Springer Series in Operations Research and Financial Engineering

Darinka Dentcheva
Andrzej Ruszczyński

# Risk-Averse Optimization and Control

## Theory and Methods

Springer

# Springer Series in Operations Research and Financial Engineering

The Springer Series in Operations Research and Financial Engineering publishes monographs and textbooks on important topics in theory and practice of Operations Research, Management Science, and Financial Engineering. The Series is distinguished by high standards in content and exposition, and special attention to timely or emerging practice in industry, business, and government. Subject areas include:

Linear, integer and non-linear programming including applications; dynamic programming and stochastic control; interior point methods; multi-objective optimization; Supply chain management, including inventory control, logistics, planning and scheduling; Game theory Risk management and risk analysis, including actuarial science and insurance mathematics; Queuing models, point processes, extreme value theory, and heavy-tailed phenomena; Networked systems, including telecommunication, transportation, and many others; Quantitative finance: portfolio modeling, options, and derivative securities; Revenue management and quantitative marketing Innovative statistical applications such as detection and inference in very large and/or high dimensional data streams; Computational economics.

Darinka Dentcheva • Andrzej Ruszczyński

# Risk-Averse Optimization and Control

## Theory and Methods

Springer

Darinka Dentcheva
Department of Mathematical Sciences
Stevens Institute of Technology
Hoboken, NJ, USA

Andrzej Ruszczyński
Department of Management Science
and Information Systems
Rutgers University
Piscataway, NJ, USA

ISSN 1431-8598 ISSN 2197-1773 (electronic)
Springer Series in Operations Research and Financial Engineering
ISBN 978-3-031-57987-5 ISBN 978-3-031-57988-2 (eBook)
https://doi.org/10.1007/978-3-031-57988-2

Mathematics Subject Classification: 90C40, 90-08, 49J52, 91G70, 49K27, 91B16, 90C15



This Springer imprint is published by the registered company Springer Nature Switzerland AG
The registered company address is: Gewerbestrasse 11, 6330 Cham, Switzerland

# Preface

This book is the result of our work during the last 25 years, dedicated to the modeling and optimization of risk in various contexts. Much of its content is based on our research during this period, and some results or entire sections are original. Our intention was to present a unified and cohesive theory—a very challenging task given the rapid and vigorous development of the field. As a result, the outcome is certainly far from perfect.

With the idea of reaching a broad audience, we have presented some of the results at a level of generality which is sufficient to cover major applications, but not necessarily in the most general versions. We hope that the book will be useful for both advanced graduate students and researchers in the areas such as optimization under uncertainty, statistics, data science, and other fields where risk and stochastic optimization are relevant.

We would like to express our gratitude for the support of the National Science Foundation, Air Force Office of Scientific Research, and the Office of Naval Research through research grants for various topics covered in this book.

Hoboken, NJ, USA — Darinka Dentcheva
Piscataway, NJ, USA — Andrzej Ruszczyński
January 2024

# Introduction

Optimization is an established area of applied mathematics that plays a significant role in engineering, economics, statistics, business, and many other areas. In the simplest possible formulation, an optimization problem is to find an element $\hat{x}$ in some subset $\mathcal{X}$ of a space $\mathfrak{X}$, such that the value of a function $f : \mathfrak{X} \to \mathbb{R}$ at the point $\hat{x}$ is the smallest (or the largest) among all possible values of $f(x)$ at points $x \in \mathcal{X}$. For different classes of spaces $\mathfrak{X}$, sets $\mathcal{X}$, and functions $f(\cdot)$, the theory provides us with the characterizations of the optimal solutions and numerical methods for their identification.

The fundamental assumption behind the theory and methods of optimization is the existence of the function $f(\cdot)$, called the objective function, that allows us to compare the decisions and choose the best one. Such functions are evident when the optimization problem has a clear mathematical or physical nature, such as finding the shortest path, the minimum mass, the lowest energy, etc. However, in many applications, the determination of such a function is not straightforward. This is especially challenging when the outcomes of decisions are uncertain. Such situations arise in most decision problems in which anticipation of future events is essential.

Several ways of modeling uncertain outcomes exist, such as stochastic models, set-valued operators, scenario representations, etc. In this book, we focus exclusively on stochastic models of uncertainty. We assume, in the simplest case, that a decision $x \in \mathfrak{X}$ results in a random outcome $Z(x)$ which is an element of a space $\mathcal{Z}$ of random variables or random vectors, such as the space of random variables having a finite expected value, or the space of random variables with a finite variance, or higher moments. In such situations, it is difficult to directly compare decisions. For example, in an investment portfolio problem, the decision $x$ may be a finite-dimensional vector representing the investments in various assets, and $Z(x)$ may represent the random profit or return rate in a fixed period following the investment. In this situation, when making the decision, we cannot directly compare $Z(x)$ and $Z(y)$ for two different values ($x$

and $y$) of the decision vector, because this would require the anticipation of the future returns. Likewise, if $x$ represents the resources deployed in various oil exploration projects, and $Z(x)$ is the vector of values of the projects a year from now, it is extremely difficult to rank different investment strategies involving many locations with various prospects, and many currencies. Similar difficulties arise in public health decisions, logistics, marketing campaigns, and many other applications.

A branch of the optimization theory, known as stochastic optimization or stochastic programming, studies problems involving random outcomes, such as the examples mentioned above. A large portion of the theory and methods developed in this area are concerned with the optimization of the expected values of the random outcomes. However, in many problems of practical relevance, the use of the expected value as the major measure guiding our decisions may not be appropriate and would not be accepted by the practitioners in the corresponding fields. For example, in the portfolio problem mentioned above, the expected value criterion would suggest concentrating all investments on the one or few assets that have the best prospects. In public health policy problems, the use of the expected value may lead to neglecting the needs of a particular group of patients, or to the promotion of actions involving a significant probability of highly undesirable effects. In oil exploration, all funds would go to one project. In these and other applications, an additional aspect must be considered: risk.

The concept of risk arises in most decision problems under uncertainty. In the simplest form, it is reflected by the first question of a person offered an investment opportunity: "How much can be lost?" The Latin proverb "Venture not all in one ship"[†] already contained a recipe for risk control. Today, we can go much further than that in the description, analysis, and optimization of risk. Specific application areas developed their own methods to formalize and treat risk. For example, finance uses the concept of the value at risk, a refined version of the amount that can be lost, involving a low probability level. In the portfolio theory, the variance of the outcome is used as the second measure, in addition to the expected value, of the quality of the decision. This has further been extended to other mean–risk models and led to the introduction of the theory of measures of risk. In engineering, the reliability theory provides tools to control the risk of construction. In economics, the theory of utility and stochastic dominance are recognized methods to model risk-averse preferences.

In this book, we aim to provide a unified view on the ways to model, analyze, and optimize risk.

We start from the utility theory in Chap. 1 with the goal of presenting a concise and unified view on the classical expected utility and dual utility models. We also introduce the concept of risk aversion in this context.

---

[†] *Uni navi ne committas omnia.*

In Chap. 2, we present the modern theory of measures of risk, provide their main theoretical properties, and discuss many examples. In this chapter, we also present the theory of systemic risk, which is essential for the operation of distributed complex systems. Our treatment is based on the understanding that we address the risk associated with vector-valued outcomes. An essential part of this chapter is the theory of risk forms, which are risk measures associated with a variable probability measure.

Chapter 3 develops the theoretical foundation of optimization problems involving measures of risk. We provide a thorough analysis of the compositions of measures of risk with various classes of stochastic operators, involving non-convex and non-differentiable operators. This is followed by the derivation of the optimality conditions and duality theory. Finally, we present new stochastic subgradient methods with special attention to non-convex and non-smooth problems arising in modern applications.

Chapter 4 discusses dynamic risk models and dynamic risk optimization. When our decisions are made in successive periods and random outcomes are observed in these periods, the question of evaluating the risk of a sequence of outcomes becomes relevant. Subtle issues of time consistency of such an evaluation led to the theory of dynamic measures of risk. We provide a modern view on this theory, use it to construct risk-averse dynamic optimization problems, and develop optimality conditions for these problems. Our techniques lead to new results even in the expected-value case. Finally, we present specialized decomposition methods for dynamic risk optimization.

In Chap. 5, we present another approach to the risk control in decision problems, which is based on stochastic orders. For a major part of it, we focus on the notion of stochastic dominance of general order, presenting its properties that are essential to risk-averse optimization. The main portion of the chapter contains an analysis of optimization problems using stochastic dominance relations as constraints. We present optimality conditions and duality theory in various forms for convex and non-convex forms of several problem formulations. These results elucidate deeper relations of stochastic dominance to utility functions, distortions (or rank-dependent utility), and risk measures.

Chapter 6 addresses the comparison of random vectors and finite sequences which can be integrated into risk-averse decision problems. Vector-valued outcomes arise in the context of systemic risk, multi-objective optimization under uncertainty, or in any problem involving high-dimensional risks. The question of constraints of distributions in a sequential decision problem is an important one as already mentioned as the decision maker needs to mind the time consistency of the model. This chapter introduces the notion of a time-consistent stochastic order for sequences and provides a theoretical foundation for constructing time-consistent multistage stochastic programming problems with sequential stochastic dominance constraints.

In Chap. 7, we present numerical methods for solving optimization problems with stochastic order constraints with a focus on the second-order stochastic dominance relation in both univariate and multivariate settings. We start with reformulations of the static problems that allow the application of available commercial solvers for optimization problems. This approach is suitable for relatively small problems, in which the probability space consists of a small number of simple events. The main focus of the chapter is on methods for solving the respective optimization problems when dealing with general probability spaces. In this case, we construct sequential approximations of the stochastic dominance constraints by inequalities, which we call event cuts. Several approximations are presented for the problems formulated in Chaps. 5 and 6 and their convergence is analyzed. Methods for evaluation of the quality of the approximation and the validity of the dominance relation at the iterates are presented as well.

Finally, Chap. 8 considers risk models for Markov dynamical systems and risk-averse control of such systems. This is a very special class of dynamic models where a deep theory and dedicated methods can be developed, which fully exploit the Markovian structure of the system. The key concept is that of a Markov risk measure, which allows us to reduce the dynamic risk evaluation into a sequence of static evaluations with transition risk mappings. This leads to specialized dynamic programming equations in the finite-horizon, infinite-horizon, absorbing, and partially observable cases.

Most of the book contains the results of our own research conducted over a period of 25 years, and several parts of the book are entirely new. We discuss the relation to earlier research and point to relevant reading in the Bibliographical Remarks at the end of the book.

# Contents

# Chapter 1
# Elements of the Utility Theory

The utility theory is concerned with analyzing the existence and the forms of numerical representations of preference relations. We start our presentation from abstract preference relations in a space $\mathcal{X}$, which we call the *prospect space*. Our considerations in this section are rather abstract, but in the applications to risk-averse decision-making that we discuss in this book, three special cases of $\mathcal{X}$ are particularly important.

The first case is the space $\mathcal{P}(\mathcal{S})$ of probability measures on a Polish space $\mathcal{S}$ equipped with the $\sigma$-algebra $\mathcal{B}$ of Borel sets (frequently, $\mathcal{S} = \mathbb{R}^n$). We will also consider its subspaces of measures satisfying additional integrability conditions: the existence of finite moments or order $p \in [1, \infty]$. In this context, we are interested in the following question: can we derive a functional $U : \mathcal{P}(\mathcal{S}) \to \mathbb{R}$ such that we prefer $\mu \in \mathcal{P}(\mathcal{S})$ over $\nu \in \mathcal{P}(\mathcal{S})$ if and only if $U(\mu) < U(\nu)$. Here, and later in this chapter, we adopt the convention that smaller values of $U(\cdot)$ are associated with preferred prospects, and thus $U(\mu)$ can be interpreted as the "fair price" of $\mu$.

The second case of $\mathcal{X}$ is the space of random vectors. For a probability space $(\Omega, \mathcal{F}, P)$, we consider the space $\mathcal{L}_0(\Omega, \mathcal{F}, P; \mathbb{R}^n)$ of random vectors in $\mathbb{R}^n$, or its subspaces $\mathcal{L}_p(\Omega, \mathcal{F}, P; \mathbb{R}^n)$ of random vectors having finite moments of order $p \in [1, \infty]$. Again, can we construct $U(\cdot)$ such that $Z$ is "better than" $V$ (both in $\mathcal{X}$) if and only if $U(Z) < U(V)$?

The third case is the space $\mathcal{Q}_{\mathrm{b}}$ of the quantile functions of scalar and bounded random variables. While one might reduce it to the first case with $\mathcal{S} = \mathbb{R}$, it is convenient to consider it separately, because of future applications to measures of risk.



D. Dentcheva, A. Ruszczyński, *Risk-Averse Optimization and Control*, Springer Series in Operations Research and Financial Engineering,
https://doi.org/10.1007/978-3-031-57988-2_1

## 1.1 Preference Relations

We assume that a preference relation in the prospect space $\mathcal{X}$ is defined by a certain *total preorder*, that is, a binary relation $\trianglelefteq$ on $\mathcal{X}$ satisfying the following two conditions:

**Transitivity:** For all $z, v, w \in \mathcal{X}$, if $z \trianglelefteq v$ and $v \trianglelefteq w$, then $z \trianglelefteq w$; and
**Completeness:** For all $z, v \in \mathcal{X}$, either $z \trianglelefteq v$, or $v \trianglelefteq z$, or both are true.

It follows from the completeness condition that the preorder is *reflexive*, that is, $z \trianglelefteq z$ for all $z \in \mathcal{X}$. The corresponding *indifference relation* $\sim$ is defined in a usual way: $z \sim v$, if $z \trianglelefteq v$ and $v \trianglelefteq z$. It is an equivalence relation. We say that $z$ is *strictly preferred* over $v$ and write it $z \triangleleft v$, if $z \trianglelefteq v$, and $v \ntrianglelefteq z$.

In decision problems, it is convenient to represent prospects by numbers, which summarize their usefulness.

**Definition 1.1.** A functional $U : \mathcal{X} \to \mathbb{R}$ is a *numerical representation* of the preference relation $\trianglelefteq$ on $\mathcal{X}$, if

$$z \triangleleft v \iff U(z) < U(v).$$

In our presentation, continuous representations play a fundamental role. To speak about continuity, we need to assume that the space of prospects $\mathcal{X}$ is a topological space.

The special cases of $\mathcal{X}$ mentioned at the beginning of this chapter can be equipped with appropriate topologies. The space $\mathcal{P}(\mathcal{S})$ of probability measures on $(\mathcal{S}, \mathcal{B})$ can be equipped with the topology of weak convergence. In $\mathcal{L}_0(\Omega, \mathcal{F}, P; \mathbb{R}^n)$ we may consider the topology of almost sure convergence, or the topology of convergence in probability. The spaces $\mathcal{L}_p(\Omega, \mathcal{F}, P; \mathbb{R}^n)$, for $p \in [1, \infty]$, as Banach spaces, may be equipped with the strong (norm) topology or the weak topology. The space $\mathcal{Q}_b$ can be considered as a subspace of the space of bounded functions and equipped with the topology of uniform convergence.

To guarantee the existence of a continuous numerical representation, we need to make additional assumptions about the preference relation and the space itself. The first one is a fundamental condition, which we use many times in our considerations.

**Definition 1.2.** The preference relation $\trianglelefteq$ on $\mathcal{X}$ is *continuous*, if for every $z \in \mathcal{X}$ the sets $\{v \in \mathcal{X} : v \trianglelefteq z\}$ and $\{v \in \mathcal{X} : z \trianglelefteq v\}$ are closed.

A classical result in topology states that a continuous total preorder on a separable and connected topological space has a continuous numerical representation.

Furthermore, if $U : \mathcal{X} \to \mathbb{R}$ is a numerical representation of a preorder $\trianglelefteq$, then for any strictly increasing function $\varphi : \mathbb{R} \to \mathbb{R}$ the functional $\varphi(U(\cdot))$ is also a numerical representation of the preorder $\trianglelefteq$. If $\varphi(\cdot)$ and $U(\cdot)$ are continuous, so is their composition. Conversely, for any two numerical representations $U(\cdot)$ and $\widetilde{U}(\cdot)$ of $\trianglelefteq$, a strictly increasing function $\varphi : U(\mathcal{X}) \to \mathbb{R}$ exists, such that $\widetilde{U}(\cdot) = \varphi(U(\cdot))$.

For specific prospect spaces and under additional conditions, we can say more about the structure of a numerical representation of a preorder $\trianglelefteq$. This is our main objective in this chapter.

## 1.2 Expected Utility Theory

The expected utility theory is concerned with comparing probability measures.

### *1.2.1 The Prospect Space of Probability Measures*

Given a Polish space $\mathcal{S}$, equipped with its $\sigma$-algebra $\mathcal{B}$ of Borel sets, we consider the prospect space $\mathcal{X} = \mathcal{P}(\mathcal{S})$ of probability measures on $\mathcal{S}$. For most applications discussed in this book, the finite-dimensional space $\mathcal{S} = \mathbb{R}^d$ is sufficient, but we present the theory in more general settings because they does not require much additional effort and broaden the scope of applications.

We assume that the preference relation $\trianglelefteq$ satisfies two additional conditions:

**Independence Axiom:** For all $\mu$, $\nu$, and $\lambda$ in $\mathcal{P}(\mathcal{S})$ one has

$$\mu \vartriangleleft \nu \implies \alpha\mu + (1-\alpha)\lambda \vartriangleleft \alpha\nu + (1-\alpha)\lambda, \quad \forall\, \alpha \in (0,1);$$

**Archimedean Axiom:** For all $\mu$, $\nu$, and $\lambda$ in $\mathcal{P}(\mathcal{S})$, satisfying the relations

$$\mu \vartriangleleft \nu \vartriangleleft \lambda,$$

there exist $\alpha, \beta \in (0,1)$ such that

$$\alpha\mu + (1-\alpha)\lambda \vartriangleleft \nu \vartriangleleft \beta\mu + (1-\beta)\lambda.$$

We can derive the following properties of a preorder satisfying these axioms.

**Lemma 1.3.** *Suppose a total preorder $\trianglelefteq$ on $\mathcal{P}(\mathcal{S})$ satisfies the independence axiom. Then for every $\mu \in \mathcal{P}(\mathcal{S})$ the indifference set $\{\nu \in \mathcal{P}(\mathcal{S}) : \nu \sim \mu\}$ is convex.*

*Proof.* Let $\nu \sim \mu$ and $\lambda \sim \mu$. Suppose $(1-\alpha)\nu + \alpha\lambda \vartriangleleft \nu$ for some $\alpha \in (0,1)$. Then also $(1-\alpha)\nu + \alpha\lambda \vartriangleleft \lambda$. Using the independence axiom with these two relations, we obtain a contradiction as follows:

$$(1-\alpha)\nu + \alpha\lambda = (1-\alpha)\big[(1-\alpha)\nu + \alpha\lambda\big] + \alpha\big[(1-\alpha)\nu + \alpha\lambda\big]$$
$$\lhd (1-\alpha)\nu + \alpha\big[(1-\alpha)\nu + \alpha\lambda\big] \lhd (1-\alpha)\nu + \alpha\lambda.$$

The case when $\nu \lhd (1-\alpha)\nu + \alpha\lambda$ is excluded similarly. We conclude that $(1-\alpha)\nu + \alpha\lambda \sim \mu$, for all $\alpha \in (0,1)$. ∎

**Lemma 1.4.** *Suppose a total preorder $\unlhd$ on $\mathcal{P}(\mathcal{S})$ satisfies the independence and Archimedean axioms. Then for all $\mu, \nu \in \mathcal{P}(\mathcal{S})$, satisfying the relation $\mu \lhd \nu$, and for all $\lambda \in \mathcal{P}(\mathcal{S})$, there exists $\bar{\alpha} > 0$ such that*

$$(1-\alpha)\mu + \alpha\lambda \lhd \nu \quad \textit{and} \quad \mu \lhd (1-\alpha)\nu + \alpha\lambda, \quad \forall\, \alpha \in [0,\bar{\alpha}]. \tag{1.1}$$

*Proof.* We focus on the left relation in (1.1) and consider three cases.
*Case 1:* $\nu \lhd \lambda$. The left relation in (1.1) is true for some $\bar{\alpha} \in (0,1)$, owing to the Archimedean axiom. If $\alpha \in (0,\bar{\alpha})$ then for $\beta = \alpha/\bar{\alpha} \in (0,1)$ the independence axiom yields

$$(1-\alpha)\mu + \alpha\lambda = (1-\beta)\mu + \beta\big[(1-\bar{\alpha})\mu + \bar{\alpha}\lambda\big] \lhd (1-\beta)\mu + \beta\nu \lhd \nu.$$

*Case 2:* $\lambda \lhd \nu$. Applying the independence axiom twice, we obtain

$$(1-\alpha)\mu + \alpha\lambda \lhd (1-\alpha)\nu + \alpha\lambda \lhd \nu, \quad \forall\, \alpha \in (0,1).$$

*Case 3:* $\lambda \sim \nu$. By virtue of Lemma 1.3, $(1-\alpha)\nu + \alpha\lambda \sim \nu$ for all $\alpha \in (0,1)$, and the left relation in (1.1) follows from the independence axiom.

The right relation in (1.1) is proved analogously. ∎

### 1.2.2 *Affine Numerical Representation*

The set $\mathcal{P}(\mathcal{S})$ is a convex subset of the vector space $\mathcal{M}(\mathcal{S})$ of signed regular finite measures on $\mathcal{S}$. It is also convenient for our derivations to consider the linear subspace $\mathcal{M}_0(\mathcal{S})$ of measures $\mu \in \mathcal{M}(\mathcal{S})$ such that $\mu(\mathcal{S}) = 0$.

**Theorem 1.5.** *Suppose the total preorder $\unlhd$ on $\mathcal{P}(\mathcal{S})$ satisfies the independence and Archimedean axioms. Then an affine numerical representation of $\unlhd$ on $\mathcal{P}(\mathcal{S})$ exists.*

*Proof.* In the space $\mathcal{M}_0(\mathcal{S})$, define the set

$$C_0 = \{\mu - \nu : \mu \in \mathcal{P}(\mathcal{S}),\ \nu \in \mathcal{P}(\mathcal{S}),\ \mu \lhd \nu\}.$$

Consider two arbitrary points $\vartheta$ and $\varkappa$ in $C_0$, that is,

$$\vartheta = \mu - \nu, \qquad \mu, \nu \in \mathcal{P}(\mathcal{S}), \qquad \mu \lhd \nu,$$
$$\varkappa = \lambda - \sigma, \qquad \lambda, \sigma \in \mathcal{P}(\mathcal{S}), \qquad \lambda \lhd \sigma.$$

For every $\alpha \in (0, 1)$, using the independence axiom twice, we obtain

$$\alpha\mu + (1-\alpha)\lambda \lhd \alpha\nu + (1-\alpha)\lambda \lhd \alpha\nu + (1-\alpha)\sigma.$$

Then,

$$\alpha\vartheta + (1-\alpha)\varkappa = \big[\alpha\mu + (1-\alpha)\lambda\big] - \big[\alpha\nu + (1-\alpha)\sigma\big]$$

is an element of $C_0$, which proves that $C_0$ is convex.

Define $C = \{\gamma\vartheta : \vartheta \in C_0,\ \gamma > 0\}$. It is evident that $C$ is a convex cone, that is, for all $\vartheta, \varkappa \in C$, and all $\alpha > 0$ and $\beta > 0$ we have $\alpha\vartheta + \beta\varkappa \in C$. Moreover, $C \subset \mathcal{M}_0$.

We shall prove that the algebraic interior of $C$ in $\mathcal{M}_0$, denoted $\text{cor}(C)$, is nonempty and that in fact $C = \text{cor}(C)$. Consider any $\vartheta \in C$, an arbitrary nonzero measure $\lambda \in \mathcal{M}_0$, and the ray

$$z(\tau) = \vartheta + \tau\lambda, \quad \tau > 0.$$

Our objective is to show that $z(\tau) \in C$ for a sufficiently small $\tau > 0$. Let $\lambda = \lambda^+ - \lambda^-$ be the Jordan decomposition of $\lambda$. With no loss of generality, we may assume that the direction $\lambda$ is normalized so that $|\lambda| = \lambda^+(\mathcal{S}) + \lambda^-(\mathcal{S}) = 2$. As $\lambda \in \mathcal{M}_0$, we have then $\lambda^+(\mathcal{S}) = \lambda^-(\mathcal{S}) = 1$. Let $\gamma > 0$ be such that the point $\vartheta_0 = \gamma\vartheta \in C_0$. Since $C$ is a cone, $z(\tau) \in C$ if and only if $\gamma z(\tau) \in C$. Setting $t = \gamma\tau$, we reformulate our question as follows:

Does $\vartheta_0 + t\lambda$ belong to $C$ for sufficiently small $t > 0$?

Since $\vartheta_0 \in C_0$, we can represent it as a difference $\vartheta_0 = \mu - \nu$, with $\mu, \nu \in \mathcal{P}(\mathcal{S})$, and $\mu \lhd \nu$. Then

$$\vartheta_0 + t\lambda = \big[(1-t)\mu + t\lambda^+\big] - \big[(1-t)\nu + t\lambda^-\big] + t\vartheta_0. \tag{1.2}$$

Both expressions in brackets are probability measures for $t \in [0, 1]$. By the independence axiom,

$$\mu \lhd \frac{1}{2}\mu + \frac{1}{2}\nu \lhd \nu.$$

According to Lemma 1.4, there exists $t_0 > 0$, such that for all $t \in [0, t_0]$ we also have

$$(1-t)\mu + t\lambda^+ \lhd \frac{1}{2}\mu + \frac{1}{2}\nu \lhd (1-t)\nu + t\lambda^-.$$

This proves that

$$\big[(1-t)\mu + t\lambda^+\big] - \big[(1-t)\nu + t\lambda^-\big] \in C_0, \quad \forall\, t \in [0, t_0].$$

For these values of $t$, the right hand side of (1.2) is a sum of two elements of $C$. As the set $C$ is a convex cone, this sum is an element of $C$ as well. Consequently, $\vartheta + \tau\lambda \in C$ for all $\tau \in [0, t_0/\alpha]$.

Summing up, $C$ is convex, $C = \operatorname{cor}(C)$, and $\mathbb{0} \notin C$, where $\mathbb{0}$ denotes the zero element in $\mathcal{M}_0(\mathcal{S})$. By virtue of the separation theorem for a convex set with a nonempty algebraic interior, the point $\mathbb{0}$ and the set $C$ can be separated strictly: a linear functional $U_0$ on $\mathcal{M}_0(\mathcal{S})$ exists, such that

$$U_0(\vartheta) < 0, \quad \forall\, \vartheta \in C. \tag{1.3}$$

We can extend the linear functional $U_0$ to the whole space $\mathcal{M}(\mathcal{S})$ by choosing a measure $\lambda \in \mathcal{P}(\mathcal{S})$ and setting

$$U(\mu) = U_0\big(\mu - \mu(\mathcal{S})\lambda\big), \quad \mu \in \mathcal{M}(\mathcal{S}).$$

It is linear and coincides with $U_0$ on $\mathcal{M}_0(\mathcal{S})$. The relation (1.3) is equivalent to the following statement: for all $\mu, \nu \in \mathcal{P}(\mathcal{S})$ such that $\mu \lhd \nu$, we have

$$U_0(\mu - \nu) = U(\mu - \nu) = U(\mu) - U(\nu) < 0.$$

It follows that $U$ restricted to $\mathcal{P}(\mathcal{S})$ is the postulated numerical representation of the preorder $\unlhd$. It is affine on $\mathcal{P}(\mathcal{S})$. □

### 1.2.3 *Integral Representation. Utility Functions*

To prove the main result of this section, we assume that the space $\mathcal{M}(\mathcal{S})$ is equipped with the topology of weak convergence of measures. This makes it a complete topological vector space. Recall that the topology of weak convergence is metrizable and $\mathcal{M}(\mathcal{S})$ is a Polish space itself.

**Theorem 1.6.** *Suppose the total preorder $\unlhd$ on $\mathcal{P}(\mathcal{S})$ is continuous and satisfies the independence axiom. Then a continuous and bounded function $u : \mathcal{S} \to \mathbb{R}$ exists, such that the functional*

$$U(\mu) = \int_{\mathcal{S}} u(z)\, \mu(\mathrm{d}z) \tag{1.4}$$

*is a numerical representation of $\unlhd$ on $\mathcal{P}(\mathcal{S})$.*

*Proof.* The continuity of the preorder $\unlhd$ implies the Archimedean axiom. Indeed, the sets $\{\pi \in \mathcal{P}(\mathcal{S}) : \pi \lhd \nu\}$ and $\{\pi \in \mathcal{P}(\mathcal{S}) : \mu \lhd \pi\}$ are open, and the mapping $\alpha \mapsto \alpha\pi + (1-\alpha)\lambda$, $\alpha \in [0, 1]$, is continuous for any $\lambda \in \mathcal{P}(\mathcal{S})$.

By virtue of Theorem 1.5, an affine numerical representation $U : \mathcal{P}(\mathcal{S}) \to \mathbb{R}$ of $\unlhd$ exists.

We shall prove that the functional $U(\cdot)$ is continuous on $\mathcal{P}(\mathcal{S})$, that is, for every $\alpha$ the sets

$$A = \{\mu \in \mathcal{P} : U(\mu) \le \alpha\} \qquad \text{and} \qquad B = \{\mu \in \mathcal{P} : U(\mu) \ge \alpha\}$$

are closed. Since $\mathcal{P}(\mathcal{S})$ is convex and $U(\cdot)$ is affine, the set $U(\mathcal{P})$ is convex. Therefore, for every $\alpha$ one of three cases may occur:

(i) $U(\mu) < \alpha$ for all $\mu \in \mathcal{P}(\mathcal{S})$;

(ii) $U(\mu) > \alpha$ for all $\mu \in \mathcal{P}(\mathcal{S})$;

(iii) $\alpha \in U\big(\mathcal{P}(\mathcal{S})\big)$.

In cases (i) and (ii) there is nothing to prove. In case (iii), let $\nu \in \mathcal{P}(\mathcal{S})$ be such that $U(\nu) = \alpha$. Since $U(\cdot)$ is a numerical representation of the preorder, we have

$$A = \{\mu \in \mathcal{P} : \mu \trianglelefteq \nu\} \qquad \text{and} \qquad B = \{\mu \in \mathcal{P} : \nu \trianglelefteq \mu\}.$$

Both sets are closed due to the continuity of the preorder $\trianglelefteq$. Thus, $U(\cdot)$ in continuous on $\mathcal{P}(\mathcal{S})$.

For $z \in \mathcal{S}$, let $\delta_z$ denote the Dirac measure supported at $z$. Define a function $u : \mathcal{S} \to \mathbb{R}$ as follows:

$$u(z) = U(\delta_z), \quad z \in \mathcal{S}.$$

The function $u(\cdot)$ is continuous, because $z_n \to z$ implies $\delta_{z_n} \xrightarrow{w} \delta_z$, and thus $u(z_n) \to u(z)$. Suppose $u(\cdot)$ is not bounded from above. Then we can find a sequence $\{z_n\}$ such that $u(z_n) > n$ for all $n$. Consider the sequence of probability measures

$$\mu_n = \Big(1 - \frac{1}{\sqrt{n}}\Big)\delta_{z_1} + \frac{1}{\sqrt{n}}\delta_{z_n}.$$

Then $\mu_n \xrightarrow{w} \delta_{z_1}$, but $U(\mu_n) \to +\infty$, as $n \to \infty$. This contradicts the continuity of $U(\cdot)$. The unboundedness from below is excluded in a similar way, and thus $u(\cdot)$ is bounded.

We now verify the representation (1.4). For a discrete distribution

$$\mu = \sum_{n=1}^{N} p_n \delta_{z_n}, \quad \sum_{n=1}^{N} p_n = 1, \quad p_n \ge 0, \tag{1.5}$$

the affinity of $U(\cdot)$ yields

$$U(\mu) = \sum_{n=1}^{N} p_n u(z_n),$$

which is exactly (1.4). Furthermore, the set of discrete distributions of form (1.5) is dense in $\mathcal{P}(\mathcal{S})$, in the topology of weak convergence. By the continuity of $U(\cdot)$, the formula (1.4) is true at all $\mu \in \mathcal{P}$. ∎

The formula (1.4) is referred to as the *expected utility representation*, and $u(\cdot)$ is called the *utility function*.

The utility function in Theorem 1.6 is bounded. If we restrict the prospect space to include only measures satisfying additional integrability conditions, representations with unbounded utility functions may occur.

Let $\psi : \mathcal{S} \to [1, \infty)$ be a continuous function (called the *gauge function* in this context), and let $\mathcal{C}_{\mathrm{b}}^{\psi}(\mathcal{S})$ be the set of functions $f : \mathcal{S} \to \mathbb{R}$, such that $f/\psi \in \mathcal{C}_{\mathrm{b}}(\mathcal{S})$, where $\mathcal{C}_{\mathrm{b}}(\mathcal{S})$ is the space of continuous and bounded functions on $\mathcal{S}$. We can define the space $\mathcal{M}^{\psi}(\mathcal{S})$ of regular signed measures $\mu$, such that

$$\left| \int_{\mathcal{S}} f(z)\, \mu(\mathrm{d}z) \right| < \infty, \quad \forall\, f \in \mathcal{C}_{\mathrm{b}}^{\psi}(\mathcal{S}).$$

Similar to the topology of weak convergence, we can define the topology of $\psi$-weak convergence on $\mathcal{M}^{\psi}(\mathcal{S})$: $\{\mu_n\}_{n\in\mathbb{N}} \xrightarrow{\psi} \mu$ if and only if $\int_{\mathcal{S}} f(z)\, \mu_n(\mathrm{d}z) \to \int_{\mathcal{S}} f(z)\, \mu(\mathrm{d}z)$ for all $f \in \mathcal{C}_{\mathrm{b}}^{\psi}(\mathcal{S})$. All continuity statements will be now made for this topology. We use the symbol $\mathcal{P}^{\psi}(\mathcal{S})$ to denote the set of probability measures in $\mathcal{M}^{\psi}(\mathcal{S})$.

**Theorem 1.7.** *Suppose the total preorder $\trianglelefteq$ on $\mathcal{P}^{\psi}(\mathcal{S})$ is continuous and satisfies the independence axiom. Then a function $u \in \mathcal{C}_{\mathrm{b}}^{\psi}(\mathcal{S})$ exists such that the functional*

$$U(\mu) = \int_{\mathcal{S}} u(z)\, \mu(\mathrm{d}z) \tag{1.6}$$

*is a numerical representation of $\trianglelefteq$ on $\mathcal{P}^{\psi}(\mathcal{S})$.*

*Proof.* The proof is almost identical to the proof of Theorem 1.6, except that we use the continuity with respect to the $\psi$-weak convergence. To verify that $u \in \mathcal{C}_{\mathrm{b}}^{\psi}(\mathcal{S})$, we argue by contradiction again. Suppose $u/\psi$ is not bounded from above. Then we can find a sequence $\{z_n\}$ such that $u(z_n) > n\psi(z_n)$ for all $n$. Consider the sequence of probability measures

$$\mu_n = \Big(1 - \frac{1}{\psi(z_n)\sqrt{n}}\Big)\delta_{z_1} + \frac{1}{\psi(z_n)\sqrt{n}}\delta_{z_n}.$$

Then $\mu_n$ converges $\psi$-weakly to $\delta_{z_1}$, but $U(\mu_n) \to +\infty$, as $n \to \infty$. This contradicts the continuity of $U(\cdot)$. The unboundedness of $u/\psi$ from below is excluded in a similar way and thus $u \in \mathcal{C}_{\mathrm{b}}^{\psi}(\mathcal{S})$. ∎

### *1.2.4 Monotonicity and Risk Aversion*

Suppose a partial order relation $\le$ on $\mathcal{S}$ is defined (our notation is motivated by the applications in which $\mathcal{S}$ is a finite-dimensional vector space). In this case, it makes sense to define the monotonicity of a preference relation $\trianglelefteq$.

**Monotonicity Axiom:** For all $z, v \in \mathcal{S}$ one has

$$z \le v \implies \delta_z \trianglelefteq \delta_v.$$

**Theorem 1.8.** *Suppose the total preorder $\trianglelefteq$ on $\mathcal{P}(\mathcal{S})$ is monotonic, continuous, and satisfies the independence axiom. Then a nondecreasing, continuous and bounded function $u : \mathcal{S} \to \mathbb{R}$ exists, such that the functional* (1.4) *is a numerical representation of $\trianglelefteq$ on $\mathcal{P}(\mathcal{S})$.*

*Proof.* In view of Theorem 1.6, it is sufficient to verify that the function $u(\cdot)$ in (1.4) is monotonic with respect to $\le$. To this end, we consider $z, v \in \mathcal{S}$ such that $z \le v$. By the monotonicity of the order, $u(z) = U(\delta_z) \le U(\delta_v) = u(v)$. ∎

We now focus on the case with the gauge function $\psi_p(z) = 1 + \|z\|^p$, where $p \ge 1$. Then for every $\mu \in \mathcal{P}^{\psi_p}(\mathcal{S})$ and for every $\sigma$-subalgebra $\mathcal{G}$ of $\mathcal{B}$ the conditional expectation $\mathcal{E}_{\mu|\mathcal{G}} : \mathcal{S} \to \mathcal{S}$ is well-defined, as a $\mathcal{G}$-measurable function satisfying the equation

$$\int_G \mathcal{E}_{\mu|\mathcal{G}}(z)\ \mu(\mathrm{d}z) = \int_G z\ \mu(\mathrm{d}z), \quad G \in \mathcal{G}. \tag{1.7}$$

The conditional expectation $\mathcal{E}_{\mu|\mathcal{G}}$ induces a probability measure on $(\mathcal{S}, \mathcal{B})$ as follows

$$\mu_{\mathcal{G}}(A) = \mu\big\{\mathcal{E}_{\mu|\mathcal{G}}^{-1}(A)\big\}, \quad A \in \mathcal{B}.$$

**Definition 1.9.** A preference relation $\trianglelefteq$ on $\mathcal{P}^{\psi_p}(\mathcal{S})$ is *risk-averse*, if $\mu_{\mathcal{G}} \trianglelefteq \mu$, for every $\mu \in \mathcal{P}^{\psi_p}(\mathcal{S})$ and every $\sigma$-subalgebra $\mathcal{G}$ of $\mathcal{B}$.

By choosing $\mathcal{G} = \{\mathcal{S}, \emptyset\}$, we observe that Definition 1.9 implies that $\delta_{\mathcal{E}_\mu} \trianglelefteq \mu$, where $\mathcal{E}_\mu = \int_{\mathcal{S}} z\ \mu(\mathrm{d}z)$ is the expected value of the measure $\mu$.

**Theorem 1.10.** *Suppose a total preorder $\trianglelefteq$ on $\mathcal{P}^{\psi_p}(\mathcal{S})$ is continuous, risk-averse, and satisfies the independence axiom. Then a convex function $u \in C_{\mathrm{b}}^{\psi_p}(\mathcal{S})$ exists such that the functional* (1.6) *is a numerical representation of $\trianglelefteq$ on $\mathcal{P}^{\psi}(\mathcal{S})$.*

*Proof.* In view of Theorem 1.7, we only need to prove the convexity of $u(\cdot)$. By the risk aversion, for every $\mu \in \mathcal{P}^{\psi_p}(\mathcal{S})$ we have

$$u\left(\int_{\mathcal{S}} z \; \mu(\mathrm{d}z)\right) \leq \int_{\mathcal{S}} u(z) \; \mu(\mathrm{d}z).$$

This is Jensen's inequality, which is equivalent to the convexity of $u(\cdot)$. ∎

**Remark 1.11.** It is clear from the proof that the convexity of $u(\cdot)$ could have been obtained by simply assuming that $\delta_{\mathcal{E}_\mu} \trianglelefteq \mu$. The convexity of $u(\cdot)$ would imply risk aversion in the sense of Definition 1.9, by virtue of Jensen's inequality for conditional expectations. Therefore, Definition 1.9 and the requirement that $\delta_{\mathcal{E}_\mu} \trianglelefteq \mu$ are equivalent within the framework of the expected utility theory. Nonetheless, we prefer to leave Definition 1.9 in its full form, because we shall use the concept of risk aversion in connection with other axioms, where such equivalence cannot be derived. ∎

## 1.3 Dual Utility Theory

The dual utility theory is formulated in a more restrictive setting: for the probability distributions on the real line.

### *1.3.1 The Prospect Space of Quantile Functions*

With every probability measure $\mu \in \mathcal{P}(\mathbb{R})$ we associate the distribution function: $F_\mu(\eta) \triangleq \mu\big((-\infty, \eta]\big)$. It is nondecreasing and right-continuous. We can, therefore, define its inverse

$$F_\mu^{-1}(p) \triangleq \inf\{\eta \in \mathbb{R} : F_\mu(\eta) \geq p\}, \quad p \in (0,1).$$

By definition, $F_\mu^{-1}(p)$ is the smallest $p$-quantile of $\mu$. We call $F_\mu^{-1}(\cdot)$ the *quantile function* associated with the probability measure $\mu$. Every quantile function is nondecreasing and left-continuous on the open interval $(0, 1)$. On the other hand, every nondecreasing and left-continuous function $\Phi(\cdot)$ on $(0, 1)$ uniquely defines the following distribution function:

$$F_\mu(\eta) = \varphi^{-1}(\eta) \triangleq \sup\{p \in (0,1) : \Phi(p) \leq \eta\}, \quad \eta \in \mathbb{R}.$$

It is nondecreasing, right-continuous, and thus corresponds to a certain probability measure $\mu \in \mathcal{P}(\mathbb{R})$.

*The set $\mathcal{Q}$ of all nondecreasing and left-continuous functions on the interval* $(0, 1)$ *will be our prospect space*. It is evident that $\mathcal{Q}$ is a convex cone in the vector space $\mathcal{L}_0(0, 1)$ of all Lebesgue measurable functions on the interval $(0, 1)$.

We assume that the preference relation $\trianglelefteq$ on $\mathcal{Q}$ is a total preorder and satisfies the following two conditions.

**Dual Independence Axiom:** For all $\Phi$, $\Psi$, and $\Upsilon$ in $\mathcal{Q}$ one has

$$\Phi \lhd \Psi \implies \alpha\Phi + (1-\alpha)\Upsilon \lhd \alpha\Psi + (1-\alpha)\Upsilon, \quad \forall\, \alpha \in (0, 1);$$

**Dual Archimedean Axiom:** For all $\Phi$, $\Psi$, and $\Upsilon$ in $\mathcal{Q}$, satisfying the relations

$$\Phi \lhd \Psi \lhd \Upsilon,$$

there exist $\alpha, \beta \in (0, 1)$ such that

$$\alpha\Phi + (1-\alpha)\Upsilon \lhd \Psi \lhd \beta\Phi + (1-\beta)\Upsilon.$$

Similar to Lemmas 1.3 and 1.4, we can easily derive the following properties of a preorder satisfying the dual axioms.

**Lemma 1.12.** *Suppose a total preorder $\trianglelefteq$ on $\mathcal{Q}$ satisfies the dual independence axiom. Then for every $\Phi \in \mathcal{Q}$ the indifference set $\{\Psi \in \mathcal{Q} : \Psi \sim \Phi\}$ is convex.*

**Lemma 1.13.** *Suppose a total preorder $\trianglelefteq$ on $\mathcal{Q}$ satisfies the dual independence and Archimedean axioms. Then for all $\Phi, \Psi \in \mathcal{Q}$, satisfying the relation $\Phi \lhd \Psi$, and for all $\Upsilon \in \mathcal{Q}$, there exists $\bar{\alpha} > 0$ such that*

$$(1-\alpha)\Phi + \alpha\Upsilon \lhd \Psi \quad \textit{and} \quad \Phi \lhd (1-\alpha)\Psi + \alpha\Upsilon, \quad \forall\, \alpha \in [0, \bar{\alpha}]. \tag{1.8}$$

The proofs of both these facts are virtually the same as the proofs of Lemmas 1.3 and 1.4.

### *1.3.2 Affine Numerical Representation*

It is convenient for our derivations to consider the linear span of $\mathcal{Q}$ defined as follows:

$$\text{lin}(\mathcal{Q}) = \left\{ \sum_{i=1}^{k} \alpha_i \Phi_i : \alpha_i \in \mathbb{R},\ \Phi_i \in \mathcal{Q},\ i = 1, \dots, k,\ k \in \mathbb{N} \right\} = \mathcal{Q} - \mathcal{Q},$$

where $\mathcal{Q} - \mathcal{Q}$ is the Minkowski sum of the sets $\mathcal{Q}$ and $-\mathcal{Q}$. The last equality follows from the fact that $\mathcal{Q}$ is a convex cone.

**Theorem 1.14.** *If a total preorder* $\trianglelefteq$ *on* $\mathcal{Q}$ *satisfies the dual independence and Archimedean axioms, then a linear functional on* $\text{lin}(\mathcal{Q})$ *exists, whose restriction to* $\mathcal{Q}$ *is a numerical representation of* $\trianglelefteq$*.*

*Proof.* Define in the space $\text{lin}(\mathcal{Q})$ the set

$$C = \{\Phi - \Psi : \Phi \in \mathcal{Q},\ \Psi \in \mathcal{Q},\ \Phi \lhd \Psi\}.$$

Exactly as in the proof of Theorem 1.5 on page 4, we can prove that $C$ is convex. We shall prove that it is a cone. Suppose $\Phi \lhd \Psi$ and let $\alpha > 0$. If $\alpha \in (0,1)$, then the independence axiom implies that[†]

$$\alpha\Phi = \alpha\Phi + (1-\alpha)\mathbb{0} \lhd \alpha\Psi + (1-\alpha)\mathbb{0} = \alpha\Psi.$$

Consider $\alpha > 1$, and suppose $\alpha\Psi \trianglelefteq \alpha\Phi$. If $\alpha\Psi \lhd \alpha\Phi$, then, owing to the independence axiom, we obtain a contradiction: $\Psi = \frac{1}{\alpha}(\alpha\Psi) \lhd \frac{1}{\alpha}(\alpha\Phi) = \Phi$. Consider the case when $\alpha\Psi \sim \alpha\Phi$. By virtue of Lemma 1.12 and the independence axiom, for any $\beta \in (0, 1/\alpha)$ we obtain a contradiction in the following way:

$$\begin{aligned}\alpha\Psi &\sim \beta(\alpha\Phi) + (1-\beta)(\alpha\Psi) = (\beta\alpha)\Phi + (1-\beta\alpha)\Big[\frac{(1-\beta)\alpha}{1-\beta\alpha}\Psi\Big]\\ &\lhd (\beta\alpha)\Psi + (1-\beta)(\alpha\Psi) = \alpha\Psi.\end{aligned}$$

Therefore, $\alpha\Phi \lhd \alpha\Psi$ for all $\alpha > 0$. We conclude that for every $\alpha > 0$ the element $\alpha(\Phi - \Psi) \in C$. Consequently, $C$ is a convex cone.

We shall prove that the algebraic interior of $C$, denoted $\text{cor}(C)$, is nonempty, and that, in fact, $C = \text{cor}(C)$. Consider any $\Gamma \in C$, a function $\Upsilon \in \text{lin}(\mathcal{Q})$, and the ray

$$Z(t) = \Gamma + t\Upsilon, \quad t > 0.$$

Our objective is to show that $Z(t) \in C$ for a sufficiently small $t > 0$. By the definition of $\text{lin}(\mathcal{Q})$, we can represent $\Upsilon = \Upsilon^+ - \Upsilon^-$, with $\Upsilon^+, \Upsilon^- \in \mathcal{Q}$.

Since $\Gamma \in C$, we can represent it as a difference $\Gamma = \Phi - \Psi$, with $\Phi, \Psi \in \mathcal{Q}$, and $\Phi \lhd \Psi$. Then

$$\Gamma + t\Upsilon = \big[(1-t)\Phi + t\Upsilon^+\big] - \big[(1-t)\Psi + t\Upsilon^-\big] + t\Gamma. \tag{1.9}$$

Both expressions in brackets are elements of $\mathcal{Q}$. By the dual independence axiom,

$$\Phi \lhd \frac{1}{2}\Phi + \frac{1}{2}\Psi \lhd \Psi.$$

---

[†] We use the symbol $\mathbb{0}$ to denote the quantile function identically equal to 0.

According to Lemma 1.13, there exists $t_0 > 0$ such that for all $t \in [0, t_0]$ we also have

$$(1-t)\Phi + t\Upsilon^{+} \lhd \frac{1}{2}\Phi + \frac{1}{2}\Psi \lhd (1-t)\Psi + t\Upsilon^{-}.$$

This proves that

$$\left[(1-t)\Phi + t\Upsilon^{+}\right] - \left[(1-t)\Psi + t\Upsilon^{-}\right] \in C,$$

provided that $t \in [0, t_0]$. Thus relation (1.9) implies that for every $t \in [0, t_0]$ the point $\Gamma + t\Upsilon$ is a sum of two elements of $C$. Since the set $C$ is a convex cone, this point is also an element of $C$.

As $C$ is convex, $C = \operatorname{cor}(C)$, and $\mathbb{0} \notin C$, the point $\mathbb{0}$ and the set $C$ can be separated strictly: a linear functional $U$ on $\operatorname{lin}(\mathcal{Q})$ exists, such that

$$U(\Gamma) < 0, \quad \forall\, \Gamma \in C.$$

Thus,

$$U(\Phi) - U(\Psi) < 0, \quad \text{whenever} \quad \Phi \lhd \Psi,$$

as required. ∎

### 1.3.3 Integral Representation with Rank Dependent Utility Functions

In order to derive an integral representation of the numerical representation $U(\cdot)$ of the preorder $\unlhd$, we need much stronger conditions, than those of Theorem 1.14. Two issues are important in this respect:

- Continuity of $U(\cdot)$ on an appropriate complete topological vector space containing the set $\mathcal{Q}$ of quantile functions; and
- Integral representation of a continuous linear functional on this space.

To address the first issue, we recall the construction of the space of bounded functions from classical functional analysis. Consider the algebra $\Sigma$ of all sets obtained by finite unions and intersections of intervals of the form $(a, b]$ in $(0, 1]$, where $0 < a < b \le 1$. With $\mathcal{S} = (0, 1]$ and with $\Sigma$, we can define simple functions on $(0, 1]$ as follows:

$$f(p) = \sum_{i=1}^{n} \alpha_i \mathbb{1}_{A_i}(p), \quad p \in \mathcal{S}, \tag{1.10}$$

where $\alpha_i \in \mathbb{R}$ for $i = 1, \ldots, n$, and $A_i$, $i = 1, \ldots, n$, are disjoint elements of the algebra $\Sigma$. In the formula above, $\mathbb{1}_A(\cdot)$ denotes the characteristic function of a set $A$. Next, we define the space $B\big((0,1],\Sigma\big)$ of all bounded functions on $(0,1]$ that can be obtained as uniform limits of sequences of simple functions. The space $B\big((0,1],\Sigma\big)$ equipped with the supremum norm:

$$\|\Phi\| = \sup_{0<p\le 1} \Phi(p),$$

is a Banach space.

From now on, we shall consider only compactly supported distributions on $\mathbb{R}$. The corresponding quantile functions form the prospect space $\mathcal{Q}_{\mathrm{b}}$ of all bounded, nondecreasing, and left-continuous functions on $(0,1]$.[†] The set $\mathcal{Q}_{\mathrm{b}}$ is contained in $B\big((0,1],\Sigma\big)$. Indeed, every monotonic function may have only countably many jumps and their sizes are summable due to the boundedness of the function. Owing to the left-continuity, it can be represented as a uniform limit of simple functions.

We make the following assumption:

**Monotonicity Axiom:** For all $\Phi, \Psi \in \mathcal{Q}_{\mathrm{b}}$ one has

$$\Phi \le \Psi \implies \Phi \trianglelefteq \Psi,$$

where the inequality between the functions is understood pointwise.

**Theorem 1.15.** *If a total preorder $\trianglelefteq$ on $\mathcal{Q}_{\mathrm{b}}$ is continuous and satisfies the dual independence and monotonicity axioms, then a linear continuous functional on $B\big((0,1],\Sigma\big)$ exists, whose restriction to $\mathcal{Q}_{\mathrm{b}}$ is a numerical representation of $\trianglelefteq$.*

*Proof.* Since the continuity axiom implies the Archimedean axiom, Theorem 1.14 implies the existence of a linear functional $U : \operatorname{lin}(\mathcal{Q}_{\mathrm{b}}) \to \mathbb{R}$ whose restriction to $\mathcal{Q}_{\mathrm{b}}$ is a numerical representation of $\trianglelefteq$. The continuity axiom implies the continuity of the functional $U(\cdot)$ on $\mathcal{Q}_{\mathrm{b}}$. We shall extend $U(\cdot)$ to a continuous functional on the entire space $B\big((0,1],\Sigma\big)$.

Every simple function can be expressed as

$$\begin{aligned}\Phi &= \sum_{i=1}^{n} \alpha_i \mathbb{1}_{(p_i,p_{i+1}]}\\ &= \sum_{i:\alpha_i>0} \alpha_i\big(\mathbb{1}_{(p_i,1]} - \mathbb{1}_{(p_{i+1},1]}\big) + \sum_{i:\alpha_i<0} |\alpha_i|\big(\mathbb{1}_{(p_{i+1},1]} - \mathbb{1}_{(p_i,1]}\big).\end{aligned}$$

with $0 = p_1 < p_2 < \cdots < p_{n+1} = 1$. Rearranging terms, we observe that every simple function is a difference of two simple functions in $\mathcal{Q}_{\mathrm{b}}$ and is thus

[†] Bounded nondecreasing functions on $(0,1)$ can be extended to $(0,1]$ by assigning their left limits as values at 1.

an element of $\text{lin}(\mathcal{Q}_\text{b})$. Consequently, the linear functional $U(\cdot)$ is well-defined on the space of simple functions.

Since the preorder $\trianglelefteq$ satisfies the monotonicity axiom, the linear functional $U(\cdot)$ is monotonic on $\mathcal{Q}_\text{b}$. We shall prove that it is also monotonic on the set of simple functions in $B\big((0,1],\Sigma\big)$. Let $\Phi$ and $\Psi$ be two simple functions, and let $\Phi \le \Psi$. Then $\Phi = \Phi_1 - \Phi_2$, $\Psi = \Psi_1 - \Psi_2$, where $\Phi_1, \Phi_2, \Psi_1, \Psi_2 \in \mathcal{Q}_\text{b}$, and

$$\Phi_1 + \Psi_2 \le \Phi_2 + \Psi_1.$$

As both sides are elements of $\mathcal{Q}_\text{b}$ and $U(\cdot)$ is nondecreasing in $\mathcal{Q}_\text{b}$ and linear, regrouping the terms we obtain

$$U(\Phi) - U(\Psi) = U(\Phi_1 - \Phi_2 - \Psi_1 + \Psi_2) = U(\Phi_1 + \Psi_2) - U(\Phi_2 + \Psi_1) \le 0.$$

This proves the monotonicity of $U(\cdot)$ on the linear subspace of simple functions.

For any function $\Gamma \in B\big((0,1],\Sigma\big)$, we construct two sequences of simple functions: $\{\Phi_n\}$ and $\{\Psi_n\}$ such that $\Phi_n \le \Gamma \le \Psi_n$, for $n = 1, 2, \dots$, and

$$\Gamma = \lim_{n\to\infty} \Phi_n = \lim_{n\to\infty} \Psi_n.$$

The sequence $\{U(\Phi_n)\}$ is bounded from above by $U(\Psi_k)$ for any $k$, due to the monotonicity of $U(\cdot)$. Similar, the sequence $\{U(\Psi_n)\}$ is bounded from below by $U(\Phi_k)$ for any $k$. Moreover,

$$\begin{aligned} 0 &\le U(\Psi_n) - U(\Phi_n) = U(\Psi_n - \Phi_n) \\ &\le U\big(\|\Psi_n - \Phi_n\|\mathbb{1}_{(0,1]}\big) = U(\mathbb{1}_{(0,1]})\|\Psi_n - \Phi_n\| \to 0. \end{aligned}$$

Therefore, both sequences $\{U(\Phi_n)\}$ and $\{U(\Psi_n)\}$ have the same limit and we can define

$$U(\Gamma) = \lim_{n\to\infty} U(\Phi_n) = \lim_{n\to\infty} U(\Psi_n).$$

We may use any sequence of simple functions $\Gamma_n \to \Gamma$ to calculate $U(\Gamma)$. Indeed, setting $\Phi_n = \Gamma_n - \|\Gamma_n - \Gamma\|$ and $\Psi_n = \Gamma_n + \|\Gamma_n - \Gamma\|$, we obtain $\Phi_n \le \Gamma \le \Psi_n$ and $\Phi_n \le \Gamma_n \le \Psi_n$. Consequently, $U(\Phi_n) \le U(\Gamma_n) \le U(\Psi_n)$ and

$$\lim_{n\to\infty} U(\Gamma_n) = U(\Gamma).$$

The functional $U : B\big((0,1],\Sigma\big) \to \mathbb{R}$ defined in this way is linear on the set of simple functions, which is a subspace of $\text{lin}\,\mathcal{Q}$. Consider two elements $\Phi$ and $\Psi$ of $B\big((0,1],\Sigma\big)$, and two sequences $\{\Phi_n\}$ and $\{\Psi_n\}$ of simple functions such that $\Phi_n \to \Phi$ and $\Psi_n \to \Psi$. For any $a, b \in \mathbb{R}$, we obtain