some basic propositions about conditional acceptability mappings.

Chapter 5 gives an introduction to multi-period acceptability functionals and analyzes two types of multi-period functionals that are based on conditional acceptability mappings: SEC-functionals and acceptability compositions. In particular acceptability compositions, their supergradients and conjugates are studied in more detail.

Part 1

Fundamentals

CHAPTER 2

$\boldsymbol{L}_p(\Omega, \mathcal{F}, \mu)$-Spaces as Partially Ordered Vector Spaces

This chapter reviews the theoretical foundations for the study of conditional acceptability mappings in chapter 4. Such mappings operate between spaces of random variables. The natural function spaces for dealing with random variables are the $\boldsymbol{L}_p(\Omega, \mathcal{F}, \mu)$-spaces. In section 2.1 we will briefly recall fundamental properties of such Banach-spaces and discuss how they are useful in the context of our investigation.

We abstain from dealing with the subject of Banach spaces in chapter 1, which just ggives an informal overview of single period acceptability functionals. Nevertheless it should be noted that $\boldsymbol{L}_p(\Omega, \mathcal{F}, \mu)$-spaces are indispensable for a mathematically sound analysis even in this case. The role of $\boldsymbol{L}_p(\Omega, \mathcal{F}, \mu)$-spaces in the theory of risk and acceptability functionals is well known and was presented e.g. by Ruszczyński [**33**] and Pflug and Römisch [**23**] .

Section 2.2 brings a second theme into the discussion: Acceptability functionals map into the real line - although their domain is some $\boldsymbol{L}_p(\Omega, \mathcal{F}, \mu)$-space. For real numbers it is absolutely clear how to order them and how to understand inequalities. Further, using inequalities we can define sub- and supergradients of functionals in the usual way, sketched in chapter 1. It is also possible to define conjugate functionals based on the notion of an infimum. This can be done in a meaningful way, because order completeness is an axiom for the real numbers: the existence of infima and suprema is guaranteed for bounded sets of real numbers.

These properties of real numbers are not guaranteed automatically for random variables. The notion of an $\boldsymbol{L}_p(\Omega, \mathcal{F}, \mu)$-space alone is not sufficient for this, an adequate partial order is needed. In Section 2.2 we will analyze the almost sure order - which is the most natural partial order for $\boldsymbol{L}_p(\Omega, \mathcal{F}, \mu)$-spaces.

Basically, we will follow [**3**], showing that it is possible to use meaningful inequalities in $\boldsymbol{L}_p(\Omega, \mathcal{F}, \mu)$-spaces and that $\boldsymbol{L}_p(\Omega, \mathcal{F}, \mu)$-spaces are order complete with respect to the almost sure order. After all, this will allow the definition of the infimum and supremum, both playing a key role in the investigation of conditional mappings, in particular for defining their conjugates.

2.1. $\boldsymbol{L}_p(\Omega, \mathcal{F}, \mu)$-Spaces: Basic Properties

Consider first measure spaces $(\Omega, \mathcal{F}, \mu)$, where Ω is a state space, $\mathcal{F}$ a $\sigma-$Algebra and μ a measure. In the context of risk-management of course we are interested particularly in the special case of probability spaces, where μ is a probability measure. Therefore, in the following we will use the language of probability theory, though many results remain valid for general measure spaces. Random variables are measurable real functions on Ω and the space of all random variables can be regarded as a vector space.

Recall now that two random variables are equivalent if they are equal $\mu-$almost surely. The p-norm of a random variable $X(\omega)$ is defined by the integral

$$\| X \|_p = \left(\int |X(\omega)|^p \, d\mu(\omega) \right)^{\frac{1}{p}},$$

where $1 \leq p < \infty$. This integral depends only on the equivalence class of the random variable X. The space $\boldsymbol{L}_p(\Omega, \mathcal{F}, \mu)$ can now be defined as a collection of equivalence classes of random variables for which the p-norm is finite. Although the distinction between random variables and equivalence classes can be neglected for many purposes, it will turn out to be critical when we want to define the infimum of a set in a space $\boldsymbol{L}_p(\Omega, \mathcal{F}, \mu)$.

In the context of probability spaces the above integral is an expectation and we will write

$$\| X \|_p = \mathbb{E}\left(|X|^p\right)^{\frac{1}{p}},$$

as is conventional in probability theory. For example, a space $\boldsymbol{L}_1(\Omega, \mathcal{F}, \mu)$ is just the collection of all $\mu-$integrable random variables, measurable with respect to the $\sigma-$Algebra $\mathcal{F}$.

For $p = \infty$ the norm is defined as the essential supremum of the absolute value and a space $\boldsymbol{L}_\infty(\Omega, \mathcal{F}, \mu)$ contains all essentially bounded random variables defined on the probability space $(\Omega, \mathcal{F}, \mu)$.

As we will see, $\boldsymbol{L}_p(\Omega, \mathcal{F}, \mu)$- spaces with $1 \leq p \leq \infty$ have many favorable properties: All of them are Banach spaces. In fact it will be clear in the next section that they are also Banach lattices. A special case are $\boldsymbol{L}_2(\Omega, \mathcal{F}, \mu)$-spaces, which are Hilbert spaces.

A key role in convex analysis for $\boldsymbol{L}_p(\Omega, \mathcal{F}, \mu)$- spaces is played by the elements of their dual spaces. They can be used to define subgradients as well as conjugate functionals and mappings and we will make use of them very frequently in later chapters. The (topological) dual space of a vector space E is given by the space of continuous linear functionals, mapping from E into $\mathbb{R}$. For the case of $\boldsymbol{L}_p(\Omega, \mathcal{F}, \mu)$- spaces it turns out that their dual spaces can be represented in a very meaningful way: The elements of the dual space can be identified with the elements of a certain $\boldsymbol{L}_q(\Omega, \mathcal{F}, \mu)$- space .

THEOREM 2.1.1. *(F. Riesz) If $1 < p, q < \infty$ are exponents such that $\frac{1}{p} + \frac{1}{q} = 1$ (conjugate exponents), then the integral*

$$H_Z(X) = \int X(\omega) \cdot Z(\omega) d\mu(\omega)$$

defines a continuous linear functional on $\boldsymbol{L}_p(\Omega, \mathcal{F}, \mu)$ for each $Z \in \boldsymbol{L}_q(\Omega, \mathcal{F}, \mu)$.

Moreover the norm dual of $\boldsymbol{L}_p(\Omega, \mathcal{F}, \mu)$ can be identified with the space $\boldsymbol{L}_q(\Omega, \mathcal{F}, \mu)$ itself[1]*.*

If $\boldsymbol{L}_1(\Omega, \mathcal{F}, \mu)$ is a $\sigma-$finite measure space then the norm dual of $\boldsymbol{L}_1(\Omega, \mathcal{F}, \mu)$ can be identified with the space $\boldsymbol{L}_\infty(\Omega, \mathcal{F}, \mu)$.

PROOF. For a stricter formulation and proofs see theorems 13.26 and 13.28 in [**3**] and theorems 31.16, 37.9 and 37.11 in [**2**]. □

As mentioned above we want to concentrate on probability spaces. This means that the continuous linear functionals are given by the expectations $\mathbb{E}(X \cdot Z)$ and can be identified with random variables $Z \in \boldsymbol{L}_q(\Omega, \mathcal{F}, \mu)$. Furthermore, because probability spaces are always $\sigma-$finite the representation of dual spaces by $\boldsymbol{L}_q(\Omega, \mathcal{F}, \mu)$ will work for $1 \leq p, q \leq \infty$.

It is possible to interpret the dual function $Z(\omega)$ as the Radon-derivative of a signed measure $\nu(A) = \int_A Z(\omega) d\mu(\omega)$. $Z(\omega)$ then defines a change of measures and the expectation $\mathbb{E}(X \cdot Z)$ is the expectation of X under the new signed measure ν.

In the theory of acceptability functionals unconditional expectation suffices for defining supergradients and conjugates. While in the unconditional case we have mappings from some $\boldsymbol{L}_p(\Omega, \mathcal{F}, \mu)$- space into the real line, conditional mappings operate between two spaces $\boldsymbol{L}_p(\Omega, \mathcal{F}, \mu)$ and $\boldsymbol{L}_{p'}(\Omega, \mathcal{F}_1, \mu)$. If we want to define supergradients and concave conjugates, we need continuous linear mappings between these spaces. It turns out that

[1]The mapping $Z \mapsto H_Z(\cdot)$ is a lattice isometry from $\boldsymbol{L}_q(\Omega, \mathcal{F}, \mu)$ onto $\boldsymbol{L}'_p(\Omega, \mathcal{F}, \mu)$.

this role is played by the conditional expectations $\mathbb{E}(X \cdot Z|\mathcal{F})$, where Z is an element of the dual space $\boldsymbol{L}_q(\Omega, \mathcal{F}, \mu)$: Theorem 2.1.1 can be generalized in the following way.

COROLLARY 2.1.2. *If $1 < p, q < \infty$ are exponents such that $\frac{1}{p} + \frac{1}{q} = 1$ (conjugate exponents), then for all $Z \in \boldsymbol{L}_q(\Omega, \mathcal{F}, \mu)$ the mapping $E_Z(\bullet) : \boldsymbol{L}_p(\Omega, \mathcal{F}, \mu) \to \boldsymbol{L}_1(\Omega, \mathcal{F}_1, \mu)$*

$$E_Z(X) = \mathbb{E}(X \cdot Z|\mathcal{F})$$

is a continuous linear mapping.

PROOF. The linearity of conditional expectation is clear. Assume now that $Z \in \boldsymbol{L}_q(\Omega, \mathcal{F}, \mu)$. The dual space contains those Z, for which $\mathbb{E}(\bullet \cdot Z)$ is a continuous mapping. That means that if the sequence X_n converges to X in the p-norm, the sequence $\mathbb{E}(X_n \cdot Z)$ will converge to $\mathbb{E}(X \cdot Z)$, or $\mathbb{E}((X_n - X) \cdot Z) \to 0$ if $n \to 0$.

As $|\bullet|$ is convex we can apply Jensen's inequality and get

$$\mathbb{E}(|\mathbb{E}((X_n - X) \cdot Z|\mathcal{F})|) \leq \mathbb{E}(\mathbb{E}(|(X_n - X) \cdot Z|\,|\mathcal{F})) = \mathbb{E}(|(X_n - X) \cdot Z|)\,.$$

Hölder's inequality then gives

$$\mathbb{E}(|(X_n - X) \cdot Z|) = \|(X_n - X) \cdot Z\|_1 \leq \|X_n - X\|_p \cdot \|Z\|_q$$

As X_n converges to X in the p-norm and Z is q-integrable, it follows that also $\|X_n - X\|_p \cdot \|Z\|_q \to 0$. Because of this the integral $\mathbb{E}(|\mathbb{E}((X_n - X) \cdot Z|\mathcal{F})|)$ will also converge to zero and the conditional expectation $E_Z(X)$ is a continuous mapping. □

Corollary 2.1.2 guarantees that the conditional mapping $\mathbb{E}(X \cdot Z|\mathcal{F})$ is continuous as a mapping into $\boldsymbol{L}_1(\Omega, \mathcal{F}_1, \mu)$ under reasonable requirements on the dual variable Z. Later we will need more, namely conditional expectations $\mathbb{E}(X \cdot Z|\mathcal{F})$, together with some restrictions on Z that guarantee the conditional expectation to be a continuous mapping into some space $\boldsymbol{L}_{p'}(\Omega, \mathcal{F}_1, \mu)$.

Such restricted conditional expectations are defined in chapter 4 and will be used to define supergradients and conjugates of conditional mappings. This approach ensures that the results regarding nonsmooth analysis for general mappings in chapter 3 will be applicable.

2.2. $\boldsymbol{L}_p(\Omega, \mathcal{F}, \mu)$-Spaces as Banach Lattices

As said above, the spaces $\boldsymbol{L}_p(\Omega, \mathcal{F}, \mu)$ are collections of equivalence classes of measurable functions: The equivalence relation is defined by the relation »equal with probability one«.

In this context we can understand all equalities in the almost sure sense. Furthermore, the natural partial order for $\boldsymbol{L}_p(\Omega, \mathcal{F}, \mu)$-spaces is based on inequalities that hold almost everywhere.

If we use this partial order we will see that $\boldsymbol{L}_p(\Omega, \mathcal{F}, \mu)$-spaces are order complete Riesz spaces with the pointwise algebraic and lattice operations. Moreover it is well known that for $1 \leq p \leq \infty$ all the $\boldsymbol{L}_p(\Omega, \mathcal{F}, \mu)$-spaces are in fact Banach lattices.

A critical precondition for defining something like conjugate mappings is the notion and the existence of infima. Basically order completeness is needed for using the infimum of sets in $\boldsymbol{L}_p(\Omega, \mathcal{F}, \mu)$-spaces in a reasonable way. This will also enable us to define Fenchel-Moreau conjugates for concave mappings between $\boldsymbol{L}_p(\Omega, \mathcal{F}, \mu)$-spaces and use them for the analysis of conditional acceptability mappings.

Definition 2.2.1. (ordered vector space) An **ordered vector space** V is a vector space with an order relation $\geq$ that is compatible with the algebraic structure of V in the following sense:

a) $x \geq y \Rightarrow x + z \geq y + z$ for any $z \in V$

b) $x \geq y \Rightarrow \alpha \cdot x \geq \alpha \cdot y$ for each $\alpha \geq 0$

It should be noted that it is not necessary that the order in definition 2.2.1 is a total order - a partial order relation (reflexive, antisymmetric and transitive) is sufficient. For $x \geq y$ we will also say informally: »x dominates y«.

From definition 2.2.1 it is clear that in an ordered vector space V, the set $\{x \in V : \; x \geq 0\}$ is a pointed convex cone. This cone is called the positive (or non-negative) cone of V, denoted V^+.

The next step is to define the infimum and the supremum for pairs of elements of a vector space:

Definition 2.2.2. (infimum of a pair) An element $z \in V$ is the **infimum** $z = \inf\{x, y\}$ of the pair of elements $x, y \in V$ if

a) z is a lower bound of the set $\{x, y\}$- that means: $x \geq z$ and $y \geq z$.

b) z is the largest such bound: i.e. $x \geq u$ and $y \geq u$ imply $z \geq u$.

The supremum of two elements is defined similarly, replacing $\geq$ by $\leq$.

We denote the infimum and the supremum of two elements by $\inf\{x, y\} = x \wedge y$ and $\sup\{x, y\} = x \vee y$. The functions $\vee$ and $\wedge$ are called »lattice operators« on V.

Based on the lattice operators we can define lattices and Riesz spaces in the following way:

DEFINITION 2.2.3. (lattice) A partially ordered set $(V, \geq)$ is a **lattice** if each pair of elements $x, y \in V$ has a supremum and an infimum in V.

From this definition it is clear that in a lattice every finite nonempty set has a supremum and an infimum. We can write $\inf\{x_1, \ldots, x_n\} = \bigwedge_{i=1}^{n} x_i$ and $\sup\{x_1, \ldots, x_n\} = \bigvee_{i=1}^{n} x_i$.

DEFINITION 2.2.4. (Riesz space) An ordered vector space $(V, \geq)$ that is also a lattice is called **Riesz space** or vector **lattice**.

Although Riesz spaces are based only on a partial order, there are many affinities between Riesz spaces and the real numbers. For instance it is possible to define the positive part x^+, the negative part x^- and the absolute value $|x|$ of a vector x in a Riesz space.

$$x^+ = x \vee 0 \tag{2.2.1}$$

$$x^- = x \wedge 0 \tag{2.2.2}$$

$$|x| = x \vee (-x) \tag{2.2.3}$$

In this context the typical equations $x = x^+ - x^-$ and $|x| = x^+ + x^-$ both hold.

A subset A of a Riesz space V is called **order bounded** from above if there exists a vector $u \in V$ that dominates each element of A. The subset is called order bounded from below if there is a vector $v \in V$ that is dominated by each element of A. The subset A is called order bounded if it is both order bounded from below and above.

DEFINITION 2.2.5. (infimum, supremum) An element u of a Riesz space $(V, \geq)$ is the **infimum** of a nonempty subset $A \subseteq V$ if

a) l is a lower bound of the set A: $a \geq l$ for all $a \in A$

b) $a \geq v$ for all $a \in A$ implies $l \geq v$.

Again the supremum of the set is defined similarly, replacing $\geq$ by $\leq$.

It is clear that any subset of a Riesz space has at most one supremum and one infimum. If the set is not bounded in one of the directions the infimum, the supremum or both do not exist. The critical question is whether there exists an infimum if the set is bounded below and whether there exists a supremum if it is bounded above. For the real numbers this is ensured by the completeness axiom: any bounded subset of real numbers has both an infimum and a supremum. For Riesz spaces in general this is not mandatory.

DEFINITION 2.2.6. (order completeness)

A Riesz space is **order complete** (Dedekind complete) if every nonempty subset that is order bounded from below has an infimum.

Equivalently the space is order complete if every nonempty subset that is order bounded from above has a supremum. Also, any order bounded subset will have both infimum and supremum.

A potential characteristic of Riesz spaces, used later on to assure order completeness for $\boldsymbol{L}_p(\Omega, \mathcal{F}, \mu)$-spaces is the Archimedean property, defined as follows:

DEFINITION 2.2.7. A Riesz space $(V, \geq)$ is **Archimedean** if whenever $0 \leq nx \leq y$ for all $n \in \mathbb{N}$ and some $y \in V^+$, then $x = 0$.

It should be noted that every order complete Riesz space is Archimedean [2] , but generally the converse is not true.

There are even more analogies between real numbers and the elements of Riesz spaces. If we look at identities and inequalities using the lattice operators $\vee$ and $\wedge$, positive and negative parts and the absolute value we have the following proposition:

PROPOSITION 2.2.8. *Every lattice identity that is true for real numbers is also true in every Archimedean Riesz space.*

If a lattice inequality is true for real numbers, then it is true in any Riesz space.

PROOF. See Theorem 8.6 and Corollary 8.7 in [**3**]. □

That means that in Riesz spaces inequalities like $|x+y| \leq |x| + |y|$, $|\alpha x| \leq |\alpha| |x|$ will hold exactly as for real numbers.

In arguing that $\boldsymbol{L}_p(\Omega, \mathcal{F}, \mu)$-spaces are order complete Riesz spaces, we have to go even deeper into mathematical details. Not only the order relation but also some special properties of the norms (p-norms) involved play a crucial role.

[2]See Lemma 8.4 in [**3**].

DEFINITION 2.2.9. (lattice norm) If a norm $\|\cdot\|$ has the property that $|x| \geq |y|$ implies $\|x\| \geq \|y\|$ it is called a **lattice norm**.

Basically a lattice norm preserves the order given by the order relation$\geq$ and builds the link between general normed (Banach) spaces and ordered (Riesz) spaces.

DEFINITION 2.2.10. (Banach lattice) A Riesz space equipped with a lattice norm is called a **normed Riesz space**. A (Cauchy-)complete normed Riesz space is called a **Banach lattice**.

It is a crucial fact that a normed Riesz space is a Banach lattice if and only if every increasing positive Cauchy sequence is norm convergent[3].

An important connection between the topological and the order structure of Banach lattices is called the »order continuity of the norm«:

A decreasing net is a net $\{x_\alpha\}_{\alpha \in D}$ - defined on some directed set $(D, \geq)$ - with $\alpha_1 \geq \alpha_2 \Longrightarrow x_{\alpha_1} \leq x_{\alpha_2}$ and an increasing net is a net with $\alpha_1 \geq \alpha_2 \Longrightarrow x_{\alpha_1} \geq x_{\alpha_2}$. We use the notation $x_\alpha \downarrow x$ to indicate that $\{x_\alpha\}_{\alpha \in D}$ is a decreasing net and x is its infimum. Just as well $x_\alpha \uparrow x$ means an increasing net with supremum x. We can also formulate this in the following way: $x_\alpha \uparrow x$ means that $\bigvee x_\alpha = x$ and $x_\alpha \downarrow x$ means that $\bigwedge x_\alpha = x$.

If there are two nets $\{y_\alpha\}_{\alpha \in D}$ and $\{z_\alpha\}_{\alpha \in D}$ with $y_\alpha \leq x_\alpha \leq z_\alpha$ and both $y_\alpha \uparrow x$ and $z_\alpha \downarrow x$, we say that x_α **order converges** to x, i.e. $x_\alpha \xrightarrow{o} x$.

DEFINITION 2.2.11. (order continuity) A lattice (semi)norm $\|\cdot\|$ on a Riesz space is order continuous if $x_\alpha \downarrow 0$ implies $\|x_\alpha\| \downarrow 0$.

REMARK. An equivalent characterization of order continuity is the following: $\|\cdot\|$ is an order continuous lattice (semi)norm if and only if $0 \leq x_\alpha \uparrow x$ implies $\|x_\alpha - x\| \downarrow 0$.

Basically this means that an order continuous norm ensures that order convergence implies norm convergence.

Under very general conditions Banach lattices with order continuous norm are order complete. This will enable us to use infima in a meaningful way in later chapters. In particular we will be able to define conjugate mappings for mappings between $\boldsymbol{L}_p(\Omega, \mathcal{F}, \mu)$-spaces. First we need the following Lemma[4]:

[3] See Theorem 9.3 in [**3**]

[4] Lemma 12.8 from [**1**]

LEMMA 2.2.12. *If* $0 \leq x_\alpha \uparrow\leq x$ *holds in an Archimedean Riesz space* E*, then the set* $D = \{y \in E : \forall\alpha : x_\alpha \leq y\}$ *is directed downward and* $y - x_\alpha \downarrow_{y,\alpha} 0$*.*

PROOF. Clearly, the set D is directed downward. Let u be such that $0 \leq u \leq y - x_\alpha$ holds for all α and $y \in D$. Then $x_\alpha \leq y - u$ holds for all α and from this it follows that $y - u \in D$ for all $y \in D$. By induction it can be shown that $y - n \cdot u \in D$ for all $y \in D$ and all $n \in \mathbb{N}$. In particular - because $x \in D$ - we have $0 \leq x_\alpha \leq x - n \cdot u$ and hence $0 \leq n \cdot u \leq x$.

Since E is Archimedean, it follows that $u = 0$ and therefore $y - x_\alpha \downarrow_{y,\alpha} 0$. □

Now it is possible to state a fundamental result regarding order completeness of Banach lattices[5]:

PROPOSITION 2.2.13. *An Archimedean Banach lattice* E *with order continuous norm is order complete.*

PROOF. Let $\{x_\alpha\}$ be a nonnegative increasing, order bounded net in E: $0 \leq x_\alpha \uparrow\leq x$. By Lemma 2.2.12 there exists a net $\{y_\lambda\} \subseteq E$ with $y_\lambda - x_\alpha \downarrow 0$. Because E has order continuous norm it follows that $\|y_\lambda - x_\alpha\| \downarrow 0$. That means that for each $\varepsilon > 0$ there exist α_0,λ_0 such that $\|y_\lambda - x_\alpha\| < \varepsilon$ holds for all $\lambda \geq \lambda_0$ and $\alpha \geq \alpha_0$.

Using the inequality $\|x_\alpha - x_\beta\| \leq \|x_\alpha - y_\lambda\| + \|x_\beta - y_\lambda\|$ we see that $\|x_\alpha - x_\beta\| < 2 \cdot \varepsilon$ holds for all $\alpha, \beta \geq \alpha_0$. Hence $\{x_\alpha\}$ is a norm Cauchy net.

Because E is complete it also contains the limit of $\{x_\alpha\}$, which we will denote by y. So $y - x_\alpha \downarrow_{y,\alpha} 0$ implies $x_\alpha \uparrow y$, y is the supremum of $\{x_\alpha\}$ and therefore E is order complete. □

If we want to use the concepts of order theory in the context of $\boldsymbol{L}_p(\Omega, \mathcal{F}, \mu)$-spaces, we first have to define an order relation. This can be done in a natural way using the concept of almost sure inequalities:

DEFINITION 2.2.14. (almost sure order) The partial order $X \geq Y$ for elements of a (probability) space $\boldsymbol{L}_p(\Omega, \mathcal{F}, \mu)$ is given by the relation

$$X \geq Y \Leftrightarrow \mu(\{\omega \in \Omega : X(\omega) < Y(\omega)\}) = 0.$$

We will also write this relation as $X \geq Y$ *a.s.* .

[5]See Theorem 12.9 in [**1**]. Our statement is in principle contained in the proof ((1) ⇒ (2)), but we emphasize the necessity of the archimedean property.

Remark. In the following we will often use the simple notation $X \geq Y$. So any inequality between random variables should be understood in the almost sure sense given by definition 2.2.14, whether $a.s.$ is stated explicitly or not. Also equations between random variables generally must be understood to hold in the almost sure sense.

With this ordering each $\boldsymbol{L}_p(\Omega, \mathcal{F}, \mu)$ is a Riesz space, and the p-norm is an order continuous norm. Order convergence is synonymous with μ-almost everywhere convergence. Moreover for $1 \leq p \leq \infty$ a $\boldsymbol{L}_p(\Omega, \mathcal{F}, \mu)$ is also a Banach lattice. This is the assertion of the following famous theorem[6]:

Theorem 2.2.15. *(Riesz-Fischer) For $1 \leq p \leq \infty$ any Riesz space $\boldsymbol{L}_p(\Omega, \mathcal{F}, \mu)$ equipped with the p-norm is a Banach Lattice.*

Proof. We give the proof for $1 \leq p < \infty$. Basically we have to show that every increasing positive Cauchy sequence is norm convergent.

Let $0 \leq X_n \uparrow$ be a Cauchy sequence in $\boldsymbol{L}_p(\Omega, \mathcal{F}, \mu)$. Then there exists some $M > 0$ such that $0 \leq \int (X_n(\omega))^p \, d\mu(\omega) = \|X_n\|_p \uparrow \leq M$. By the monotone convergence theorem there exists a random variable $0 \leq X \in \boldsymbol{L}_1(\Omega, \mathcal{F}, \mu)$ such that $X_n^p \uparrow X$ a.s.. then $0 \leq Y = X^{\frac{1}{p}} \in \boldsymbol{L}_p(\Omega, \mathcal{F}, \mu)$. From the Lebesgue Dominated Convergence theorem $\|X_n - Y\|_p \to 0$ follows. □

The infimum in the context of $\boldsymbol{L}_p(\Omega, \mathcal{F}, \mathbb{P})$-spaces is defined as a special case of definition 2.2.5 in the following way:

Definition 2.2.16. (infimum) Let $V = \boldsymbol{L}_p(\Omega, \mathcal{F}, \mathbb{P})$ be a random space and $\geq$ the almost sure order from definition 2.2.14. A random variable $X_0 \in V$ is the **infimum** of a nonempty subset $A \subseteq V$, if

a) X_0 is a lower bound of the set A, i.e. $Y \geq X_0 \, a.s.$ for all $Y \in A$

b) $Y \geq Z$ for all $Y \in A$ implies $X_0 \geq Z \, a.s.$

We denote the infimum of A as $\inf A$.

Again the supremum of the set is defined similarly, replacing $\geq$ by $\leq$.

It is critical to remember that the elements of $\boldsymbol{L}_p(\Omega, \mathcal{F}, \mu)$-spaces are equivalence-classes of random variables and not concrete random variables: A pointwise infimum, operating on an uncountable set of concrete measurable random variables would not make much

[6] See Theorem 13.5 in [**3**]

sense, because such an infimum easily could be not measurable: Take for example an uncountable set of measurable random functions which are equal to one for each $\omega \in \Omega$ and equal to zero for a single ω , different for each function. The pointwise minimum of such a set would be a function that equals one for $\omega \in \Omega$ where all the individual functions are equal to one and that equals zero where any of the individual functions is equal to zero. Although the individual functions are measurable, it will be possible to construct the zeros in a way such that the set of zero points is not contained in the relevant σ-algebra $\mathcal{F}$. That means that the infimum will not be measurable.

The usage of representatives avoids this problem, resulting in measurable infima for measurable representants. Because all the individual functions equal one almost everywhere, they are members of the same equivalence class, represented by the same random variable and their infimum will be represented also by a function that is one almost everywhere.

The correct approach is therefore to use representants at first and switch over to concrete individual random functions only after taking the infimum.

Given this definition it is time to address the issue of order completeness for $\boldsymbol{L}_p(\Omega, \mathcal{F}, \mu)$-spaces. Although there are fundamental differences between $\boldsymbol{L}_\infty$-spaces and $\boldsymbol{L}_p$-spaces with finite p - for instance the ∞-norm is not order continuous - it turns out that in both cases we have order complete spaces. To show order completeness we use proposition 2.2.13. The first question is, whether $\boldsymbol{L}_p$-spaces have the Archimedean property:

LEMMA 2.2.17. *For $1 \leq p \leq \infty$ any Riesz space $\boldsymbol{L}_p(\Omega, \mathcal{F}, \mu)$ equipped with the p-norm is an Archimedean Banach Lattice.*

PROOF. $\boldsymbol{L}_p(\Omega, \mathcal{F}, \mu)$ is an Banach lattice by theorem 2.2.15. Consider $X \in \boldsymbol{L}_p(\Omega, \mathcal{F}, \mu)$ with $0 \leq n \cdot X \leq Y$ for all $n \in \mathbb{N}$ and some $0 < Y \in \boldsymbol{L}_p(\Omega, \mathcal{F}, \mu)$. Assume now that $X(\omega) > 0$ for $\omega \in A$, A being some set with positive probability. Taking (lattice) norms we get $0 \leq n \cdot \left(\int X(\omega)^p d\mu(\omega)\right)^{1/p} \leq \left(\int Y(\omega)^p d\mu(\omega)\right)^{1/p}$. Because $X(\omega) = 0$ for $\omega \in \Omega \setminus A$ and $X(\omega) > 0$ for $\omega \in A$ it follows that $\left(\int X(\omega)^p d\mu(\omega)\right)^{1/p} > 0$. But that would mean that $\mathbb{R}$ is not Archimedean, which is not the case. Hence A must have zero probability. That shows that $\boldsymbol{L}_p(\Omega, \mathcal{F}, \mu)$ -spaces must also be Archimedean[7]. □

THEOREM 2.2.18. *(order completeness) The Banach lattices $\boldsymbol{L}_p(\Omega, \mathcal{F}, \mu)$ are order complete.*

[7]See Theorem 13.7 in [**3**].

PROOF. We will give the proof for $1 \le p < \infty$. All spaces $\boldsymbol{L}_p(\Omega, \mathcal{F}, \mu)$ are Archimedean Banach lattices by lemma 2.2.17. Furthermore, p-norms are order continuous for $1 \le p < \infty$:

Assume $X_\alpha \downarrow 0$ in $\boldsymbol{L}_p(\Omega, \mathcal{F}, \mu)$ and let $\int |X_\alpha|^p \, d\mu \downarrow s$. We have to show that $s = 0$.

Select a sequence of indices $\{\alpha_n\}$ with $\alpha_{n+1} \ge \alpha_n$ and $\int |X_{\alpha_n}|^p \, d\mu \downarrow s$. Let now $|X_{\alpha_n}|^p \downarrow X \ge 0$ and fix some index α. For each n there exists some index β_n such that $\beta_n \ge \alpha$ and $\beta_n \ge \alpha_n$ and $\beta_{n+1} \ge \beta_n$ for all n.

If $|X_{\beta_n}|^p \downarrow Y \ge 0$, then $X \ge Y$ and $\int |X|^p \, d\mu = \int |Y|^p \, d\mu$, hence $X = Y$. This means that $X = Y \le |X_{\beta_n}|^p \le |X_\alpha|^p$ must hold. Because of $X_\alpha \downarrow 0$ we infer that $X \downarrow 0$. Therefore $\int |X_{\alpha_n}|^p \, d\mu \downarrow 0$ and hence $\lim_{n \to \infty} \int |X_\alpha|^p \, d\mu = 0$.

Together - using proposition 2.2.13 - it follows that the spaces $\boldsymbol{L}_p(\Omega, \mathcal{F}, \mu)$ are order complete. $\square$

From theorem 2.2.18 it is clear that it makes sense to use the supremum or infimum with respect to the almost sure ordering $\ge$ in the context of $\boldsymbol{L}_p$-spaces in a similar way as the supremum or infimum of real numbers. Summing up the argument goes as follows: $\boldsymbol{L}_p(\Omega, \mathcal{F}, \mu)$-spaces are Banach lattices by theorem 2.2.15 and have order continuous norm by theorem 2.2.18. Therefore - because Banach lattices with order continuous norm are order complete by proposition 2.2.13, $\boldsymbol{L}_p(\Omega, \mathcal{F}, \mu)$-spaces are order complete.

The infimum for $\boldsymbol{L}_p(\Omega, \mathcal{F}, \mathbb{P})$-spaces has the useful property that the sequence of conditional expectation and infimum can be interchanged if the infimum is attained:

PROPOSITION 2.2.19. *Let $S \subset \boldsymbol{L}_p(\Omega, \mathcal{F}, \mathbb{P})$ be an order bounded set of random variables. Then*

$$\inf \{\mathbb{E}(X|\mathcal{F}) : X \in S\} = \mathbb{E}(\inf \{X : X \in S\} | \mathcal{F})$$

holds almost surely, if $\inf S = \inf \{X : X \in S\} \in S$.

PROOF. The infimum of the set S is a lower bound of S: $\inf \{X : X \in S\} \le X$ for all $X \in S$. From the monotonicity of conditional expectation it follows that

$$\mathbb{E}(\inf \{X : X \in S\} | \mathcal{F}) \le \mathbb{E}(X|\mathcal{F})$$

for all $X \in S$. This means that $\mathbb{E}(\inf \{X : X \in S\} | \mathcal{F})$ is a lower bound for $\{\mathbb{E}(X|\mathcal{F}) : X \in S\}$.

Because $\{\mathbb{E}(X|\mathcal{F}) : X \in S\}$ is order bounded from below, the infimum for this set must exist. Let us denote it by $Y = \inf\{\mathbb{E}(X|\mathcal{F}) : X \in S\}$ and assume that it is different from $\mathbb{E}(\inf\{X : X \in S\}|\mathcal{F})$.

To be the infimum of $\{\mathbb{E}(X|\mathcal{F}) : X \in S\}$, first of all Y must be a lower bound: $Y \leq \mathbb{E}(X|\mathcal{F})$ for all $X \in S$. But then, $Y \leq \mathbb{E}(\inf\{X : X \in S\}|\mathcal{F})$ must hold, because $\inf\{X : X \in S\} \in S$. Second Y has to be the greatest lower bound of $\{\mathbb{E}(X|\mathcal{F}) : X \in S\}$, and as $\mathbb{E}(\inf\{X : X \in S\}|\mathcal{F})$ is a lower bound, this also means that

$$Y \geq \mathbb{E}(\inf\{X : X \in S\}|\mathcal{F})$$

.

Together this results in $Y = \mathbb{E}(\inf\{X : X \in S\}|\mathcal{F})$, which contradicts the assumption that Y is a lower bound, different from $\mathbb{E}(\inf\{X : X \in S\}|\mathcal{F})$. $\square$

CHAPTER 3

Nonsmooth Convex Analysis on Partially Ordered Vector Spaces

This chapter reviews relevant results from Papageorgiou [**21**]. We concentrate on those theorems that are most useful for our study of supergradients and concave conjugates of conditional acceptability mappings below and want to lay a solid basis for the concepts and definitions of chapter 4. On the other hand we will neglect a lot of theorems - for instance all propositions about Gateaux differentiability, Lipschitz-continuity and quasiconcavity.

In his original work Papageorgiou writes in terms of convex mappings and subgradients and we will follow this convention in this chapter. Later on in chapter 4 we will give in detail all relevant definitions for the case of concave mappings between $\boldsymbol{L}_p\left(\Omega, \mathcal{F}, \mu\right)$-spaces: They are a special case of the theory developed by Papageorgiou and basically whether concavity or convexity is used is only a matter of sign - of course altering the role of concepts like properness and semicontinuity. While in the convex case the notion of subdifferentials is the relevant concept for local approximation, in the concave case this role will be taken by the superdifferential. This said it should be kept in mind that statements about convex mappings can easily be translated into propositions about concave mappings, which we will do implicitly later on.

Finally it should be noted that in chapter 3 we follow the original notation of Papageorgiou as closely as possible, while in chapter 4 we will use the usual notation of probability theory again. The main difference will be that in chapter 3 X and Y denote spaces, whereas in chapter 4 capital letters will denote random variables. Moreover Papageorgiou denotes linear mappings by $A(\bullet)$, whilst in $\boldsymbol{L}_p\left(\Omega, \mathcal{F}, \mu\right)$-spaces the linear mappings are given by the conditional expectations $\mathbb{E}\left(\bullet \cdot Z|\mathcal{F}\right)$ with Z from the dual space.

3.1. Basic Assumptions

Papageorgiou considers proper convex mappings between spaces X and Y, having the following properties:

- X and Y are Hausdorff topological vector spaces and locally convex.

- Y is normal for its topology, which means that any disjoint closed sets E, F can be separated by neighborhoods.
- Y is partially ordered based on a proper convex cone. Moreover Y is order complete.
- Order completeness and some additional conditions ensure that order convergence[1] and relative uniform convergence are equal.
- For Y an extended space $\overline{Y} = Y \cup \{+\infty, -\infty\}$ is defined, where $+\infty, -\infty$ are a greatest and a smallest element with respect to the order.

These conditions are applicable in the case of $\boldsymbol{L}_p(\Omega, \mathcal{F}, \mu)$ spaces with $1 \leq p \leq \infty$: They are normed and metric spaces hence also Hausdorff, normal and locally convex. As we have seen in section 2.2, they are order complete with their natural order based on the cone of almost sure nonnegative random variables. For such $\boldsymbol{L}_p(\Omega, \mathcal{F}, \mu)$-spaces order convergence is equivalent to μ-a.e. convergence. Also the so called diagonal property[2] holds and together with order completeness this ensures the equivalence of order convergence and relative uniform convergence for $\boldsymbol{L}_p(\Omega, \mathcal{F}, \mu)$ spaces. It is also possible to extend them by infinite values, which we will do in chapter 4.

Convexity, properness and the effective domain are defined in the usual way:

Definition 3.1.1. (convexity) A mapping $f : X \to \overline{Y}$ is called **convex** if and only if $f(\lambda x + (1-\lambda)z) \leq \lambda f(x) + (1-\lambda) f(z)$ for $0 \leq \lambda \leq 1$.

The effective domain of a convex mapping f is given by $\operatorname{dom} f = \{x \in X : f(x) < +\infty\}$

A convex mapping f is called proper if $\operatorname{dom} f \neq \emptyset$ and it does not take the value $-\infty$.

3.2. Subgradients

Papageorgiou distinguishes between algebraic subdifferentials and topological subdifferentials. When we analyze conditional mappings in chapter 4, we will concentrate on the

[1] A sequence $\{x_n\}_{n\in\mathbb{N}}$ c*onverges relatively uniformly* to x if and only if there is an element $z \in \mathbf{K}_Y^+$ such that $|x_n - x| \leq \lambda_n z$, where $\lambda_n \in \mathbb{R}^+$ and $\lambda_n \downarrow 0$.

[2] A vector lattice Y is said to have the diagonal property if whenever

1. $\{x_{nm}\}_{n,m\in\mathbb{N}} \subseteq Y$
2. $x_{nm} \overset{o}{\to} x_n \ \forall n \in \mathbb{N}$
3. $x_n \overset{o}{\to} x$

then there is a diagonal subsequence $\{x_{nm_n}\}_{n\in\mathbb{N}}$ which order converges to x.

latter notion, which is based on continuous linear mappings. Nevertheless the more general algebraic subdifferential is a useful tool for proving the existence of the topological subgradient in proposition 3.2.2 below.

DEFINITION 3.2.1. (subdifferential, superdifferential) Let $f : X \to \overline{Y}$ be a convex mapping. The set

$$\partial^\alpha f(x_0) := \{A \in L(X,Y) : A(x - x_0) \leq f(x) - f(x_0), \forall x \in \mathrm{dom} f\},$$

where $L(X,Y)$ are the linear mappings between X and Y, is called the *algebraic subdifferential* of f at x_0.

Similarly the set

$$\partial f(x_0) := \{A \in \mathcal{L}(X,Y) : A(x - x_0) \leq f(x) - f(x_0), \forall x \in \mathrm{dom} f\},$$

where $\mathcal{L}(X,Y)$ denotes the space of continuous linear mappings between X and Y is called the *(topological) subdifferential* of f at x_0.

The elements of $\partial^\alpha f(x_0)$ and $\partial f(x_0)$ are called the *algebraic subgradients* and the s*ubgradients* of f at x_0.

Algebraic and topological superdifferentials can be defined easily by replacing $\leq$ by $\geq$.

Concerning our analysis of conditional mappings, one of the key assertions in [**21**] is the following:

PROPOSITION 3.2.2. *If Y is normal and f is continuous at some $x_0 \in \mathrm{dom}\, f$ then $\partial^\alpha f(x) = \partial f(x)$ for all $x \in X$. Under these conditions $\partial f(x_0) \neq \emptyset$ holds for all $x \in \mathrm{int}\,\mathrm{dom}\, f$.*

PROOF. For a complete proof see lemma 3.2 in [**21**].

The idea of the proof is the following: Generally $\partial^\alpha f(x) \supseteq \partial f(x)$ holds and for all $x \in \mathrm{int}\,\mathrm{dom}\, f$: $\partial^\alpha f(x) \neq \emptyset$. Subsequently it can be shown that any $A \in \partial^\alpha f(x)$ is continuous at x_0 and therefore - by linearity - $A \in \mathcal{L}(X,Y)$, which is equivalent to exact equality: $\partial^\alpha f(x) = \partial f(x)$.

From this proposition - and from the fact that $\partial^\alpha f(x) \neq \emptyset$ for all $x \in \mathrm{int}\,\mathrm{dom}\, f$ - it can be concluded that $\partial f(x_0) \neq \emptyset$ for all $x \in \mathrm{int}\,\mathrm{dom}\, f$. □

Based on this general notion of subgradients, some rules of subdifferential calculus are developed in the following. First there is a generalization of the Moreau-Rockafellar formula for the gradient of a sum of functions:

Proposition 3.2.3. *If Y is a normal space and $f_1, f_2 : X \to \overline{Y}$ are convex mappings with f_1 continuous at some $x_0 \in \operatorname{dom} f_2$ then*

$$\partial\left(f_1(x) + f_2(x)\right) = \partial f_1(x) + \partial f_2(x)\ \forall x \in X$$

holds.

Proof. See theorem 4.1 in [**21**]. □

Remark. In general, only $\partial\left(f_1(x) + f_2(x)\right) \supseteq \partial f_1(x) + \partial f_2(x)\ \forall x \in X$ holds.

Papageorgiou also gives a chain rule for the concatenation $g \circ f$ under some technical assumptions. This chain rule is based on a chain rule for the concatenation of an affine mapping with a convex mapping:

Proposition 3.2.4. *Let $\mathbf{A} \in \mathcal{L}(X, Y)$ with $\operatorname{im} \mathbf{A} = Y$ ($\mathbf{A}$ is surjective) and consider the affine mapping $\alpha(x) = \mathbf{A}x + y$ for some (fixed) $y \in Y$. Additionally let $f : Y \to \overline{Z}$ be a convex mapping. Then for $f \circ \alpha : X \to \overline{Z}$ the following chain-rule holds:*

$$\partial\left(f \circ \alpha\right)(x) = \bigcup_{\mathrm{R} \in \partial f(\alpha(x))} \mathbf{R} \circ \mathbf{A}\ \forall x \in X$$

Proof. See theorem 4.2 from [**21**]. □

Proposition 3.2.5. *Let now $f : X \to \overline{Y}$ be a convex mapping and $g : Y \to \overline{Z}$ be a convex monotonic increasing mapping with $g(\pm\infty) = \pm\infty$. If f is continuous for some $x_0 \in (\operatorname{dom} f)^{\alpha i}$ within the algebraic interior*[3] *and g is completely continuous*[4] *at $f(x_0)$ then the subdifferential for the concatenation $g \circ f$ is given by*

$$\partial\left(g \circ f\right)(x) = \bigcup_{\mathbf{A} \in \partial g(f(x))} \partial\left(\mathbf{A} \circ f\right)(x)\ \forall x \in (\operatorname{dom} f)^{\alpha i}$$

.

Proof. This is theorem 4.3 from [**21**]. □

[3] αi means the algebraic interior: Consider A to be a convex set. We say that $v \in A$ **lies in the algebraic interior** of A if for any straight line l passing through v and the point v lies in the interior of the intersection $A \cap l$. The set of all points that lie in the algebraic interior of A is called the algebraic interior $(A)^{\alpha i}$of A.

[4] A bounded linear mapping $f : X \to Y$ - where X, Y are Banach spaces - is called **completely continuous** if, for every weakly convergent sequence (x_n) from X, the sequence $(f(x_n))$ is norm-convergent in Y.

Unfortunately there is no way to use those propositions in our context, where acceptability mappings will not be necessarily completely continuous and in most cases not even linear at all. Theorem 5.2.5 - which gives a chain rule for conditional mappings - will not make any assumptions about complete continuity: We will derive a chain rule for nested acceptability mappings, based on monotonicity and convexity, which are necessary attributes of acceptability mappings.

3.3. Fenchel-Moreau Conjugates

A basic notion for duality theory is lower semicontinuity. Papageorgiou defines it as follows:

DEFINITION 3.3.1. (lower and upper semicontinuity) Let $f : X \to \overline{Y}$ be a mapping where $\operatorname{dom} f$ is closed in X. If f is finite at $x \in X$ then it is said to be **lower semicontinuous** (l.s.c.) at x_0 if and only if for every $y \in \operatorname{int} K_Y^+$ there is a neighborhood U of x_0 in X such that $f(z) + y - f(x_0) \in \operatorname{int} K_Y^+$ for every $z \in U$. If $f(x_0) = -\infty$ then f is l.s.c. at x_0.

A mapping f is said to be **upper semicontinuous** (u.s.c.) at x_0 if and only if $-f$ is l.s.c. at x_0.

Later on we will define upper semicontinuity (and lower semicontinuity) in a similar manner for the special case of $\boldsymbol{L}_p(\Omega, \mathcal{F}, \mu)$-spaces. It should be noted that if a mapping is both l.s.c. and u.s.c. at a point x_0 then it is continuous at this point. Moreover if a mapping is finite and continuous at some point then it is both l.s.c. and u.s.c..[5]

In partially ordered vector spaces convex conjugates - which Papageorgiou calls Fenchel transforms - are defined in the following way:

DEFINITION 3.3.2. (Fenchel Transform, convex conjugate) Let $f : X \to \overline{Y}$ be a mapping. Then the **Fenchel transform** is defined to be the mapping $f^* : \mathcal{L}(X, Y) \to \overline{Y}$ given by

$$f^*(A) = \sup \{A(x) - f(x) : x \in \operatorname{dom} f\}.$$

We will also call Fenchel transforms "convex conjugates" and later define and use "concave conjugates" in the context of $\boldsymbol{L}_p(\Omega, \mathcal{F}, \mu)$-spaces for the analysis of acceptability mappings, which are concave. Convex conjugates are suprema over the function $A(x) - f(x)$, which is

[5] see Theorem 5.3 and 5.4 from [**21**].

convex both in x and in A. Then the conjugate f^* is convex under very weak conditions[6], which are fulfilled for $\boldsymbol{L}_p(\Omega, \mathcal{F}, \mu)$-spaces.

It is also possible to iterate the process of conjugation and define the Fenchel transform of the Fenchel transform or biconjugate as follows:

Definition 3.3.3. (convex biconjugate) Let $f : X \to \overline{Y}$ be a mapping, f^* its Fenchel transform. Then the convex biconjugate is defined to be the mapping $f^* : \mathcal{L}(\mathcal{L}(X, Y), Y) \to \overline{Y}$

$$f^{**}(x) = \sup\{A(x) - f^*(A) : A \in \operatorname{dom} f^*\}.$$

Again this will be a convex function under very weak conditions.

Generally X is only a subspace of $\mathcal{L}(\mathcal{L}(X, Y), Y)$. So a function $\iota : X \to \mathcal{L}(\mathcal{L}(X, Y), Y)$ with $[\iota(x)](A) = A(x)$ is used to restrict f^{**} on $\iota(x)$ and denote the restricted version $f^{**}(\iota(x))$ again by $f^{**}(x)$.

For us, the main result about conjugates, which we will use in chapter 4 is the following:

Proposition 3.3.4. *In general, $f^{**}(x) \le f(x)$ holds for any x. Moreover if $\partial f(x) \neq \emptyset$, then the equation $f^{**}(x) = f(x)$ holds.*

Proof. See lemma 5.4 and proposition 5.8 in [**21**]. □

There remains the question of when an infimum is attained, being in fact a minimum. Luckily it is possible to give a sufficient condition using the terms and definitions of the current section:

Proposition 3.3.5. *If $f : K \to \bar{Y}$ is a convex, lower semicontinuous mapping where $K \subseteq X$ is a compact set and $\overline{\operatorname{Im} f}$ is order complete then f attains its infimum on K.*

Proof. This is theorem 5.2 from [**21**]. □

In our context this would mean that not only the $\boldsymbol{L}_p-$space into which the operator maps must be order complete, but also the closure of the image set of the mapping must be an order complete subset of the $\boldsymbol{L}_p-$space under consideration. Further, the effective domain of the mapping would have to be a compact set.

[6]see Proposition 5.7 in [**21**].

But compactness is hard to achieve in infinite dimensional spaces like $\boldsymbol{L}_p\left(\Omega, \mathcal{F}, \mu\right)$-spaces. For this reason we will base our considerations about when the infimum for the conjugate mappings is attained not on proposition 3.3.5 but on the close relationship between subgradients and convex conjugates given by proposition 3.3.4.

Part 2

Conditional Acceptability Mappings and Multi-Period Acceptability Functionals

CHAPTER 4

Conditional Acceptability Mappings

Acceptability functionals evaluate in terms of a real number, how favorable a random variable (or distribution) - representing financial return - is. Basically they are modeled as functions, mapping from a $\boldsymbol{L}_p(\Omega, \mathcal{F}, \mu)$ $-$space into $\mathbb{R}$. The valuation is done at the beginning of a single period, assuming that there is no nontrivial information available that could be used for calculating the acceptability. A natural generalization of such functionals are conditional risk and acceptability mappings. Also measuring the acceptability of a random variable connected to a single period, in principle such mappings assume that there is additional nontrivial information available at the beginning of the period under consideration.

Mathematically, information is expressed in terms of $\sigma-$algebras and so we can say that unconditional acceptability functionals measure acceptability relative to the trivial $\sigma-$Algebra $\{\emptyset, \Omega\}$, while conditional acceptability mappings calculate acceptability relative to some $\sigma-$Algebra $\mathcal{F}_1 \supset \{\emptyset, \Omega\}$. This means: conditional acceptability mappings are mappings from $\boldsymbol{L}_p(\Omega, \mathcal{F}, \mu)$ into $\boldsymbol{L}_{p'}(\Omega, \mathcal{F}_1, \mu)$, resulting not in a real number but in a random variable.

Conditional risk and acceptability mappings were defined and analyzed by Detlefsen [**9**], Ruszczyński [**32**] and Pflug and Römisch [**23**]. By their defining axioms such mappings have some reasonable properties closely related to the properties of acceptability functionals from definition 1.1.1.

Conditional mappings can be useful in different ways:

- If they have the property that more information results in a higher acceptability (information monotonicity) one can compare the acceptability of a random variable under different information sets using the natural partial order ($\geq$ $a.s.$)

from section 2.2[1]. As stated before we will use the notation $\geq$ without explicitly stating that it is meant in the almost sure sense.

- They can be used to track how the (conditional) acceptability of a future random variable changes over time, as additional information becomes available. Such sequences of acceptability mappings are called dynamic acceptability functionals ([**23**]).
- In the context of random processes they can be used to define multi-period acceptability functionals, measuring the desirability of a sequence of random variables relative to some filtration. We will study some possibilities in chapter 5.

The above authors analyzed conditional mappings in very different ways: While Detlefsen and Scandolo use the essential infimum for defining conjugate functionals[2]- which is feasible only for spaces $\boldsymbol{L}_\infty(\Omega, \mathcal{F}, \mu)$, Ruszczyński and Shapiro base their analysis on pointwise arguments, neglecting the question of measurability of infima in $\boldsymbol{L}_p-$spaces. They define the (convex) conjugate - dependent on individual $\omega \in \Omega$ - as $\sup\{\mathbb{E}(\mathrm{YZ}) - \mathcal{A}(Y)\}(\omega)$. But the pointwise supremum of an uncountable set of measurable random variables easily could be not measurable oneself.

Generalizing the definition of conditional expectation, Pflug and Römisch use a trick from probability theory and get a representation theorem for conditional acceptability mappings[3]: For any u.s.c. conditional acceptability mapping $\mathcal{A}(Y|\mathcal{F}_1)$ and for every $Y \in \boldsymbol{L}_p(\Omega, \mathcal{F}, \mu)$ and $B \in \mathcal{F}_1$ the conjugate of the restricted expectation $\mathbb{E}(\mathcal{A}(Y|\mathcal{F}_1) \cdot \mathbf{1}_B)$ can be written as

$$\mathbb{E}(\mathcal{A}(Y|\mathcal{F}_1) \cdot \mathbf{1}_B) = \inf_{Z \in \boldsymbol{L}_q(\Omega, \mathcal{F}, \mu)} \{\mathbb{E}(Y \cdot Z) - A_B(Z) : Z \geq 0,\, \mathbb{E}(Z|\mathcal{F}_1) = \mathbf{1}_B\},$$

where $A_B(Z) = \inf\limits_{Y \in \boldsymbol{L}_p(\Omega, \mathcal{F}, \mu)} \{\mathbb{E}(\mathrm{YZ}) - \mathbb{E}(\mathcal{A}(Y|\mathcal{F}_1) \cdot \mathbf{1}_B)\}$.

This meets the concerns for measurability and allows to use the full duality theory for real valued functionals and in this way Pflug and Römisch are able to derive some crucial propositions. However such a representation is an implicit one, which sometimes makes the reasoning more complicated.

[1]Such a direct comparison is different from the usual value of information (see e.g. Pflug and Römisch [**23**] p 177) which takes into account the possible actions of optimizing agents with different decision spaces, relative to their information sets.
[2]This approach was also recently used in [**13**].
[3]See theorem 2.5.1 in [**23**].

In the following, another notion of infimum will be used. Based on the results derived in section 2.2 - especially the order completeness of $\boldsymbol{L}_p-$spaces - we will use the infimum with respect to the almost sure order[4] to define conjugate mappings in a very natural way. It is crucial to remember that this definition is based on equivalence classes of random variables and not on individual random variables.

Moreover, it is also possible to define sub- and supergradients based on the almost sure order. These definitions will be used to derive propositions about acceptability functionals, their conjugates and supergradients, generalizing the usual results for real valued functionals.

In this context conditional acceptability mappings map into a partially ordered vector space. That means that the usual arguments for dealing with subgradients and conjugates - based on functional separability - break down. It will not be possible to rely on separating hyperplane theorems any more. Furthermore, nonsmooth convex analysis was already generalized in the past to work on partially ordered vector spaces. We will base our arguments on the work of N. Papageorgiou **[21]**, which was summarized in chapter 3.

4.1. Definition and Basic Properties

We saw that acceptability functionals $\mathcal{A}(Y)$ are mappings between a function space of random variables and the real line. If an acceptability functional is version independent, we can also understand it as a mapping $\mathcal{A}\{F_Y\}$ between the distribution functions of the random variables under consideration and the real line. But then it also makes sense to apply the acceptability functionals on some conditional distributions, which can be written as $\mathcal{A}\left\{F_{Y|X}\right\}$.

More generally, conditional acceptability mappings are a special class of operators, mapping from one function space $\boldsymbol{L}_p(\Omega, \mathcal{F}, \mathbf{P})$ into another space $\boldsymbol{L}_{p'}(\Omega, \mathcal{F}_1, \mathbf{P})$ with $\mathcal{F}_1 \subseteq \mathcal{F}$ and $1 \leq p' \leq p$. We can think of the $\sigma-$Algebra $\mathcal{F}$ as the information available at the end of a time period under consideration and of $\mathcal{F}_1$ as the information available at the beginning.

Again we will adjoin a greatest element $+\infty$ and a smallest element $-\infty$ to each vector space that contains the image of a mapping. The greatest element can be interpreted as a random variable that takes the value $+\infty$ with probability one. The extended space

[4] see definition 2.2.16

$\boldsymbol{L}_p(\Omega, \mathcal{F}_1, \mathbf{P}) \cup \{+\infty, -\infty\}$ will be denoted as $\overline{\boldsymbol{L}}_p(\Omega, \mathcal{F}_1, \mathbf{P})$ in the following. The domain space is modeled by a space of ordinary random variables with values in $\mathbb{R}$.

In this context we should also modify slightly the definition of bounded sets.

DEFINITION 4.1.1. (boundedness) A set S in $\overline{\boldsymbol{L}}_p(\Omega, \mathcal{F}, \mathbf{P})$ is called **bounded below** if there exists an element $X \in \overline{\boldsymbol{L}}_p(\Omega, \mathcal{F}, \mathbf{P})$ with $X \leq Y$ for all $Y \in \overline{\boldsymbol{L}}_p(\Omega, \mathcal{F}, \mathbf{P})$ and $X > -\infty$ *a.s.*.

A set S in $\overline{\boldsymbol{L}}_p(\Omega, \mathcal{F}, \mathbf{P})$ is called **bounded above** if there exists an element $X \in \overline{\boldsymbol{L}}_p(\Omega, \mathcal{F}, \mathbf{P})$ with $X \geq Y$ for all $Y \in \overline{\boldsymbol{L}}_p(\Omega, \mathcal{F}, \mathbf{P})$ and $X < +\infty$ *a.s.*.

A set S in $\overline{\boldsymbol{L}}_p(\Omega, \mathcal{F}, \mathbf{P})$ is called **bounded** if it is bounded both below and above.

As usual in the case of functionals we will also concentrate on proper mappings in the following:

DEFINITION 4.1.2. (properness and effective domain) The effective domain of a concave mapping $\mathcal{A}(\cdot|\mathcal{F}_1) : \boldsymbol{L}_p(\Omega, \mathcal{F}, \mathbf{P}) \rightarrow \overline{\boldsymbol{L}}_{p'}(\Omega, \mathcal{F}_1, \mathbf{P})$ is given by the set dom $f = \{X \in \boldsymbol{L}_p(\Omega, \mathcal{F}, \mathbf{P}) : \mathcal{A}(X|\mathcal{F}_1) > -\infty \; a.s.\}$

A concave mapping $\mathcal{A}(\cdot|\mathcal{F}_1) : \boldsymbol{L}_p(\Omega, \mathcal{F}, \mathbf{P}) \rightarrow \overline{\boldsymbol{L}}_{p'}(\Omega, \mathcal{F}_1, \mathbf{P})$ is called **proper** if the effective domain is nonempty and the mapping does not take the value $+\infty$.

REMARK. In section 3 we defined convex mappings and the appropriate notions of properness and effective domain. The reason for the differences in sign is that generally we want to minimize convex functions while we want to maximize concave functions. Because the domain of a conditional mapping is understood as its effective domain it is possible to restrict it by some conditions, including some implicit restrictions (e.g. $\mathcal{A}(X|\mathcal{F}_1) = -\infty$ if $\mathbb{E}(X|\mathcal{F}_1) \leq a$). Proposition 4.2.3 will remain applicable in this case.

While it is not feasible to define conditional acceptability mappings in a similar way as the conditional expectation, following ([**23, 33**]) it is possible to define their critical features in an axiomatic way. The properties of acceptability functionals have their counterpart in the following definition for conditional acceptability mappings.

DEFINITION 4.1.3. (conditional acceptability mapping)

A proper mapping $\mathcal{A}(\cdot|\mathcal{F}_1) : \boldsymbol{L}_p(\Omega, \mathcal{F}, \mathbf{P}) \rightarrow \overline{\boldsymbol{L}}_{p'}(\Omega, \mathcal{F}_1, \mathbf{P})$ is called *conditional acceptability mapping with observable information* $\mathcal{F}_1$ if the following conditions are satisfied for all $X, Y \in \boldsymbol{L}_p(\Omega, \mathcal{F}, \mathbf{P})$, $Y_1 \in \boldsymbol{L}_{p'}(\Omega, \mathcal{F}_1, \mathbf{P})$:

(CA1) ***Predictable Translation Equivariance***. $\mathcal{A}(Y + Y_1|\mathcal{F}_1) = \mathcal{A}(Y|\mathcal{F}_1) + Y_1$ holds a.s. for every $Y \in \boldsymbol{L}_p(\Omega, \mathcal{F}, \mathbf{P})$ and $Y_1 \in \boldsymbol{L}_{p'}(\Omega, \mathcal{F}_1, \mathbf{P})$.

(CA2) ***Concavity***. $\mathcal{A}(\lambda \cdot X + (1-\lambda) \cdot Y|\mathcal{F}_1) \geq \lambda \cdot \mathcal{A}(X|\mathcal{F}_1) + (1-\lambda) \cdot \mathcal{A}(Y|\mathcal{F}_1)$ holds a.s. for $\lambda \in [0,1]$.

(CA3) ***Monotonicity***. $X \leq Y\ a.s. \Rightarrow \mathcal{A}(X|\mathcal{F}_1) \leq \mathcal{A}(Y|\mathcal{F}_1)\ a.s.$

NOTATION 4.1.4. Sometimes we will write $\mathcal{A}_{\mathcal{F}_1}(\cdot)$ or even $\mathcal{A}_1(\cdot)$ for $\mathcal{A}(\cdot|\mathcal{F}_1)$.

Mappings from $\boldsymbol{L}_p(\Omega, \mathcal{F}, \mathbf{P}) \to \overline{\boldsymbol{L}}_{p'}(\Omega, \mathcal{F}_1, \mathbf{P})$ without the properties of conditional acceptability mappings we will call *conditional mappings*. If they are at least monotonic we will call them acceptability type conditional mappings. A conditional mapping $\mathcal{A}$ is called positive homogeneous, if for every $\lambda > 0$ the condition $\mathcal{A}_{\mathcal{F}_1}(\lambda \cdot X) = \lambda \cdot \mathcal{A}_{\mathcal{F}_1}(X)$ holds.

Although conditional acceptability functionals are defined in an »axiomatic« way completely independent from conditional expectation, it is clear that conditional expectation is a prototype for definition 4.1.3: The conditional expectation is a conditional acceptability mapping fulfilling (CA1) - (CA3). Additionally it is positive homogeneous and upper semicontinuous.

REMARK. We have defined acceptability mappings in an abstract, axiomatic way and this will be very useful in the following for analyzing general properties of conditional acceptability mappings and some multi-period functionals based on conditional mappings. However in practical application we would use a special case of conditional mappings: If random variables are represented by their distribution function, we can apply version independent single period acceptability functionals on conditional distribution functions to account for conditional information.

As stated above, information in terms of $\sigma-$Algebras is a key point for conditional acceptability mappings and will be of interest also for the multi-period acceptability functionals defined below. In definition 4.1.3 the mapping is defined with respect to a fixed $\sigma-$Algebra. We might be interested in "generic" acceptability mappings, where the same "calculation principle" is applied to map into similar random spaces with different $\sigma-$Algebras. As a result a second form of monotonicity is relevant: Information monotonicity guarantees that a finer $\sigma-$Algebra, representing better information, will result in a higher acceptability.

DEFINITION 4.1.5. (Information Monotonicity) Let $C = (\mathcal{A}(\cdot|\mathcal{F}_i))_{i \in I}$ be a collection of acceptability mappings from a space $\boldsymbol{L}_p(\Omega, \mathcal{F}, \mathbf{P})$ into different spaces $\boldsymbol{L}_{p'}(\Omega, \mathcal{F}_i, \mathbf{P})$ with

$\mathcal{F}_i \subseteq \mathcal{F}$ and $p' \leq p$. We will call the collection **information monotonic** if for any $\mathcal{A}(\cdot|\mathcal{F}_i), \mathcal{A}(\cdot|\mathcal{F}_j) \in C$ the implication

$$\mathcal{F}_i \subseteq \mathcal{F}_j \cup \mathcal{N} \Rightarrow \mathcal{A}(\cdot|\mathcal{F}_i) \leq \mathcal{A}(\cdot|\mathcal{F}_j),$$

where $\mathcal{N}$ denotes the set of $\mathbf{P}-$null sets in $\mathcal{F}$, holds for any $X \in \boldsymbol{L}_p(\Omega, \mathcal{F}, \mathbf{P})$.

We will not distinguish between an actual mapping and a collection of mappings into different spaces in the following. That means that we always assume that for a meaningful acceptability mapping there is a definition or calculation rule that allows it to be calculated for different image spaces.

For information monotonic mapping we can evaluate the effect of additional information just by using the relevant norm:

Definition 4.1.6. (increase in acceptability) Let $C = (\mathcal{A}(\cdot|\mathcal{F}_i))_{i \in I}$ an information monotonic collection of acceptability mappings into the spaces $\boldsymbol{L}_{p'}(\Omega, \mathcal{F}_i, \mathbf{P})$, which are subspaces of $\boldsymbol{L}_p(\Omega, \mathcal{F}, \mathbf{P})$. Then for given data X the **increase in acceptability** between $\mathcal{F}_i$ and $\mathcal{F}_j$ with $\mathcal{F}_i \supseteq \mathcal{F}_j$ is given by the distance

$$\left\|\mathcal{A}(X|\mathcal{F}_i) - \mathcal{A}(X|\mathcal{F}_j)\right\|_{p'}.$$

4.2. Superdifferentials and Concave Conjugates of Conditional Acceptability Mappings

In the following let $\mathcal{A}(\cdot|\mathcal{F}_1) : \boldsymbol{L}_p(\Omega, \mathcal{F}, \mathbf{P}) \to \overline{\boldsymbol{L}}_{p'}(\Omega, \mathcal{F}_1, \mathbf{P})$ with $\mathcal{F}_1 \subseteq \mathcal{F}$ be a conditional acceptability mapping. Because such functionals are concave in the almost sure sense it is possible to define supergradients and concave conjugates in the usual way, but taking care of the fact that $\mathcal{A}(\cdot|\mathcal{F}_1)$ is a random variable and not a real number. This feasible if we use the almost sure order and the results from section 2.2 and chapter 3.

From definitions 3.2.1 and 3.3.2 we see that continuous linear mappings $\boldsymbol{L}_p(\Omega, \mathcal{F}, \mathbf{P}) \to \boldsymbol{L}_{p'}(\Omega, \mathcal{F}_1, \mathbf{P})$ are needed to define supergradients and conjugates for mappings between $\boldsymbol{L}_p(\Omega, \mathcal{F}, \mathbf{P}) \to \overline{\boldsymbol{L}}_{p'}(\Omega, \mathcal{F}_1, \mathbf{P})$. While continuity and linearity is guaranteed by corollary 2.1.2 if conditional expectations $\mathbb{E}(\bullet \cdot Z|\mathcal{F}_1)$ with $Z \in \boldsymbol{L}_q(\Omega, \mathcal{F}, \mathbf{P})$ are used, it is not clear generally, that such mappings also map into the correct spaces $\boldsymbol{L}_{p'}(\Omega, \mathcal{F}_1, \mathbf{P})$, which is required if we want to define supergradients and conjugates mappings for acceptability mappings into $\boldsymbol{L}_{p'}(\Omega, \mathcal{F}, \mathbf{P})$.

For this reason we have to restrict the set of possible Z by the condition $Z \in \boldsymbol{L}_s(\Omega, \mathcal{F}, \mathbf{P})$ with $s = \frac{p \cdot p'}{p - p'}$.

Theorem 4.2.1. *For $p < \infty$ the conditional expectations $X \mapsto \mathbb{E}(X \cdot Z|\mathcal{F}_1)$ with arguments $X \in \boldsymbol{L}_p(\Omega, \mathcal{F}, \mathbf{P})$ and dual variable $Z \in \boldsymbol{L}_s(\Omega, \mathcal{F}, \mathbf{P}) \subseteq \boldsymbol{L}_q(\Omega, \mathcal{F}, \mathbf{P})$ are linear, continuous and map into $\boldsymbol{L}_{p'}(\Omega, \mathcal{F}, \mathbf{P})$, if $s \geq \frac{p \cdot p'}{p - p'}$. For mappings $\boldsymbol{L}_p(\Omega, \mathcal{F}, \mathbf{P}) \to \boldsymbol{L}_{p'}(\Omega, \mathcal{F}, \mathbf{P})$ with $p = \infty$ the conditional expectations above have these properties, if $Z \in \boldsymbol{L}_{p'}(\Omega, \mathcal{F}, \mathbf{P})$.*

Proof. Linearity of a conditional expectation $\mathbb{E}(\bullet \cdot Z|\mathcal{F}_1)$ is clear.

Moreover Z are dual variables such that $\mathbb{E}(\bullet \cdot Z|\mathcal{F}_1)$ maps into $\boldsymbol{L}_{p'}(\Omega, \mathcal{F}, \mathbf{P})$, if $Z \in \boldsymbol{L}_q(\Omega, \mathcal{F}, \mathbf{P}) \cap \{Y : X \cdot Y \in \boldsymbol{L}_{p'}(\Omega, \mathcal{F}_1, \mathbf{P}), \forall X \in \boldsymbol{L}_p(\Omega, \mathcal{F}_1, \mathbf{P})\}$. This means that $X \cdot Z$ must be p'-integrable.

For $s = \frac{p \cdot p'}{p - p'}$, which is equivalent to $\frac{1}{p} + \frac{1}{s} = \frac{1}{p'}$ a generalized Hölder inequality can be used[5]:

$$\|X \cdot Z\|_{p'} \leq \|X\|_p \cdot \|Z\|_s .$$

As X is p-integrable by assumption this means that $X \cdot Z$ will be p' integrable if Z is s-integrable.

For $p = \infty$ one could also remember that the product $X \cdot Z$ is p' integrable if X is $\infty-$integrable and X is p'-integrable. This also holds for $p' = \infty$.

To ensure continuity as a mapping into $\boldsymbol{L}_{p'}(\Omega, \mathcal{F}, \mathbf{P})$, assume now that X_n converges to X in the p-norm, which means that $\|X_n - X\|_p \to 0$, as $n \to \infty$. Because $|\bullet|$ is convex we

[5] To see this for $p < \infty$ define $r_1 = \frac{p}{p - p'}$ and $r_2 = \frac{p}{p'}$. It easily can be seen that r_1 and r_2 are Hölder conjugates. Therefore we can apply Hölder's inequality:

$$\left\| |X|^{p'} \cdot |Z|^{p'} \right\|_1 \leq \left\| |X|^{p'} \right\|_{r_2} \cdot \left\| |Z|^{p'} \right\|_{r_1} ,$$

which is equivalent to

$$\|X \cdot Z\|_{p'} \leq \|X\|_{p' \cdot r_2} \cdot \|Z\|_{p' \cdot r_1} .$$

But $p' \cdot r_2 = p$ and $r_1 \cdot p' = s$, because of $s = \frac{p \cdot p'}{p - p'}$ and the definitions of r_1 and r_2. Together this gives

$$\|X \cdot Z\|_{p'} \leq \|X\|_p \cdot \|Z\|_s .$$

For $p = \infty$we have $\frac{1}{s} = \frac{1}{p'}$. Using this fact and separating the essential supremum of X, we get

$$\|X \cdot Z\|_{p'} \leq \|X\|_\infty \cdot \|Z\|_s .$$

can apply Jensen's inequality and get
(4.2.1)

$$\mathbb{E}\left(\left|\mathbb{E}\left(\left(X_n - X\right)\cdot Z|\mathcal{F}\right)\right|^{p'}\right) \leq \mathbb{E}\left(\mathbb{E}\left(\left|\left(X_n - X\right)\cdot Z\right|^{p'}|\mathcal{F}\right)\right) = \mathbb{E}\left(\left|\left(X_n - X\right)\cdot Z\right|^{p'}\right).$$

From this - using the generalized Hölder inequality again - we have

$$\left\|\mathbb{E}\left(\left(X_n - X\right)\cdot Z|\mathcal{F}\right)\right\|_{p'} \leq \left\|\left(X_n - X\right)\cdot Z\right\|_{p'} \leq \left\|\left(X_n - X\right)\right\|_{p}\cdot\left\|Z\right\|_{s}.$$

If Z is s-integrable and X_n converges to X in the p-norm, it follows that $\mathbb{E}\left(X_n \cdot Z|\mathcal{F}\right)$ converges to $\mathbb{E}\left(X \cdot Z|\mathcal{F}\right)$.

□

Example. For mappings $\boldsymbol{L}_p\left(\Omega,\mathcal{F},\mathbf{P}\right) \to \overline{\boldsymbol{L}}_1\left(\Omega,\mathcal{F}_1,\mathbf{P}\right)$ the Z must be from the natural dual space $\boldsymbol{L}_q\left(\Omega,\mathcal{F},\mathbf{P}\right)$ because $q = s$ in this cases. Until now the main focus in literature was on mappings $\boldsymbol{L}_p\left(\Omega,\mathcal{F},\mathbf{P}\right) \to \overline{\boldsymbol{L}}_p\left(\Omega,\mathcal{F}_1,\mathbf{P}\right)$. From theorem 4.2.1 we see that the dual variables must be from $\boldsymbol{L}_\infty\left(\Omega,\mathcal{F},\mathbf{P}\right)$ in this case. This is a strong restriction and coincides with the natural dual only for the case $p = 1$.

With this preparations it is possible to define superdifferentials and conjugates as follows:

Definition 4.2.2. (Superdifferential, Subdifferential) The **superdifferential** of a proper concave conditional mapping $\mathcal{A}\left(\cdot|\mathcal{F}_1\right) : \boldsymbol{L}_p\left(\Omega,\mathcal{F},\mathbf{P}\right) \to \overline{\boldsymbol{L}}_{p'}\left(\Omega,\mathcal{F}_1,\mathbf{P}\right)$ at $X_0 \in \boldsymbol{L}_p\left(\Omega,\mathcal{F},\mathbf{P}\right)$ is given by the set

$$\partial\mathcal{A}\left(X_0|\mathcal{F}_1\right) =$$
$$= \left\{Z \in \boldsymbol{L}_s\left(\Omega,\mathcal{F},\mathbf{P}\right) : \mathcal{A}\left(X|\mathcal{F}_1\right) \leq \mathcal{A}\left(X_0|\mathcal{F}_1\right) + \mathbb{E}\left(\left(X - X_0\right)\cdot Z|\mathcal{F}_1\right), \forall X \in \mathrm{dom}\mathcal{A}\right\},$$

where $\frac{1}{p} + \frac{1}{s} = \frac{1}{p'}$.

The **subdifferential** of a proper convex conditional mapping $\mathcal{D}\left(\cdot|\mathcal{F}_1\right)$ at $X_0 \in \boldsymbol{L}_p\left(\Omega,\mathcal{F},\mathbb{P}\right)$ is given by the set

$$\partial\mathcal{D}\left(X_0|\mathcal{F}_1\right) =$$
$$= \left\{Z \in \boldsymbol{L}_s\left(\Omega,\mathcal{F},\mathbf{P}\right) : \mathcal{A}\left(X|\mathcal{F}_1\right) \geq \mathcal{A}\left(X_0|\mathcal{F}_1\right) + \mathbb{E}\left(\left(X - X_0\right)\cdot Z|\mathcal{F}_1\right), \forall X \in \mathrm{dom}\mathcal{A}\right\},$$

where$\frac{1}{p} + \frac{1}{s} = \frac{1}{p'}$.

The elements of super- and subdifferentials are called *super*- and s*ubgradients*.

Remark. We use the same symbol for super- and subdifferential, assuming that it will be clear from the context whether the concave or the convex version is relevant. In particular we will concentrate on concave mappings and their supergradients exclusively for the rest of the text.

In terms of chapter 3 our definition refers to topological superdifferentials for the special case of $\boldsymbol{L}_p(\Omega, \mathcal{F}, \mathbf{P})$- spaces. We neglect algebraic differentials here, which would be based on the expectation using any random variable - even outside the natural dual space.

Using definition 4.2.2 and based on the order properties of $\boldsymbol{L}_p(\Omega, \mathcal{F}, \mathbf{P})$ spaces from section 2.2, there are a lot of similarities between supergradients of functionals and supergradients of conditional mappings. In particular, as in the case of real functionals the superdifferential will exist for a broad range of concave mappings at least at the interior of their domains. Of course, at the boundary it is easily possible that the superdifferential is empty.

Proposition 4.2.3. *If a concave mapping $\mathcal{A}(\cdot|\mathcal{F}_1) : \boldsymbol{L}_p(\Omega, \mathcal{F}, \mathbf{P}) \rightarrow \overline{\boldsymbol{L}}_{p'}(\Omega, \mathcal{F}_1, \mathbf{P})$ is continuous at some point $\hat{X} \in \mathrm{dom}\mathcal{A}$ then $\partial\mathcal{A}(X|\mathcal{F}_1) \neq \emptyset$ for all $X \in \mathrm{int}\,\mathrm{dom}\mathcal{A}$.*

Proof. $\boldsymbol{L}_p$-spaces are metric spaces hence they are normal spaces, which basically means that disjoint closed sets can be separated by neighborhoods. Using this fact as a premise we can apply proposition 3.2.2, restated for supergradients[6]. □

Proposition 4.2.4. *A conditional acceptability mapping $\mathcal{A} : \boldsymbol{L}_p(\Omega, \mathcal{F}, \mathbf{P}) \rightarrow \overline{\boldsymbol{L}}_{p'}(\Omega, \mathcal{F}_1, \mathbf{P})$ is continuous if it is locally bounded at some element $X_0 \in \boldsymbol{L}_p(\Omega, \mathcal{F}, \mathbf{P})$.*

Proof. See [**20**], theorem 4. □

From chapter 3 we know that there are a lot of rules for generalized subdifferential calculus very similar to ordinary subdifferential calculus, including a rule for sums of mappings[7] and a chain rule[8].

As a next step we will develop a theory of duality for conditional acceptability mappings which is based on concave Fenchel conjugates. First we have to define semicontinuity for conditional mappings.

[6]Another way would be to base the argument on theorem 4 in [**19**].
[7]Theorem 4.1 from [**21**]
[8]Theorem 4.3 from [**21**]

Definition 4.2.5. (upper semicontinuity) A conditional mapping $\mathcal{A}(\cdot|\mathcal{F}_1)$ with closed domain is called **continuous from above** or **upper semicontinuous** *(u.s.c.)* **at** X_0, if and only if for every $\varepsilon > 0$ $a.s.$ there exists a neighborhood U of X_0 such that $\mathcal{A}(X|\mathcal{F}_1) \leq \mathcal{A}(X_0|\mathcal{F}_1) + \varepsilon$ $a.s.$ for all $X \in U$ or if $\mathcal{A}(X_0|\mathcal{F}_1) = +\infty$.

A mapping $\mathcal{A}(\cdot|\mathcal{F}_1)$ is called lower semicontinuous (l.s.c.) at a point X_0 if $-\mathcal{A}(\cdot|\mathcal{F}_1)$ is u.s.c. at the point X_0.

If the mapping is l.s.c. (u.s.c.) at each point in its domain it is called l.s.c. (u.s.c.).

It should be noted again that a mapping is continuous if it is both lower and upper semicontinuous[9].

Now it is possible to define concave conjugates for conditional mappings in a straightforward way using the generalized infimum and supremum from section 5.22.2.

It should be remembered that such a definition is meaningful only because $\boldsymbol{L}_p(\Omega,\mathcal{F},\mathbf{P})-$spaces are order complete.

Definition 4.2.6. (concave conjugate, biconjugate) The **concave conjugate** of a mapping $\mathcal{A}(\cdot|\mathcal{F}_1):\ \boldsymbol{L}_p(\Omega,\mathcal{F},\mathbf{P}) \to \overline{\boldsymbol{L}}_{p'}(\Omega,\mathcal{F}_1,\mathbf{P})$ is given by a mapping $\mathcal{A}^*:\boldsymbol{L}_s(\Omega,\mathcal{F},\mathbf{P}) \to \overline{\boldsymbol{L}}_{p'}(\Omega,\mathcal{F}_1,\mathbf{P})$ with $\frac{1}{p}+\frac{1}{s}=\frac{1}{p'}$:

$$\mathcal{A}^*(Z|\mathcal{F}_1) = \inf_{X\in\boldsymbol{L}_p(\Omega,\mathcal{F},\mathbf{P})} \left\{\mathbb{E}(X\cdot Z|\mathcal{F}_1) - \mathcal{A}(X|\mathcal{F}_1)\right\}.$$

The **concave biconjugate** is a mapping $\mathcal{A}^{**}:\boldsymbol{L}_p(\Omega,\mathcal{F},\mathbf{P}) \to \overline{\boldsymbol{L}}_{p'}(\Omega,\mathcal{F}_1,\mathbf{P})$ defined by

$$\mathcal{A}^{**}(X|\mathcal{F}_1) = \inf_{Z\in\boldsymbol{L}_s(\Omega,\mathcal{F},\mathbf{P})} \left\{\mathbb{E}(X\cdot Z|\mathcal{F}_1) - \mathcal{A}^*(Z|\mathcal{F}_1)\right\}.$$

Remark. The domain of the biconjugate mapping should be random variables from $\boldsymbol{L}_p(\Omega,\mathcal{F},\mathbf{P})$, such that $\mathbb{E}(X\cdot Z|\mathcal{F}_1) \in \boldsymbol{L}_{p'}(\Omega,\mathcal{F}_1,\mathbf{P})$ holds for any $Z \in \boldsymbol{L}_s(\Omega,\mathcal{F},\mathbf{P})$. But $\boldsymbol{L}_s(\Omega,\mathcal{F},\mathbf{P})$ was constructed in a way such that $\mathbb{E}(X\cdot Z|\mathcal{F}_1) \in \boldsymbol{L}_{p'}(\Omega,\mathcal{F}_1,\mathbf{P})$ for any $X \in \boldsymbol{L}_p(\Omega,\mathcal{F},\mathbf{P})$ and so we can consider the whole space $\boldsymbol{L}_p(\Omega,\mathcal{F},\mathbf{P})$ as the "dual" of the restricted dual $\boldsymbol{L}_s(\Omega,\mathcal{F},\mathbf{P})$.

Conjugates and biconjugates for conditional (deviation) risk[10] functionals can be defined as convex conjugates in similar manner using the almost sure version of the supremum.

[9] Theorem 5.3 in [**21**].

[10] See [**23**] for a concise treatment of deviation risk functionals and the generalization to conditional risk mappings.

Because we concentrate exclusively on concave mappings, we will often just use "conjugate" instead of "concave conjugate".

In section 2.2 we agreed the convention to set the infimum equal to $-\infty$ if a set is not bounded below. If there are arguments Z that lead to an unbounded set in the calculation of the conjugate, this dual variable will not be selected when the bidual is calculated. Alternatively we can use the effective domain of the conjugate to restrict the feasible set for the calculation of the biconjugate:

$$\mathcal{A}^{**}(X|\mathcal{F}_1) = \inf_Z \{\mathbb{E}(X \cdot Z|\mathcal{F}_1) - \mathcal{A}^*(Z|\mathcal{F}_1) : \; Z \in \mathrm{dom}\mathcal{A}^*\}.$$

Based on definition 4.2.6 we can generalize the main results for (unconditional) acceptability functionals and get similar statements for conditional acceptability mappings. The main difference is again that the relations are valid in the almost sure sense.

The equality of mapping and bidual mapping (Fenchel-Moreau-Rockafellar theorem) can not easily be proved for concave/convex mappings in a way similar to the case of concave/convex functionals. This would require some kind of separation theorems based on almost sure separating hyperplanes.

Anyway, we will be able to identify a proper, concave mapping with its bidual at points where the subdifferential is nonempty. So again the term »supergradient-representation« is justifiable for the biconjugate.

THEOREM 4.2.7. *Let $\mathcal{A}(\cdot|\mathcal{F}_1)$ be a proper, concave mapping. Then for all $X \in \boldsymbol{L}_p(\Omega, \mathcal{F}, \mathbb{P})$, $Z \in \boldsymbol{L}_s(\Omega, \mathcal{F}, \mathbf{P})$,*

$$\mathcal{A}(X|\mathcal{F}_1) \leq \mathbb{E}(X \cdot Z|\mathcal{F}_1) - \mathcal{A}^*(Z|\mathcal{F}_1), \tag{4.2.2}$$

holds.

Moreover if $\partial\mathcal{A}(X_0|\mathcal{F}_1) \neq \emptyset$ then

$$\mathcal{A}(X_0|\mathcal{F}_1) = \mathcal{A}^{**}(X_0|\mathcal{F}_1) \tag{4.2.3}$$

In this case the infimum is attained and

$$\mathcal{A}(X|\mathcal{F}_1) = \mathbb{E}(X \cdot Z|\mathcal{F}_1) - \mathcal{A}^*(Z|\mathcal{F}_1) \; a.s. \Leftrightarrow Z \in \partial\mathcal{A}(X|\mathcal{F}_1). \tag{4.2.4}$$

If in addition $\mathcal{A}$ is continuous at some point inside $\text{dom}\mathcal{A}$ *equality 4.2.3 holds for each point in* $\text{int}\,\text{dom}\mathcal{A}^*$.

PROOF. The inequality follows directly from the definition of conjugates for conditional mappings: As $\mathcal{A}^*(Z|\mathcal{F}_1) := \inf_X \{\mathbb{E}(X \cdot Z|\mathcal{F}_1) - \mathcal{A}(X|\mathcal{F}_1)\}$ it follows that $\mathcal{A}^*(Z|\mathcal{F}_1) \leq \mathbb{E}(X \cdot Z|\mathcal{F}_1) - \mathcal{A}(X|\mathcal{F}_1)$ *a.s.* or $\mathcal{A}(X|\mathcal{F}_1) \leq \mathbb{E}(X \cdot Z|\mathcal{F}_1) - \mathcal{A}^*(Z|\mathcal{F}_1)$.

The second assertion is theorem 5.8 from Papageorgiou [**21**], cited in proposition 3.3.4 above.

The equation $\mathcal{A}(X|\mathcal{F}_1) = \mathbb{E}(X \cdot Z|\mathcal{F}_1) - \mathcal{A}^*(Z|\mathcal{F}_1)$ holds almost sure if and only if $\mathcal{A}(X|\mathcal{F}_1) + \mathbb{E}(Y \cdot Z|\mathcal{F}_1) - \mathcal{A}(Y|\mathcal{F}_1) \geq \mathbb{E}(X \cdot Z|\mathcal{F}_1)$ holds almost sure for any $Y \in \boldsymbol{L}_p(\Omega, \mathcal{F}, \mathbb{P})$.

a) Assume $\mathcal{A}(X|\mathcal{F}_1) = \mathbb{E}(X \cdot Z|\mathcal{F}_1) - \mathcal{A}^*(Z|\mathcal{F}_1)$ *a.s.*. Then we have

$$\begin{aligned}\mathcal{A}(X|\mathcal{F}_1) + \mathbb{E}(Y \cdot Z|\mathcal{F}_1) - \mathcal{A}(Y|\mathcal{F}_1) &\geq \mathcal{A}(X|\mathcal{F}_1) + \inf_Z \{\mathbb{E}(Y \cdot Z|\mathcal{F}_1) - \mathcal{A}(Y|\mathcal{F}_1)\} \\ &= \mathcal{A}(X|\mathcal{F}_1) + \mathcal{A}^*(Z|\mathcal{F}_1).\end{aligned}$$

Using the assumption we get for all Y

$$\mathcal{A}(X|\mathcal{F}_1) + \mathbb{E}(Y \cdot Z|\mathcal{F}_1) - \mathcal{A}(Y|\mathcal{F}_1) \geq \mathbb{E}(X \cdot Z|\mathcal{F}_1).$$

b) For the other direction assume $\mathcal{A}(X|\mathcal{F}_1) + \mathbb{E}(Y \cdot Z|\mathcal{F}_1) - \mathcal{A}(Y|\mathcal{F}_1) \geq \mathbb{E}(X \cdot Z|\mathcal{F}_1)$ *a.s.*. Remember that this inequality should hold for all Y - which means that $\mathbb{E}(Y \cdot Z|\mathcal{F}_1) - \mathcal{A}(Y|\mathcal{F}_1)$ has a lower bound. Then the infimum exists and it follows that

$$\mathcal{A}(X|\mathcal{F}_1) + \inf_Z \{\mathbb{E}(Y \cdot Z|\mathcal{F}_1) - \mathcal{A}(Y|\mathcal{F}_1)\} \geq \mathbb{E}(X \cdot Z|\mathcal{F}_1)$$

or

$$\mathcal{A}(X|\mathcal{F}_1) + \mathcal{A}^*(Z|\mathcal{F}_1) \geq \mathbb{E}(X \cdot Z|\mathcal{F}_1)$$

Together with inequality 4.2.2 the equation $\mathcal{A}(X|\mathcal{F}_1) + \mathcal{A}^*(Z|\mathcal{F}_1) = \mathbb{E}(X \cdot Z|\mathcal{F}_1)$ follows.

Moreover $\mathcal{A}(X|\mathcal{F}_1) + \mathbb{E}(Y \cdot Z|\mathcal{F}_1) - \mathcal{A}(Y|\mathcal{F}_1) \geq \mathbb{E}(X \cdot Z|\mathcal{F}_1)$a.s. holds if and only if $\mathcal{A}(Y|\mathcal{F}_1) \leq \mathcal{A}(X|\mathcal{F}_1) + E((Y - X) \cdot Z|\mathcal{F}_1)$ or $Z \in \partial\mathcal{A}(X|\mathcal{F}_1)$ holds.

From this equivalence together with inequality 4.2.2 equation 4.2.4 is proved for proper concave mappings.

The last statement follows from proposition 4.2.3. $\square$

Remark. It is also possible to give conditions under which $\mathcal{A}(X|\mathcal{F}_1) = \mathcal{A}^{**}(X|\mathcal{F}_1)$ on the whole space $\boldsymbol{L}_p(\Omega,\mathcal{F},\mathbb{P})$. This is stated by theorem 5.9 in Papageorgiou[**21**].

Because acceptability mappings are proper concave u.s.c., theorem 4.2.7 can be applied. Again - as for the unconditional case - we can characterize the dual representation of conditional acceptability mappings more precisely, identifying the set $A_{\mathcal{A}}$.

Theorem 4.2.8. *Assume that the conditions of proposition 4.2.3 or proposition 4.2.4 are fulfilled. Then a concave conditional mapping* $\mathcal{A}(\cdot|\mathcal{F}_1) : \boldsymbol{L}_p(\Omega,\mathcal{F},\mathbb{P}) \to \boldsymbol{L}_{p'}(\Omega,\mathcal{F}_1,\mathbb{P})$ *is an acceptability mapping if and only if the dual representation*

$$\mathcal{A}(X|\mathcal{F}_1) = \inf_{Z\in\boldsymbol{L}_s(\Omega,\mathcal{F},\mathbf{P})} \{\mathbb{E}(X\cdot Z|\mathcal{F}_1) - \mathcal{A}^*(Z|\mathcal{F}_1) :\ Z \geq 0;\ \mathbb{E}(Z|\mathcal{F}_1) = 1\, a.s., Z \in \mathcal{Z}\}$$

holds for each point in $\operatorname{int}\operatorname{dom}\mathcal{A}$. *The set* $\mathcal{Z}$ *represents additional constraints on* Z, *restricting e.g. the conjugate* $\mathcal{A}^*$ *to be finite.*

Proof. By theorem 4.2.7 , because $\mathcal{A}$ is concave and continuous at some point in $\operatorname{dom}\mathcal{A}$ we have for each point in $\operatorname{int}\operatorname{dom}\mathcal{A}$:

$$\mathcal{A}(X|\mathcal{F}_1) = \inf_Z \{\mathbb{E}(X\cdot Z|\mathcal{F}_1) - \mathcal{A}^*(Z|\mathcal{F}_1)\} = \mathbb{E}(X\cdot Z|\mathcal{F}_1) - \mathcal{A}^*(Z|\mathcal{F}_1)$$

where Z is a supergradient.

If $\mathbb{E}(Z|\mathcal{F}_1) = 1$ a.s. we have for a $\mathcal{F}_1$-measurable X_1

$$\begin{aligned}
\mathcal{A}(X+X_1|\mathcal{F}_1) &= \mathbb{E}((X+X_1)\cdot Z|\mathcal{F}_1) - \mathcal{A}^*(Z|\mathcal{F}_1)\\
&= \mathbb{E}(X\cdot Z|\mathcal{F}_1) + \mathbb{E}(X_1\cdot Z|\mathcal{F}_1) - \mathcal{A}^*(Z|\mathcal{F}_1)\\
&= \mathbb{E}(X\cdot Z|\mathcal{F}_1) + X_1\cdot\mathbb{E}(Z|\mathcal{F}_1) - \mathcal{A}^*(Z|\mathcal{F}_1)\\
&= \mathbb{E}(X\cdot Z|\mathcal{F}_1) + X_1 - \mathcal{A}^*(Z|\mathcal{F}_1)\\
&= \mathcal{A}(X|\mathcal{F}_1) + X_1
\end{aligned}$$

almost sure.

On the other hand let $\mathbb{E}(Z|\mathcal{F}_1) \neq 1$ on a set S with positive probability. Then we have

$$\begin{aligned}
\mathcal{A}(X|\mathcal{F}_1) + X_1 &= \mathbb{E}(X\cdot Z|\mathcal{F}_1) + X_1 - \mathcal{A}^*(Z|\mathcal{F}_1)\\
&\neq \mathbb{E}(X\cdot Z|\mathcal{F}_1) + X_1\cdot\mathbb{E}(Z|\mathcal{F}_1) - \mathcal{A}^*(Z|\mathcal{F}_1)\\
&= \mathcal{A}(X+X_1|\mathcal{F}_1)
\end{aligned}$$

on this set. This would contradict the assumption of predictable translation equivariance.

Assume now that $Z \geq 0$ holds a.s. and let Y be a random variable in the cone of almost sure nonnegative random variables: $Y \geq 0\, a.s.$. Then for any random variable X we have $X + Y \geq X$. Because Y and Z both are nonnegative it follows that $\mathbb{E}(Y \cdot Z|\mathcal{F}_1) \geq 0$. Using again theorem 4.2.7 we get

$$\begin{aligned} \mathcal{A}(X+Y|\mathcal{F}_1) &= \mathbb{E}((X+Y) \cdot Z|\mathcal{F}_1) - \mathcal{A}^*(Z|\mathcal{F}_1) \\ &= \mathbb{E}(X \cdot Z|\mathcal{F}_1) + \mathbb{E}(Y \cdot Z|\mathcal{F}_1) - \mathcal{A}^*(Z|\mathcal{F}_1) \\ &\geq \mathbb{E}(X \cdot Z|\mathcal{F}_1) - \mathcal{A}^*(Z|\mathcal{F}_1) \\ &\geq \inf_Z \{\mathbb{E}(X \cdot Z|\mathcal{F}_1) - \mathcal{A}^*(Z|\mathcal{F}_1)\} \\ &= \mathcal{A}(X|\mathcal{F}_1) \end{aligned}$$

For the other direction assume now that $Z < 0$ on a set S with positive probability, or $\mathbb{E}(Y \cdot Z|\mathcal{F}_1) < 0$ on this set. Then we have

$$\begin{aligned} \mathcal{A}(X|\mathcal{F}_1) &= \mathbb{E}(X \cdot Z|\mathcal{F}_1) - \mathcal{A}^*(Z|\mathcal{F}_1) \\ &> \mathbb{E}(X \cdot Z|\mathcal{F}_1) + \mathbb{E}(Y \cdot Z|\mathcal{F}_1) - \mathcal{A}^*(Z|\mathcal{F}_1) \\ &\geq \inf_Z \{\mathbb{E}(X \cdot Z|\mathcal{F}_1) + \mathbb{E}(Y \cdot Z|\mathcal{F}_1) - \mathcal{A}^*(Z|\mathcal{F}_1)\} \\ &= \mathcal{A}(X+Y|\mathcal{F}_1) \end{aligned}$$

on the set S.

But this would mean that the mapping can not be monotonic. □

Theorem 4.2.8 can also be used to define acceptability mappings and we want to close the chapter with a simple example for this, generalizing the $AV@R$ to the conditional $AV@R$. In doing so we use the dual results from theorem 1.2.1.

Definition 4.2.9. (Conditional Average Value-at-Risk) The **conditional average value at risk** is defined as a mapping $AV@R_\alpha(\bullet|\mathcal{F}_1) : \boldsymbol{L}_p(\Omega, \mathcal{F}, \mathbb{P}) \to \overline{\boldsymbol{L}}_1(\Omega, \mathcal{F}_1, \mathbb{P})$ by a generalized LP:

$$AV@R_\alpha(X|\mathcal{F}_1) = \inf_Z \left\{ \mathbb{E}(X \cdot Z|\mathcal{F}_1) : \ 0 \leq Z \leq \frac{1}{\alpha}\, a.s.,\ \mathbb{E}(Z|\mathcal{F}_1) = 1\, a.s., Z \in \boldsymbol{L}_q(\Omega, \mathcal{F}, \mathbb{P}) \right\}$$

The "dual variables" Z in this definition are bounded almost surely and therefore they also are in $\boldsymbol{L}_\infty(\Omega, \mathcal{F}, \mathbb{P})$. This means that the conditional $AV@R_\alpha$ maps the space $\boldsymbol{L}_p(\Omega, \mathcal{F}_1, \mathbb{P})$ into itself. Furthermore, the conditional $AV@R$ is the infimum of linear mappings, hence concave.

Corollary 4.2.10. *The $AV@R_\alpha$ is a continuous mapping and the superdifferential is a nonempty set for any $X \in \boldsymbol{L}_p(\Omega, \mathcal{F}, \mathbb{P})$.*

Proof. Because of theorem 4.2.1 the $AV@R_\alpha$ is continuous. Then from proposition 4.2.3 the nonemptyness of the mapping follows. □

Because the conditional $AV@R_\alpha$ is continuous it follows from theorem 4.2.8 that the mapping is monotonic and predictable translation-equivariant.

CHAPTER 5

Multi-Period Acceptability Functionals

Until now we have considered acceptability functionals and conditional mappings for random variables, both of them connected to a single period. Acceptability functionals express the desirability of a random variable in terms of a characteristic real number, based on the assumption that no nontrivial information is available. Conditional mappings take into account that there could be some additional information about a random variable at some point of time and result not in a real number but in a random variable.

In this chapter these concepts are applied to a multi-period setup. In doing so we have to look at stochastic processes $\{X_t(\omega) : t \in S\}$ where S is some index set. We stick to a finite framework and assume the set S to be a set of discrete points of time, e.g. $S = \{0, 1, \ldots, T\}$. In economic applications such a process could represent flow quantities like cash flows as well as stock quantities like reserves or firm values.

In principle such "stochastic processes" with finite index sets can be considered as random vectors $(X_1, \ldots, X_T) \in \boldsymbol{L}_{p_1}(\Omega, \mathcal{F}, \mathbb{P}) \times \ldots \times \boldsymbol{L}_{p_T}(\Omega, \mathcal{F}, \mathbb{P})$. Functionals $\mathcal{A}(X_1, \ldots, X_T)$, mapping random vectors into the real line could be defined easily and the whole theory of subgradients and conjugate functionals is applicable, if we only take into account that the dual pairing for random vectors is given by $\sum_{t=1}^{T} \mathbb{E}(X_t \cdot Z_t)$ where the Z_t are dual to the X_t.

Therefore at first glance there seems to be little difference between the case of one period acceptability functionals and the multi-period functionals defined in this chapter: Is it not sufficient to just define a multi-period functional as $\mathcal{A}(X_1, \ldots, X_T)$?

In the last chapter we already saw that information plays a key role for conditional mappings. This is the case even more for multi-period functionals: A famous example by Philippe Artzner illustrates the role of information for the valuation of multi-period risk or acceptability.

Example. A fair coin is tossed three times. Two payoff functions are given:

A) One unit is paid at the end, if head was shown at least two times.

B) One unit is paid at the end, if head was shown at the last throw.

It is easy to verify that both payoff functions have the same (multivariate) distribution. But for case A) the final payoff can in some scenarios be predicted earlier than for payoff function B). That means: the information structure is not the same for both games and A) should be preferable to B).

Basically the total risk of a process is independent of any decision regarding risk management. If it is possible to take some action - like hedging, insurance or pooling risks - the risk can be reduced to some level depending on the information available. This remaining part of the risk is called intrinsic risk. Because we have in view the application of risk functionals for optimization, intrinsic risk is the appropriate risk-concept and we have to analyze risk functionals that depend on the available information[1].

In probability theory the evolution of information is modeled by filtrations and meaningful acceptability measures should reflect relevant differences in information structure. The first section of the chapter will deal with the basic properties of multi-period acceptability functionals, and the role of information is one of the key themes here. The terminology will be based mainly on [**23**], chap. 3 and on [**24**].

Pflug and Römisch ([**23**]) define different types of multi-periodic acceptability functionals: Separable functionals as sums of univariate acceptability functionals; the value of information of a multi-stage decision problem; compositions of conditional acceptability mappings and polyhedral multi-period acceptability functionals.

We do not want to discuss all of these approaches in depth or contribute any new approach. Instead we just apply the results of the last chapters to those already known types of multi-period functionals that use conditional acceptability measures as their building blocks. Unfortunately by now there is no generally accepted way of combining conditional mappings to acceptability (type) functionals. It seems that there is the need for future research on this issue. In the last two sections of the chapter we discuss two constructions for multi-period acceptability (type) functionals and their properties, using the results of chapter 4.

Compositions of conditional acceptability mappings are investigated first. Such functionals were defined by Ruszczyński and Shapiro ([**31, 32**]) and are constructed by applying

[1]([**23**], p 133)

conditional acceptability mappings and acceptability functionals at each period under consideration: they are given by

$$\mathcal{CA}\left(Y;\left[\mathcal{A}_0(\cdot),\mathcal{A}_1(\cdot),\ldots,\mathcal{A}_{t-1}(\cdot)\right];\mathbf{F}\right):=\mathcal{A}_0\left(Y_1+\mathcal{A}_1\left(Y_2+\mathcal{A}_2\left(Y_3+\ldots\mathcal{A}_{T-1}\left(Y_T\right)\right)\right)\right).$$

Because of their nested structure we will also call such mappings "nested mappings".

In the last section the simple case of sums $\sum_{t=1}^{T}\mathbb{E}\left(\mathcal{A}\left(X_t|\mathcal{F}_{t-1}\right)\right)$ are discussed. Such functionals are special cases of separable multi-period functionals and are called *separable expected conditional (SEC)* functionals ([**23**]). They are much easier to handle than composed functionals, but take into consideration only one conditional mapping for each period.

5.1. Multi-Period Acceptability Functionals - Basic Definitions

Multi-period functionals are mappings from spaces $\times_{t=1}^{T}\boldsymbol{L}_p\left(\Omega,\mathcal{F}_t,\mathbb{P}\right)$ into the extended real line $\overline{R}$. Such spaces endowed with a p-norm $\|X\|_p=\sum_{t=1}^{T}\mathbb{E}\left(|X_t|^p\right)^{\frac{1}{p}}$, $1\leq p\leq\infty$ are Banach spaces.

Basically the idea of multi-period functionals is, to jointly valuate a random vector $X=(X_1,\ldots,X_T)'$ together with an information structure which represents the gain in information over time, by a real number. The development of information is modeled by filtrations: A filtration $\boldsymbol{\mathcal{F}}=(\mathcal{F}_0,\mathcal{F}_1,\ldots,\mathcal{F}_T)$ is an increasing sequence of $\sigma-$Algebras $\mathcal{F}_t$, where $\mathcal{F}_t\subseteq\mathcal{F}_{t+1}$. $\mathcal{F}_0$ represents the trivial $\sigma-$Algebra $\mathcal{F}_0=\{\Omega,\emptyset\}$.

It is possible to define multi-period acceptability functionals as generalization of single-period acceptability functionals with the additional requirement of information monotonicity in the following way:

Definition 5.1.1. (Multi-Period Acceptability Functional) We will call a multi-period functional $\mathcal{A}\left(X;\boldsymbol{\mathcal{F}}\right)$ **multi-period acceptability functional**, if it is proper and satisfies

(MA0) **Information Monotonicity**: If $\mathcal{F}_t\subseteq\mathcal{F}_t'$ for all t, then

$$\mathcal{A}\left(X;\boldsymbol{\mathcal{F}}\right)\leq\mathcal{A}\left(X;\boldsymbol{\mathcal{F}}'\right)$$

(MA1) **Predictable Translation Equivariance**: For all periods t

$$\mathcal{A}\left(X_1,\ldots,X_t+C,X_T;\boldsymbol{\mathcal{F}}\right)=\mathcal{A}\left(X_1,\ldots,X_t,X_T;\boldsymbol{\mathcal{F}}\right)+\mathbb{E}\left(C\right)$$

holds, if C is a $\mathcal{F}_{t-1}-$measurable function.

(MA2) **Concavity**: The mapping $X \mapsto \mathcal{A}(X;\boldsymbol{\mathcal{F}})$ is concave.

(MA3) **Monotonicity**: $X_t \leq X_t^{\prime}\, a.s.$for all t implies $\mathcal{A}(X;\boldsymbol{\mathcal{F}}) \leq \mathcal{A}\left(X^{\prime};\boldsymbol{\mathcal{F}}\right)$

The condition (MA1) is very strong and it is not easy to construct acceptability measures that fulfill it. Therefore a couple of weaker conditions have been formulated in the literature ([**7, 35, 12**]). A reasonable condition is weak translation-equivariance:

(MA1') **Weak Translation Equivariance**:

$$\mathcal{A}(X_1,\ldots,X_t+c,X_T;\boldsymbol{\mathcal{F}}) = \mathcal{A}(X_1,\ldots,X_t,X_T;\boldsymbol{\mathcal{F}}) + c$$

for all constants c.

We will also refer to functionals that fulfill (MA0), (MA1'), (MA2), (MA3) and (MA4) as *weak multi-period acceptability functionals.*

Similar to the case of one-period functionals there are additional properties of multi-period acceptability functionals that can be interesting:

Definition 5.1.2. (positive homogeneity) A multi-period (weak) acceptability functional $\mathcal{A}(X;\boldsymbol{\mathcal{F}})$ is **positively homogeneous,** if

(MA4) $\mathcal{A}(\lambda\cdot X_1,\ldots,\lambda\cdot X_t,\lambda\cdot X_T;\boldsymbol{\mathcal{F}}) = \lambda\cdot\mathcal{A}(X_1,\ldots,X_t,X_T;\boldsymbol{\mathcal{F}})$ holds for all $\lambda > 0$.

A multi-period (weak) acceptability is **strict,** if it satisfies

(MA5) $\mathcal{A}(X_1,\ldots,X_t,X_T;\boldsymbol{\mathcal{F}}) \leq \sum_{t=1}^{T}\mathbb{E}(X_t)$.

The dual space of a space $\times_{t=1}^{T}\boldsymbol{L}_p(\Omega,\mathcal{F}_t,\mathbb{P})$ can be identified with the space $\times_{t=1}^{T}\boldsymbol{L}_q(\Omega,\mathcal{F}_t,\mathbb{P})$, where $\frac{1}{p}+\frac{1}{q}=1$ and the dual pairing between elements of these spaces is given by $\langle X,Z\rangle = \sum_{t=1}^{T}\mathbb{E}(X_tZ_t)$. If $\mathcal{A}(X;\boldsymbol{\mathcal{F}})$is a functional mapping from $\times_{t=1}^{T}\boldsymbol{L}_q(\Omega,\mathcal{F}_t,\mathbb{P})$ to the extended real line $\bar{R}$, it is possible to define conjugate and biconjugate functionals in the usual way, similar to the case of single-period functionals.

Definition 5.1.3. (concave conjugate) The **(concave) conjugate** of an multi-period functional $\mathcal{A}(X;\boldsymbol{\mathcal{F}})$ is given by

$$\mathcal{A}^*(Z;\boldsymbol{\mathcal{F}}) = \inf\left\{\langle X,Z\rangle - \mathcal{A}(X;\boldsymbol{\mathcal{F}}):\ X\in\times_{t=1}^{T}\boldsymbol{L}_p(\Omega,\mathcal{F}_t,\mathbb{P})\right\}.$$

The (concave) biconjugate is given by

$$\mathcal{A}^{**}(X;\boldsymbol{\mathcal{F}}) = \inf\left\{\langle X,Z\rangle - \mathcal{A}^*(Z;\boldsymbol{\mathcal{F}}):\ Z\in\times_{t=1}^{T}\boldsymbol{L}_q(\Omega,\mathcal{F}_t,\mathbb{P})\right\}.$$

If $\mathcal{A}(X;\mathbf{F})$ is proper, concave and upper semicontinuous the generalization of the Fenchel-Moreau-Rockafellar theorem[2] ensures that the dual representation

$$\mathcal{A}(X;\boldsymbol{\mathcal{F}}) = \mathcal{A}^{**}(X;\boldsymbol{\mathcal{F}})$$

holds.

This fact can be used to extract the representation for the special case of (weak) multi-period acceptability functionals.

PROPOSITION 5.1.4. *Let $\mathcal{A}(\bullet;\boldsymbol{\mathcal{F}})$ be an upper semicontinuous multi-period acceptability functional. Then the representation*

$$\mathcal{A}(X;\boldsymbol{\mathcal{F}}) = \inf_{Z}\left\{\sum_{t=1}^{T}\mathbb{E}(X_t \cdot Z_t) - \mathcal{A}^{*}(Z;\boldsymbol{\mathcal{F}}) :\ Z_t \geq 0;\ \mathbb{E}(Z_t|\mathcal{F}_{t-1}) = 1\right\} \tag{5.1.1}$$

holds for any $X \in \times_{t=1}^{T}\boldsymbol{L}_p(\Omega,\mathcal{F}_t,\mathbb{P})$.

Conversely - if $\mathcal{A}(\bullet;\boldsymbol{\mathcal{F}})$can be represented by a dual representation 5.1.1 and the conjugate $\mathcal{A}^*$ is proper, then $\mathcal{A}$ is proper, upper semicontinuous and satisfies (MA0)-(MA3).

PROOF. This is theorem 3.20 from **[23]**. □

PROPOSITION 5.1.5. *Let $\mathcal{A}(\bullet;\boldsymbol{\mathcal{F}})$ be an upper semicontinuous multi-period functional satisfying (MA1'), (MA2) and (MA3). Then the representation*

$$\mathcal{A}(X;\boldsymbol{\mathcal{F}}) = \inf_{Z}\left\{\sum_{t=1}^{T}\mathbb{E}(X_t \cdot Z_t) - \mathcal{A}^{*}(Z;\boldsymbol{\mathcal{F}}) :\ Z_t \geq 0;\ \mathbb{E}(Z_t) = 1\right\} \tag{5.1.2}$$

holds for every $X \in \times_{t=1}^{T}\boldsymbol{L}_p(\Omega,\mathcal{F}_t,\mathbb{P})$.

Conversely - if $\mathcal{A}(\bullet;\boldsymbol{\mathcal{F}})$ can be represented by a dual representation 5.1.2 and the conjugate $\mathcal{A}^*$ is proper, then $\mathcal{A}$ is proper, upper semicontinuous and satisfies (MA1'),(MA2) and (MA3).

PROOF. This is theorem 3.21 from **[23]**. □

It should be noted that theorem 5.1.5 does not say anything about information monotonicity (MA0), so this property must be verified separately for a given functional, to assess the weak multiperiod acceptability property.

[2]**[28]**, Theorem 5

5.2. Acceptability Compositions

In calculating the expected present value of some discounted cash-flows $\{X_i\}_{i\in\{1,\ldots,T\}}$over time, in some algorithms it is useful procedure to do the evaluation in a recursive manner, going back from the last period to the first one, using the projection property of conditional expectation:

$$\mathbb{E}\left(\mathbb{E}\left(X_t|\mathcal{F}_{t-1}\right)|\mathcal{F}_{t-2}\right) = \mathbb{E}\left(X_t|\mathcal{F}_{t-2}\right),$$

if $\mathcal{F}_{t-2} \subseteq \mathcal{F}_{t-1}$.

First $\mathbb{E}(X_T|\mathcal{F}_{T-1})$ is calculated, then $\mathbb{E}(X_{T-1}+\mathbb{E}(X_T|\mathcal{F}_{T-1})|\mathcal{F}_{T-2}) = \mathbb{E}(X_{T-1}+X_T|\mathcal{F}_{T-2})$. This procedure can be iterated backwards until the unconditional expectation for the first period can be calculated: $\mathbb{E}(X_1 + \mathbb{E}(\sum_{t=2}^{T} X_t|\mathcal{F}_1)) = \mathbb{E}(\sum_{t=1}^{T} X_t)$.

Although the projection property generally does not hold for conditional acceptability mappings, the principle of composition can be used to define multi-period acceptability functionals. This idea was used first by Ruszczyński in [**31, 33**].

If acceptability mappings should be nested it is necessary to ensure that at each step the mapping is done into the right subspace:

Definition 5.2.1. Let $\bar{p}$ be a sequence of real numbers $\bar{p} = (p_0 \leq p_1 \leq \ldots \leq p_T)$, with $1 \leq p_t \leq \infty$. We will call a sequence of mappings $\{\mathcal{A}_t(\cdot)\}_{t\in\{0,\ldots,T-1\}}$ $\bar{p}-$**integrability adapted** if $\mathcal{A}_{t-1}(\cdot)$ maps from $\boldsymbol{L}_{p_t}(\Omega, \mathcal{F}_t, \mathbb{P})$ into $\boldsymbol{L}_{p_{t-1}}(\Omega, \mathcal{F}_{t-1}, \mathbb{P})$ for all $t \in \{1, \ldots, T-1\}$.

For a concise definition, initially we define nested conditional mappings as compositions of future acceptability mappings, relative to a starting period. If an unconditional acceptability functional is applied to a nested conditional mapping the result is an acceptability composition:

Definition 5.2.2. (acceptability composition) Let $\mathcal{A}_0(\cdot)$ be an acceptability functional and $\{\mathcal{A}_t\}_{t\in\{1,\ldots,T\}}$ a collection of $\bar{p}-$integrability adapted conditional acceptability mappings. Moreover let $Y = \{Y_t\}_{t\in\{1,\ldots,T\}}$ be a sequence of random variables adapted to a filtration $\boldsymbol{\mathcal{F}}$. Then we define **nested conditional acceptability functionals** $\mathcal{CA}_t$ for time-points t recursively as

$$\mathcal{CA}_t\left(Y_{t+1},...,Y_T;[\mathcal{A}_t(\cdot),\ldots,\mathcal{A}_{T-1}(\cdot)];\boldsymbol{\mathcal{F}}\right):=\begin{cases}0,\text{ if } t\geq T\\ \mathcal{A}_t\left(Y_{t+1}+\mathcal{CA}_{t+1}\left(Y_{t+2},...,Y_T\right)\right),\text{ otherwise}\end{cases}.$$

.

An **acceptability composition** is an (unconditional) multi-period acceptability functional defined as

$$\mathcal{CA}\left(Y;[\mathcal{A}_0(\cdot),\mathcal{A}_1(\cdot),\ldots,\mathcal{A}_{T-1}(\cdot)];\boldsymbol{\mathcal{F}}\right):=\mathcal{CA}_0\left(Y;[\mathcal{A}_0(\cdot),\mathcal{A}_1(\cdot),\ldots,\mathcal{A}_{T-1}(\cdot)];\boldsymbol{\mathcal{F}}\right),$$

where $\mathcal{A}_0(\cdot)$ is an (unconditional) acceptability functional.

Sometimes we will write the composition informally as

$$\mathcal{CA}\left(Y;[\mathcal{A}_0(\cdot),\mathcal{A}_1(\cdot),\ldots,\mathcal{A}_{t-1}(\cdot)];\boldsymbol{\mathcal{F}}\right):=\mathcal{A}_0\left(Y_1+\mathcal{A}_1\left(Y_2+\mathcal{A}_2\left(Y_3+...\mathcal{A}_{T-1}\left(Y_T\right)\right)\right)\right).$$

If the filtration and the acceptability functionals are clear from the context we may write $\mathcal{CA}\left(Y\right)$ for the composition and $\mathcal{CA}_t\left(Y\right)$ for the nested functional. Using the components of Y explicitly we also may write $\mathcal{CA}\left(Y_1,\ldots,Y_T\right)$.

Because of the monotonicity of conditional properties it is possible to state a chain rule for the supergradient of compositions of acceptability functionals. Such a supergradient is determined by the supergradients of the individual conditional acceptability functionals and the supergradient of the unconditional acceptability functional under consideration. We do the formulation in two steps, proving the following lemma first:

LEMMA 5.2.3. *Let $\{\mathcal{A}_t\}_{t\in\{1,...,T-1\}}$ be a collection of $\bar{p}-$integrability adapted conditional acceptability functionals with $T\geq 2$. Given supergradients*

$$\bar{Z}_T\in\partial\mathcal{A}_{T-1}(X_T)$$

and

$$\bar{Z}_k\in\partial\mathcal{A}_{k-1}(X_k+CA_k\left(X_{k+1},...,X_T\right))$$

for $k\in\{t_0+1,...,T\}$, a supergradient for the nested conditional acceptability functional $\mathcal{CA}_{t_0}\left(\cdot;[\mathcal{A}_{t_0}(\cdot),,\ldots,\mathcal{A}_{T-1}(\cdot)];\boldsymbol{\mathcal{F}}\right)$ at the base points $X_{t_0+1},...,X_T$ is given by a $(T-t)$-tupel $W=(W_{t_0+1},...,W_T)$ with

$$W_{t_0+1}\equiv\bar{Z}_{t_0+1}$$

and

$$W_{k+1} = W_k \cdot \bar{Z}_{k+1}$$

for $t_0 < k \leq T-1$.

PROOF. We use backward induction to prove the result:

For $t = T-1$ we have $\mathcal{CA}_{T-1}(X_T; [\mathcal{A}_{\mathrm{T}-1}(\cdot)]) = \mathcal{A}_{T-1}(X_T)$ and $\bar{Z}_T$ is a supergradient by the premises of the lemma.

Assume now that the proposition is shown for all $t \geq t_0$. Then subgradients for

$\mathcal{CA}_{t_0}(Y_{t_0+1}, ..., Y_T; [\mathcal{A}_{t_0}(\cdot), ..., \mathcal{A}_{\mathrm{T}-1}(\cdot)])$ are given by $(W_{t_0+1}, ..., W_T)$. That means:

$\mathcal{CA}_{t_0}(X_{t_0+1} + Y_{t_0+1}, \ldots, X_T + Y_T) \leq$

$\mathcal{CA}_{t_0}(X_{t_0+1}, \ldots, X_T) + \mathbb{E}\left(Y_{t_0+1} \cdot \bar{Z}_{t_0+1} | \mathcal{F}_{t_0}\right) + \ldots + \mathbb{E}\left(Y_T \cdot \bar{Z}_T \cdot \bar{Z}_{T-1} \cdot \ldots \cdot \bar{Z}_{t_0+1} | \mathcal{F}_{t_0}\right)$

Because of monotonicity we have for $t_0 - 1$:

$$\begin{aligned}
&\mathcal{CA}_{t_0-1}(X_{t_0} + Y_{t_0}, X_{t_0+1} + Y_{t_0+1}, \ldots, X_T + Y_T) = \\
&\qquad = \mathcal{A}_{t_0-1}(X_{t_0} + Y_{t_0} + \mathcal{CA}_{t_0}(X_{t_0+1} + Y_{t_0+1}, \ldots, X_T + Y_T)) \\
&\leq \mathcal{A}_{t_0-1}\left(X_{t_0} + \mathcal{CA}_{t_0}(X_{t_0+1}, \ldots, X_T) + Y_{t_0} + \mathbb{E}(Y_{t_0+1} \cdot \bar{Z}_{t_0+1} | \mathcal{F}_{t_0}) + \ldots + \mathbb{E}\left(Y_T \cdot \prod_{t=t_0+1}^{T} \bar{Z}_t | \mathcal{F}_{t_0}\right)\right)
\end{aligned}$$

As $\bar{Z}_{t_0}$ is a supergradient of $\mathcal{A}_{t_0-1}$ at $X_{t_0} + \mathcal{CA}_{t_0}(X_{t_0+1}, \ldots, X_T)$ it follows that

$$\begin{aligned}
&\mathcal{CA}_{t_0-1}(X_{t_0} + Y_{t_0}, X_{t_0+1} + Y_{t_0+1}, \ldots, X_T + Y_T) \leq \\
&\qquad \leq \mathcal{A}_{t_0-1}(X_{t_0} + \mathcal{CA}_{t_0}(X_{t_0+1}, \ldots, X_T)) + \\
&\quad + \mathbb{E}\left(\bar{Z}_{t_0} \cdot \left[Y_{t_0} + \mathbb{E}\left(Y_{t_0+1} \cdot \bar{Z}_{t_0+1} | \mathcal{F}_{t_0}\right) + \ldots + \mathbb{E}\left(Y_T \cdot \prod_{t=t_0+1}^{T} \bar{Z}_t | \mathcal{F}_{t_0}\right)\right] | \mathcal{F}_{t_0-1}\right) = \\
&= \mathcal{CA}_{t_0-1}(X_{t_0}, X_{t_0+1}, \ldots, X_T) + \mathbb{E}\left(Y_{t_0} \cdot \bar{Z}_{t_0} | \mathcal{F}_{t_0-1}\right) + \mathbb{E}\left(Y_{t_0+1} \cdot \bar{Z}_{t_0+1} \cdot \bar{Z}_{t_0} | \mathcal{F}_{t_0-1}\right) + \ldots \\
&\qquad + \mathbb{E}\left(Y_T \cdot \bar{Z}_T \cdot \bar{Z}_{T-1} \cdot \ldots \cdot \bar{Z}_{t_0+1} \cdot \bar{Z}_{t_0} | \mathcal{F}_{t_0-1}\right)
\end{aligned}$$

It should be noted that because the mappings are integrability adapted and the $\bar{Z}_t$ are supergradients, the conditional expectations $\mathbb{E}\left(X_t \cdot \bar{Z}_t | \mathcal{F}_{t-1}\right)$ are also p_{t-1}-integrable by theorem 4.2.1. This means all the (conditional) expectations involved in each nesting-step are p_t-integrable.

Because we made no restriction on T it was shown by backward induction that $W = (W_{t+1}, ..., W_T)$ as defined above is a supergradient for $\mathcal{CA}_t(\cdot)$ for any $t \leq T$. □

Interestingly the supergradients W_t in Lemma 5.2.3 above are martingales. This is shown in the following corollary:

COROLLARY 5.2.4. *Let $\{\mathcal{A}_t\}_{t\in\{1,...,T\}}$ be a collection of u.s.c. $\bar{p}-$integrability adapted conditional acceptability mappings again. Then the process $\{W_t\}_{t\in\{t_0+1,...,T\}}$ of the supergradients defined in Lemma 5.2.3 is a martingale.*

PROOF. From theorem 4.2.8 we know $\mathbb{E}\left(\bar{Z}_t|\mathcal{F}_{t-1}\right) = 1$ because of predictable translation equivariance. Moreover as the process W_t is adapted to $\mathcal{F}_t$, W_{t-1} is $\mathcal{F}_{t-1}$-measurable. Therefore

$$\mathbb{E}\left(W_t|\mathcal{F}_{t-1}\right) = \mathbb{E}\left(W_{t-1} \cdot \bar{Z}_t|\mathcal{F}_{t-1}\right) = W_{t-1} \cdot \mathbb{E}\left(\bar{Z}_t|\mathcal{F}_{t-1}\right) = W_{t-1}$$

holds. □

Using Lemma 5.2.3, the chain rule for compositions can be formulated in the following theorem:

THEOREM 5.2.5. *Let $\mathcal{A}(\cdot)$ be an acceptability functional and $\{\mathcal{A}_{t-t}\}_{t\in\{1,...,T\}}$, $T \geq 2$ a collection of integrability adapted conditional acceptability functionals. Choose supergradients $\bar{Z}_T \in \partial\mathcal{A}_{T-1}(\cdot)|_{X_T}$ and $\bar{Z}_t \in \partial\mathcal{A}_{t-1}(\cdot)|_{X_t+CA_t(X_{t+1},...,X_T)}$ for $t \in \{0, ..., T-1\}$. Then a supergradient for the acceptability composition $\mathcal{CA}(\cdot; [\mathcal{A}_0(\cdot), \mathcal{A}_1(\cdot), \ldots, \mathcal{A}_{T-1}(\cdot)]; \boldsymbol{\mathcal{F}})$ at the base points $X_1, ..., X_T$ is given by a T-tuple $W = (W_1, ..., W_T)$ with*

$$W_1 \equiv Z_1$$

and

$$W_{t+1} = W_t \cdot \bar{Z}_{t+1}.$$

PROOF. From Lemma 5.2.3 it follows that the composition of the integrability adapted acceptability functionals $\{\mathcal{A}_t\}_{t\in\{1,...,T\}}$ has a supergradient given by $W^1 = (W_2, ..., W_T)$. Applying the same argument as above, using monotonicity and the definition of subgradients this time for the unconditional acceptability functional $\mathcal{A}_0$ we get

$$\mathcal{CA}_0\left(X_1 + Y_1, X_2 + Y_2, \ldots, X_T + Y_T\right) = \mathcal{A}_0\left(X_1 + Y_1 + \mathcal{CA}_1\left(X_2 + Y_2, \ldots, X_T + Y_T\right)\right)$$

$$\leq \mathcal{A}_0\left(X_1+\mathcal{CA}_1\left(X_2,\ldots,X_T\right)+Y_1+\mathbb{E}\left(Y_2\cdot\bar{Z}_2|\mathcal{F}_1\right)+\ldots+\mathbb{E}\left(Y_T\cdot\bar{Z}_T\cdot\bar{Z}_{T-1}\cdot\ldots\cdot\bar{Z}_2|\mathcal{F}_1\right)\right)$$

and therefore

$$\begin{aligned}
&\mathcal{CA}_0\left(X_1+Y_1,X_2+Y_2,\ldots,X_T+Y_T\right)\leq\\
&\qquad\leq \mathcal{A}_0\left(X_1+\mathcal{CA}_1\left(X_2,\ldots,X_T\right)\right)+\\
&\qquad+\mathbb{E}\left(\bar{Z}_1\cdot\left[Y_1+\mathbb{E}\left(Y_2\cdot\bar{Z}_2|\mathcal{F}_1\right)+\ldots+\mathbb{E}\left(Y_T\cdot\bar{Z}_T\cdot\bar{Z}_{T-1}\cdot\ldots\cdot\bar{Z}_2|\mathcal{F}_1\right)\right]\right)\\
&=\mathcal{CA}_0\left(X_1,X_2,\ldots,X_T\right)+\mathbb{E}\left(\mathrm{Y}_1\cdot\bar{\mathrm{Z}}_1\right)+\mathbb{E}\left(Y_2\cdot\bar{Z}_2\cdot\bar{Z}_1\right)+\ldots+\mathbb{E}\left(Y_T\cdot\bar{Z}_T\cdot\bar{Z}_{T-1}\cdot\ldots\cdot\bar{Z}_2\cdot\bar{Z}_1\right)
\end{aligned}$$

□

Remark. The critical conditions for theorem 5.2.5 are monotonicity and concavity. While concavity assures that something like a supergradient makes sense, by monotonicity inequalities for later periods can be translated into inequalities for earlier periods. Translation equivariance is not a necessary condition for theorem 5.2.5, so it could be stated for concave monotonic conditional acceptability mappings.

As pointed out before, we abstain from using theorem 4.3 in Papageorgiou [**21**], which is valid for completely continuous operators. Nevertheless at least continuity of the operators at some point inside the domain would - in the light of theorem 4.2.3 - be a useful additional property, assuring the existence of subgradients in the interior of the domain.

So far we have characterized supergradients of an acceptability composition. Using theorem 4.2.8 we can also calculate the dual representation of the composition using the concave conjugate functions of the constituent conditional mappings.

Theorem 5.2.6. *Under the assumptions of theorem 5.2.5 an acceptability composition can be represented by*

$$\mathcal{CA}_0\left(X_1,X_2,\ldots,X_T\right)=\inf_{Z_1,\ldots,Z_T}\left\{\sum_{t=1}^{T}\mathbb{E}\left(X_t\cdot M_t\right)-\mathcal{A}_0^*\left(Z_1\right)-\sum_{t=1}^{T-1}\mathbb{E}\left(\mathcal{A}_t^*\left(Z_{t+1}\right)\cdot M_t\right):\ Z\in\mathcal{Z}\right\}\tag{5.2.1}$$

with $\mathcal{Z}=\{Z|\,\forall t:\ M_1=Z_1;\ M_{t+1}=M_t\cdot Z_{t;};\ Z_t\geq 0;\ \mathbb{E}\left(Z_t|\mathcal{F}_{t-1}\right)=1,\ Z_t\in\mathcal{Z}_t\}$ *being a subset of the dual space.*

Proof. Theorem 4.2.8 can be used recursively to replace the conditional mapping $\mathcal{A}_t$ in the definition of conditional acceptability mappings (definition 5.2.2). In this way the theorem can be proved by backward induction. Here we will only show one step in this procedure:

By theorem 4.2.8 and because the $\mathcal{A}_t$ are conditional acceptability mappings by theorem 4.2.8 we have

$$\mathcal{A}_{T-2}\left(Y_{T-1}\right) = \inf_{Z_{T-1}} \left\{\mathbb{E}_{T-2}\left(Y_{T-1} \cdot Z_{T-1}\right) - \mathcal{A}^*_{T-2}\left(Z_{T-1}\right) : \; Z_{T-1} \geq 0; \; \mathbb{E}_{T-2}\left(Z_{T-1}\right) = 1\right\}.$$

This holds also for time $t = 0$, where the appropriate σ-Algebra becomes the trivial σ-Algebra and the mapping, its dual and the expectation can be understood as unconditional. If we apply these identities to the definition of acceptability compositions we get :

$$\begin{aligned}
\mathcal{A}_{T-2}\left(X_{T-1} + \mathcal{A}_{T-1}\left(X_T\right)\right) &= \inf_{Z_{T-1}} \left\{\mathbb{E}_{T-2}\left(\left[X_{T-1} + \mathcal{A}_{T-1}\left(X_T\right)\right] \cdot Z_{T-1}\right) - \mathcal{A}^*_{T-2}\left(Z_{T-1}\right)\right\} \\
&= \inf_{Z_{T-1}} \left\{\mathbb{E}_{T-2}\left(\left[X_{T-1} + \inf_{Z_T}\left\{\mathbb{E}_{T-1}\left(Y_T \cdot Z_T | \mathcal{F}_{T-1}\right) - \mathcal{A}^*_{T-1}\left(Z_T\right)\right\}\right] \cdot Z_{T-1}\right) - \mathcal{A}^*_{T-2}\left(Z_{T-1}\right)\right\} \\
&= \inf_{Z_{T-1}} \left\{\mathbb{E}_{T-2}\left(X_{T-1} \cdot Z_{T-1}\right) + \mathbb{E}_{T-2}\left(\inf_{Z_T}\left\{\mathbb{E}_{T-1}\left(Y_T \cdot Z_T\right) - \mathcal{A}^*_{T-1}\left(Z_T\right)\right\} \cdot Z_{T-1}\right) - \mathcal{A}^*_{T-2}\left(Z_{T-1}\right)\right\}
\end{aligned}$$

We know a supergradient $\bar{Z}_T$ of $\mathcal{A}_{T-1}\left(X_T\right)$ by assumption and this supergradient is also a minimizer of $\left\{\mathbb{E}_{T-1}\left(Y_T \cdot Z_T | \mathcal{F}_{T-1}\right) - \mathcal{A}^*_{T-1}\left(Z_T\right)\right\}$ by theorem 4.2.7. Because of monotonicity all the Z_t are nonnegative and hence $\bar{Z}_T$ is also a minimizer of $\left[\mathbb{E}_{T-1}\left(Y_T \cdot Z_T\right) - \mathcal{A}^*_{T-1}\left(Z_T\right)\right] \cdot Z_{T-1}$. Therefore we can interchange infimum and conditional expectation by proposition 2.2.19.

$$\begin{aligned}
&\mathcal{A}_{T-2}\left(X_{T-1} + \mathcal{A}_{T-1}\left(X_T\right)\right) = \\
&= \inf_{Z_{T-1}} \left\{\mathbb{E}_{T-2}\left(X_{T-1} \cdot Z_{T-1}\right) + \inf_{Z_T}\left\{\mathbb{E}_{T-2}\left(\left[\mathbb{E}_{T-1}\left(Y_T \cdot Z_T\right) - \mathcal{A}^*_{T-1}\left(Z_T\right)\right] \cdot Z_{T-1}\right)\right\} - \mathcal{A}^*_{T-2}\left(Z_{T-1}\right)\right\} \\
&= \inf_{Z_{T-1}, Z_T} \left\{\mathbb{E}_{T-2}\left(X_{T-1} \cdot Z_{T-1}\right) + \mathbb{E}_{T-2}\left(\left[\mathbb{E}_{T-1}\left(Y_T \cdot Z_T\right) - \mathcal{A}^*_{T-1}\left(Z_T\right)\right] \cdot Z_{T-1}\right) - \mathcal{A}^*_{T-2}\left(Z_{T-1}\right)\right\} \\
&= \inf_{Z_{T-1}, Z_T} \left\{\mathbb{E}_{T-2}\left(X_{T-1} \cdot Z_{T-1}\right) + \mathbb{E}_{T-2}\left(Y_T \cdot Z_T \cdot Z_{T-1}\right) - \mathbb{E}_{T-2}\left(\mathcal{A}^*_{T-1}\left(Z_T\right) \cdot Z_{T-1}\right) - \mathcal{A}^*_{T-2}\left(Z_{T-1}\right)\right\}
\end{aligned}$$

All the infima in this derivation must be understood with respect to the constraints $Z_T \geq 0 \wedge \mathbb{E}\left(Z_T | \mathcal{F}_1\right) = 1$ and $Z_{T-1} \geq 0 \wedge \mathbb{E}\left(Z_{T-1} | \mathcal{F}_1\right) = 1$.

In this derivation we make use of proposition 2.2.19. This is possible because the inner infimum, representing a conjugate mapping must be attained if the superdifferential is not empty - which is the case, since we know the supergradients by assumption.

Iterating these steps and defining the variable M_t as $M_1 = Z_1$ and $M_{t+1} = M_t \cdot Z_{t+1}$ we get the statement of the theorem. $\square$

Evidently there is a tight connection between theorems 5.2.6 and 5.2.5: The Z_t can be interpreted as supergradients of the conditional mappings $\mathcal{A}_t$ and then - in the light of theorem 5.2.5 - the M_t constitute a supergradient of the composition. So the equation

$$\mathcal{C}\mathcal{A}_0\left(X_1, X_2, \ldots, X_T\right) =$$

$$\inf_{Z_1,\ldots,Z_T} \left\{ \sum_{t=1}^{T} \mathbb{E}\left(X_t \cdot M_t\right) - \mathcal{A}_0^*\left(Z_1\right) - \textstyle\sum_{t=1}^{T-1} \mathbb{E}\left(\mathcal{A}_t^*\left(Z_{t+1}\right) \cdot M_t\right) :\ Z \in \mathcal{Z} \right\}$$

basically gives the conjugate representation of the composition and the mapping $\mathcal{A}_0^*\left(Z_1\right) + \sum_{t=1}^{T} \mathbb{E}\left(\mathcal{A}_t^*\left(Z_{t+1}\right) \cdot M_t\right)$ is its conjugate $\mathcal{C}\mathcal{A}_0^*$. We only have to eliminate the Z_t from the equation.

This can be done easily, if we assume $Z_t \neq 0\, a.s.$ for all t- leading to the inequality $Z_t > 0$, which means that all the mappings and functionals are strictly monotonic. In this case the Z_t can be easily replaced by $Z_t = \frac{M_t}{M_{t-1}}$. Using this substitution on equation 5.2.6, we get

$$\text{(5.2.2)}\quad \mathcal{C}\mathcal{A}_0\left(X_1, X_2, \ldots, X_T\right) =$$

$$\inf_{M_1,M_2,\ldots,M_T} \left\{ \sum_{t=1}^{T} \mathbb{E}\left(X_t \cdot M_t\right) - \mathcal{A}_0^*\left(M_1\right) - \sum_{t=1}^{T} \mathbb{E}\left(\mathcal{A}_t^*\left(\frac{M_{t+1}}{M_t}\right) \cdot M_t\right) :\ M_t \in \mathcal{M} \right\},$$

with

$$\mathcal{M} = \left\{ M_t |\ M_t > 0;\ \mathbb{E}\left(\frac{M_t}{M_{t-1}} | \mathcal{F}_{t-1}\right) = 1 \right\}.$$

Here the conjugate of the composition is given by $\mathcal{A}_0^*\left(M_1\right) + \sum_{t=1}^{T} \mathbb{E}\left(\mathcal{A}_t^*\left(\frac{M_{t+1}}{M_t}\right) \cdot M_t\right)$.

If the $Z_t(\omega)$ can be zero on some set Q with positive probability, equation 5.2.1 can be rewritten in the following way:

$$(5.2.3) \quad \mathcal{CA}_0 \left(X_1, X_2, \ldots, X_T\right) =$$
$$\inf_{M_1, M_2, \ldots, M_T} \left\{ \sum_{t=1}^{T} \mathbb{E}\left(X_t \cdot M_t\right) - \mathcal{A}_0^* \left(M_1\right) - \sum_{t=1}^{T} \psi(M_t, M_{t+1}) : \; M_t \in \mathcal{M}' \right\},$$

where

$$\mathcal{M}' = \left\{M | \; M_t \geq 0; \; \mathbb{E}\left(M_t | \mathcal{F}_{t-1}\right) = M_{t-1}\right\}$$

and

$$\psi(M_t, M_{t+1}) = \sup_{Z_{t+1}} \left\{\mathbb{E}\left(\mathcal{A}_t^* \left(Z_{t+1}\right) \cdot M_t\right) : M_{t+1} = M_t \cdot Z_{t+1}, \; Z_{t+1} \in \mathcal{Z}_t\right\}.$$

$\mathcal{M}'$ restricts the process M_t to be a nonnegative martingale and the functions $\psi(M_t, M_{t+1})$ are defined for them:

If we would assume a set of ω where M_{t-1} equals zero and M_t is positive, there must be another set with positive probability, where M_t is negative - otherwise the martingale restriction $\mathbb{E}\left(M_t | \mathcal{F}_{t-1}\right) = M_{t-1}$ can not be fulfilled. But for a nonnegative martingale such a set must have probability zero.

It follows that 5.2.1 or 5.2.3 are in fact the dual representations of acceptability compositions and $\mathcal{A}_0^* \left(M_1\right) + \sum_{t=1}^{T} \psi(M_t, M_{t+1})$ is the associated conjugate mapping.

We formulate this result as a corollary:

Corollary 5.2.7. *Under the assumptions of theorem 5.2.5 and with $\mathcal{M}'$ and $\psi(M_t, M_{t+1})$ defined as above, the supergradient representation of an acceptability composition is given by*

$$\mathcal{CA}_0 \left(X_1, X_2, \ldots, X_T\right) =$$

$$\inf_{M_1, M_2, \ldots, M_T} \left\{ \sum_{t=1}^{T} \mathbb{E}\left(X_t \cdot M_t\right) - \mathcal{A}_0^* \left(M_1\right) - \textstyle\sum_{t=1}^{T} \psi(M_t, M_{t+1}) : \; M_t \in \mathcal{M}' \right\},$$

with $\mathcal{M}' = \left\{M_t | \; M_t \geq 0; \; \mathbb{E}\left(M_t | \mathcal{F}_{t-1}\right) = M_{t-1}\right\}$.

Proof. Using the arguments above this follows from theorem 5.2.6 . ☐

It is easy to see that any acceptability composition must be concave and monotonic: As a nesting of monotonic concave mappings it must be concave itself and it is monotonic because of the inequalities $M_t \geq 0\, a.s.$ for all t. Furthermore, a composition of information monotonic acceptability mappings will have the property of (multi-period) information monotonicity: This is assured by the information monotonicity of each conditional mapping and the monotonicity of the unconditional functional together with the monotonicity of each mapping and functional.

It should be noted that acceptability compositions are not translation-equivariant: The equation $\mathbb{E}\left(M_t|\mathcal{F}_{t-1}\right) = M_{t-1}$ states that the process of dual variables is a martingale, but without additional assumptions there is no way to get the equation $\mathbb{E}\left(M_t|\mathcal{F}_{t-1}\right) = 1$ which would be required for translation equivariance . However from $M_0 = 1$ and $\mathbb{E}\left(M_t|\mathcal{F}_{t-1}\right) = M_{t-1}$ at least the equation $\mathbb{E}\left(M_t\right) = 1$ follows, which - by theorem 5.1.5 - is the criterion for weak translation equivariance.

We state these results as a corollary:

Corollary 5.2.8. *An acceptability composition is a concave (MA2), monotonic (MA3) multi-period functional, but generally not an acceptability functional. Only weak translation equivariance (MA1') holds. If the conditional mappings involved are information monotonic, the composition will be information monotonic (M0) as well. Under this condition an acceptability composition is a weak multi-period acceptability functional and also proper and upper semicontinuous.*

Proof. See the argumentation above. Properness and upper semicontinuity follows from 5.1.5. An alternative proof for (MA1'), (MA2) and (MA3), which is independent of the dual representation 5.2.2 is given in [**23**], theorem 3.33. □

Example 5.2.9. Based on theorem 5.2.7 the nested average value at risk - constructed by nesting conditional $AV@R_\alpha$s for the later periods with an unconditional $AV@R_\alpha$for the first period - has the following representation:

$$nAV@R_\alpha(X;\mathbf{F}) =$$

$$\inf_{M_1,M_2,\ldots,M_T}\left\{\sum_{t=1}^{T}\mathbb{E}\left(X_t\cdot M_t\right):\ 0\leq M_t\leq \tfrac{1}{\alpha}\cdot M_{t-1},\, \mathbb{E}\left(M_t|\mathcal{F}_{t-1}\right)=M_{t-1}, M_0=1\right\}$$

5.3. Separable Expected Conditional Functionals

A seemingly obvious way for defining multi-period functionals consists of applying single-period functionals to the random variables X_t of the process under consideration and sum up the results:

$$\mathcal{A}(X;\boldsymbol{\mathcal{F}}) = \sum_{t=1}^{T} \mathcal{A}^{[t]}(X_t)$$

Such multi-period functionals are called *separable functionals*. If the single-period functionals used are concave, the resulting separable functional - as a weighted sum of concave functionals with nonnegative, nonzero weights - is also concave (MA2). Also it is easily seen that separable functionals are weakly translation-equivariant (MA1') and monotonic (MA3) if this is also true for all their constituents.

Unfortunately it is not automatically guaranteed that a multi-period functional is information monotonic for arbitrary single period acceptability functionals. That means that generally such multi-period functionals are not weak acceptability functionals without additional measures to assure this property. For example sums of expectations or sums of $AV@R$s as well as sums of other typical single-period functionals are not information monotonic.

It is clear why this problem arises: If a multi-period functional should be information monotonic, the constituent single-period functionals must account for the information structure. But normal single-period functionals only rely on the information available at the beginning - namely the trivial σ-Algebra. Any information that might become known after the beginning is not relevant for them.

Happily there is a way for defining separable functionals that account for information in the right way: As we have seen above, information monotonicity is also an issue for conditional acceptability mappings (4.1.5). So, a better way of constructing multi-period functionals consists in taking sums of expectations of conditional acceptability mappings. Such functionals are called separable expected conditional (SEC, **[23]**, p 145).

DEFINITION 5.3.1. (SEC-functional) A multi-period acceptability functional is called **separable expected conditional (SEC)** if it is of the form

$$\mathcal{A}(X;\boldsymbol{\mathcal{F}}) = \sum_{t=1}^{T} \mathbb{E}\left(\mathcal{A}^{[t]}(X_t|\mathcal{F}_{t-1})\right),$$

where the $\mathcal{A}^{[t]}(\bullet|\mathcal{F}_{t-1})$ are conditional u.s.c. acceptability mappings.

The most important and best known SEC-functional is the multi-period average value at risk ([**25**]) $\sum_{t=1}^{T}\mathbb{E}\left(AV@R_\alpha\left(X_t|\mathcal{F}_{t-1}\right)\right)$.

SEC-functionals are weak multi-period acceptability functionals: As separable functionals they fulfill (MA1'), (MA2) and (MA3) and are information monotonic, if their constituent conditional mappings are information monotonic.

A multi-period functional $\mathcal{A}$ is separable if and only if its dual is separable, and it is SEC if and only if its dual is SEC (proposition 3.27 in [**23**]). That means that the concave conjugate of a SEC-functional can be represented in the form

$$\mathcal{A}^*\left(Z;\boldsymbol{\mathcal{F}}\right)=\sum_{t=1}^{T}\mathbb{E}\left(\mathcal{A}^{[t]*}\left(Z_t|\mathcal{F}_{t-1}\right)\right).$$

Building on the chain-rule from theorem 5.2.5 we can characterize the supergradients of SEC-functionals:

THEOREM 5.3.2. *Let* $\mathcal{A}\left(X;\boldsymbol{\mathcal{F}}\right)=\sum_{t=1}^{T}\mathbb{E}\left(\mathcal{A}^{[t]}\left(X_t|\mathcal{F}_{t-1}\right)\right)$ *be a SEC-functional and* $\overline{Z}=\left(\overline{Z}_1,\ldots,\overline{Z}_T\right)$ *a vector of supergradients of the constituent conditional acceptability mappings* $\mathcal{A}^{[t]}$. *Then* $\overline{Z}$ *is a supergradient of the SEC-functional.*

PROOF. Each summand $\mathbb{E}\left(\mathcal{A}^{[t]}\left(X_t|\mathcal{F}_{t-1}\right)\right)$ can be interpreted as an acceptability composition, with no action between the beginning and period t. The conjugate of the expectation equals zero at one and is unbounded elsewhere. That means that the (super)gradient of the expectation must be one (almost sure). Supergradients for the conditional acceptability mappings $\mathcal{A}^{[t]}$ are given by $\overline{Z}_t$.

Applying theorem 5.2.5 we see that for each t the product $\overline{Z}_t\cdot 1=\overline{Z}_t$ must be a supergradient of the associated acceptability composition $\mathbb{E}\left(\mathcal{A}^{[t]}\left(X_t|\mathcal{F}_{t-1}\right)\right)$, which means that

$$\mathbb{E}\left(\mathcal{A}^{[t]}\left(X_t+Y_t|\mathcal{F}_{t-1}\right)\right)\leq\mathbb{E}\left(\mathcal{A}^{[t]}\left(X_t|\mathcal{F}_{t-1}\right)\right)+\mathbb{E}\left(Y_t\cdot\overline{Z}_t\right).$$

Summing over all t we get

$$\sum_{t=1}^{T}\mathbb{E}\left(\mathcal{A}^{[t]}\left(X_t+Y_t|\mathcal{F}_{t-1}\right)\right)\leq\sum_{t=1}^{T}\mathbb{E}\left(\mathcal{A}^{[t]}\left(X_t|\mathcal{F}_{t-1}\right)\right)+\sum_{t=1}^{T}\mathbb{E}\left(Y_t\cdot\overline{Z}_t\right),$$

or

$$\mathcal{A}\left(X+Y;\boldsymbol{\mathcal{F}}\right) \leq \mathcal{A}\left(X;\boldsymbol{\mathcal{F}}\right) + \sum_{t=1}^{T} \mathbb{E}\left(Y_t \cdot \overline{Z}_t\right).$$

This shows that $\overline{Z}$ is a supergradient for the whole SEC-functional. □

REMARK. Because the constituents of a SEC functional are conditional acceptability mappings, the restrictions on the supergradients $\overline{Z}_t$ from theorem 4.2.8 - $\overline{Z}_t \geq 0$ and $\mathbb{E}\left(\overline{Z}_t|\mathcal{F}_{t-1}\right) = 1$ - hold.

5.4. Concluding Remarks

Acceptability compositions have some nice features e.g. the intuitive nesting structure, similar to the case of nested conditional expectations, the martingale property of their supergradients and information monotonicity. Additionally they are at least weak acceptability functionals. The main drawback comes from the requirement of integrability adaptedness - especially for the general case: It is tedious to keep track of which mapping goes into which space at which period and what is the related allowed space for the supergradients.

Of course there are possibilities for simplification: First of all it is possible to use conditional mappings $\boldsymbol{L}_p\left(\Omega, \mathcal{F}_t, \mathbb{P}\right) \rightarrow \overline{\boldsymbol{L}}_p\left(\Omega, \mathcal{F}_{\mathrm{t}-1}, \mathbb{P}\right)$. This seems to be the usual idea in literature [**23, 24, 32**]. But it might not not be easy to find such conditional mappings for $p > 1$.

So the most common special case are mappings $\boldsymbol{L}_1\left(\Omega, \mathcal{F}_t, \mathbb{P}\right) \rightarrow \overline{\boldsymbol{L}}_1\left(\Omega, \mathcal{F}_{\mathrm{t}-1}, \mathbb{P}\right)$ like the nested $AV@R$ above.

On the other hand we have seen that SEC-functionals have the same favorable properties as acceptability compositions, while their construction seems to be a lot easier. Also the requirements on the mappings involved and their supergradients are more modest. This is because the expectation will always work, as long as the conditional mappings used for each period at least map into $\overline{\boldsymbol{L}}_1\left(\Omega, \mathcal{F}_{\mathrm{t}-1}, \mathbb{P}\right)$.

Interestingly there is a connection between SEC-functionals and acceptability compositions: Assume that we apply any conditional acceptability mappings on a process, projecting the process back one period. On this "projected" process we apply an acceptability composition, using conditional expectations as the conditional acceptability mappings and expectation as the unconditional functional. What we then get is a SEC-functional.

This leads to a possible generalization of SEC -functionals: Apply conditional acceptability mappings on the process, projecting one period back and then apply any acceptability compositions with mappings $\boldsymbol{L}_1(\Omega, \mathcal{F}_t, \mathbb{P}) \to \overline{\boldsymbol{L}}_1(\Omega, \mathcal{F}_{\mathrm{t}-1}, \mathbb{P})$. One option would be again the nested $AV@R$.

Such a functional is not separable any longer. On the other hand

- Those aspects of risk that are connected with higher moments can be accounted for by the conditional mapping at the first step, which would not the case for simple acceptability composition that use exclusively mappings into $\overline{\boldsymbol{L}}_1(\Omega, \mathcal{F}_{\mathrm{t}-1}, \mathbb{P})$.
- The functional is simpler than a general acceptability composition.
- It is more informative than a SEC-functional, because for each period the whole evolvement of available information up to that period is used to measure the acceptability.

Finally the analysis of such and other possible multi-period acceptability functionals - based on conditional mappings - should be the subject of further research.

Bibliography

[1] C. Aliprantis and O. Burkinshaw. *Positive Operators.* Number 119 in Pure and Applied Mathematics. Academic Press, 1985.

[2] C. Aliprantis and O. Burkinshaw. *Principles of real analysis.* San Diego: Academic Press, 3d. edition, 1998.

[3] C. D. Aliprantis and K. C. Border. *Infinite Dimensional Analysis - A Hitchhiker's Guide.* Springer, 3d edition, 2006.

[4] K. J. Arrow. *Aspects of the theory of risk-bearing.* Helsinki, 1965.

[5] P. Artzner, F. Delbaen, J. M. Eber, and D. Heath. Coherent measures of risk. *Mathematical Finance*, 9:203–228, 1999.

[6] P. Artzner, F. Delbaen, and D. Heath. Thinking coherently. *Risk*, 10:203–228, November 1997.

[7] P. Artzner, F. Delbaen, D. Heath, and H. Ku. Coherent multiperiod risk adjusted values and bellman's principle. *Annals of Operations Research*, 152(1):5–22, 2007.

[8] F. Black and M. Scholes. The pricing of options and corporate liabilities. *Journal of Political Economics*, 81:637–654, 1973.

[9] K. Detlefsen and G. Scandolo. Conditional and dynamic convex risk measures. Sfb 649 discussion papers sfb649dp2005-006, Humboldt University, Berlin, Germany, http://sfb649.wiwi.hu-berlin.de/papers/pdf/SFB649DP2005-006.pdf, 2005.

[10] H. Föllmer and A. Schied. *Stochastic Finance - An Introduction in Discrete Time.* Studies in Mathematics. de Gruyter, 2002.

[11] M. Frittelli and G. E. Rosazza. Putting order in risk measures. *Journal of Banking and Finance*, 26, 2002.

[12] M. Frittelli and G. Scandolo. Risk measures and capital requirements for processes. *Mathematical Finance*, 16:589–612, 2006.

[13] S. Jacka and A. Berkaoui. On representing and hedging claims for coherent risk measures. Preprint, http://www2.warwick.ac.uk/fac/sci/statistics/staff/academic/jacka/representfin.pdf, August 2008.

[14] P. Krokhmal. Higher moment coherent risk measures. Working Paper, 2007.

[15] H. M. Markowitz. Portfolio selection. *The Journal of Finance*, March 1952.

[16] H. M. Markowitz. Foundations of portfolio theory. Nobel Lecture, December 1990.

[17] R. Merton. Theory of rational option pricing. *Bell Journal of Economic and Management Sciences*, 4:141–183, 1973.

[18] J.-J. Moreau. *Fonctions convexes en dualité.* Univ. Montpellier, 1962.

[19] A. Nemeth. On the subdifferentiability of convex operators. *J. London Math. Soc.*, 2(34):552–558, 1986.

[20] K. Nikodem. Continuity properties of convex-type set-valued mappings. *Journal of Inequalities in Pure and Applied Mathematics*, 4(3):Article 52 (http://jipam.vu.edu.au/), 2003.

[21] N. S. Papageorgiou. Nonsmooth analysis on partially ordered vector spaces: Part 1 - convex case. *Pacific Journal of Mathematics*, 107(2):403–458, 1983.

[22] G. Pflug. Subdifferential representation of risk measures. *Mathematical Programming*, 108 (2-3):339–354, 2007.

[23] G. Pflug and W. Römisch. *Modeling, Measuring and Managing Risk*. World Scientific, August 2007.

[24] G. Pflug and W. Römisch. Multi-period risk functionals. Working Paper, 2008.

[25] G. Pflug and A. Ruszczyński. Measuring risk for income streams. *Computational Optimization and Applications*, 32:161–178, 2005.

[26] J. W. Pratt. Risk aversion in the small and in the large. *Econometrica*, 1964.

[27] R. Rockafellar. *Convex Analysis*. Princeton University Press, 1970.

[28] R. Rockafellar. Conjugate duality and optimization. In *CBMS-NSF Regional Conference Series in Applied Mathematics*, volume 16. SIAM, Philadelphia, 1974.

[29] R. Rockafellar and S. Uryasev. Optimization of conditional value-at-risk. *Journal of Risk*, 2:21–41, 2000.

[30] R. Rockafellar and R. J.-B. Wets. Stochastic convex programming: basic duality. *Pacific Journal of Mathematics*, 62(1):173–195, 1977.

[31] A. Ruszczyński and A. Shapiro. Optimization of risk measures. In G. Calafiore and F. Dabbene, editors, *Probabilistic and Randomized Methods for Design under Uncertainty*, pages 17–158. Springer, 2005.

[32] A. Ruszczyński and A. Shapiro. Conditional risk mappings. *Mathematics of Operations Research*, 31(3):544–561, August 2006.

[33] A. Ruszczyński and A. Shapiro. Optimization of convex risk functions. *Mathematics of Operations Research*, 31, 3:433–452, August 2006.

[34] J. Von Neumann and O. Morgenstern. *Theory of Games and Economic Behavior*. Princeton University Press, 3d edition, 1953.

[35] S. Weber. Distribution-invariant dynamic risk measures, information and dynamic consistency. *Mathematical finance*, 16(2):419–441, 2006.

Anhang

Zusammenfassung

Conditional Acceptability Mappings beschreiben die Akzeptanz von Zufallsvariablen bedingt auf die verfügbare nichttriviale Information. Sie können als Abbildungen von Räumen $\boldsymbol{L}_p(\Omega, \mathcal{F}, \mu)$ nach Räumen $\boldsymbol{L}_{p'}(\Omega, \mathcal{F}_1, \mu)$ modelliert werden, wobei die $\sigma-$Algebra $\mathcal{F}_1$ die zur Bewertung verfügbare Information beschreibt. Zusätzlich wird von derartigen Abbildungen Konkavität, Translationsequivarianz und Monotonie gefordert.

Basierend auf den Ordnungseigenschaften - insbesondere der Ordnungsvollständigkeit - von $\boldsymbol{L}_p(\Omega, \mathcal{F}, \mu)$-Räumen, die als Banachverbände interpretierbar sind, werden das Superdifferential und die Fenchel-Moreau Konjugierte von konkaven bedingten Abbildungen definiert, sowie deren Eigenschaften untersucht. Die konsequente Nutzung der fast sicheren Halbordnung zu diesem Zweck ist neu in der Literatur und vereinfacht im Folgenden Argumentation und Beweisführung bei gleichzeitiger Rücksichtnahme auf alle Bedenken hinsichtlich Stetigkeit, Integrierbarkeit und Meßbarkeit der resultierenden Supergradienten und Konjugierten.

Abschließend werden die Ergebnisse über bedingte Abbildungen herangezogen, um Aussagen über jene bisher in der Literatur beschriebenen Ansätze für mehrperiodige Akzeptanzmaße zu gewinnen, die sich in ihrer Konstruktion auf Conditional Acceptability Mappings stützen: SEC-Funktionale und verkettete Acceptability Mappings. Insbesondere wird für letztere eine Kettenregel für das Superdifferential, sowie eine einfache Darstellung der konjugierten Abbildung hergeleitet.

Abstract

Conditional Acceptability Mappings quantify the degree of desirability of random variables modeling financial returns, accounting for available, non-trivial information. They are defined as mappings from spaces $\boldsymbol{L}_p(\Omega, \mathcal{F}, \mu)$ to spaces $\boldsymbol{L}_{p'}(\Omega, \mathcal{F}_1, \mu)$, where the $\sigma-$algebra $\mathcal{F}_1 \subseteq \mathcal{F}$ describes the available information. Additionally, such mappings have to be concave, translation- equivariant and monotonically increasing.

Based on the order characteristics (in particular the order completeness) of $\boldsymbol{L}_p(\Omega, \mathcal{F}, \mu)$-spaces, superdifferentials and concave conjugates for conditional acceptability mappings are defined and analyzed. The novelty of this work is that the almost sure partial order is consequently used for this purpose, which results in simpler definitions and proofs, but also accounts for all requirements concerning continuity, integrability and measurability of the supergradients and conjugates.

Furthermore, the results about conditional mappings are used to show properties of multiperiod acceptability functionals that are based on conditional acceptability mappings, such as SEC-functionals and acceptability compositions. A chain rule for superdifferentials as well as the conjugate of multiperiod functionals and their properties are derived.

Lebenslauf

Raimund Kovacevic, *4.4.1968 in Wien.

- **Schule**: Volksschule (1974-1978) und Gymnasium (1978-1986, Akademisches Gymnasium) in Wien
- **Studium:**
  - Konservatorium der Stadt Wien (1982-1995): Violine und Lehramt Violine
  - Universität Wien (1995-1999): Statistik, Mag. rer. soc. oec.
  - Universität Wien (2006-2008): Doktoratsstudium der Sozial und Wirtschaftswissenschaften (Statistik)
- **Berufliche und wissenschaftliche Tätigkeiten:**
  - **Institut für höhere Studien**, Wien (1998-1999): Stipendiat der Abteilung Soziologie
  - **Siemens AG Austria** (1999-2000): Anwendungsentwickler und Berater - Kraftwerkseinsatzplanung
  - **Feilmeier & Junker GesmbH.**, Wien, München (2000-2004): Senior Consultant - Asset/Liability Management
  - **Technische Universität Wien**, Institut für Wirtschaftsmathematik (2004-2006): Projektassistent - Entwicklung von Risikomanagementsystemen für Hedgefondsstrategien in enger Zusammenarbeit mit den Firmen Financial Soft Computing und Siemens.
  - **Financial Soft Computing/Siemens** (2007-2008): Risikomanager, Researcher Portfolio Optimization
  - **Universität Wien**, Institut für Statistik und Decision Support Systems (2008 -): Projektassistent - Multiperiodische Risikomaße in der Portfoliooptimierung.
- **Veröffentlichungen:**
  - Kovacevic, Willomitzer : Überlegungen zur Effizienz der Veranlagungsportefeuilles von Lebensversicherern, *Versicherungswirtschaft*, 56 (2001).

- Kovacevic, Willomitzer: Ein einfaches Optimierungsmodell für Erlebensversicherungen, *Verlautbarungen der Aktuarvereinigung Österreichs*, 11 (2001), S 59-67.
- Kovacevic, Willomitzer: Zur Bewertung der Stornooption für Lebensversicherungsverträge, *Versicherungswirtschaft*, 7 (2005).
- Brandner, Kovacevic, Anwendung der Zinsstrukturkurve bei Bewertungen von Sozialkapital nach internationalen Standards, *Zeitschrift für Recht und Rechnungswesen*, 4 (2007).

* Konferenzbeiträge:
  - Dragoti-Cela, Haumer, Kovacevic, Accounting for estimation risk in dynamic optimal portfolio selection, EURO 2006.
  - Dragoti-Cela, Kovacevic, Kührer, Mean-variance optimization with CVaR constraints for a multi-sector portfolio selection problem, EURO 2007.