VECTOR OPTIMIZATION

Andreas Löhne

Vector Optimization with Infimum and Supremum

Springer

Vector Optimization

Series Editor:

Johannes Jahn
University of Erlangen-Nürnberg
Department of Mathematics
Martensstr. 3
81058 Erlangen
Germany
jahn@am.uni-erlangen.de

Vector Optimization

The series in Vector Optimization contains publications in various fields of optimization with vector-valued objective functions, such as multiobjective optimization, multi criteria decision making, set optimization, vector-valued game theory and border areas to financial mathematics, biosystems, semidefinite programming and multiobjective control theory. Studies of continuous, discrete, combinatorial and stochastic multiobjective models in interesting fields of operations research are also included. The series covers mathematical theory, methods and applications in economics and engineering. These publications being written in English are primarily monographs and multiple author works containing current advances in these fields.

Andreas Löhne

Vector Optimization with Infimum and Supremum

Springer

Andreas Löhne
Martin-Luther-Universität Halle-Wittenberg
NWF II - Institut für Mathematik
Theodor-Lieser-Str. 5
06120 Halle
Germany
andreas.loehne@mathematik.uni-halle.de

ISSN 1867-8971 e-ISSN 1867-898X
ISBN 978-3-642-18350-8 e-ISBN 978-3-642-18351-5
DOI 10.1007/978-3-642-18351-5
Springer Heidelberg Dordrecht London Newyork

Cover design: Spi Publisher Services

Printed on acid-free paper

Springer is part of Springer Science+Business Media (www.springer.com)

To Jana and Pascal

Preface

Infimum and supremum are indispensable concepts in optimization. Nevertheless their role in vector optimization has been rather marginal. This seems to be due the fact that their existence in partially ordered vector spaces is connected with restrictive assumptions. The key to an approach to vector optimization based on infimum and supremum is to consider set-valued objective functions and to extend the partial ordering of the original objective space to a suitable subspace of the power set. In this new space the infimum and supremum exist under the usual assumptions.

These ideas lead to a novel exposition of vector optimization. The reader is not only required to familiarize with several new concepts, but also a change of philosophy is suggested to those being acquainted with the classical approaches. The goal of this monograph is to cover the most important concepts and results on vector optimization and to convey the ideas, which can be used to derive corresponding variants of all the remaining results and concepts. This selection ranges from the general theory including solution concepts and duality theory, through to algorithms for the linear case.

Researchers and graduate-level students working in the field of vector optimization belong to the intended audience. In view of many facts and notions that are recalled, the book is also addressed to those who are not familiar with classical approaches to vector optimization. However, it should be taken into account that a fundamental motivation of vector optimization and applications are beyond the scope of this book.

Some basic knowledge in (scalar) optimization, convex analysis and general topology is necessary to understand the first part, which deals with general and convex problems. The second part is a self-contained exposition of the linear case. Infimum and supremum are not visible but present in the background. The connections to the first part are explained at several places, but they are not necessary to understand the results for the linear case. Some knowledge on (scalar) linear programming is required.

The results in this book arose from several research papers that have been published over the last five years. The results and ideas of this exposition are contributed by Andreas Hamel, Frank Heyde and Christiane Tammer concerning the first part as well as Frank Heyde, Christiane Tammer and Matthias Ehrgott concerning the second part. A first summary, extension and consolidation of these results has been given in the author's habilitation thesis, which appeared in 2010. This book is an extension. It contains one new chapter with extended variants of algorithms and more detailed explanations.

I thank all persons who supported me to write this book. In particular, I'm greatly indebted to Matthias Ehrgott, Gabriele Eichfelder, Andreas Hamel, Frank Heyde, Johannes Jahn and Christiane Tammer for their valuable comments, important corrections and all their advice that entailed a considerable increase of quality.

Halle (Saale), *Andreas Löhne*
November 2010

Contents

Introduction

From a mathematical point of view, vector optimization is the theory of optimization problems with a vector-valued objective function. Instead of the extended real numbers $\overline{\mathbb{R}} := \mathbb{R} \cup \{+\infty, -\infty\}$, one considers an extended partially ordered vector space as the image space of the objective map. One of the main difficulties is the lack of a suitable infimum and supremum. For many instances of extended partially ordered vector spaces, even in finite dimensions, an infimum does not exist at all. But even if the infimum in the sense of a greatest lower bound exists, it is usually not related to the typical optimality notions which are motivated by applications in multiobjective optimization.

The idea of multiobjective optimization is to present a decision maker all or at least a representative selection of minimal or efficient vectors. The decision maker's job is to choose one of these vectors. An infimum in an extended partially ordered vector space, if it exists, is of course a vector. But the requirement from an applicational point of view is to evaluate a set of efficient points in order to present them to the decision maker.

The infimum of a fixed subset is generally changing when the partially ordered set, say the universal set, is extended to a larger partially ordered set. The reason is that more candidates for greatest lower bounds are available in a larger set. This basic idea is applied to vector optimization as we create a suitable notion of infimum by embedding the extended partially ordered vector space into a larger partially ordered set, in fact, into a subset of the power set. This allows us to develop a theory of vector optimization which is based on infimum and supremum. This leads to new insights and a high degree of analogy to the scalar optimization theory.

Vector optimization has its origin in economics, in particular, in welfare theory and utility theory. The foundations are connected with the names Vilfredo Pareto (1848-1923) and Francis Ysidro Edgeworth (1845-1926). Independently vector optimization also arose from game theory which was initiated by Émile Borel (1871-1956), Maurice René Fréchet (1878-1973) and

A. Löhne, *Vector Optimization with Infimum and Supremum*, Vector Optimization,
DOI 10.1007/978-3-642-18351-5_1, © Springer-Verlag Berlin Heidelberg 2011

John von Neumann (1903-1957). For more details the reader is referred to the survey paper by Stadler (1979). From a theoretical perspective the foundations of vector optimization were laid by Georg Cantor (1845-1918) by his famous intersection theorem; by Felix Hausdorff (1868-1942), who showed the existence of utility functions in the context of partially ordered sets; and by Max Zorn (1906-1993), who gave conditions for the existence of maximal elements without using a utility function (see Göpfert *et al.*, 2009). What is today considered to be vector optimization, multiobjective optimization or multicriteria optimization has its origin in the 1950s. The notion of *efficient points* was introduced (compare Stadler, 1979) by Koopmans (1951, Definition 4.2): *"A possible point in the commodity is called efficient whenever an increase in the one of its coordinates (the net output of one good) can be achieved only at the cost of a decrease in some other coordinate (the net output of another good)"*. Kuhn and Tucker (1951) introduced (compare Stadler, 1979) the term *vector maximum problem*. Today there exists a number of textbooks on vector optimization, among them (Sawaragi *et al.*, 1985; Jahn, 1986, 2004; Luc, 1988; Göpfert and Nehse, 1990; Ehrgott, 2000, 2005; Göpfert *et al.*, 2003; Chen *et al.*, 2005; Eichfelder, 2008; Boţ *et al.*, 2009). There are several thousands of research papers on this subject.

This monograph differs from the literature as it is based on the complete lattice $(\mathcal{I}, \preccurlyeq)$ of self-infimal subsets of the original objective space $(Y, \leq)$ of a given extended vector-valued objective function. Starting with a vector optimization problem

$$\text{minimize } f : X \to \overline{Y} \text{ with respect to } \leq \text{ over } S \subseteq X, \tag{V}$$

we assign to f an $\mathcal{I}$-valued objective function

$$\bar{f} : X \to \mathcal{I}, \qquad \bar{f}(x) := \operatorname{Inf}\{f(x)\}$$

and consider the related problem

$$\text{minimize } \bar{f} : X \to \mathcal{I} \text{ with respect to } \preccurlyeq \text{ over } S \subseteq X. \tag{$\mathcal{V}$}$$

There is a close connection between the values of f and $\bar{f}$; that is, for all $x^1, x^2 \in X$ we have

$$f(x^1) \leq f(x^2) \qquad \Longleftrightarrow \qquad \bar{f}(x^1) \preccurlyeq \bar{f}(x^2).$$

Since the objective space $\mathcal{I}$ in Problem $(\mathcal{V})$ is a complete lattice, the latter correspondence allows us to develop the theory of vector optimization based on infimum and supremum.

This approach was firstly pointed out in (Löhne and Tammer, 2007; Heyde *et al.*, 2009a), but it is based on a couple of pre-investigations, such as (Hamel *et al.*, 2004; Hamel, 2005; Löhne, 2005a,b). It turned out that it is possible to formulate and prove vectorial duality theorems very similar to the corresponding scalar results if the vectorial image space is replaced by the complete lattice $\mathcal{I}$. But, the space $\mathcal{I}$ of self-infimal sets does not only provide a complete lattice; the infimum with respect to this complete lattice is also closely related to the standard solution concepts in vector optimization. Even though infimal sets were used before (see e.g. Nieuwenhuis, 1980; Sawaragi *et al.*, 1985; Tanino, 1988, 1992; Song, 1997, 1998), in particular in duality theory, the deeper context was not pointed out: the complete lattice $\mathcal{I}$.

An approach to vector optimization based on infimum and supremum leads to the question how to integrate conventional solution concepts into the theory. It turned out that there is no standard way to say what is a solution to a vector optimization problem from a mathematical point of view. On the one hand this is concerned with the question whether a solution is a set of vectors or just a single vector (see also the introduction to Chapter 2). On the other hand there are different types of efficient vectors depending on different possible interpretations of "less than" when the ordering relation is more complex than the one in $\mathbb{R}$.

It might be worth noting that the solution concept proposed in this work involves two different types of minimality notions: weakly minimal and minimal vectors. This could shed a new light on the role of weakly efficient solutions in vector optimization. Jahn (2004, p. 110) writes that *"the concept of weak minimality is of theoretical interest, and it is not an appropriate notion for applied problems."* This is in accordance with the fact that weak minimality is essential to construct the complete lattice $\mathcal{I}$, but our solution concept itself is based on minimality. One can say that the theoretical benefits of weak minimality and the application-oriented properties of minimality are involved in one concept.

This monograph is organized as follows. Part I is devoted to the general ideas and to convex problems. In Chapter 1 we introduce the complete lattice $\mathcal{I}$, which is the basis of this exposition. We also provide several concepts and facts from the literature as far as they are needed in this book. Chapter 2 is devoted to solution concepts and Chapter 3 is concerned with duality. Part II deals with linear problems. Even though the connections to the first part are often discussed, the second part is a self-contained exposition of the linear theory. In Chapter 4 we focus on solution concepts and duality. The concepts from Part I are adapted and special features of the linear duality theory are shown. Chapter 5 is devoted to algorithms to solve linear problems. Each chapter begins with a specific introduction and ends with several notes on the literature; in particular, the origin of the results is discussed.

This book offers a systematic introduction and a summary of recent developments in the theory of vector optimization with infimum and supremum. It is based on the cited papers, but the theory is presented in a more general setting and with several extensions. This book aims to be a self-contained summary and an extension of recent results.

Part I
General and Convex Problems

Chapter 1
A complete lattice for vector optimization

Extended real-valued objective functions are characteristic for scalar optimization problems. The space of extended real numbers $\overline{\mathbb{R}} := \mathbb{R} \cup \{+\infty\} \cup \{-\infty\}$ enjoys several properties which are quite important for optimization:

(i) $\mathbb{R}$ is a vector space, but $\overline{\mathbb{R}}$ is not. The linear operations can be partially extended to $\overline{\mathbb{R}}$.
(ii) The linear operations on $\mathbb{R}$ are continuous, i.e., the topology is compatible with the linear structure.
(iii) $\overline{\mathbb{R}}$ is totally ordered by the usual ordering $\leq$. The ordering on $\mathbb{R}$ is compatible with the linear operations.
(iv) $\overline{\mathbb{R}}$ is a complete lattice, i.e., every subset has an infimum and a supremum.

In vector optimization we have to replace $\mathbb{R}$ and $\overline{\mathbb{R}}$ by a more general space. Certain properties can be maintained, others must be abandoned. Underlying a partially ordered topological vector space Y and its extension $\overline{Y} := Y \cup \{\pm\infty\} := Y \cup \{-\infty, +\infty\}$, we obtain all the mentioned properties up to the following two exceptions: First, Y is not totally but partially ordered only. Secondly, a complete lattice is obtained by $\overline{Y}$ only in special cases. This depends on the choice of Y and the choice of the partial ordering. But even in the special cases where $\overline{Y}$ is a complete lattice (e.g. $Y = \mathbb{R}^q$ equipped with the "natural" componentwise ordering), the infimum is different to the typical vectorial minimality notions, which arise from applications. This is illustrated in Figure 1.1.

As a consequence, infimum and supremum (at least in the sense of greatest lower and least upper bounds) do not occur in the standard literature on vector optimization. Some authors, among them Nieuwenhuis (1980); Tanino (1988, 1992), used a generalization of the infimum in $\overline{\mathbb{R}}$. Although the same notion is also involved into this work, the new idea is that we provide an appropriate complete lattice. To this end we work with a subset of the power set of a given partially ordered topological vector space. The construction and the properties of this complete lattice are the subject of this chapter. We

A. Löhne, *Vector Optimization with Infimum and Supremum*, Vector Optimization,
DOI 10.1007/978-3-642-18351-5_2, © Springer-Verlag Berlin Heidelberg 2011

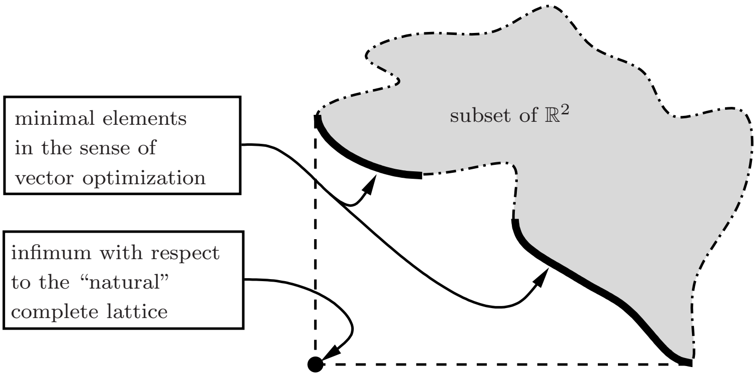

Fig. 1.1 $\overline{\mathbb{R}^2}$ equipped with the natural ordering provides a complete lattice. But, the infimum can be far away from the minimal elements.

recall in this chapter several standard notions and results but we also present the basics of a set-valued approach to vector optimization.

1.1 Partially ordered sets and complete lattices

This section is a short summary of several concepts and results related to ordered sets as they are required for this exposition.

Definition 1.1. Let Z be a nonempty set. A relation $R \subseteq Z \times Z$ is called a *partial ordering* on Z if the following properties are satisfied:

(i) R is reflexive: $\forall z \in Z : (z,z) \in R$,
(ii) R is transitive: $[(z^1,z^2) \in R \wedge (z^2,z^3) \in R] \implies (z^1,z^3) \in R$,
(iii) R is antisymmetric: $[(z^1,z^2) \in R \wedge (z^2,z^1) \in R] \implies z^1 = z^2$.

Instead of $(z^1,z^2) \in R$, we write $z^1 \leq_R z^2$.

The index R is usually omitted or replaced (for instance, if the ordering is generated by a cone C, we write $z^1 \leq_C z^2$ whenever $z^2 - z^1 \in C$) and we just say that $\leq$ is a partial ordering. A nonempty set Z equipped with a partial ordering on Z is called a *partially ordered set*. It is denoted by $(Z,\leq)$. The following convention is used throughout: If $(Z,\leq)$ is a partially ordered set and $A \subseteq Z$, we speak about a subset of the partially ordered set $(Z,\leq)$.

Definition 1.2. Let $(Z,\leq)$ be a partially ordered set and let $A \subseteq Z$. An element $l \in Z$ is called a *lower bound* of A if $l \leq z$ for all $z \in A$. An upper bound is defined analogously.

Next we define an *infimum* and a *supremum* for a subset A of a partially ordered set $(Z, \leq)$.

Definition 1.3. Let $(Z, \leq)$ be a partially ordered set and let $A \subseteq Z$. An element $k \in Z$ is called a *greatest lower bound* or *infimum* of $A \subseteq Z$ if k is a lower bound of A and for every other lower bound l of A we have $l \leq k$. We use the notation $k = \inf A$ for the infimum of A, if it exists.

The *least upper bound* or *supremum* is defined analogously and is denoted by $\sup A$. The lower (upper) bound of Z, if it exists, is called *least (greatest)* element.

Proposition 1.4. *Let $(Z, \leq)$ be a partially ordered set and let $A \subseteq Z$. If the infimum of A exists, then it is uniquely defined.*

Proof. Let both k and l be infima of A. Then, l and k are lower bounds of A. The definition of the infimum yields $l \leq k$ and $k \leq l$. As $\leq$ is antisymmetric, we get $l = k$. □

Definition 1.5. A partially ordered set $(Z, \leq)$ is called a *complete lattice* if the infimum and supremum exist for every subset $A \subseteq Z$.

Note that a one-sided condition is already sufficient to characterize a complete lattice.

Proposition 1.6. *A partially ordered set $(Z, \leq)$ is a complete lattice if and only if the infimum exists for every subset $A \subseteq Z$.*

Proof. Let $A \subseteq Z$ be a given set and let $B \subseteq Z$ be the set of all upper bounds of A. By assumption, $p := \inf B$ exists. As p is a lower bound of B, $z \geq p$ holds for every upper bound z of A. Every $z \in A$ is a lower bound of B. By the definition of the infimum we get $p \geq z$ for every $z \in A$. Together we have $p = \sup A$. □

Example 1.7. The extended real numbers $\overline{\mathbb{R}} := \mathbb{R} \cup \{\pm\infty\}$ equipped with the usual ordering $\leq$ provide a complete lattice.

Example 1.8. Let $\leq$ be the componentwise ordering relation in $\mathbb{R}^q$. If the ordering relation $\leq$ is extended to $Z := \mathbb{R}^q \cup \{\pm\infty\}$ by setting $-\infty \leq z \leq +\infty$ for all $z \in Z$, $(Z, \leq)$ provides a complete lattice. The infimum of a subset $A \subseteq Z$ is

$$\inf A = \begin{cases} \left(\inf\limits_{z \in A} z_1, \dots, \inf\limits_{z \in A} z_q\right)^T & \text{if } \exists b \in \mathbb{R}^q, \forall z \in A : b \leq z \\ +\infty & \text{if } A = \emptyset \\ -\infty & \text{otherwise.} \end{cases}$$

Example 1.9. Let $Z = \mathbb{R}^3$ and let C be the polyhedral (convex) cone which is spanned by the vectors $(0,0,1)^T$, $(0,1,1)^T$, $(1,0,1)^T$, $(1,1,1)^T$. Then $(Z, \leq_C)$ is not a complete lattice. For instance, there is no supremum of the finite set $\{(0,0,0)^T, (1,0,0)^T\}$.

Note that the previous example is a special case of the following result by Peressini (1967): $\mathbb{R}^n$ *is an Archimedean vector lattice with respect to the order generated by a cone* C *if and only if there are* n *linearly independent vectors* v^i *such that*

$$C := \left\{x \in \mathbb{R}^n \,|\, \forall i = 1, \dots, n : \left\langle x, v^i \right\rangle \geq 0\right\}. \tag{1.1}$$

Note further that, as pointed out by Anderson and Annulis (1973), the word "Archimedean" is inadvertently omitted in (Peressini, 1967). A vector lattice is Archimedean if

$$(\forall n \in \mathbb{N} : 0 \leq nx \leq z) \implies x = 0.$$

The pair $(\mathbb{R}^2, L)$, where

$$L := \left\{x \in \mathbb{R}^2 \,|\, x_1 > 0 \vee [x_1 = 0 \wedge x_2 \geq 0]\right\}$$

is the lexicographic ordering cone provides an example of a vector lattice, which is not Archimedean. As demonstrated in (Anderson and Annulis, 1973), L cannot be expressed as in (1.1).

Example 1.10. Let X be a nonempty set and let $\mathcal{P}(X) = 2^X$ be the power set of X. $(\mathcal{P}(X), \supseteq)$ provides a complete lattice. The infimum and supremum of a nonempty subset $\mathcal{A} \subseteq \mathcal{P}(X)$ are given as

$$\inf \mathcal{A} = \bigcup_{A \in \mathcal{A}} A \qquad \sup \mathcal{A} = \bigcap_{A \in \mathcal{A}} A.$$

Note that $X \in \mathcal{P}(X)$ is the least element and $\emptyset \in \mathcal{P}(X)$ is the greatest element in $(\mathcal{P}(X), \supseteq)$. If $\mathcal{A}$ is empty, we set $\sup \mathcal{A} = X$ and $\inf \mathcal{A} = \emptyset$.

Example 1.11. Let X be a vector space and let $\mathcal{C}(X)$ be the family of all convex subsets of X. $(\mathcal{C}(X), \supseteq)$ provides a complete lattice. The infimum and supremum of a nonempty subset $\mathcal{A} \subseteq \mathcal{C}(X)$ are given as

$$\inf \mathcal{A} = \operatorname{co} \bigcup_{A \in \mathcal{A}} A \qquad \sup \mathcal{A} = \bigcap_{A \in \mathcal{A}} A.$$

If $\mathcal{A}$ is empty, we set again $\sup \mathcal{A} = X$ and $\inf \mathcal{A} = \emptyset$.

Example 1.12. Let X be a topological space and let $\mathcal{F}(X)$ be the family of all closed subsets of X. $(\mathcal{F}(X), \supseteq)$ provides a complete lattice. The infimum and supremum of a nonempty subset $\mathcal{A} \subseteq \mathcal{F}(X)$ are given as

$$\inf \mathcal{A} = \operatorname{cl} \bigcup_{A \in \mathcal{A}} A \qquad \sup \mathcal{A} = \bigcap_{A \in \mathcal{A}} A.$$

If $\mathcal{A}$ is empty, we set again $\sup \mathcal{A} = X$ and $\inf \mathcal{A} = \emptyset$.

Example 1.13. Let X be a set, L a complete lattice and $\mathcal{L}(X)$ be the set of all L-valued functions on X. A partial ordering on $\mathcal{L}(X)$ is defined by

$$l_1 \leq l_2 :\Longleftrightarrow \forall x \in X : l_1(x) \leq l_2(x).$$

Then $(\mathcal{L}(X), \leq)$ provides a complete lattice. The infimum and supremum are given by the pointwise infimum and supremum.

1.2 Conlinear spaces

Convexity is one of the most important concepts in optimization. A *convex set* C is usually defined to be a subset of a vector space X satisfying the condition

$$[\lambda \in [0,1] \wedge x, y \in C] \implies \lambda x + (1-\lambda) y \in C. \tag{1.2}$$

A very important special case of a convex set is a *convex cone*, where a *cone* is defined to be a set K satisfying

$$[\lambda \in \mathbb{R}_+, x \in K] \implies \lambda x \in K, \tag{1.3}$$

where $\mathbb{R}_+ := \{\lambda \in \mathbb{R} |\ \lambda \geq 0\}$.

We observe that neither of the definitions require a multiplication by a negative real number. It is therefore consistent to define convexity on more general spaces. This can be motivated by the examples below showing convex sets and convex cones which cannot be embedded into a linear space (vector space). The natural framework for convexity seems to be a *conlinear space* rather than a linear one.

Definition 1.14. A nonempty set Z equipped with an addition $+ : Z \times Z \to Z$ and a multiplication $\cdot : \mathbb{R}_+ \times Z \to Z$ is said to be a *conlinear space* with the *neutral element* $\theta \in Z$ if for all $z, z^1, z^2 \in Z$ and all $\alpha, \beta \geq 0$ the following axioms are satisfied:

(C1) $z^1 + \left(z^2 + z\right) = \left(z^1 + z^2\right) + z$,
(C2) $z + \theta = z$,
(C3) $z^1 + z^2 = z^2 + z^1$,
(C4) $\alpha \cdot (\beta \cdot z) = (\alpha\beta) \cdot z$,
(C5) $1 \cdot z = z$,
(C6) $0 \cdot z = \theta$,
(C7) $\alpha \cdot \left(z^1 + z^2\right) = \left(\alpha \cdot z^1\right) + \left(\alpha \cdot z^2\right)$.

An instance of a conlinear space is given in Example 1.31 below. The axioms of a conlinear space $(Z, +, \cdot)$ are appropriate to deal with convexity. A convex set and a cone in a conlinear space are defined, respectively, by (1.2) and (1.3). The *convex hull* $\operatorname{co} A$ of a subset A of a conlinear space $(Z, +, \cdot)$

is the intersection of all convex sets in Z containing A. The convex hull of a set A coincides with set of all (finite) convex combinations of elements of A (Hamel, 2005, Theorem 3).

Additionally to the axioms (C1) to (C7), it is sometimes useful to use a second distributive law, that is, for all $z \in Z$ and all $\alpha, \beta \geq 0$ we can suppose additionally that

(C8) $\quad \alpha \cdot z + \beta \cdot z = (\alpha + \beta) \cdot z.$

In a conlinear space, singleton sets are not necessarily convex, see Example 1.17 below. Indeed this requirement is equivalent to (C8).

Proposition 1.15. *For every conlinear space, the following statements are equivalent:*

(i) *Every singleton set is convex,*
(ii) *(C8) holds.*

Proof. This is obvious. □

Definition 1.16. An element $\bar{z}$ of a conlinear space Z is said to be *convex*, if the set $\{\bar{z}\}$ is convex.

Example 1.17. An element of a conlinear space can be nonconvex. Indeed, let $Z = \mathcal{P}(\mathbb{R})$ and consider the element $A := \{0, 1\} \in Z$. We have $\frac{1}{2}A + \frac{1}{2}A = \left\{0, \frac{1}{2}, 1\right\} \neq A$.

If $(Z, +, \cdot)$ is a conlinear space, we denote by Z_{co} the subset of all $z \in Z$ satisfying (C8).

Proposition 1.18. $(Z_{\mathrm{co}}, +, \cdot)$ *is a conlinear space.*

Proof. Let $z^1, z^2 \in Z_{\mathrm{co}}$. For $\alpha, \beta \geq 0$, we have

$$\begin{aligned} \alpha(z^1 + z^2) + \beta(z^1 + z^2) &\overset{\text{(C7), (C3)}}{=} \alpha z^1 + \beta z^1 + \alpha z^2 + \beta z^2 \\ &\overset{\text{(C8)}}{=} (\alpha + \beta) z^1 + (\alpha + \beta) z^2 \\ &\overset{\text{(C7)}}{=} (\alpha + \beta)(z^1 + z^2), \end{aligned}$$

i.e., $z^1 + z^2 \in Z_{\mathrm{co}}$. Similarly we obtain $\gamma \cdot z^1 \in Z_{\mathrm{co}}$ for $\gamma \geq 0$. □

If a conlinear space is equipped with an ordering relation, it is useful to require that this ordering relation is compatible with the conlinear structure. The same procedure is well-known for partially ordered vector spaces.

Definition 1.19. Let $(Z, +, \cdot)$ be a conlinear space and let $\leq$ be a partial ordering on the set Z. $(Z, +, \cdot, \leq)$ is called a *partially ordered conlinear space* if for every $z^1, z^2, z \in Z$ and every $\alpha \in \mathbb{R}_+$ the following conditions hold:

(O1) $z^1 \leq z^2 \implies z^1 + z \leq z^2 + z,$
(O2) $z^1 \leq z^2 \implies \alpha z^1 \leq \alpha z^2.$

Of course, a partially ordered vector space is a special case of a partially ordered conlinear space. Let us define convex functions in the general setting of conlinear spaces.

Definition 1.20. Let W be a conlinear space and let Z be a partially ordered conlinear space. A function $f : W \to Z$ is said to be *convex* if

$$\begin{aligned} &\forall \lambda \in [0,1], \forall w^1, w^2 \in W : \\ &\quad f\big(\lambda \cdot w^1 + (1-\lambda) \cdot w^2\big) \leq \lambda \cdot f(w^1) + (1-\lambda) \cdot f(w^2). \end{aligned}$$

Concave functions are defined likewise.

1.3 Topological vector spaces

A well-known concept is that of a topological vector space, also called topological linear space or linear topological space (see e.g. Kelley *et al.*, 1963; Köthe, 1966; Schaefer, 1980). The idea is to equip a linear space with a topology and to require that the topology is compatible with the linear structure. Many useful results depend on this compatibility.

Definition 1.21. Let Y be a real linear space (vector space) and let τ be a topology on Y. The pair (Y, τ) is called a *topological vector space* (or *linear topological space*) if the following two axioms are satisfied:

(L1) $(y^1, y^2) \longmapsto y^1 + y^2$ is continuous on $Y \times Y$ into Y,
(L2) $(\lambda, y) \longmapsto \lambda y$ is continuous on $\mathbb{R} \times Y$ into Y.

If there is no risk of confusion, a topological vector space (Y, τ) is simply denoted by Y. We write $\operatorname{int} A$ and $\operatorname{cl} A$, respectively, for the interior and closure of a subset A of a topological vector space Y. The boundary of $A \subseteq Y$ is the set $\operatorname{bd} A := \operatorname{cl} A \setminus \operatorname{int} A$.

Proposition 1.22. *Let Y be a topological vector space.*

(i) *For any subset A of Y and any base $\mathcal{U}$ of the neighborhood filter of $0 \in Y$, the closure of A is given by*

$$\operatorname{cl} A = \bigcap_{U \in \mathcal{U}} A + U.$$

(ii) *If A is an open subset of Y and B is any subset of Y, then $A + B$ is open.*

Proof. See e.g. Schaefer (1980). □

A collection of useful results concerning the closure of the convex hull is next provided.

Proposition 1.23. *Let Y be a topological vector space, let $A_i, A, B \subseteq Y$ and let I be an arbitrary index set, then*

(i) $\operatorname{cl}\operatorname{co} A \supseteq \operatorname{co}\operatorname{cl} A$,
(ii) $\operatorname{cl}\operatorname{co} A = \operatorname{cl}\operatorname{co}\operatorname{cl} A$,
(iii) $\operatorname{cl}\operatorname{co} \bigcap_{i\in I} A_i \subseteq \bigcap_{i\in I} \operatorname{cl}\operatorname{co} A_i$,
(iv) $\operatorname{cl}\operatorname{co} \bigcup_{i\in I} A_i \supseteq \bigcup_{i\in I} \operatorname{cl}\operatorname{co} A_i$,
(v) $\operatorname{cl}\operatorname{co} \bigcup_{i\in I} A_i = \operatorname{cl}\operatorname{co} \bigcup_{i\in I} \operatorname{cl}\operatorname{co} A_i$,
(vi) $\operatorname{cl}(A+B) \supseteq \operatorname{cl} A + \operatorname{cl} B$,
(vii) $\operatorname{cl}(A+B) = \operatorname{cl}(\operatorname{cl} A + \operatorname{cl} B)$,
(viii) $\operatorname{cl}\operatorname{co}(A+B) \supseteq \operatorname{cl}\operatorname{co} A + \operatorname{cl}\operatorname{co} B$,
(ix) $\operatorname{cl}\operatorname{co}(A+B) = \operatorname{cl}(\operatorname{cl}\operatorname{co} A + \operatorname{cl}\operatorname{co} B)$.

Proof. These statements are standard in the literature. A collection of detailed proofs can be found in (Löhne, 2005a). □

Proposition 1.24. *Let Y be a topological vector space, $A, B \subseteq Y$ and* $\operatorname{int} B \neq 0$. *Then*

$$A + \operatorname{int} B = \operatorname{cl} A + \operatorname{int} B.$$

Proof. Of course, $A + \operatorname{int} B \subseteq \operatorname{cl} A + \operatorname{int} B$. To show the opposite inclusion let $y \in \operatorname{cl} A + \operatorname{int} B$. We have $y - a \in \operatorname{int} B$ for some $a \in \operatorname{cl} A$. For all neighborhoods U of 0 there exists some $\bar{a} \in A$ such that $-\bar{a} \in \{-a\} + U$. Since $\operatorname{int} B$ is nonempty and open, there exists some neighborhood U of 0 such that $y - \bar{a} \in \{y - a\} + U \subseteq \operatorname{int} B$. □

The following statements require convexity.

Proposition 1.25. *Let Y be a topological vector space and let $A \subseteq Y$ be convex. Then*

(i) $\operatorname{cl} A$ *is convex,*
(ii) $\operatorname{int} A$ *is convex,*
(iii) $\operatorname{int} A \neq \emptyset$ *implies that* $\operatorname{cl} A = \operatorname{cl}(\operatorname{int} A)$ *and* $\operatorname{int}(\operatorname{cl} A) = \operatorname{int} A$.

Proof. See e.g. (Schaefer, 1980, page 38). □

The next result is useful to show the existence of boundary points, in particular, it is used in Section 1.4 to show the existence of weakly minimal points and to prove several results on infimal sets.

Theorem 1.26. *Let A and B be subsets of a topological vector space Y. Let B be convex and assume that* $\operatorname{cl} A \cap B \neq \emptyset$. *Then*

$$\operatorname{cl} A \cap B \subseteq \operatorname{int} A \quad \Longrightarrow \quad A \supseteq B.$$

Proof. Let $a \in \operatorname{cl} A \cap B$ and assume there is some $b \in B \setminus A$. We consider the expression

$$\bar{\lambda} := \inf\{\lambda \geq 0 |\ \lambda a + (1-\lambda)b \in A\}.$$

There exists a sequence $(\lambda_n) \to \bar{\lambda}$ such that $\lambda_n a + (1-\lambda_n)b \in A$ for all $n \in \mathbb{N}$. As B is convex and $\bar{\lambda} \in [0,1]$, we get

$$\bar{\lambda}a + (1-\bar{\lambda})b \in \operatorname{cl} A \cap B \subseteq \operatorname{int} A.$$

In particular, we see that $\bar{\lambda} > 0$. On the other hand, there is a sequence $(\mu_n) \to \bar{\lambda}$ such that $\mu_n a + (1-\mu_n)b \notin A$ for all $n \in \mathbb{N}$. This yields

$$\bar{\lambda}a + (1-\bar{\lambda})b \notin \operatorname{int} A.$$

This is a contradiction. □

A topological vector space is now equipped with a partial ordering $\leq$ on Y. It is required that the linear structure is compatible with both the topology and the ordering.

Definition 1.27. Let Y be a vector space, τ a topology and $\leq$ a partial ordering on Y. The triple $(Y, \tau, \leq)$ is called a *partially ordered topological vector space* if the axioms (L1), (L2), (O1) and (O2) are satisfied.

It is well-known that in a partially ordered vector space $(Y, \leq)$ (this notion involves the axioms (O1) and (O2) from Definition 1.19) the cone $C := \{y \in Y |\ y \geq 0\}$ is convex and pointed (i.e. $C \cap (-C) = \{0\}$). Vice versa, given a convex, pointed cone $C \subseteq Y$, we obtain by the relation

$$x \leq_C y \quad :\Longleftrightarrow \quad y - x \in C$$

a partially ordered space $(Y, \leq_C)$ with $C := \{y \in Y |\ y \geq_C 0\}$, see e.g. Peressini (1967). The situation in conlinear spaces is more complicated, see Hamel (2005).

Let Y be a partially ordered vector space. We consider the extension by two new elements $+\infty$ and $-\infty$. The extended space is denoted by $\overline{Y} := Y \cup \{\pm\infty\}$. The ordering is extended as

$$\forall y \in Y: \quad -\infty \leq y \leq +\infty. \tag{1.4}$$

The linear operations are extended by the following calculus rules:

$$0 \cdot (+\infty) = 0, \quad 0 \cdot (-\infty) = 0, \tag{1.5}$$
$$\forall \alpha > 0: \quad \alpha \cdot (+\infty) = +\infty, \tag{1.6}$$
$$\forall \alpha > 0: \quad \alpha \cdot (-\infty) = -\infty, \tag{1.7}$$
$$\forall y \in \overline{Y}: \quad y + (+\infty) = +\infty + y = +\infty, \tag{1.8}$$
$$\forall y \in Y \cup \{-\infty\}: \quad y + (-\infty) = -\infty + y = -\infty. \tag{1.9}$$

The extended space $\overline{Y}$ is a conlinear space, but not a linear space. This can be seen as there is no inverse element for $+\infty$ and $-\infty$. Of course, $+\infty$ and $-\infty$ play the role of the greatest and least element in $\left(\overline{Y}, \leq\right)$, respectively. The following result shows that a linear extension by a greatest or least element is generally not possible.

Theorem 1.28. *There is no partially ordered vector space besides the trivial case $Y = \{0\}$ having a greatest element.*

Proof. Let Y be a partially ordered vector space and let l be the greatest element. Then we have $l + l \leq l$. It follows $l \leq 0$ and hence $l = 0$. Let $y \in Y$. Then

$$\left.\begin{array}{rcr} y + y \leq l = 0 & \Longrightarrow & y \leq -y \\ -y - y \leq l = 0 & \Longrightarrow & -y \leq y \end{array}\right\} \Longrightarrow y = -y \Longrightarrow y = 0.$$

Hence, $Y = \{0\}$. □

In contrast to many authors, we do not avoid the sum of $-\infty$ and $+\infty$. Our rules imply the so-called *inf-addition* (see Rockafellar and Wets, 1998)

$$+\infty + (-\infty) = -\infty + (+\infty) = +\infty.$$

The inf-addition is the preferable choice for minimization problems. Analogously, one can define the sup-addition, where we have to permute the role of $-\infty$ and $+\infty$ with respect to the addition. The calculus rules are analogous to the above rules.

Definition 1.29. Let $(Y, \tau, \leq)$ be a partially ordered topological vector space and let $\overline{Y} := Y \cup \{\pm\infty\}$. If the linear operations on Y are extended to $\overline{Y}$ by the rules (1.5) to (1.9) and if the partial ordering on Y is extended to $\overline{Y}$ by (1.4), then the triple $\left(\overline{Y}, \tau, \leq\right)$ is called an *extended partially ordered topological vector space*. We also say that $\overline{Y}$ is the *extension of the partially ordered topological vector space Y by the inf-addition*, shortly, the *extension* of Y.

We next consider the convex hull of subsets of an extended partially ordered (topological) vector space.

Proposition 1.30. *Let $\overline{Y}$ be an extended partially ordered (topological) vector space. The convex hull of a subset $A \subseteq \overline{Y}$ can be expressed as*

$$\operatorname{co} A = \operatorname{co}\left(A \setminus \{\pm\infty\}\right) \cup \left(A \cap \{\pm\infty\}\right).$$

Proof. This is elementary. □

Example 1.31. Let $\mathcal{P}(Y)$ be the power set of a vector space Y. The *Minkowski-addition* is defined as

$$+ \; : \; \mathcal{P}(Y) \times \mathcal{P}(Y) \to \mathcal{P}(Y),$$

$$A + B = \{y \in Y | \; \exists a \in A, \; \exists b \in B : \; y = a + b\}.$$

We define a multiplication by nonnegative real numbers by

$$\cdot \; : \; \mathbb{R}_+ \times \mathcal{P}(Y) \to \mathcal{P}(Y),$$

$$\alpha \cdot A = \begin{cases} \{y \in Y | \; \exists a \in A : y = \alpha a\} & \text{if} \;\; \alpha > 0 \\ \{0\} & \text{if} \;\; \alpha = 0. \end{cases}$$

In particular, this implies the rule $0 \cdot \emptyset = \{0\}$. The triple $(\mathcal{P}(Y), +, \cdot)$ provides a conlinear space. Moreover, the set inclusion $\supseteq$ is a partial ordering on $\mathcal{P}(Y)$ and $(\mathcal{P}(Y), +, \cdot, \supseteq)$ provides a partially ordered conlinear space.

Example 1.32. Let $\mathcal{P}\left(\overline{Y}\right)$ be the power set of an extended (partially ordered) vector space $\overline{Y}$, whereas the ordering relation is not relevant for the moment. We underlie the rules (1.5) to (1.9) (i.e. the inf-addition). Using the operations

$$+ \; : \; \mathcal{P}\left(\overline{Y}\right) \times \mathcal{P}\left(\overline{Y}\right) \to \mathcal{P}\left(\overline{Y}\right),$$

$$A + B = \left\{y \in \overline{Y} | \; \exists a \in A, \; \exists b \in B : \; y = a + b\right\},$$

and

$$\cdot \; : \; \mathbb{R}_+ \times \mathcal{P}\left(\overline{Y}\right) \to \mathcal{P}\left(\overline{Y}\right),$$

$$\alpha \cdot A = \begin{cases} \{y \in Y | \; \exists a \in A : y = \alpha a\} & \text{if} \;\; \alpha > 0 \\ \{0\} & \text{if} \;\; \alpha = 0, \end{cases}$$

we obtain a conlinear space $\left(\mathcal{P}\left(\overline{Y}\right), +, \cdot\right)$. Moreover, $\left(\mathcal{P}\left(\overline{Y}\right), +, \cdot, \supseteq\right)$ provides a partially ordered conlinear space with the greatest element $\emptyset$ and the least element $\overline{Y}$. In particular we have the rules $0 \cdot \emptyset = \{0\}$, $0 \cdot \{+\infty, -\infty\} = \{0\}$, $\{+\infty\} + \{-\infty\} = \{+\infty\}$.

We continue with a well-known separation theorem for topological vector spaces.

Definition 1.33. Let Y be a linear space. We say a nonzero linear functional $y^* : Y \to \mathbb{R}$ *properly separates* two subsets A and B of Y if

$$\sup_{x \in A} y^*(x) \leq \inf_{y \in B} y^*(y) \quad \text{and} \quad \inf_{x \in A} y^*(x) < \sup_{y \in B} y^*(y).$$

Theorem 1.34. *Let Y be a topological vector space. Two nonempty disjoint convex subsets of Y can be properly separated by a nonzero continuous linear functional, provided one of them has an interior point.*

Proof. See e.g. (Aliprantis and Border, 1994, Theorem 4.46).

Let Y^* be the topological dual of the topological vector space (Y, τ), i.e., Y^* is the set of all linear continuous functionals on Y. The *polar cone* of a cone $C \subseteq Y$ is defined by

$$C^\circ := \{y^* \in Y^* |\ \forall y \in C : y^*(y) \le 0\}\,. \tag{1.10}$$

We will assume frequently that Y is a partially ordered topological vector space, where the ordering is generated by a cone C (which is consequently convex and pointed) such that $\emptyset \ne \operatorname{int} C \ne Y$. Theorem 1.34 implies that the polar cone of C contains nonzero elements.

We continue with an important subclass of topological vector spaces. The essential idea is that a stronger hypothesis leads to a stronger separation property.

Definition 1.35. A topological vector space (Y, τ) is said to be a *locally convex space* if every neighborhood of zero includes a convex neighborhood of zero.

Definition 1.36. Let Y be a vector space. We say a nonzero linear functional $y^* : Y \to \mathbb{R}$ *strongly separates* two subsets A and B of Y if

$$\sup_{x \in A} y^*(x) < \inf_{y \in B} y^*(y).$$

Theorem 1.37. *Let Y be a locally convex space. Two disjoint nonempty convex subsets of Y can be strongly separated by a nonzero continuous linear functional, provided one is compact and the other closed.*

Proof. See e.g. (Aliprantis and Border, 1994, Theorem 4.54).

Let (Y, τ) be a locally convex space and let Y^* be the topological dual. The support function $\sigma_A : Y^* \longrightarrow \overline{\mathbb{R}}$ with respect to $A \subseteq Y$ is defined by

$$\sigma_A(y^*) := \sigma(y^* |\, A) := \sup_{y \in A} y^*(y),$$

where $\overline{\mathbb{R}}$ is equipped with the sup-addition, in particular this means

$$-\infty + (+\infty) = +\infty + (-\infty) = -\infty.$$

Moreover, this ensures that the expression

$$\forall y^* \in Y^* : \quad \sigma(y^* |\, A + B) = \sigma(y^* |\, A) + \sigma(y^* |\, B)$$

is valid for all (not necessarily nonempty) sets $A, B \subseteq Y$. Of course,

$$\forall y^* \in Y^* : \quad \sigma(y^* |\, \alpha A) = \alpha \sigma(y^* |\, A)$$

holds for all sets $A \subseteq Y$ and all nonnegative real numbers $\alpha \ge 0$, where we use the conventions $0 \cdot \emptyset = \{0\}$ and $0 \cdot (\pm\infty) = 0$. From Theorem 1.37, we immediately obtain for all $A, B \subseteq Y$ the important assertion

$$A \subseteq \operatorname{cl} \operatorname{co} B \quad \Longleftrightarrow \quad \forall y^* \in Y^* :\ \sigma_A(y^*) \le \sigma_B(y^*).$$

If $A_i \subseteq Y$ for all $i \in I$, where I is an arbitrary index set, then

$$\forall y^* \in Y^* : \quad \sigma\left(y^* \middle| \bigcup_{i\in I} A_i\right) = \sup_{i\in I} \sigma\left(y^* \middle| A_i\right)$$

and

$$\forall y^* \in Y^* : \sigma\left(y^* \middle| \bigcap_{i\in I} A_i\right) \le \inf_{i\in I} \sigma\left(y^* \middle| A_i\right).$$

A net (y_α) in Y is convergent to $y \in Y$ with respect to the *weak topology* w if $y^*(y_\alpha)$ converges to $y^*(y)$ in $\mathbb{R}$ for all $y^* \in Y^*$. A net (y^*_α) in Y^* is convergent to $y^* \in Y^*$ with respect to the *weak* topology* w^* if $y^*_\alpha(y)$ converges to $y^*(y)$ for all $y \in Y$. Both (Y, w) and (Y^*, w^*) are Hausdorff (i.e., different elements have different neighborhoods) locally convex spaces. A linear functional ϕ on Y^* is weak*-continuous if and only if there exists some $y \in Y$ such that $\phi(y^*) = y^*(y)$ for all $y^* \in Y^*$ (see e.g. Kelley *et al.*, 1963, Theorem 7.6).

Definition 1.38. A *dual pair* $\langle Y, Y^* \rangle$ is a pair of vector spaces (Y, Y^*) together with a function $(y, y^*) \mapsto \langle y, y^* \rangle$, from $Y \times Y^*$ into $\mathbb{R}$ (called the *duality* of the pair), satisfying:

(i) The mapping $y^* \mapsto \langle y, y^* \rangle$ is linear for each $y \in Y$;
(ii) The mapping $y \mapsto \langle y, y^* \rangle$ is linear for each $y^* \in Y^*$;
(iii) If $\langle y, y^* \rangle = 0$ for each $y^* \in Y^*$, then $y = 0$;
(iv) If $\langle y, y^* \rangle = 0$ for each $y \in Y$, then $y^* = 0$.

If (Y, τ) is a Hausdorff locally convex space, and Y^* is its topological dual, i.e., Y^* is the set of all linear continuous functionals on Y, then $\langle Y, Y^* \rangle$ provides a dual pair (follows from Theorem 1.37). Vice versa, every dual pair $\langle Y, Y^* \rangle$ is obtained from a locally convex Hausdorff space (Y, τ) and its topological dual Y^* (compare Aliprantis and Border, 1994, Theorem 4.69). Locally convex spaces are naturally obtained by the weak topology, which is induced by a family of linear functionals. A locally convex Hausdorff topology τ is said to be *consistent* with the dual pair $\langle Y, Y^* \rangle$ if the topological dual of (Y, τ) is just Y^*. All topologies consistent with a given dual pair have the same closed convex sets. If (Y, τ) is a locally convex Hausdorff space and Y^* is its topological dual, then both τ and the weak topology w are consistent with the dual pair $\langle Y, Y^* \rangle$.

1.4 Infimal and supremal sets

This section is devoted to infimal and supremal sets of subsets of a topological vector space Y and its extension $\overline{Y}$. These concepts were introduced by Nieuwenhuis (1980) as a generalization of the infimum and supremum in

$\overline{\mathbb{R}}$ to a partially ordered vector space Y. An infimal subset of Y cannot be an infimum in Y. This already follows from the fact that an infimum is an element and not a subset of the underlying partially ordered set. However, it turned out that infimal and supremal sets are essential for the construction of a complete lattice, which can be used in vector optimization. This complete lattice is not $\overline{Y}$ but a subset of the power set of $\overline{Y}$.

Definition 1.39. Let $\overline{Y}$ be an extended partially ordered topological vector space and let the ordering cone C of Y satisfy $\emptyset \neq \operatorname{int} C \neq Y$. The *upper closure* of a subset $A \subseteq \overline{Y}$ (with respect to C) is defined by

$$\mathrm{Cl}_+ A := \begin{cases} Y & \text{if } -\infty \in A \\ \emptyset & \text{if } A = \{+\infty\} \\ \{y \in Y |\ \{y\} + \operatorname{int} C \subseteq A \setminus \{+\infty\} + \operatorname{int} C\} & \text{otherwise.} \end{cases}$$

The upper closure is illustrated in Figure 1.2. It is clear from the definition that $\mathrm{Cl}_+ A$ is always a subset of Y even if A is a subset of the extended space $\overline{Y}$. The following characterization is useful and can serve as a definition, when the ordering cone has an empty interior.

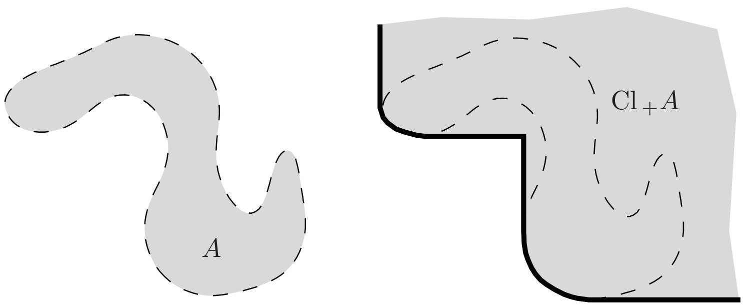

Fig. 1.2 The upper closure of a set $A \subseteq \mathbb{R}^2$ with respect to the ordering cone $C = \mathbb{R}^2_+$.

Proposition 1.40. *Let $\overline{Y}$ be an extended partially ordered topological vector space with an ordering cone C such that $\emptyset \neq \operatorname{int} C \neq Y$ and let $A \subseteq \overline{Y}$. Then*

$$\mathrm{Cl}_+ A = \begin{cases} Y & \text{if } -\infty \in A \\ \emptyset & \text{if } A = \{+\infty\} \\ \mathrm{cl}\,\big(A \setminus \{+\infty\} + C\big) & \text{otherwise.} \end{cases}$$

Proof. Without loss of generality we can assume that $+\infty \notin A$. It remains to show that

$$B^1 := \{y \in Y |\ \{y\} + \operatorname{int} C \subseteq A + \operatorname{int} C\} = \mathrm{cl}\,\big(A + C\big) =: B^2.$$

(i) Let $y \in B^1$ and let $c \in \operatorname{int} C$. We have $y + \frac{1}{n}c \in A + \operatorname{int} C$ for all $n \in \mathbb{N}$. Taking the limit for $n \to \infty$, we obtain $y \in B^2$.

(ii) Let $y \in B^2$. Using Proposition 1.24 we obtain

$$\{y\} + \operatorname{int} C \subseteq \operatorname{cl}\big(A + C\big) + \operatorname{int} C = A + C + \operatorname{int} C = A + \operatorname{int} C,$$

and hence $y \in B^1$. □

Definition 1.41. Let $\overline{Y}$ be an extended partially ordered topological vector space with an ordering cone C such that $\emptyset \neq \operatorname{int} C \neq Y$. The set of *weakly minimal* points of a subset $A \subseteq Y$ (with respect to C) is defined by

$$\operatorname{wMin} A := \{y \in A |\ (\{y\} - \operatorname{int} C) \cap A = \emptyset\}.$$

The next lemma shows the existence of weakly minimal elements. It is based on Theorem 1.26.

Lemma 1.42. *Let $\overline{Y}$ be an extended partially ordered topological vector space, ordered by a cone C with $\emptyset \neq \operatorname{int} C \neq Y$. Let $A \subseteq \overline{Y}$ be an arbitrary set and let $B \subseteq Y$ be a convex set. Let $\operatorname{Cl}_+ A \cap B \neq \emptyset$ and $B \setminus \operatorname{Cl}_+ A \neq \emptyset$. Then*

$$\operatorname{wMin}(\operatorname{Cl}_+ A \cap B) \neq \emptyset.$$

Proof. Assuming that $\operatorname{wMin}(\operatorname{Cl}_+ A \cap B)$ is empty, we get

$$\forall y \in \operatorname{Cl}_+ A \cap B,\ \exists z \in \operatorname{Cl}_+ A \cap B: \quad y \in \{z\} + \operatorname{int} C.$$

This implies

$$\operatorname{Cl}_+ A \cap B \subseteq (\operatorname{Cl}_+ A \cap B) + \operatorname{int} C.$$

It follows

$$\begin{aligned} \operatorname{Cl}_+ A \cap B &\subseteq (\operatorname{Cl}_+ A \cap B) + \operatorname{int} C \\ &\subseteq (\operatorname{Cl}_+ A + \operatorname{int} C) \cap (B + \operatorname{int} C) \\ &\subseteq \operatorname{Cl}_+ A + \operatorname{int} C \\ &\subseteq \operatorname{int} \operatorname{Cl}_+ A. \end{aligned}$$

Since B is convex, Theorem 1.26 yields $\operatorname{Cl}_+ A \supseteq B$, which contradicts the assumption $B \setminus \operatorname{Cl}_+ A \neq \emptyset$. □

We continue with a further lemma concerning weakly minimal elements.

Lemma 1.43. *Let $\overline{Y}$ be an extended partially ordered topological vector space ordered by a cone C with $\emptyset \neq \operatorname{int} C \neq Y$. Let $A \subseteq \overline{Y}$ be an arbitrary set and let $B \subseteq Y$ be an open set. Then*

$$\operatorname{wMin}(\operatorname{Cl}_+ A \cap B) = (\operatorname{wMin} \operatorname{Cl}_+ A) \cap B.$$

Proof. In order to prove the inclusion $\mathrm{wMin}(\mathrm{Cl}_+ A \cap B) \subseteq (\mathrm{wMin}\,\mathrm{Cl}_+ A) \cap B$, let $y \in \mathrm{wMin}(\mathrm{Cl}_+ A \cap B)$. Of course, this implies $y \in B$ and $y \in \mathrm{Cl}_+ A$. It remains to show that $(\{y\} - \operatorname{int} C) \cap \mathrm{Cl}_+ A = \emptyset$. Assuming the contrary, we get some $z \in \mathrm{Cl}_+ A$ such that $c := y - z \in \operatorname{int} C$. As B is open, there exists some $\varepsilon \in (0,1)$ such that $w := y - \varepsilon c \in B$. From $z \in \mathrm{Cl}_+ A$ we get $w \in \mathrm{Cl}_+ A + \operatorname{int} C \subseteq \mathrm{Cl}_+ A$. Thus, we have $w \in (\{y\} - \operatorname{int} C) \cap (\mathrm{Cl}_+ A \cap B)$ and hence $y \notin \mathrm{wMin}(\mathrm{Cl}_+ A \cap B)$, a contradiction.

The opposite inclusion $\supseteq$ follows directly from the definition. □

The following conclusion of Lemma 1.42 plays a crucial role in the following.

Corollary 1.44. *Let $\overline{Y}$ be an extended partially ordered topological vector space with an ordering cone C such that $\emptyset \neq \operatorname{int} C \neq Y$. For every set $A \subseteq \overline{Y}$ the following statements are equivalent:*

(i) $\emptyset \neq \mathrm{Cl}_+ A \neq Y$,
(ii) $\mathrm{wMin}\,\mathrm{Cl}_+ A \neq \emptyset$.

Proof. (i) $\Rightarrow$ (ii). Follows from Lemma 1.42 for the choice $B = Y$.

(ii) $\Rightarrow$ (i). This is a direct consequence of the definition of wMin. □

We now define a central concept of this book, an infimal set for a subset of the extended space $\overline{Y}$.

Definition 1.45. Let $\overline{Y}$ be an extended partially ordered topological vector space with an ordering cone C such that $\emptyset \neq \operatorname{int} C \neq Y$. The *infimal set* of $A \subseteq \overline{Y}$ (with respect to C) is defined by

$$\mathrm{Inf}\, A := \begin{cases} \mathrm{wMin}\,\mathrm{Cl}_+ A & \text{if } \emptyset \neq \mathrm{Cl}_+ A \neq Y \\ \{-\infty\} & \text{if } \mathrm{Cl}_+ A = Y \\ \{+\infty\} & \text{if } \mathrm{Cl}_+ A = \emptyset. \end{cases}$$

An infimal set is illustrated in Figure 1.3. If A is a nonempty subset of Y and $\mathrm{cl}\,(A+C) \neq Y$, then $\mathrm{Inf}\, A = \mathrm{wMin}\,\mathrm{cl}\,(A+C)$. By Corollary 1.44, $\mathrm{Inf}\, A$ is always a nonempty set. Clearly, if $-\infty$ belongs to A, we have $\mathrm{Inf}\, A = \{-\infty\}$, in particular, $\mathrm{Inf}\,\{-\infty\} = \{-\infty\}$. Moreover, we have $\mathrm{Inf}\,\emptyset = \mathrm{Inf}\,\{+\infty\} = \{+\infty\}$. Furthermore, $\mathrm{Cl}_+ A = \mathrm{Cl}_+(A \cup \{+\infty\})$ holds and hence $\mathrm{Inf}\, A = \mathrm{Inf}(A \cup \{+\infty\})$ for all $A \subseteq \overline{Y}$.

We close this section with several useful results concerning the infimal set and the upper closure. We derive them from Lemma 1.42 (based on Theorem 1.26) and Lemma 1.43.

Lemma 1.46. *Let $\overline{Y}$ be an extended partially ordered topological vector space with an ordering cone C such that $\emptyset \neq \operatorname{int} C \neq Y$. For $A \subseteq \overline{Y}$ with $\emptyset \neq \mathrm{Cl}_+ A \neq Y$, it is true that*

$$\mathrm{Cl}_+ A + \operatorname{int} C \subseteq \mathrm{Inf}\, A + \operatorname{int} C.$$

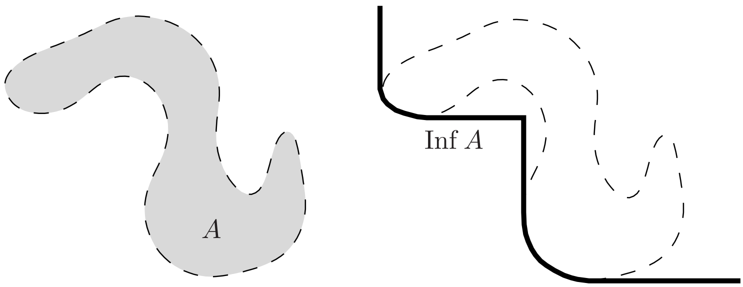

Fig. 1.3 The infimal set of a set $A \subseteq \mathbb{R}^2$ with respect to the ordering cone $C = \mathbb{R}^2_+$.

Proof. Let $y \in \operatorname{Cl}_+ A + \operatorname{int} C$, then $\big(\{y\} - \operatorname{int} C\big) \cap \operatorname{Cl}_+ A \neq \emptyset$. We set $B := \{y\} - \operatorname{int} C$. As B is convex and open, Lemma 1.42 and Lemma 1.43 imply that

$$\emptyset \neq \operatorname{wMin}(\operatorname{Cl}_+ A \cap B) = (\operatorname{wMin} \operatorname{Cl}_+ A) \cap B.$$

Thus, there exists some $z \in \operatorname{wMin} \operatorname{Cl}_+ A = \operatorname{Inf} A$ such that $y \in \{z\} + \operatorname{int} C$, whence $y \in \operatorname{Inf} A + \operatorname{int} C$. □

Lemma 1.47. *Let $\overline{Y}$ be an extended partially ordered topological vector space with an ordering cone C such that $\emptyset \neq \operatorname{int} C \neq Y$. For $A \subseteq \overline{Y}$ with $\emptyset \neq \operatorname{Cl}_+ A \neq Y$, the following statement holds true:*

$$\operatorname{Cl}_+ A \;\cup\; (\operatorname{Inf} A - \operatorname{int} C) = Y.$$

Proof. We have $\operatorname{Cl}_+ A - \operatorname{int} C \supseteq \{z\} + \operatorname{int} C - \operatorname{int} C = Y$ for every $z \in \operatorname{Cl}_+ A$. Let $y \in Y \setminus \operatorname{Cl}_+ A$. The set $B := \{y\} + \operatorname{int} C$ is open and convex. Moreover, we have $\operatorname{Cl}_+ A \cap B \neq \emptyset$, since otherwise we get the contradiction $y \notin \operatorname{Cl}_+ A - \operatorname{int} C = Y$. We show that $B \setminus \operatorname{Cl}_+ A \neq \emptyset$. Indeed, assuming the contrary, we obtain $B \subseteq \operatorname{Cl}_+ A$ which implies the contradiction $y \in \operatorname{cl} B \subseteq \operatorname{Cl}_+ A$. Lemma 1.42 and Lemma 1.43 imply

$$\emptyset \neq \operatorname{wMin}(\operatorname{Cl}_+ A \cap B) = (\operatorname{wMin} \operatorname{Cl}_+ A) \cap B.$$

Consequently, there exists some $z \in \operatorname{wMin} \operatorname{Cl}_+ A = \operatorname{Inf} A$ such that $z \in B = \{y\} + \operatorname{int} C$, whence $y \in \operatorname{Inf} A - \operatorname{int} C$. □

The next corollary provides several useful consequences of the preceding two lemmas.

Corollary 1.48. *Let $\overline{Y}$ be an extended partially ordered topological vector space with an ordering cone C such that $\emptyset \neq \operatorname{int} C \neq Y$. If $A, B \subseteq \overline{Y}$ with $\emptyset \neq \operatorname{Cl}_+ A \neq Y$ and $\emptyset \neq \operatorname{Cl}_+ B \neq Y$, then*

(i) $\qquad \operatorname{Cl}_+ A + \operatorname{int} C = \operatorname{Inf} A + \operatorname{int} C,$

(ii) $\operatorname{Inf} A = \{y \in Y |\ \{y\} + \operatorname{int} C \subseteq \operatorname{Cl}_+ A + \operatorname{int} C \ \wedge\ y \notin \operatorname{Cl}_+ A + \operatorname{int} C\}$,
(iii) $\operatorname{int} \operatorname{Cl}_+ A = \operatorname{Cl}_+ A + \operatorname{int} C$,
(iv) $\operatorname{Inf} A = \operatorname{bd} \operatorname{Cl}_+ A$,
(v) $\operatorname{Inf} A = \operatorname{Cl}_+ A \setminus (\operatorname{Cl}_+ A + \operatorname{int} C)$,
(vi) $\operatorname{Cl}_+ A = \operatorname{Cl}_+ B \iff \operatorname{Inf} A = \operatorname{Inf} B$,
(vii) $\operatorname{Cl}_+ A = \operatorname{Cl}_+ B \iff \operatorname{Cl}_+ A + \operatorname{int} C = \operatorname{Cl}_+ B + \operatorname{int} C$,
(viii) $\operatorname{Inf} A = \operatorname{Inf} B \iff \operatorname{Inf} A + \operatorname{int} C = \operatorname{Inf} B + \operatorname{int} C$,
(ix) $\operatorname{Cl}_+ A = \operatorname{Inf} A \cup (\operatorname{Inf} A + \operatorname{int} C)$,
(x) $\operatorname{Inf} A$, $(\operatorname{Inf} A - \operatorname{int} C)$ *and* $(\operatorname{Inf} A + \operatorname{int} C)$ *are pairwise disjoint,*
(xi) $\operatorname{Inf} A \cup (\operatorname{Inf} A - \operatorname{int} C) \cup (\operatorname{Inf} A + \operatorname{int} C) = Y$.

Proof. (i) We have $\operatorname{Inf} A = \operatorname{wMin} \operatorname{Cl}_+ A \subseteq \operatorname{Cl}_+ A$ and hence $\operatorname{Inf} A + \operatorname{int} C \subseteq \operatorname{Cl}_+ A + \operatorname{int} C$. The opposite inclusion is just the statement of Lemma 1.46.

(ii) Follows from the definitions of upper closure and weakly minimal points.

(iii) Let $y \in \operatorname{int} \operatorname{Cl}_+ A$ and let $c \in \operatorname{int} C$. There is some $t > 0$ such that $y - tc \in \operatorname{Cl}_+ A$ and hence $y \in \operatorname{Cl}_+ A + \operatorname{int} C$. On the other hand, we have $\operatorname{Cl}_+ A + \operatorname{int} C \subseteq \operatorname{Cl}_+ A$. The set $\operatorname{Cl}_+ A + \operatorname{int} C$ is open, whence $\operatorname{Cl}_+ A + \operatorname{int} C \subseteq \operatorname{int} \operatorname{Cl}_+ A$.

(iv) Follows from (ii) and (iii).

(v) Follows from (ii).

(vi) Taking the closure, $\operatorname{Cl}_+ A + \operatorname{int} C = \operatorname{Cl}_+ B + \operatorname{int} C$ implies $\operatorname{Cl}_+ A = \operatorname{Cl}_+ B$. By definition this yields $\operatorname{Inf} A = \operatorname{Inf} B$.

On the other hand, $\operatorname{Inf} A = \operatorname{Inf} B$ implies $\operatorname{Inf} A + \operatorname{int} C = \operatorname{Inf} B + \operatorname{int} C$ which is by (i) equivalent to $\operatorname{Cl}_+ A + \operatorname{int} C = \operatorname{Cl}_+ B + \operatorname{int} C$.

(vii) By Proposition 1.40, we have

$$\begin{aligned} \operatorname{cl}\left(\operatorname{Cl}_+ A + \operatorname{int} C\right) &= \operatorname{cl}\left(\operatorname{cl}\left(A \setminus \{+\infty\} + C\right) + \operatorname{int} C\right) \\ &= \operatorname{cl}\left(A \setminus \{+\infty\} + C\right) = \operatorname{Cl}_+ A. \end{aligned}$$

The statement is now obvious.

(viii) Follows from (i), (vi) and (vii).

(ix) Let $y \in \operatorname{Cl}_+ A$. In the case where $y \in \operatorname{Cl}_+ A + \operatorname{int} C$, (i) implies $y \in \operatorname{Inf} A + \operatorname{int} C$. Otherwise, if $y \notin \operatorname{Cl}_+ A + \operatorname{int} C$, we get $y \in \operatorname{wMin} \operatorname{Cl}_+ A = \operatorname{Inf} A$. On the other hand, it is obvious that $\operatorname{Inf} A = \operatorname{wMin} \operatorname{Cl}_+ A \subseteq \operatorname{Cl}_+ A$ and $\operatorname{Inf} A + \operatorname{int} C \subseteq \operatorname{Cl}_+ A + \operatorname{int} C \subseteq \operatorname{Cl}_+ A$.

(x) Let $y \in \operatorname{Inf} A - \operatorname{int} C$. There is some $z \in \operatorname{Inf} A$ such that $y \in \{z\} - \operatorname{int} C$. We have $(\{z\} - \operatorname{int} C) \cap \operatorname{Cl}_+ A = \emptyset$ and hence $y \notin \operatorname{Cl}_+ A$. By (ix) we get $(\operatorname{Inf} A - \operatorname{int} C) \cap \operatorname{Inf} A = \emptyset$ and $(\operatorname{Inf} A - \operatorname{int} C) \cap (\operatorname{Inf} A + \operatorname{int} C) = \emptyset$. From (i) and the definition of an infimal set we get $\operatorname{Inf} A \cap (\operatorname{Inf} A + \operatorname{int} C) = \emptyset$.

(xi) Follows from (ix) and Lemma 1.47. □

In the following result we use in (ii) and (iii) the calculus rules (1.5) to (1.9) from Section 1.3, in particular we use the inf-addition, which involves the rule $+\infty + (-\infty) = +\infty$.

Corollary 1.49. *Let* $A \subseteq \overline{Y}$. *Then*

(i) $\operatorname{Inf}\operatorname{Inf} A = \operatorname{Inf} A$, $\operatorname{Cl}_+\operatorname{Cl}_+ A = \operatorname{Cl}_+ A$, $\operatorname{Inf}\operatorname{Cl}_+ A = \operatorname{Inf} A$, $\operatorname{Cl}_+\operatorname{Inf} A = \operatorname{Cl}_+ A$,
(ii) $\operatorname{Inf}(\operatorname{Inf} A + \operatorname{Inf} B) = \operatorname{Inf}(A + B)$,
(iii) $\alpha \operatorname{Inf} A = \operatorname{Inf}(\alpha A)$ *for* $\alpha > 0$.

Proof. This follows from the definitions and results of this section. □

The sets $\operatorname{wMax} A$ of *weakly maximal elements* of $A \subseteq Y$, as well as the lower closure $\operatorname{Cl}_- A$ and the *supremal set* $\operatorname{Sup} A$ of a subset $A \subseteq \overline{Y}$ are defined likewise. One has

$$\operatorname{Sup} A = -\operatorname{Inf}(-A) \tag{1.11}$$

and analogous results hold true, where the sup-addition has to be used.

1.5 Hyperspaces of upper closed sets and self-infimal sets

In this section we introduce two complete lattices. They are shown to be partially ordered conlinear spaces which are isomorphic and isotone. It turns out that they are suitable to act as image spaces for vector optimization problems. We start with the definition of the elements these spaces.

Let $\overline{Y}$ be an extended partially ordered topological vector space with an ordering cone C such that $\emptyset \neq \operatorname{int} C \neq Y$. Note that C is automatically convex and pointed because we assume that C generates a partial ordering on Y.

Definition 1.50. A set $A \subseteq Y$ is called an *upper closed set* if $\operatorname{Cl}_+ A = A$. A subset $B \subseteq \overline{Y}$ is called *self-infimal* if $\operatorname{Inf} B = B$ holds.

Let $\mathcal{F} := \mathcal{F}_C(Y)$ be the family of all upper closed subsets of Y. In $\mathcal{F}$ we introduce an addition $\oplus : \mathcal{F} \times \mathcal{F} \to \mathcal{F}$ and a multiplication by nonnegative real numbers $\odot : \mathbb{R}_+ \times \mathcal{F} \to \mathcal{F}$ as

$$\begin{aligned} A^1 \oplus A^2 &:= \operatorname{cl}(A^1 + A^2), \\ \alpha \odot A &:= \operatorname{Cl}_+(\alpha \cdot A). \end{aligned}$$

Note that the power set $\mathcal{P}(Y)$ is supposed to be a conlinear space with respect to the Minkowski-addition and the usual multiplication by nonnegative numbers, compare Example 1.31. In particular, we use the rule $0 \cdot \emptyset = \{0\}$, which implies that $0 \odot \emptyset = \operatorname{Cl}_+\{0\} = \operatorname{cl} C$.

Proposition 1.51. *The space* $(\mathcal{F}, \oplus, \odot, \supseteq)$ *is a partially ordered conlinear space with the neutral element* $\operatorname{cl} C$.

Proof. Of course, the relation $\supseteq$ provides a partial ordering on $\mathcal{F}$. The axioms of a conlinear space can be verified directly. □

Let $\mathcal{I} := \mathcal{I}_C\left(\overline{Y}\right)$ be the family of all self-infimal subsets of $\overline{Y}$. In $\mathcal{I}$ we introduce an addition $\oplus : \mathcal{I} \times \mathcal{I} \to \mathcal{I}$, a multiplication by nonnegative real numbers $\odot : \mathbb{R}_+ \times \mathcal{I} \to \mathcal{I}$ and an order relation $\preccurlyeq$ by

$$\begin{aligned} B^1 \oplus B^2 &:= \operatorname{Inf}(B^1 + B^2), \\ \alpha \odot B &:= \operatorname{Inf}(\alpha \cdot B), \\ B^1 \preccurlyeq B^2 &:\iff \operatorname{Cl}_+ B^1 \supseteq \operatorname{Cl}_+ B^2. \end{aligned}$$

Note that the definition of the addition $\oplus$ in $\mathcal{I}$ is based on the inf-addition in $\overline{Y}$. As a consequence we obtain $\{-\infty\} \oplus \{+\infty\} = \{+\infty\}$. Moreover, we get $0 \odot B = \operatorname{Inf}\{0\} = \operatorname{bd} C$ for all $B \in \mathcal{I}$. The addition and the ordering in $\mathcal{I}$ are illustrated in Figure 1.4.

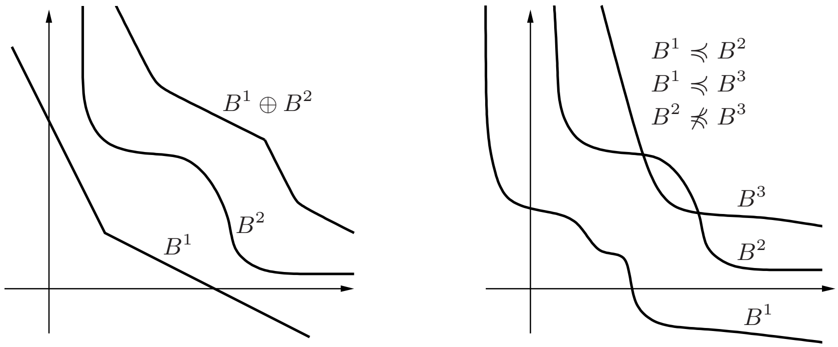

Fig. 1.4 The addition and the ordering in $\mathcal{I}$ for $C = \mathbb{R}^2_+$.

Proposition 1.52. *The space* $(\mathcal{I}, \oplus, \odot, \preccurlyeq)$ *is a partially ordered conlinear space with the neutral element* $\operatorname{Inf}\{0\} = \operatorname{bd} C$. *The spaces* $(\mathcal{F}, \oplus, \odot, \supseteq)$ *and* $(\mathcal{I}, \oplus, \odot, \preccurlyeq)$ *are isomorphic and isotone. The corresponding bijection is given by*

$$j : \mathcal{F} \to \mathcal{I}, \quad j(\,\cdot\,) = \operatorname{Inf}(\,\cdot\,), \quad j^{-1}(\,\cdot\,) = \operatorname{Cl}_+(\,\cdot\,).$$

Proof. By Corollary 1.49 (i), j is a bijection between $\mathcal{F}$ and $\mathcal{I}$. From Corollary 1.49 (ii) we obtain that $j\left(A^1\right) \oplus j\left(A^2\right) = j\left(A^1 \oplus A^2\right)$ for all $A^1, A^2 \in \mathcal{F}$. It can easily be verified that $\alpha \odot j(A) = j(\alpha \odot A)$ for all $\alpha \leq 0$ and all $A \in \mathcal{F}$. It follows from the definition of the ordering $\preccurlyeq$ in $\mathcal{I}$ that we have $A^1 \supseteq A^2$ if and only if $j\left(A^1\right) \preccurlyeq j\left(A^2\right)$. □

Proposition 1.53. $(\mathcal{F}, \supseteq)$ *and* $(\mathcal{I}, \preccurlyeq)$ *are complete lattices. For nonempty subsets* $\mathcal{A} \subseteq \mathcal{F}$ *and* $\mathcal{B} \subseteq \mathcal{I}$ *the infimum and supremum can be expressed by*

$$\inf \mathcal{A} = \operatorname{cl} \bigcup_{A \in \mathcal{A}} A, \qquad \sup \mathcal{A} = \bigcap_{A \in \mathcal{A}} A,$$

$$\inf \mathcal{B} = \operatorname{Inf} \bigcup_{B \in \mathcal{B}} \operatorname{Cl}_+ B, \qquad \sup \mathcal{B} = \operatorname{Inf} \bigcap_{B \in \mathcal{B}} \operatorname{Cl}_+ B.$$

Proof. For the space $(\mathcal{F}, \supseteq)$ the statements are obvious and for $(\mathcal{I}, \preccurlyeq)$ they follow from Proposition 1.52. □

As usual, if $\mathcal{A} \subseteq \mathcal{F}$ and $\mathcal{B} \subseteq \mathcal{I}$ are empty, we define the infimum (supremum) to be the greatest (least) element in the corresponding complete lattice, i.e., $\inf \mathcal{A} = \emptyset$, $\sup \mathcal{A} = Y$, $\inf \mathcal{B} = \{+\infty\}$ and $\sup \mathcal{B} = \{-\infty\}$.

It follows the main result of this section, which is illustrated in Figure 1.5.

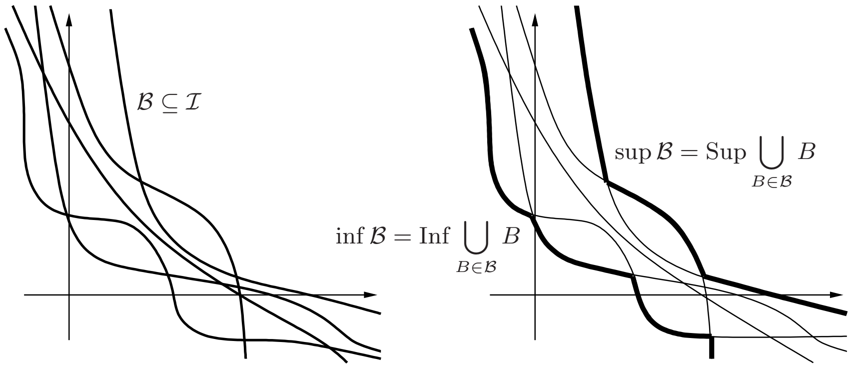

Fig. 1.5 The infimum and supremum in $\mathcal{I}$ for $C = \mathbb{R}^2_+$

Theorem 1.54. *For nonempty sets* $\mathcal{B} \subseteq \mathcal{I}$*, we have*

$$\inf \mathcal{B} = \operatorname{Inf} \bigcup_{B \in \mathcal{B}} B, \qquad \sup \mathcal{B} = \operatorname{Sup} \bigcup_{B \in \mathcal{B}} B.$$

Proof. (i) The expression for the infimum can be shown as follows:

$$\begin{aligned} \inf \mathcal{B} &= \operatorname{Inf} \bigcup_{B \in \mathcal{B}} \operatorname{Cl}_+ B = \operatorname{Inf} \operatorname{Cl}_+ \bigcup_{B \in \mathcal{B}} \operatorname{Cl}_+ B \\ &= \operatorname{Inf} \operatorname{Cl}_+ \bigcup_{B \in \mathcal{B}} B = \operatorname{Inf} \bigcup_{B \in \mathcal{B}} B. \end{aligned}$$

(ii) Let us prove the expression for the supremum. By Proposition 1.53, it remains to show that $\operatorname{Sup}\bigcup_{B\in\mathcal{B}} B = \operatorname{Inf}\bigcap_{B\in\mathcal{B}} \operatorname{Cl}_+ B$. We distinguish three cases:

a) If $\{+\infty\} \in \mathcal{B}$, we have $+\infty \in \bigcup_{B\in\mathcal{B}} B$ and hence $\operatorname{Sup}\bigcup_{B\in\mathcal{B}} B = \{+\infty\}$. On the other hand, since $\operatorname{Cl}_+\{+\infty\} = \emptyset$, we have $\operatorname{Inf}\bigcap_{B\in\mathcal{B}} \operatorname{Cl}_+ B = \operatorname{Inf}\emptyset = \{+\infty\}$.

b) Let $\{+\infty\} \notin \mathcal{B}$ but $\{-\infty\} \in \mathcal{B}$. If $\{-\infty\}$ is the only element in $\mathcal{B}$, the assertion is obvious. Otherwise we can omit this element without changing the expressions.

c) Let $\{+\infty\} \notin \mathcal{B}$ and $\{-\infty\} \notin \mathcal{B}$. Then, $B \subseteq Y$ and $\emptyset \neq \operatorname{Cl}_+ B \neq Y$ for all $B \in \mathcal{B}$, i.e., we can use the statements of Corollary 1.48. We define the sets

$$V := \bigcup_{B\in\mathcal{B}} (B - \operatorname{int} C) = \left(\bigcup_{B\in\mathcal{B}} B\right) - \operatorname{int} C, \qquad W := \bigcap_{B\in\mathcal{B}} \operatorname{Cl}_+ B.$$

We show that $V \cap W = \emptyset$ and $V \cup W = Y$. Assume there exists some $y \in V \cap W$. Then there is some $\bar{B} \in \mathcal{B}$ such that $y \in \left(\bar{B} - \operatorname{int} C\right) \cap \operatorname{Cl}_+ \bar{B} = \emptyset$, a contradiction. Let $y \in Y \setminus W$ (we have $W \neq Y$, because otherwise $\operatorname{Cl}_+ B = Y$ holds for all $B \in \mathcal{B}$). Then there exists some $\bar{B} \in \mathcal{B}$ such that $y \notin \operatorname{Cl}_+ \bar{B}$. By Corollary 1.48 (ix), (xi), we obtain $y \in \bar{B} - \operatorname{int} C \subseteq V$.

If $\operatorname{Cl}_- V = Y$, we get (using Proposition 1.24) $V = \operatorname{Cl}_- V - \operatorname{int} C = Y$, whence $W = \emptyset$. It follows

$$\operatorname{Sup}\bigcup_{B\in\mathcal{B}} B = \operatorname{Sup} V = \{+\infty\} = \operatorname{Inf}\emptyset = \operatorname{Inf} W.$$

Otherwise, we have $\emptyset \neq \operatorname{Cl}_- V \neq Y$ and $\emptyset \neq \operatorname{Cl}_+ W \neq Y$. By Corollary 1.48, we obtain

$$\begin{aligned} \operatorname{Sup}\bigcup_{B\in\mathcal{B}} B &= \{y \in Y |\; y \notin V,\; \{y\} - \operatorname{int} C \subseteq V\} \\ &= \{y \in Y |\; y \in W, (\{y\} - \operatorname{int} C) \cap W = \emptyset\} \\ &= \operatorname{wMin} W = \operatorname{wMin} \operatorname{Cl}_+ W = \operatorname{Inf} W, \end{aligned}$$

which completes the proof. □

Even though $\mathcal{I}$ is not a linear space we have the following result.

Corollary 1.55. *The following assertion is true:*

$$A \in \mathcal{I} \quad \iff \quad -A \in \mathcal{I}.$$

Proof. Let $A \in \mathcal{I}$. Of course, we have $\sup\{A\} = A$ and Theorem 1.54 yields $\operatorname{Sup} A = A$. It follows $-A = -\operatorname{Sup} A = \operatorname{Inf}(-A)$ and hence $-A \in \mathcal{I}$. □

Note that the last statement is not true for $A \in \mathcal{F}$. Nevertheless, it is sometimes easier to work with the complete lattice $\mathcal{F}$ in the proofs. Then, the corresponding results for the space $\mathcal{I}$ can be obtained by Proposition 1.52.

In the following proposition we use a generalization of the Minkowski sum. For $\mathcal{A}, \mathcal{B} \subseteq \mathcal{I}$, we set

$$\mathcal{A} \oplus \mathcal{B} := \{I \in \mathcal{I} |\ \exists A \in \mathcal{A}, \exists B \in \mathcal{B} :\ I = A \oplus B\}.$$

Proposition 1.56. *Let $\mathcal{A}, \mathcal{B} \subseteq \mathcal{I}$, then*

(i) $\inf \mathcal{A} \oplus \mathcal{B} = \inf \mathcal{A} \oplus \inf \mathcal{B}$,
(ii) $\sup \mathcal{A} \oplus \mathcal{B} \preccurlyeq \sup \mathcal{A} \oplus \sup \mathcal{B}$.

Proof. (i) If $\mathcal{A} = \emptyset$, we have $\inf \mathcal{A} \oplus \mathcal{B} = \inf \mathcal{A} = \{+\infty\}$ and thus $\inf \mathcal{A} \oplus \mathcal{B} = \inf \mathcal{A} \oplus \inf \mathcal{B} = \{+\infty\}$. Otherwise, we get

$$\begin{aligned}
\inf \mathcal{A} \oplus \mathcal{B} = \operatorname{Inf} \bigcup_{A \in \mathcal{A}, B \in \mathcal{B}} A \oplus B &= \operatorname{Inf} \bigcup_{A \in \mathcal{A}, B \in \mathcal{B}} A + B \\
&= \operatorname{Inf} \left(\bigcup_{A \in \mathcal{A}} A + \bigcup_{B \in \mathcal{B}} B \right) = \operatorname{Inf} \bigcup_{A \in \mathcal{A}} A \oplus \operatorname{Inf} \bigcup_{B \in \mathcal{B}} B \\
&= \inf \mathcal{A} \oplus \inf \mathcal{B}.
\end{aligned}$$

(ii) For all $A \in \mathcal{A}$, $B \in \mathcal{B}$ we have $A \oplus B \preccurlyeq \sup \mathcal{A} \oplus \sup \mathcal{B}$. Taking the supremum, we obtain the desired statement. □

The following example shows that Proposition 1.56 (ii) does not hold with equality.

Example 1.57. We consider the space $\mathcal{I}$ for $Y = \mathbb{R}^2$ and $C = \mathbb{R}^2_+$. Let $\mathcal{A} = \left\{A^1, A^2\right\}$ and $\mathcal{B} = \{B\}$, where we set

$$A^1 = \operatorname{Inf} \left\{(0,1)^T\right\}, \quad A^2 = \operatorname{Inf} \left\{(1,0)^T\right\}, \quad B = \left\{y \in \mathbb{R}^2 |\ y_1 + y_2 = 0\right\}.$$

Then we have

$$A^1 \oplus B = A^2 \oplus B = \left\{y \in \mathbb{R}^2 |\ y_1 + y_2 = 1\right\},$$

$$\sup \mathcal{A} \oplus \mathcal{B} = \sup \left\{A^1 \oplus B, A^2 \oplus B\right\} = \left\{y \in \mathbb{R}^2 |\ y_1 + y_2 = 1\right\},$$

$$\sup \mathcal{A} \oplus \sup \mathcal{B} = \sup \left\{A^1, A^2\right\} \oplus B = \left\{y \in \mathbb{R}^2 |\ y_1 + y_2 = 2\right\}.$$

Whence $\sup \mathcal{A} \oplus \mathcal{B} \neq \sup \mathcal{A} \oplus \sup \mathcal{B}$.

1.6 Subspaces of convex elements

We next investigate subspaces of $\mathcal{F}$ and $\mathcal{I}$ which turned out to be useful, in particular, for convex and linear problems. As shown in Section 1.2, the subsets $\mathcal{F}_{\text{co}}$ and $\mathcal{I}_{\text{co}}$ of all convex elements are again conlinear spaces. Recall that

$$\mathcal{F}_{\text{co}} = \{A \in \mathcal{F} |\ \forall \lambda \in (0,1) : \lambda \odot A \oplus (1-\lambda) \odot A = A\}$$

and

$$\mathcal{I}_{\text{co}} = \{B \in \mathcal{I} |\ \forall \lambda \in (0,1) : \lambda \odot B \oplus (1-\lambda) \odot B = B\}.$$

Proposition 1.58. *The spaces* $(\mathcal{F}_{\text{co}}, \oplus, \odot, \supseteq)$ *and* $(\mathcal{I}_{\text{co}}, \oplus, \odot, \preccurlyeq)$ *are isomorphic and isotone.*

Proof. This is an immediate consequence of the fact that $(\mathcal{F}, \oplus, \odot, \subseteq)$ and $(\mathcal{I}, \oplus, \odot, \preccurlyeq)$ are isomorphic and isotone, and the fact that $\mathcal{F}_{\text{co}}$ and $\mathcal{I}_{\text{co}}$ are defined by the additional axiom (C8) from Section 1.2. □

The conlinear space $\mathcal{F}_{\text{co}}$ can be characterized using the convex hull.

Proposition 1.59. *Let Y be a partially ordered topological vector space with an ordering cone C such that $\emptyset \neq \operatorname{int} C \neq Y$ and let $\mathcal{F} = \mathcal{F}_C(Y)$, then*

$$\mathcal{F}_{\text{co}} = \{A \subseteq Y |\ \mathrm{Cl}_+ \mathrm{co}\, A = A\}.$$

Proof. We have

$$\begin{aligned}
&A = \mathrm{Cl}_+ \mathrm{co}\, A \iff \\
&A = \mathrm{co}\, A \ \wedge\ A = \mathrm{Cl}_+ A \iff \\
&\forall \lambda \in [0,1]:\ A = \mathrm{Cl}_+\big(\lambda A + (1-\lambda) A\big) \iff \\
&\forall \lambda \in [0,1]:\ A = \lambda \odot A \oplus (1-\lambda) \odot A.
\end{aligned}$$

The result now follows from Proposition 1.15. □

Taking into account the relation

$$\mathrm{co}\, B = \mathrm{co}\,\big(B \setminus \{\pm\infty\}\big) \cup \big(B \cap \{\pm\infty\}\big)$$

(compare Proposition 1.30), we can also work with the convex hull of a subset $B \subseteq \overline{Y}$ and we obtain a similar characterization of the conlinear space $\mathcal{I}_{\text{co}}$. To this end we need the following result.

Proposition 1.60. *Let $\overline{Y}$ be an extended partially ordered topological vector space with an ordering cone C such that $\emptyset \neq \operatorname{int} C \neq Y$. For every subset $B \subseteq \overline{Y}$, the following assertion is true:*

$$\mathrm{Inf}\, \mathrm{co}\, \mathrm{Inf}\, \mathrm{co}\, B = \mathrm{Inf}\, \mathrm{co}\, \mathrm{Inf}\, B = \mathrm{Inf}\, \mathrm{co}\, B.$$

Proof. The cases $\mathrm{Cl}_+ \mathrm{co}\, B = Y$ and $\mathrm{Cl}_+ \mathrm{co}\, B = \emptyset$ can be verified directly. Therefore, let $\emptyset \neq \mathrm{Cl}_+ \mathrm{co}\, B \neq Y$, which implies $\emptyset \neq \mathrm{Cl}_+ B \neq Y$, $-\infty \notin B$ and $B \neq \{+\infty\}$. By Corollary 1.48 (i), we have

$$\begin{aligned}\operatorname{Inf} \operatorname{co} B + \operatorname{int} C &= \operatorname{Cl}_{+}\operatorname{co} B + \operatorname{int} C = \operatorname{co}\left(\operatorname{Cl}_{+}\operatorname{co} B + \operatorname{int} C\right)\\ &= \operatorname{co}\left(\operatorname{Inf} \operatorname{co} B + \operatorname{int} C\right) = \operatorname{co} \operatorname{Inf} \operatorname{co} B + \operatorname{int} C\end{aligned}$$

and similarly

$$\begin{aligned}\operatorname{Inf} \operatorname{co} B + \operatorname{int} C &\overset{\text{Cor. 1.48 (i)}}{=} \operatorname{Cl}_{+}\operatorname{co} B + \operatorname{int} C\\ &\overset{\text{Pr. 1.40}}{=} \operatorname{cl}\left(\operatorname{co} B \setminus \{+\infty\} + C\right) + \operatorname{int} C\\ &\overset{\text{Pr. 1.24}}{=} \operatorname{co} B \setminus \{+\infty\} + \operatorname{int} C\\ &= \operatorname{co}\left(B \setminus \{+\infty\}\right) + \operatorname{int} C\\ &= \operatorname{co}\left(\left(B \setminus \{+\infty\}\right) + \operatorname{int} C\right)\\ &= \operatorname{co}\left(\operatorname{Cl}_{+}B + \operatorname{int} C\right)\\ &\overset{\text{Cor. 1.48 (i)}}{=} \operatorname{co}\left(\operatorname{Inf} B + \operatorname{int} C\right) = \operatorname{co} \operatorname{Inf} B + \operatorname{int} C.\end{aligned}$$

Taking the closure in both equalities and using Proposition 1.40, we get $\operatorname{Cl}_{+}\operatorname{Inf} \operatorname{co} B = \operatorname{Cl}_{+}\operatorname{co} \operatorname{Inf} \operatorname{co} B$ and $\operatorname{Cl}_{+}\operatorname{Inf} \operatorname{co} B = \operatorname{Cl}_{+}\operatorname{co} \operatorname{Inf} B$. The statement now follows from Corollary 1.49 (i) . □

Proposition 1.61. *Let $\overline{Y}$ be an extended partially ordered topological vector space with an ordering cone C such that $\emptyset \neq \operatorname{int} C \neq Y$ and let $\mathcal{I} = \mathcal{I}_C(\overline{Y})$, then*

$$\mathcal{I}_{\text{co}} = \left\{B \subseteq \overline{Y} \middle|\ \operatorname{Inf} \operatorname{co} B = B\right\}.$$

Proof. The spaces $\mathcal{F}_{\text{co}}$ and $\mathcal{I}_{\text{co}}$ are isomorphic by the bijection $j : \mathcal{F}_{\text{co}} \to \mathcal{I}_{\text{co}}$, $j(\cdot) = \operatorname{Inf}(\cdot)$. By Proposition 1.59,

$$A \in \mathcal{F}_{\text{co}} \iff A = \operatorname{Cl}_{+}\operatorname{co} A \iff [A = \operatorname{co} A \ \wedge\ A = \operatorname{Cl}_{+}A].$$

Thus, we have $A \in \mathcal{F}_{\text{co}}$ if and only if

$$\begin{aligned}\operatorname{Inf} A &= \operatorname{Inf} \operatorname{Cl}_{+}\operatorname{co} A \overset{\text{Cor. 1.49}}{=} \operatorname{Inf} \operatorname{co} A\\ &\overset{\text{Pr. 1.60}}{=} \operatorname{Inf} \operatorname{co} \operatorname{Inf} \operatorname{co} A \overset{A \text{ convex}}{=} \operatorname{Inf} \operatorname{co} \operatorname{Inf} A.\end{aligned}$$

This means $j(A) = \operatorname{Inf} \operatorname{co} j(A)$. Since $j : \mathcal{F}_{\text{co}} \to \mathcal{I}_{\text{co}}$ is a bijection, the statement follows. □

An element of $\mathcal{I}_{\text{co}}$ is in general not a convex subset of $\overline{Y}$, for instance, $B := \operatorname{Inf}\{0\} = \operatorname{bd} C$ is usually non-convex. However, it can be easily seen that we have $B \in \mathcal{I}_{\text{co}}$ if and only if $B = \operatorname{Inf} B$ and $\operatorname{Cl}_{+}B$ is a convex set.

We continue with properties of $\mathcal{F}_{\text{co}}$ and $\mathcal{I}_{\text{co}}$ with respect to the ordering.

Proposition 1.62. *The space $(\mathcal{F}_{\text{co}}, \supseteq)$ is a complete lattice. The infimum and supremum of a nonempty subset $\mathcal{A} \subseteq \mathcal{F}_{\text{co}}$ are given by*

$$\inf \mathcal{A} = \operatorname{cl} \operatorname{co} \bigcup_{A \in \mathcal{A}} A \qquad \textit{and} \qquad \sup \mathcal{A} = \bigcap_{A \in \mathcal{A}} A.$$

Proof. Let $\mathcal{A} \subseteq \mathcal{F}_{\mathrm{co}}$ be nonempty. By Proposition 1.59 we have $A = \operatorname{Cl}_+ \operatorname{co} A$ for all $A \in \mathcal{A}$. We obtain

$$\begin{aligned}
\operatorname{cl} \operatorname{co} \bigcup_{A \in \mathcal{A}} A = \operatorname{cl} \operatorname{co} \bigcup_{A \in \mathcal{A}} \operatorname{Cl}_+ \operatorname{co} A &= \operatorname{cl} \operatorname{co} \bigcup_{A \in \mathcal{A}} \operatorname{cl} \operatorname{co} (A + C) \\
&= \operatorname{cl} \operatorname{co} \bigcup_{A \in \mathcal{A}} (A + C) = \operatorname{cl} \left(\left(\operatorname{co} \bigcup_{A \in \mathcal{A}} A \right) + C \right) \\
&= \operatorname{Cl}_+ \operatorname{co} \bigcup_{A \in \mathcal{A}} A.
\end{aligned}$$

Thus $\operatorname{cl} \operatorname{co} \bigcup_{A \in \mathcal{A}} A \in \mathcal{F}_{\mathrm{co}}$. As $\bigcap_{A \in \mathcal{A}} A$ is convex and upper closed, we get $\bigcap_{A \in \mathcal{A}} A \in \mathcal{F}_{\mathrm{co}}$. Now the statements are easy to verify. □

The next theorem is similar to Theorem 1.54, but it refers to $\mathcal{I}_{\mathrm{co}}$ rather than $\mathcal{I}$. The only difference is the convex hull in the expression of the infimum. The infimum and supremum in the space $\mathcal{I}_{\mathrm{co}}$ are illustrated in Figure 1.6.

Theorem 1.63. *The space $(\mathcal{I}_{\mathrm{co}}, \preccurlyeq)$ is a complete lattice. The infimum and supremum of a nonempty subset $\mathcal{B} \subseteq \mathcal{I}_{\mathrm{co}}$ are given by*

$$\inf \mathcal{B} = \operatorname{Inf} \operatorname{co} \bigcup_{B \in \mathcal{B}} B \qquad \textit{and} \qquad \sup \mathcal{B} = \operatorname{Sup} \bigcup_{B \in \mathcal{B}} B.$$

Proof. The conlinear spaces $\mathcal{F}_{\mathrm{co}}$ and $\mathcal{I}_{\mathrm{co}}$ are isomorphic and isotone, where the bijection is $j : \mathcal{F}_{\mathrm{co}} \to \mathcal{I}_{\mathrm{co}}$ with $j(\cdot) = \operatorname{Inf}(\cdot)$ and $j^{-1}(\cdot) = \operatorname{Cl}_+(\cdot)$. In view of Proposition 1.62, it remains to show

$$j^{-1} \left(\operatorname{Inf} \operatorname{co} \bigcup_{B \in \mathcal{B}} B \right) = \operatorname{cl} \operatorname{co} \bigcup_{B \in \mathcal{B}} j^{-1}(B),$$

which follows from

$$\begin{aligned}
j \left(\operatorname{Cl}_+ \operatorname{co} \textstyle\bigcup_{B \in \mathcal{B}} j^{-1}(B) \right) &\overset{\text{Cor. 1.49}}{=} \operatorname{Inf} \operatorname{co} \bigcup_{B \in \mathcal{B}} j^{-1}(B) \\
&\overset{\text{Pr. 1.62}}{=} \operatorname{Inf} \operatorname{co} \operatorname{Inf} \bigcup_{B \in \mathcal{B}} j^{-1}(B) \\
&\overset{\text{Cor. 1.49}}{=} \operatorname{Inf} \operatorname{co} \operatorname{Inf} \bigcup_{B \in \mathcal{B}} B \\
&\overset{\text{Pr. 1.62}}{=} \operatorname{Inf} \operatorname{co} \bigcup_{B \in \mathcal{B}} B.
\end{aligned}$$

For the supremum it remains to show that $\operatorname{Sup}\bigcup_{B\in\mathcal{B}} B \in \mathcal{I}_{\mathrm{co}}$. But in Theorem 1.54 (see also Proposition 1.53) it is shown that

$$\operatorname{Sup} \bigcup_{B\in\mathcal{B}} B = \operatorname{Inf} \bigcap_{B\in\mathcal{B}} \operatorname{Cl}_+ B.$$

As $\operatorname{Cl}_+ B$ is convex for $B \in \mathcal{I}_{\mathrm{co}}$, we get

$$\operatorname{Inf} \bigcap_{B\in\mathcal{B}} \operatorname{Cl}_+ B = \operatorname{Inf} \operatorname{co} \bigcap_{B\in\mathcal{B}} \operatorname{Cl}_+ B.$$

The statement follows from Proposition 1.61 and Proposition 1.60. □

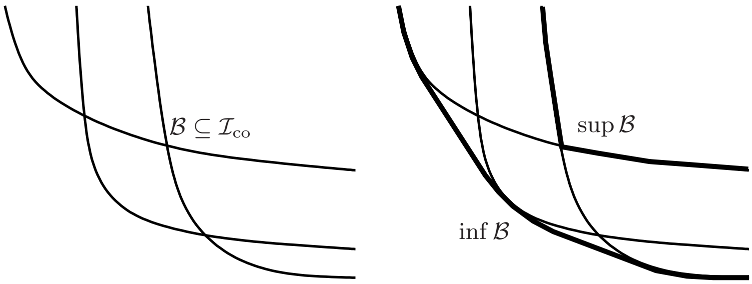

Fig. 1.6 Illustration of Theorem 1.63 for $Y = \mathbb{R}^2$ and $C = \mathbb{R}^2_+$.

The following result is important for convex minimization problems. It states that the space $\mathcal{I}_{\mathrm{co}}$ rather than $\mathcal{I}$ is also adequate for these problems. We set

$$\inf_{x\in S} f(x) := \inf \{f(x)|\ x \in S\}.$$

Proposition 1.64. *Let X be a vector space, $S \subseteq X$ a convex subset of X and $f : X \to \mathcal{I}$ a convex function. Then*

(i) $f : X \to \mathcal{I}_{\mathrm{co}}$,
(ii) $\inf_{x\in S} f(x) \in \mathcal{I}_{\mathrm{co}}$.

Proof. Since (i) is a special case (set $S = \{x\}$), it remains to show (ii). For all $\lambda \in [0,1]$, we have

$$\begin{aligned}\inf_{x\in S} f(x) &= \inf_{y,z\in S} f(\lambda y + (1-\lambda)z)\\ &\preccurlyeq \inf_{y,z\in S} (\lambda \odot f(y) \oplus (1-\lambda) \odot f(z))\\ &\overset{\text{Pr. 1.56 (i)}}{=} \lambda \odot \inf_{x\in S} f(x) \oplus (1-\lambda) \odot \inf_{x\in S} f(x).\end{aligned}$$

Hence $\inf_{x\in S} f(x) \in \mathcal{I}_{\text{co}}$. □

1.7 Scalarization methods

Scalarization is an important tool in vector optimization. In the conventional theory of vector optimization, convex problems are usually scalarized by linear functionals. As we consider set-valued problems (more precisely $\mathcal{I}$-valued problems), we use (a modification of) the support function rather than linear functionals for scalarization.

Let $\overline{Y}$ be an extended partially ordered locally convex space with an ordering cone C such that $\emptyset \neq \operatorname{int} C \neq Y$ and let Y^* be its topological dual. Furthermore, let $\mathcal{F} = \mathcal{F}_C(Y)$ and $\mathcal{I} = \mathcal{I}_C(\overline{Y})$. Based on the support function as introduced in Section 1.3, we define a scalarizing functional depending on a parameter $y^* \in C^\circ \setminus \{0\}$. For $A \in \mathcal{I}$, we set

$$\varphi_A(y^*) := \varphi(y^* | \, A) := -\sigma(y^* | \operatorname{Cl}_+ A). \tag{1.12}$$

For fixed y^*, we get by (1.12) a functional from $\mathcal{I}$ into $\overline{\mathbb{R}}$. For fixed $A \in \mathcal{I}$, we consider φ_A to be a function from $C^\circ \setminus \{0\}$ into $\overline{\mathbb{R}}$, that is,

$$\varphi_A : \; C^\circ \setminus \{0\} \to \overline{\mathbb{R}}.$$

For some $\gamma \in \overline{\mathbb{R}}$ we write $\varphi_A \equiv \gamma$ whenever $\varphi_A(y^*) = \gamma$ for all $y^* \in C^\circ \setminus \{0\}$. The addition, the multiplication by positive real numbers, the ordering relation, the infimum and the supremum for the extended real-valued function φ_A are defined pointwise for all $y^* \in C^\circ \setminus \{0\}$. We use the inf-addition, i.e., $-\infty + (+\infty) = +\infty + (-\infty) = +\infty$.

We continue with a collection of useful properties of the functional φ_A.

Theorem 1.65. *Let $A, B \in \mathcal{I}$ and $\alpha > 0$, then*

(i) $[A \in \mathcal{I}_{\text{co}} \;\wedge\; \varphi_A \equiv -\infty] \iff A = \{-\infty\}$,
(ii) $\varphi_A \equiv +\infty \iff \left[\exists y^* \in C^\circ \setminus \{0\} : \varphi_A(y^*) = +\infty\right] \iff A = \{+\infty\}$,
(iii) $A \preccurlyeq B \implies \varphi_A \leq \varphi_B$,
(iv) $\left[A \in \mathcal{I}_{\text{co}} \;\wedge\; \varphi_A \leq \varphi_B\right] \implies A \preccurlyeq B$,
(v) $\varphi_{A \oplus B} = \varphi_A + \varphi_B$,
(vi) $\alpha \cdot \varphi_A = \varphi_{\alpha \odot A}$.

Let $\mathcal{A} \subseteq \mathcal{I}$ be nonempty, then

(vii) $\varphi_{\inf \mathcal{A}} = \inf\limits_{A \in \mathcal{A}} \varphi_A$,
(viii) $\varphi_{\sup \mathcal{A}} \geq \sup\limits_{A \in \mathcal{A}} \varphi_A$.

Proof. The proof is based on the properties of the support function, see Section 1.3.

(i) As $\mathrm{Cl}_+\{-\infty\} = Y$, we get $\varphi_{\{-\infty\}} = -\sigma(y^* \,|\, Y) = -\infty$ for all $y^* \in C^\circ \setminus \{0\}$. On the other hand, $\varphi_A \equiv -\infty$ implies that $\sigma(y^* \,|\, \mathrm{Cl}_+ A) = +\infty$ for all $y^* \in C^\circ \setminus \{0\}$. Moreover, $\sigma(y^* \,|\, C) = +\infty$ for all $y^* \in Y \setminus C^\circ$. We get

$$+\infty = \sigma(y^* \,|\, \mathrm{Cl}_+ A) + \sigma(y^* \,|\, C) = \sigma(y^* \,|\, \mathrm{Cl}_+ A + C) = \sigma(y^* \,|\, \mathrm{Cl}_+ A)$$

for all $y^* \in Y \setminus \{0\}$. It follows $\mathrm{Cl}_+ A = \mathrm{cl}\,\mathrm{co}\,\mathrm{Cl}_+ A = Y$ and hence $A = \{-\infty\}$.

(ii) We have $\mathrm{Cl}_+\{+\infty\} = \emptyset$ and so $A = \{+\infty\}$ implies $\varphi_A(y^*) = -\sigma(y^*|\,\emptyset\,) = +\infty$ for all $y^* \in C^\circ \setminus \{0\}$. If $\sigma(y^*|\mathrm{Cl}_+ A) = -\infty$ for some $y^* \in C^\circ \setminus \{0\}$, then $\mathrm{Cl}_+ A = \emptyset$. Since $A \in \mathcal{I}$, this implies $A = \{+\infty\}$.

(iii) Let $A \preccurlyeq B$. We get $\mathrm{Cl}_+ A \supseteq \mathrm{Cl}_+ B$ and hence

$$\varphi_A(y^*) = -\sigma(y^*|\mathrm{Cl}_+ A) \leq -\sigma(y^*|\mathrm{Cl}_+ B) = \varphi_B(y^*)$$

for all $y^* \in \mathbb{R}^q$, in particular, for all $y^* \in C^\circ \setminus \{0\}$.

(iv) Let $\varphi_A \leq \varphi_B$, i.e., for all $y^* \in C^\circ \setminus \{0\}$, $-\sigma(y^*|\mathrm{Cl}_+ A) \leq -\sigma(y^*|\mathrm{Cl}_+ B)$ holds. By similar arguments as in the proof of (i), the latter inequality is valid for all $y^* \in Y$. As $\mathrm{Cl}_+ A$ is convex and closed we get $\mathrm{Cl}_+ A = \mathrm{cl}\,\mathrm{co}\,\mathrm{Cl}_+ A \supseteq \mathrm{cl}\,\mathrm{co}\,\mathrm{Cl}_+ B \supseteq \mathrm{Cl}_+ B$ and thus $A \preccurlyeq B$.

In order to prove the statements (v) to (vii), let $y^* \in C^\circ \setminus \{0\}$ be arbitrarily given.

(v) If A or B equals $\{+\infty\}$, then $A \oplus B = \{+\infty\}$ and the statement follows as $\mathrm{Cl}_+\{+\infty\} = \emptyset$. If A and B are not $\{+\infty\}$ but one of them or both equal $\{-\infty\}$, then the result follows from the fact $\mathrm{Cl}_+\{-\infty\} = Y$. Thus we can assume $A, B \subseteq Y$. In this case we have $\mathrm{Cl}_+ A = \mathrm{cl}\,(A + C)$. It follows

$$\begin{aligned}\varphi_{A+B}(y^*) &= -\sigma(y^*|\,\mathrm{cl}\,(A + B + C)) \\ &= -\sigma(y^*|\,\mathrm{cl}\,(A + C)) - \sigma(y^*|\,\mathrm{cl}\,(B + C)) \\ &= \varphi_A(y^*) + \varphi_B(y^*).\end{aligned}$$

(vi) If $A = \{+\infty\}$, then $\alpha \odot A = \{+\infty\}$ and hence

$$\alpha \cdot \varphi_A(y^*) = \varphi_{\alpha \odot A}(y^*) = +\infty.$$

If $A = \{-\infty\}$, then $\alpha \odot A = \{-\infty\}$ and thus

$$\alpha \cdot \varphi_A(y^*) = \varphi_{\alpha \odot A}(y^*) = -\infty.$$

If $A \subseteq Y$, then we have

$$\alpha \cdot \varphi_A(y^*) = -\alpha\sigma(y^*|\,\mathrm{cl}\,(A + C)) = -\sigma(y^*|\,\mathrm{cl}\,(\alpha A + C)) = \varphi_{\alpha \odot A}(y^*).$$

(vii) It remains to show the statement for the case $\{+\infty\} \notin \mathcal{A}$, because omitting $\{+\infty\}$ does not change anything. If $\{-\infty\} \in \mathcal{A}$ the equality can be easily shown. Therefore, let $A \subseteq Y$ for all $A \in \mathcal{A}$. We obtain

$$\varphi_{\inf \mathcal{A}}(y^*) = -\sigma\left(y^*\middle|\, \mathrm{Cl}_+ \bigcup_{A\in\mathcal{A}} A\right) = -\sigma\left(y^*\middle|\, \mathrm{cl} \bigcup_{A\in\mathcal{A}} (A+C)\right)$$
$$= \inf_{A\in\mathcal{A}} -\sigma\big(y^*|\,\mathrm{cl}\,(A+C)\big) = \inf_{A\in\mathcal{A}} \varphi_A(y^*).$$

(viii) We have $\sup \mathcal{A} \succcurlyeq A$ and hence $\varphi_{\sup \mathcal{A}} \geq \varphi_A$ for all $A \in \mathcal{A}$. Taking the supremum we obtain the desired statement. □

Statement (viii) in the last theorem does not hold with equality as the following example shows.

Example 1.66. Let $Y = \mathbb{R}^2$ and $C = \mathbb{R}^2_+$. If we set $A := \mathrm{Inf}\left\{(0,1)^T\right\}$ and $B := \mathrm{Inf}\left\{(1,0)^T\right\}$, then $\sup\{A,B\} = \mathrm{Inf}\left\{(1,1)^T\right\}$. For $y^* = (-1,-1)^T$, we get $\varphi_A(y^*) = \varphi_B(y^*) = 1$ but $\varphi_{\sup\{A,B\}}(y^*) = 2$.

It is well-known that a convex extended real-valued function $\xi : X \to \overline{\mathbb{R}}$ is said to be *proper* if

$$\forall x \in X:\ \xi(x) \neq -\infty \quad \wedge \quad \exists \bar{x} \in X:\ \xi(\bar{x}) \neq +\infty.$$

Similarly, a concave extended real-valued function $\eta : X \to \overline{\mathbb{R}}$ is called *proper* if the convex function $-\eta$ is proper, that is

$$\forall x \in X:\ \eta(x) \neq +\infty \quad \wedge \quad \exists \bar{x} \in X:\ \eta(\bar{x}) \neq -\infty.$$

The domain of a convex extended real-valued function $\xi : X \to \overline{\mathbb{R}}$ is the set

$$\mathrm{dom}\,\xi := \{x \in X|\ \xi(x) \neq +\infty\},$$

whereas the domain of the concave extended real-valued function $\eta : X \to \overline{\mathbb{R}}$ is the set $\mathrm{dom}\,(-\eta)$, i.e.,

$$\mathrm{dom}\,\eta := \{x \in X|\ \eta(x) \neq -\infty\}.$$

We introduce similar notions for $\mathcal{I}$-valued functions.

Definition 1.67. A convex function $f : X \to \mathcal{I}$ is said to be *proper* if

$$\forall x \in X:\ f(x) \neq \{-\infty\} \quad \wedge \quad \exists \bar{x} \in X:\ f(\bar{x}) \neq \{+\infty\}.$$

The *domain* of the convex function $f : X \to \mathcal{I}$ is the set

$$\mathrm{dom}\, f := \{x \in X|\ f(x) \neq \{+\infty\}\}.$$

Taking into account that $\varphi_A : C^\circ \setminus \{0\} \to \overline{\mathbb{R}}$ is a concave function, we get the following statement.

Corollary 1.68. *Let $A \in \mathcal{I}_{\mathrm{co}}$. Then*

$$A \in \mathcal{I}_{\mathrm{co}} \setminus \{\{-\infty\}, \{+\infty\}\} \iff \varphi_A \text{ is proper.}$$

Proof. This follows from Theorem 1.65 (i), (ii). □

Corollary 1.69. *Let $f : X \to \mathcal{I}$ be a function. The following statements are equivalent:*

(i) $f : X \to \mathcal{I}$ *is convex;*
(ii) *For all* $y^* \in C^\circ \setminus \{0\}$, $\varphi_{f(\cdot)}(y^*) : X \to \overline{\mathbb{R}}$ *is convex.*

Moreover, for all $y^* \in C^\circ \setminus \{0\}$, *we have* $\operatorname{dom} f = \operatorname{dom} \varphi_{f(\cdot)}(y^*)$.

Proof. The equivalence of (i) and (ii) follows from Theorem 1.65 (v), (vi). The second statement follows from Theorem 1.65 (i), (ii). □

Note that in the preceding result, both spaces $\mathcal{I}$ and $\overline{\mathbb{R}}$ are equipped with the inf-addition.

1.8 A topology on the space of self-infimal sets

In this section we introduce a topology for the space $\mathcal{I}$ which will be used to define continuous $\mathcal{I}$-valued functions. If the reader is not interested in the topology the continuity notion is based on, the characterization of continuous $\mathcal{I}$-valued functions in Theorem 1.75 can be used as a definition.

First, we recall some facts on uniform spaces. Further information on the notions used in this section can be found, for instance, in (Kelley, 1955).

Definition 1.70. Let Z be a nonempty set and let $\mathcal{N}$ be a system of subsets N of $Z \times Z := \{(z,y) :\ z,y \in Z\}$. For $N \subseteq Z \times Z$ we set

$$N^{-1} := \{(y,z) :\ (z,y) \in N\}$$

and

$$N \circ N := \{(z,y) \in Z \times Z :\ \exists w \in Z :\ (z,w),(w,y) \in N\}\,.$$

The set $\Delta := \{(z,z) \in Z \times Z\}$ is called the *diagonal.* The set Z is said to be a *uniform space* if there exists a filter $\mathcal{N}$ on $Z \times Z$ satisfying

(N1) $\forall N \in \mathcal{N} :\ \Delta \subseteq N$,
(N2) $N \in \mathcal{N} \implies N^{-1} \in \mathcal{N}$,
(N3) $\forall N \in \mathcal{N},\ \exists M \in \mathcal{N} :\ M \circ M \subseteq N$.

The system $\mathcal{N}$ is called a *uniformity* on Z. By the sets $\mathcal{U}(z) := \{U_N(z) : N \in \mathcal{N}\}$ where $U_N(z) := \{y \in Z :\ (z,y) \in N\}$, a topology is given, called the *uniform topology* on Z.

Of course, a uniform space is already well-defined by a base of its uniformity $\mathcal{N}$, i.e., a filter base $\mathcal{B}$ of the uniformity $\mathcal{N}$.

Proposition 1.71. *The topology of a uniform space is separated (or Hausdorff) if and only if*

(N4) $\bigcap_{N \in \mathcal{N}} N = \Delta.$

Proof. See e.g. (Köthe, 1966, p. 32). □

We recall a well-established result, the characterization of uniform spaces using *families of pseudo-metrics*, see e.g. Kelley (1955).

Definition 1.72. Let Z be a nonempty set. A function $p : Z \times Z \to [0, \infty)$ is called *pseudo-metric* on Z if and only if for all $z, y, w \in Z$ the following conditions are satisfied:

(P1) $p(z, z) = 0,$
(P2) $p(z, y) = p(y, z),$
(P3) $p(z, y) \leq p(z, w) + p(w, y).$

Moreover, let $(\Lambda, \prec)$ be a directed set. A system $\{p_\lambda\}_{\lambda \in \Lambda}$ of pseudo-metrics $p_\lambda : Z \times Z \to [0, \infty)$ satisfying

(P4) $\lambda \prec \mu \implies (\forall z, y \in Z : p_\lambda(z, y) \leq p_\mu(z, y))$

is called a *family of pseudo-metrics.* If additionally the condition

(P5) $(\forall \lambda \in \Lambda : p_\lambda(z, y) = 0) \implies z = y$

holds, the family of pseudo-metrics is said to be *separating.*

Proposition 1.73. *A topological space (Z, τ) is a (separated) uniform space if and only if its topology τ can be generated by a (separating) family of pseudo-metrics.*

Proof. See (Kelley, 1955, p. 187f.) and take into account that by (P4) we deal with bases instead of subbases. □

We now consider the case $Z = \mathcal{I} = \mathcal{I}_C\left(\overline{Y}\right)$, where $\left(\overline{Y}, \leq_C\right)$ is an extended partially ordered locally convex space with an ordering cone C such that $\emptyset \neq \operatorname{int} C \neq Y$. Furthermore, let Y^* be the topological dual space of Y. Using the scalarization functional $\varphi_A : C^\circ \setminus \{0\} \to \overline{\mathbb{R}}$ as defined in Section 1.7, we introduce a pseudo-metric $p_{y^*} : \mathcal{I} \times \mathcal{I} \to \mathbb{R}$ depending on a parameter $y^* \in C^\circ \setminus \{0\}$ by

$$p_{y^*}(A, B) := \left| \frac{\varphi_A(y^*)}{1 + |\varphi_A(y^*)|} - \frac{\varphi_B(y^*)}{1 + |\varphi_B(y^*)|} \right|,$$

where we set

$$\frac{+\infty}{1 + |+\infty|} := 1 \quad \text{and} \quad \frac{-\infty}{1 + |-\infty|} := -1.$$

Note that the extended real numbers $\overline{\mathbb{R}}$ are typically topologized by the metric

$$m : \mathbb{R} \times \mathbb{R} \to \mathbb{R} : \qquad m(r, s) := \left| \frac{r}{1 + |r|} - \frac{s}{1 + |s|} \right|,$$

where m generates on $\mathbb{R}$ the usual topology (see e.g. Bourbaki, 1989, p. 342).

For a given vector $c \in \operatorname{int} C$ we consider the index set

$$\Lambda := \big\{ W \subseteq Y^* |\; W \text{ is a finite subset of } \{ y^* \in C^\circ |\; y^*(c) = -1 \} \big\}.$$

Of course, $(\Lambda, \subseteq)$ is a directed set. For each $\lambda \in \Lambda$ we obtain by

$$p_\lambda : \mathcal{I} \times \mathcal{I} \to \mathbb{R}, \qquad p_\lambda(A, B) := \max\{ p_{y^*}(A, B) |\; y^* \in \lambda \}$$

again a pseudo-metric. It can be easily verified that the family of pseudo-metrics $\{p_\lambda\}_{\lambda \in \Lambda}$ satisfies the axioms (P1) to (P4). Therefore, the family of pseudo-metrics $\{p_\lambda\}_{\lambda \in \Lambda}$ generates a uniformity $\mathcal{N}$ on $\mathcal{I}$, where the sets

$$N(\lambda, t) := \{ (A, B) \in \mathcal{I} \times \mathcal{I} |\; p_\lambda(A, B) < t \}$$

for $\lambda \in \Lambda$ and $t > 0$ form a base for $\mathcal{N}$. Denoting the uniform topology of $(\mathcal{I}, \mathcal{N})$ by τ, $(\mathcal{I}, \tau)$ is a topological space. In the following, τ is simply called the *uniform topology* on the space $\mathcal{I}$ of self-infimal subsets of $\overline{Y}$.

Proposition 1.74. *Let $\mathcal{I}$ be equipped with the uniform topology τ and let $\tau' \subseteq \tau$ be the relative topology on the subspace $\mathcal{I}_{\mathrm{co}} \subseteq \mathcal{I}$. The topological space $(\mathcal{I}_{\mathrm{co}}, \tau')$ is Hausdorff.*

Proof. Let $A, B \in \mathcal{I}_{\mathrm{co}}$ and let $p_\lambda(A, B) = 0$ for all $\lambda \in \Lambda$. We obtain $\varphi_A(y^*) = \varphi_B(y^*)$ for all $y^* \in C^\circ \setminus \{0\}$. By Theorem 1.65 (iv), we get $A = B$, i.e., (P5) is satisfied. □

By the proof of the latter result it becomes clear that the topological space $(\mathcal{I}, \tau)$ is not Hausdorff, at least for a nontrivial choice of the space Y. To see this take $A, B \in \mathcal{I}$, $A \neq B$ such that $\operatorname{co} A = \operatorname{co} B$, which implies that $\varphi_A \equiv \varphi_B$.

The next theorem shows that continuity of an $\mathcal{I}$-valued function with respect to the uniform topology on $\mathcal{I}$ can be characterized by continuity of the scalarizations.

Theorem 1.75. *Let (X, σ) be a topological space and let $\mathcal{I}$ be equipped with the uniform topology τ. Furthermore, let $\overline{\mathbb{R}}$ be equipped with the topology generated by the metric m. A function $f \colon X \to \mathcal{I}$ is continuous at $\bar{x} \in X$ if and only if $\varphi_{f(\cdot)}(y^*) \colon X \to \overline{\mathbb{R}}$ is continuous at $\bar{x}$ for all $y^* \in C^\circ \setminus \{0\}$.*

Proof. Let $\mathcal{V}$ be a subbase of a neighborhood system of $f(\bar{x})$ and let $\mathcal{U}$ be a neighborhood system of $\bar{x}$. The function $f : X \to \mathcal{I}$ is continuous at $\bar{x}$ if and only if

$$\forall V \in \mathcal{V},\; \exists U \in \mathcal{U},\; \forall u \in U : f(u) \in V.$$

As $p_{y^*}(A, B) = m\left(\varphi_A(y^*), \varphi_B(y^*)\right)$, this is equivalent to

$$\forall y^* \in C^\circ \setminus \{0\}, \forall \varepsilon > 0,\ \exists U \in \mathcal{U},\ \forall u \in U : m\left(\varphi_{f(u)}(y^*), \varphi_{f(\bar{x})}(y^*)\right) < \varepsilon.$$

Of course, this is equivalent to $\varphi_{f(\cdot)}(y^*) \colon X \to \overline{\mathbb{R}}$ being continuous at $\bar{x}$ for all $y^* \in C^\circ \setminus \{0\}$. □

The following situation is of interest in vector optimization. We consider a function $f : X \to Y$, where X is a topological space and Y is a partially ordered locally convex space with ordering cone C such that $\emptyset \neq \operatorname{int} C \neq Y$. Furthermore, let Y^* be the topological dual space of Y and let w be the weak topology on Y generated by Y^*.

We are interested in a relationship between continuity of f and continuity of the extension

$$\bar{f} : X \to \mathcal{I}, \quad \bar{f}(x) := \operatorname{Inf}\{f(x)\}.$$

To this end, let us recall the concept of a cone being *normal* with respect to a topology. For more details the reader is referred to (Göpfert *et al.*, 2003).

Definition 1.76. Let (Y, τ) be a topological vector space and let $C \subseteq Y$ be a convex cone. Then, C is called τ-*normal* if the origin $0 \in Y$ has a neighborhood base formed by *full* sets with respect to C, i.e., sets A of the form

$$A = (A + C) \cap (A - C).$$

Theorem 1.77. *If $f : X \to Y$ is continuous at $\bar{x} \in X$ with respect to the weak topology w on Y, then $\bar{f} : X \to \mathcal{I}$ is continuous. If the ordering cone C is w-normal, then the opposite implication also holds true.*

Proof. The function $f : X \to Y$ is continuous at $\bar{x} \in X$ with respect to the weak topology on Y if and only if

$$\forall y^* \in Y^* : \quad y^*(f(\cdot)) \text{ is continuous at } \bar{x}. \tag{1.13}$$

The function $\bar{f} : X \to \mathcal{I}$ is continuous at $\bar{x} \in X$ if and only if

$$\forall y^* \in C^\circ : \quad \varphi_{f(\cdot)}(y^*) \text{ is continuous at } \bar{x}.$$

Because of the special form of $\bar{f}$, this is equivalent to

$$\forall y^* \in -C^\circ : \quad y^*(f(\cdot)) \text{ is continuous at } \bar{x}. \tag{1.14}$$

Of course (1.13) implies (1.14). The ordering cone C is w-normal if and only if $C^\circ - C^\circ = Y^*$ (compare e.g. Isac, 1987; Göpfert *et al.*, 2003). In this case (1.14) implies (1.13). □

Of course, $\bar{f} : X \to \mathcal{I}$ is continuous at $\bar{x}$ whenever $f : X \to Y$ is continuous at $\bar{x}$ with respect to the given topology τ of the locally convex space Y. This is due to the fact that continuity of $f : X \to Y$ with respect to τ implies continuity $f : X \to Y$ with respect to the weak topology w.

1.9 Notes on the literature

This chapter is on the one hand a collection of well-known facts and standard notions. On the other hand it contains several recent and even new concepts and results. It provides the theoretical foundations and the basic tools for dealing with infimum and supremum in the field of vector optimization.

Section 1.1 contains several well-known facts on ordered sets and complete lattices, which can be found in the literature, see e.g. Birkhoff (1979).

The notion of a conlinear space, which is introduced in Section 1.2, is due to Hamel (2005). It is related to other similar concepts in the literature, see Hamel (2005) for an overview and a comparison. This concept seems to be fundamental for a set-valued convex analysis as it was developed by Hamel (2005, 2009a,b) and Schrage (2009). Some related studies on convex functions with values in a conlinear space can be found in (Löhne, 2006).

A topological vector space, which is the subject in Section 1.3, is a standard concept (see e.g. Kelley *et al.*, 1963; Köthe, 1966; Schaefer, 1980; Aliprantis and Border, 1994; Zălinescu, 2002). Theorem 1.28 appeared in (Löhne, 2006) and Theorem 1.26, which is used to derive properties of the upper closure and the infimal set, might be new in this form. The end of Section 1.3 is concerned to well-known facts on locally convex spaces, which can be found, for instance, in (Aliprantis and Border, 1994).

The concept of infimal sets, introduced in Section 1.4, is due to Nieuwenhuis (1980). A similar but different notion was introduced by Gros (1978). An infimal set was considered in (Nieuwenhuis, 1980) to be a generalization of the infimum, which is no longer an infimum in the context of a complete lattice. But an infimal set reduces to the ordinary infimum in the special case $Z = \overline{\mathbb{R}}$. Nieuwenhuis also observed that infimal sets are useful for the duality theory in vector optimization. This idea was captured, for instance, by Tanino (1992); Song (1997, 1998); Chen and Li (2009); Li *et al.* (2009), see also Boţ *et al.* (2009). Tanino (1988) extended the concept of infimal sets. He worked with an extended space $Y \cup \{\pm\infty\}$ and demonstrated several advantages of this extension. It turned out, however, that Tanino's calculus rules for the elements $\pm\infty$ are not adequate for our reasons. Therefore, we proposed to adapt them in (Löhne and Tammer, 2007). The properties of infimal sets in the Lemma 1.46, Lemma 1.47, Corollary 1.48 and Corollary 1.49 are essentially due to Nieuwenhuis (1980). Here they are derived by short proofs from Lemma 1.42 (based on Theorem 1.26) and Lemma 1.43. Upper closed sets can be defined and used in a more general framework, see e.g. Dolecki and Malivert (1993) and Hamel (2005).

Optimization of functions with values in ordered hyperspaces seems to have its origin in (Kuroiwa, 1998a,b). This approach was also continued by Hernández and Rodríguez-Marín (2007); Hernández *et al.* (2009). Kuroiwa extended the ordering of a partially ordered vector space to its power set. It turned out to be beneficial to work with suitable equivalence classes of the power set. This is equivalent to the usage of the space $(\mathcal{F}, \supseteq)$ of upper closed sets as introduced in Section 1.5 (see e.g. Löhne, 2005a,b; Hamel, 2009a,b; Schrage, 2009; Löhne, 2010). These authors used, in contrast to Kuroiwa, the lattice structure of the space. The space $\mathcal{I}$ of self-infimal sets and the corresponding main result in Theorem 1.54 in a finite dimensional framework are due to Löhne and Tammer (2007). Hamel (2009a,b) and Schrage (2009) considered upper-closed-valued optimization problems under weaker assumptions. There is also a considerable amount of literature on set optimization, which is based on a vectorial ordering relation rather than an ordering relation on the power set, see e.g. Part V of the book by Jahn (2004). An interesting application of $\mathcal{F}$-valued optimization problems in mathematical finance can be found in (Hamel and Rudloff, 2008; Hamel and Heyde, 2010).

The definition of the space $\mathcal{F}_{\text{co}}$ and the related results in Section 1.6 can be found in Hamel (2005). The space $\mathcal{I}_{\text{co}}$ as well as the corresponding statements seem to be new.

The scalarization methods in Section 1.7 were used in the framework of set-valued optimization by Löhne (2005a,b); Löhne and Tammer (2007) and Schrage (2009). Hamel (2009a,b) used different methods in order to prove duality results which are based on separation, but he gave alternative proofs using this scalarization method. Another scalarization method for functions into hyperspaces can be found in (Hamel and Löhne, 2006)

The concepts and results related to the uniform topology on $\mathcal{I}$, as presented in Section 1.8, are new insofar as the space $\mathcal{I}$ is new. Nevertheless, they are related to the so-called scalar convergence of closed convex sets, see e.g. Beer (1993); Sonntag and Zălinescu (1992); Löhne and Zălinescu (2006); Löhne (2008).

Chapter 2
Solution concepts

For more than five decades, vector optimization has been a subject of intensive research. A common notation for a vector optimization problem is

$$\operatorname*{Min}_{x \in S} f(x), \qquad \text{(VOP)}$$

where f is a vector-valued function and S is a feasible set. The central question of this chapter is the following:

What is a *solution to* (VOP)?

It is rather surprising that there is no standard answer to this fundamental question in textbooks on vector optimization. Luc (1988) states that (VOP) *"amounts to finding a point $x \in S$, called an optimal solution of"* (VOP), where $f(x)$ is required to be minimal in the set $f[S] := \{f(x) |\, x \in S\}$ for such a point x. Similarly, Jahn (2004, p. 105) writes that (VOP) *"is to be interpreted in the following way: Determine a minimal solution $x \in S$ which is defined as the inverse image of a minimal element $f(x)$ of the image set $f(S)$."* Ehrgott (2000) writes in the same situation that *"a solution $x \in S$ is called Pareto optimal"*, which means that the term solution seems to refer to a *feasible solution* rather than a solution to (VOP). In the recent textbook by Boţ *et al.* (2009) it is stated that (VOP) *"consists in determining the minimal [...] elements of the image set of S"* and that one is *"also interested in finding the so-called efficient [...] solutions to"* (VOP), where an efficient solution is what Luc called "optimal solution"[1]. It is also stated by Boţ *et al.* (2009) that *"in practice a decision maker is only interested to have a subset or even a single element"* of the set of efficient solutions.

Therefore, it is not clear whether a single efficient solution, a subset or even the set of all efficient solutions is a "solution to (VOP)". This dilemma is underlined by the following lines, taken from an online encyclopedia[2]:

[1] It is not relevant in this discussion that there are different types of efficient solutions.

[2] Wikipedia, the free online encyclopedia, "Multiobjective optimization", english version, 2010-10-10

A. Löhne, *Vector Optimization with Infimum and Supremum*, Vector Optimization,
DOI 10.1007/978-3-642-18351-5_3, © Springer-Verlag Berlin Heidelberg 2011

"The solution to [a multiobjective optimization problem[3]] is a set of Pareto points. Pareto solutions are those for which improvement in one objective can only occur with the worsening of at least one other objective. Thus, instead of a unique solution to the problem (which is typically the case in traditional mathematical programming), the solution to a multiobjective problem is a (possibly infinite) set of Pareto points.

Even though this definition gives the precise statement that a solution to (VOP) is a set of efficient (or Pareto) points there is no further requirement to this set; a singleton set is therefore also a solution. For typical vector optimization problems, however, a single efficient point can be already obtained by solving a scalarized optimization problem. Only a fraction of the theory on vector optimization would be necessary for this reason.

The main idea of vector optimization is that a decision maker chooses an efficient solution from the set of all efficient solutions. This decision is supported by the solution to the vector optimization problem. This means, the problem must be solved prior to the decision.

We prepend this chapter two postulates.

(1) The goal of vector optimization is to provide a decision maker with a sufficient amount of information on the problem in terms of efficient elements.
(2) A solution concept for a vector optimization problem should provide a specification of the term "sufficient" in (1).

The second hypothesis consists of two aspects.

(a) Does the set of all efficient elements provide enough information?
(b) If so, are there proper subsets of the set of efficient elements that already contain enough information?

The first aspect (a) is a question of existence. The second question (b) is concerned with uniqueness, i.e., if the set of all efficient elements is the only choice, we can say the solution is unique.

Scalar optimization is of course a special case of vector optimization, so that a solution concept should reduce to the standard concept in this special case. To this end, let us first consider a general scalar optimization problem. Let X be a nonempty set and let $f : X \to \overline{\mathbb{R}}$ be a proper function on X, i.e., $f(x) \neq -\infty$ for all $x \in X$ and $f \not\equiv +\infty$. We denote by $S \subseteq X$ the set of feasible elements. Let us

$$\text{minimize } f : X \to \overline{\mathbb{R}} \text{ with respect to } \leq \text{ over } S. \tag{2.1}$$

The following statements are equivalent characterizations of $\bar{x} \in X$ being a solution to (2.1):

[3] The term "multiobjective optimization" is stands for optimization problems with more than one real-valued objective functions. These functions can be interpreted as a single vector-valued objective function.

(i) $\bar{x} \in S$ and $f(\bar{x}) \leq f(x)$ for all $x \in S$,
(ii) $\bar{x} \in S$ and $f(\bar{x}) \not> f(x)$ for all $x \in S$,
(iii) $\bar{x} \in S$ and $f(\bar{x}) = \inf_{x \in S} f(x)$.

Since in vector optimization the ordering relation is more complex than in scalar optimization, the latter conditions do not coincide any longer. While condition (i) is obviously too restrictive for vector optimization problems (utopia points), the common "solution concepts" in the literature are mainly based on (ii). There are several possibilities to interpret the relation $\not>$ ("not greater than"), which leads to a variety of different notions, such as efficient, weakly efficient and properly efficient elements. All these concepts don't take into account the infimum and supremum, which is quite important in scalar optimization. The usage of infimal sets in the literature is related to condition (iii), but the complete lattice has not been pointed out.

The solution concept for vector optimization problems, which is introduced in the next two sections, involves all the conditions (i), (ii) and (iii).

2.1 A solution concept for lattice-valued problems

A complete-lattice-valued optimization problem provides the abstract framework for solution concepts based on the *attainment of the infimum or supremum.*

Let $f : X \to Z$, where X is an arbitrary nonempty set and, unless otherwise indicated, $(Z, \leq)$ is a complete lattice. For a nonempty subset $S \subseteq X$, called *feasible set*, we consider the optimization problem

$$\text{minimize } f : X \to Z \text{ with respect to } \leq \text{ over } S. \qquad (\mathcal{L})$$

A standard concept is the following, where $(Z, \leq)$ is only supposed to be a partially ordered set in the following definition.

Definition 2.1. An element $\bar{x} \in S$ is called an *efficient solution* to $(\mathcal{L})$ if

$$[x \in S \quad \wedge \quad f(x) \leq f(\bar{x})] \quad \Longrightarrow \quad f(x) = f(\bar{x}).$$

The set of all efficient solutions to $(\mathcal{L})$ is denoted by $\operatorname{Eff}(\mathcal{L})$.

For $A \subseteq Z$ we denote by

$$\operatorname{Min} A := \{z \in A |\ (y \in A \wedge y \leq z) \Rightarrow y = z\}$$

the set of *minimal elements* of A. Using the notation

$$f[V] := \{f(x) |\ x \in V\},$$

we obtain

$$\text{Min}\, f[S] = f[\text{Eff}\,(\mathcal{L})].$$

It is demonstrated by the following two examples that the set Eff $(\mathcal{L})$ without any further requirement is unsatisfactory as a solution concept for vector optimization problems.

Example 2.2. Let $X = Z = \mathbb{R}^2$ and let Z be partially ordered by the natural ordering cone $\mathbb{R}^2_+$. Let f be the identity map and

$$S = \left\{x \in \mathbb{R}^2 |\, x_1 > 0,\ x_2 > 0,\ x_1 + 10x_2 > 10\right\} \cup \left\{(13, 0)^T\right\}.$$

We have Eff $(\mathcal{L}) = \left\{(13, 0)^T\right\}$. But the nonempty set Eff $(\mathcal{L})$ does not yield a sufficient amount of information about the problem. From a practical point of view, for instance, the feasible, non-efficient point $(1, 1)^T$ could be more interesting than the set of efficient solutions, see Figure 2.1.

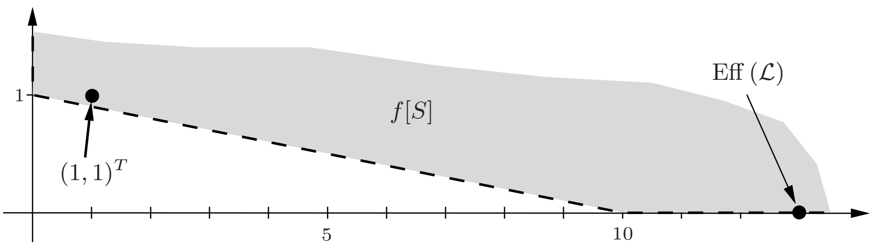

Fig. 2.1 Illustration of Example 2.2. The set of efficient points is not a useful solution concept.

On the other hand, there are vector optimization problems where it is already sufficient for the decision maker to know a proper subset of Eff $(\mathcal{L})$.

Example 2.3. Let $X = Z = \mathbb{R}^2$, Z partially ordered by $\mathbb{R}^2_+$, and

$$S = \left\{x \in \mathbb{R}^2 |\, x_1 \geq 0,\ x_2 \geq 0,\ 2x_1 + x_2 \geq 2,\ x_1 + 2x_2 \geq 2\right\}.$$

The objective $f : X \to Z$ is given as $f(x) = (0, x_2)^T$. Then

$$\text{Eff}\,(\mathcal{L}) = \left\{x \in \mathbb{R}^2 |\, x_1 \geq 2, x_2 = 0\right\}.$$

In typical applications the decision maker selects a point in the image $f[\text{Eff}\,(\mathcal{L})]$ of Eff $(\mathcal{L})$ with respect to f. We have $f[\text{Eff}\,(\mathcal{L})] = \left\{(0, 0)^T\right\}$. But, the same image is already obtained by any nonempty subset of Eff $(\mathcal{L})$, see Figure 2.2.

Example 2.3 indicates that the condition $f[\bar{X}] = \text{Min}\, f[S]$ could be one suitable requirement for a set $\bar{X} \subseteq S$ to be a solution. But additionally, the

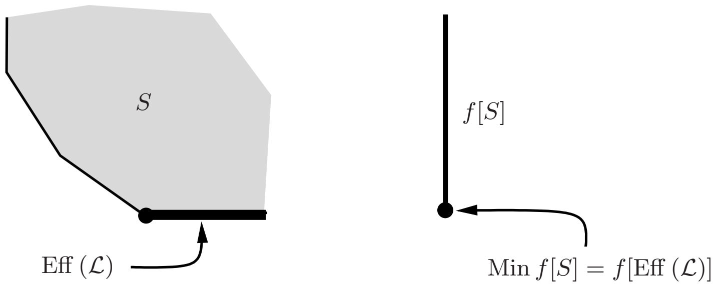

Fig. 2.2 Illustration of Example 2.3. Every nonempty subset of Eff $(\mathcal{L})$ generates the same image.

situation in Example 2.2 must be avoided. This would be possible by assuming the well-known domination property, which is recalled and discussed below. We choose, however, a weaker condition, which is connected with the *attainment of the infimum*. To ensure the existence of the infimum, we need to assume $(Z, \leq)$ to be a complete lattice.

The infimum of f over a set $S \subseteq X$ is defined by

$$\inf_{x \in S} f(x) := \inf \{f(x) |\; x \in S\} = \inf f[S].$$

Definition 2.4. Let $S \subseteq X$ and $\bar{x} \in X$. We say the *infimum of f over S is attained at $\bar{x}$* if

$$\bar{x} \in S \quad \wedge \quad f(\bar{x}) = \inf_{x \in S} f(x).$$

In case such an element $\bar{x}$ exists (does not exist), we say the infimum of f over S is (not) attained.

The attainment of the infimum is an important concept in optimization. In vector optimization it is, however, very hard to fulfill as the following example shows.

Example 2.5. Let $X = \mathbb{R}^2$, $Z = \mathbb{R}^2 \cup \{\pm\infty\}$, $\mathbb{R}^2$ partially ordered by the cone $\mathbb{R}^2_+$. The ordering is denoted by $\leq$ and extended to Z by setting $-\infty \leq z \leq +\infty$ for all $z \in Z$. Then, $(Z, \leq)$ is a complete lattice. Let

$$S = \left\{x \in \mathbb{R}^2 |\; x_1 \geq 0,\; x_2 \geq 0,\; 2x_1 + x_2 \geq 2,\; x_1 + 2x_2 \geq 2\right\}$$

and let f be the identity map. Then the infimum of f over S is not attained. Indeed, we have $\inf_{x \in S} f(x) = \{0, 0\}^T$, but there is no $\bar{x} \in S$ with $f(\bar{x}) = \{0, 0\}^T$, see Figure 2.3.

A further aspect can be observed in the previous example. We enforce that the infimum is attained in a single vector. In vector optimization we intend

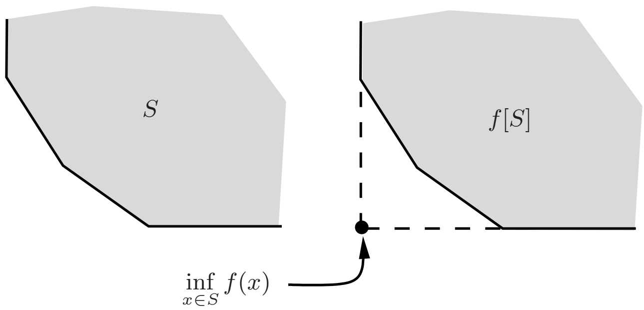

Fig. 2.3 Illustration of Example 2.5. The infimum is not attained.

to present the decision maker all or at least a representative choice of efficient vectors. Therefore, we expect a solution to be a set of feasible vectors.

This requirement is taken into account by a concept that we call *canonical extension*.

Definition 2.6. The *canonical extension* of the objective function $f : X \to Z$ in the complete-lattice-valued optimization problem $(\mathcal{L})$ is the function

$$F : 2^X \to Z, \quad F(A) := \inf_{x\in A} f(x).$$

Of course, we have $f(x) = F(\{x\})$ for all $x \in X$. Working with the canonical extension F instead of f, we make the following two observations: First, we see that attainment of the infimum is easier to realize. The second difference is that the infimum is attained in a set rather than in a single element of X.

We now give a characterization of the attainment of the infimum of the canonical extension F in terms of the given function f.

Proposition 2.7. *Let $S \subseteq X$. The following statements are equivalent.*

(i) *The infimum of F over 2^S is attained at $\bar{X}$, i.e.,*

$$\bar{X} \in 2^S \quad \wedge \quad F(\bar{X}) = \inf_{A\in 2^S} F(A).$$

(ii) $\bar{X} \subseteq S \quad \wedge \quad \inf_{x\in\bar{X}} f(x) = \inf_{x\in S} f(x).$

Proof. It remains to prove the equality

$$\inf_{A\in 2^S} F(A) = \inf_{x\in S} f(x). \tag{2.2}$$

For all $x \in S$ we have

$$\inf_{A\in 2^S} F(A) \le F(\{x\}) = f(x).$$

The infimum over $x \in S$ yields $\leq$ in (2.2). For all $A \subseteq S$ we have

$$F(A) = \inf_{x \in A} f(x) \geq \inf_{x \in S} f(x).$$

Taking the infimum over all $A \in 2^S$ we get $\geq$ in (2.2). □

Next we define a solution concept for the complete-lattice-valued problem $(\mathcal{L})$.

Definition 2.8. A nonempty set $\bar{X}$ with $f[\bar{X}] = \operatorname{Min} f[S]$ is called a *solution* to $(\mathcal{L})$ if the infimum of the canonical extension F over 2^S is attained in $\bar{X}$.

In terms of f a solution can be characterized as follows.

Corollary 2.9. *A nonempty set $\bar{X}$ is a solution to $(\mathcal{L})$ if and only if the following conditions hold:*

(i) $\bar{X} \subseteq S$,

(ii) $f[\bar{X}] = \operatorname{Min} f[S]$,

(iii) $\inf_{x \in \bar{X}} f(x) = \inf_{x \in S} f(x)$.

Proof. Follows from Proposition 2.7. □

It can easily be seen that, if a solution to $(\mathcal{L})$ exists, then $\operatorname{Eff}(\mathcal{L})$ is a solution to $(\mathcal{L})$. Of course, if $\operatorname{Eff}(\mathcal{L})$ is a solution to $(\mathcal{L})$, every subset $\bar{X}$ of $\operatorname{Eff}(\mathcal{L})$ with $f[\bar{X}] = \operatorname{Min} f[S]$ is a solution to $(\mathcal{L})$, too.

Definition 2.10. If $\bar{X} = \operatorname{Eff}(\mathcal{L})$ is the only solution to $(\mathcal{L})$, we say $\bar{X}$ is a *unique* solution.

Example 2.11. Let $X = \mathbb{R}^2$, $Y = \mathbb{R}^2$ partially ordered by the cone $\mathbb{R}^2_+$ and $Z = \overline{Y}$. Then, $(Z, \leq)$ is a complete lattice. Let

$$S = \left\{ x \in \mathbb{R}^2 |\; x_1 \geq 0,\; x_2 \geq 0,\; 2x_1 + x_2 \geq 2,\; x_1 + 2x_2 \geq 2 \right\}$$

and let f be the identity map. Then, $\bar{X} := \operatorname{Eff}(\mathcal{L}) = \operatorname{Min} f[S]$ is the unique solution to $(\mathcal{L})$. For the same problem with the choice $f(x) := (0, x_2)^T$, we get $\operatorname{Eff}(\mathcal{L}) = \left\{ x \in \mathbb{R}^2 |\; x_1 \geq 2, x_2 = 0 \right\}$. But, every nonempty subset of $\operatorname{Eff}(\mathcal{L})$ is also a solution. Thus the solution is not unique in this case. Both cases are illustrated in Figure 2.4. Note that this example is not based on a useful solution concept for vector optimization, because the current complete lattice is not suitable.

Let us consider Problem $(\mathcal{L})$ for the special case $(Z, \leq) = (\overline{\mathbb{R}}, \leq)$.

Proposition 2.12. *Let $(Z, \leq) = (\overline{\mathbb{R}}, \leq)$. For a nonempty set $\bar{X}$, the following is equivalent:*

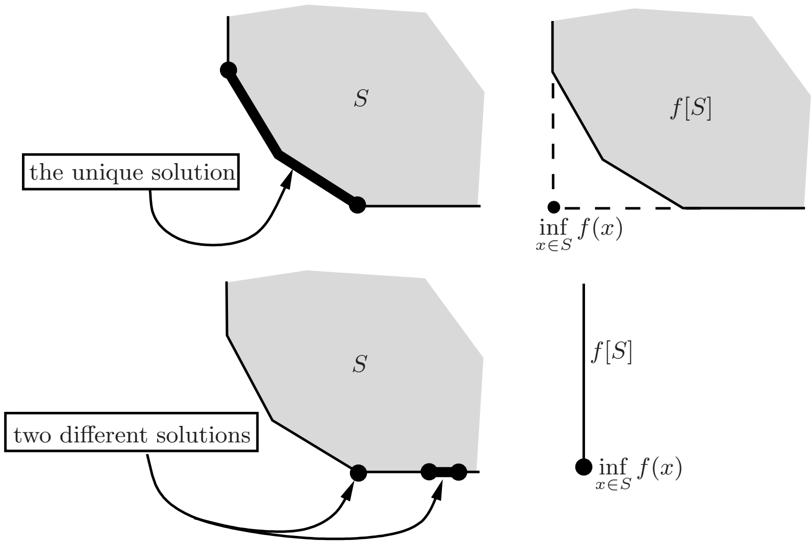

Fig. 2.4 Illustration of Example 2.11. Unique and non-unique solutions.

(i) $f[\bar{X}] = \operatorname{Min} f[S]$,
(ii) $\forall \bar{x} \in \bar{X} :\ \{f(\bar{x})\} = \operatorname{Min} f[S]$.

Proof. (i) $\Rightarrow$ (ii). This follows from the fact that $\operatorname{Min} f[S]$ is a singleton set, which is a consequence of $\leq$ being a total ordering in $\overline{\mathbb{R}}$ (i.e. arbitrary elements y^1, y^2 satisfy either $y^1 \leq y^2$ or $y^2 \leq y^1$).

(ii) $\Rightarrow$ (i). By (ii), f is constant on $\bar{X}$. Hence we have $\{f(\bar{x})\} = f[\bar{X}]$ for all $\bar{x} \in \bar{X}$. □

We next show the connection between a solution to the complete-lattice-valued problem $(\mathcal{L})$ for the case $(Z, \leq) = (\overline{\mathbb{R}}, \leq)$ and solutions to the classical extended real-valued optimization problem (2.1).

Theorem 2.13. *Consider Problem* $(\mathcal{L})$ *for the special case* $(Z, \leq) = (\overline{\mathbb{R}}, \leq)$ *and the corresponding real-valued optimization problem* (2.1)*. For a nonempty set* $\bar{X}$*, the following is equivalent:*

(i) $\bar{X}$ *is a solution to* $(\mathcal{L})$*,*
(ii) $\bar{x}$ *is a solution to* (2.1) *for every* $\bar{x} \in \bar{X}$*.*

An element $\bar{x}$ *is a unique solution to* (2.1) *if and only if* $\{\bar{x}\}$ *is a unique solution to* $(\mathcal{L})$*.*

Proof. (i) is equivalent to

$$\bar{X} \subseteq S \quad \wedge \quad \inf_{x \in \bar{X}} f(x) = \inf_{x \in S} f(x) \quad \wedge \quad f[\bar{X}] = \operatorname{Min} f[S].$$

By Proposition 2.12, this is equivalent to

$$\forall \bar{x} \in \bar{X}: \qquad \bar{x} \in S \quad \wedge \quad f(\bar{x}) = \inf_{x \in S} f(x) \quad \wedge \quad \{f(\bar{x})\} = \operatorname{Min} f[S],$$

which is an alternative way to express (ii). □

In Example 2.2 (where a complete lattice Z is obtained by extending $\mathbb{R}^2$ by two elements $\pm\infty$), $\operatorname{Eff}(\mathcal{L})$ is not a solution to $(\mathcal{L})$; whence a solution does not exist. A natural condition ensuring that $\operatorname{Eff}(\mathcal{L})$ is a solution is the well-known *domination property* (see e.g. Dolecki and Malivert, 1993).

Definition 2.14. Let $(Z, \leq)$ be a partially ordered set. We say that the *domination property* holds for Problem $(\mathcal{L})$ if

$$\forall x \in S, \;\; \exists \bar{x} \in \operatorname{Eff}(\mathcal{L}): \;\; f(\bar{x}) \leq f(x). \tag{2.3}$$

Proposition 2.15. *The set* $\operatorname{Eff}(\mathcal{L})$ *is a solution to* $(\mathcal{L})$ *if the domination property holds.*

Proof. Set $\bar{X} := \operatorname{Eff}(\mathcal{L})$. Of course, we have $\bar{X} \in 2^S$. According to Proposition 2.7, the attainment of the infimum of the canonical extension F over 2^S in $\bar{X}$ is equivalent to

$$\inf_{x \in \bar{X}} f(x) = \inf_{x \in S} f(x). \tag{2.4}$$

From (2.3) we get $\inf_{x \in \bar{X}} f(x) \leq \inf_{x \in S} f(x)$ and the opposite inequality in (2.4) follows from $\bar{X} \subseteq S$. □

The domination property is not necessary for the existence of a solution. An example is given below (Example 2.23).

2.2 A solution concept for vector optimization

A vector optimization problem is now transformed such that it becomes a special case of the complete-lattice-valued problem $(\mathcal{L})$. One can infer from the examples in the previous section that the choice of a suitable complete lattice $(Z, \leq)$ is rather essential. Originally, the image space of a vector optimization problem is a partially ordered vector space $(Y, \leq)$. In some cases, Y can be extended to a complete lattice by setting $Z := Y \cup \{\pm\infty\}$, where the ordering is extended in the usual way by setting $-\infty \leq z \leq +\infty$ for all z. We already mentioned the two drawbacks of this procedure. On the one hand, in many (even finite dimensional) cases we do not obtain a complete lattice in this way, see Example 1.9. On the other hand, even if a complete lattice is acquired in this way, our solution concept is unsatisfactory with this choice of Z. This is demonstrated by the following example.

Example 2.16. Let $X = \mathbb{R}^2$, $(Z, \le)$ the complete lattice from Example 2.5, f the identity map and

$$S = \left\{x \in \mathbb{R}^2 |\ x_1 > 0,\ x_2 > 0,\ x_1 + x_2 > 1\right\} \\ \cup \ \{x \in \mathbb{R}^2 |\ x_1 \ge 0, x_2 \ge 0, x_1 + x_2 \ge 2\}.$$

Then $\bar{X} := \left\{(0,2)^T, (2,0)^T\right\}$ is a solution. This is unsatisfactory from the viewpoint of vector optimization, because this set does not contain enough information, see Figure 2.5.

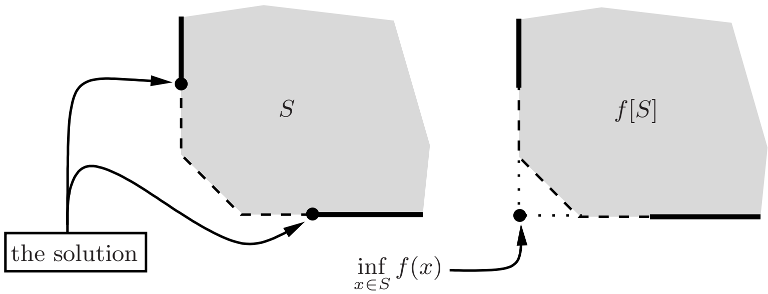

Fig. 2.5 Illustration of Example 2.16. The extended vector space $\mathbb{R}^2 \cup \{\pm\infty\}$ is a complete lattice, but not suitable for vector optimization.

The loophole is the usage of the complete lattice $\mathcal{I}$ of self-infimal subsets of $\overline{Y}$ instead of the space $\overline{Y}$ as the image space. The space $\mathcal{I}$ was introduced in Section 1.5. Recall further that we denote by $\operatorname{Inf} A$ the infimal set of a set $A \subseteq \overline{Y}$, see Section 1.4. We can identify a vector y in Y by the element $\operatorname{Inf}\{y\}$ of $\mathcal{I}$. In this way the ordering relation in $\mathcal{I}$ is an extension of the ordering relation in $\overline{Y}$. Note that the partial ordering on Y is generated by a pointed, convex cone C with $\emptyset \neq \operatorname{int} C \neq Y$, which is involved in the definition of infimal sets. The new image space $\mathcal{I}$ is a complete lattice even if $\overline{Y}$ is not a complete lattice. Moreover, an infimum is now an element of $\mathcal{I}$, which contains more information than a single vector. In particular, an infimum contains the information which is required by a solution concept based on the above postulates.

Let X be a nonempty set and $S \subseteq X$. Let $\overline{Y}$ be an extended partially ordered topological vector space, let the ordering cone C of Y be closed and let $\emptyset \neq \operatorname{int} C \neq Y$. Note that C is automatically pointed and convex, compare the remark after Definition 1.27. We consider the vector optimization problem

$$\text{minimize } f : X \to \overline{Y} \text{ with respect to } \le_C \text{ over } S. \qquad \text{(V)}$$

We assign to (V) a corresponding $\mathcal{I}$-valued-problem, i.e., a problem of type $(\mathcal{L})$, where the complete lattice $(Z, \le) = (\mathcal{I}, \preccurlyeq)$ is used. Note that $(\mathcal{I}, \preccurlyeq)$

is defined with respect to the ordering cone C of the vector optimization problem.

Given a function $f : X \to \overline{Y}$, we set

$$\bar{f} : X \to \mathcal{I}, \quad \bar{f}(x) := \operatorname{Inf}\{f(x)\}$$

and we assign to (V) the problem

$$\text{minimize } \bar{f} : X \to \mathcal{I} \text{ with respect to } \preccurlyeq \text{ over } S. \tag{$\mathcal{V}$}$$

Problem ($\mathcal{V}$) is said to be the *lattice extension*, or more precisely the *$\mathcal{I}$-extension*, of the vector optimization problem (V). This terminology can be motivated by the fact that the lattice extension of the vector optimization problem allows us to handle the problem in the framework of complete lattices. The ordering relation of the original objective space $\overline{Y}$ is extended to the complete lattice $\mathcal{I}$ as shown in the following proposition. Note that this extension is the reason for the assumption of C being closed.

Proposition 2.17. *For all $x, v \in X$ we have*

$$f(x) \leq_C f(v) \qquad \Longleftrightarrow \qquad \bar{f}(x) \preccurlyeq \bar{f}(v).$$

Proof. Let $\operatorname{Inf}\{y\} \preccurlyeq \operatorname{Inf}\{z\}$, then $\operatorname{Cl}_+\{y\} \supseteq \operatorname{Cl}_+\{z\}$. By Proposition 1.40, we get $z \in \operatorname{cl}(\{z\}+C) \subseteq \operatorname{cl}(\{y\}+C)$. Since C is closed, we obtain $z \in \{y\}+C$. This means $y \leq_C z$. The opposite inclusion is obvious. □

We next see that both problems (V) and ($\mathcal{V}$) are related as they have the same efficient solutions.

Proposition 2.18. *A feasible element $\bar{x} \in S$ is an efficient solution to the vector optimization problem* (V) *if and only if it is an efficient solution to its lattice extension* ($\mathcal{V}$)*.*

Proof. This is a direct consequence of Proposition 2.17. □

Proposition 2.19. *The domination property holds for the vector optimization problem* (V) *if and only it holds for its lattice extension* ($\mathcal{V}$)*.*

Proof. Follows from Proposition 2.17. □

We now define a solution concept for the vector optimization problem (V).

Definition 2.20. A nonempty set $\bar{X} \subseteq X$ is called a *solution* to the vector optimization problem (V) if $\bar{X}$ is a solution to its lattice extension ($\mathcal{V}$).

The next theorem provides a characterization of a solution to the vector optimization problem (V) by standard notations.

Theorem 2.21. *A nonempty set $\bar{X} \subseteq X$ is a solution to the vector optimization problem* (V) *if and only if the following three conditions are satisfied:*

(i) $\bar{X} \subseteq S$,
(ii) $f[\bar{X}] = \operatorname{Min} f[S]$,
(iii) $\operatorname{Inf} f[\bar{X}] = \operatorname{Inf} f[S]$.

Proof. This is a direct consequence of Proposition 2.7 and Theorem 1.54. □

Example 2.22. Consider the vector optimization problem (V) with a linear objective function f and a polyhedral convex feasible set S. Then, the set Eff $(\mathcal{L})$ is a solution whenever it is nonempty. As shown in (Hamel *et al.*, 2004, Lemma 2.1) (note that the cone has to be pointed there) the domination property is fulfilled in this case. Thus Proposition 2.15 yields that Eff $(\mathcal{L})$ is a solution.

Example 2.23. Consider the vector optimization problem (V) with $f : \mathbb{R}^2 \to \mathbb{R}^2$ being the identity map, let $C = \mathbb{R}^2_+$ and

$$S = \{x \in \mathbb{R}^2 |\; x_1 \geq 0,\; x_2 \geq 0,\; x_1 + x_2 \geq 1\} \setminus \{(0,1)^T\}.$$

Then $\bar{X} := \operatorname{Eff}(\mathcal{L}) = \{\lambda\,(0,1)^T + (1-\lambda)\,(1,0)^T |\; 0 \leq \lambda < 1\}$ is a solution, but the domination property is not satisfied.

As we will see in Chapter 3 the solution concept of Definition 2.8 is also relevant for problems which are not a lattice extension of a given vector optimization problem. There we consider a set-valued dual problem of a given vector optimization problem. In special cases, the values of the dual objective map are self-infimal hyperplanes.

Another lattice extension will be of interest in this work. The $\mathcal{I}$-valued extension $\bar{f} : X \to \mathcal{I}$ of a vectorial objective (as introduced above) is actually $\mathcal{I}_{\rm co}$-valued, see Section 1.6. Therefore, we also consider the lattice extension

$$\text{minimize } \bar{f} : X \to \mathcal{I}_{\rm co} \text{ with respect to } \preccurlyeq \text{ over } S. \qquad (\mathcal{V}_{\rm co})$$

If $\bar{f}$ is regarded to be $\mathcal{I}_{\rm co}$-valued, we have a different infimum and thus a different solution concept. Problem $(\mathcal{V}_{\rm co})$ is said to be the *convex lattice extension*, or more precisely, the $\mathcal{I}_{\rm co}$*-extension* of the vector optimization problem (V).

Definition 2.24. A nonempty set $\bar{X} \subseteq X$ is called a *convexity solution* or $\mathcal{I}_{\rm co}$*-solution* to the vector optimization problem (V) if $\bar{X}$ is a solution to the corresponding convex lattice extension $(\mathcal{V}_{\rm co})$.

Convexity solutions can be characterized as follows.

Theorem 2.25. *A nonempty set $\bar{X} \subseteq X$ is a convexity solution to the vector optimization problem* (V) *if and only if the following three conditions are satisfied:*

(i) $\bar{X} \subseteq S$,

(ii) $f[\bar{X}] = \operatorname{Min} f[S]$,
(iii) $\operatorname{Inf} \operatorname{co} f[\bar{X}] = \operatorname{Inf} \operatorname{co} f[S]$.

Proof. This is a direct consequence of Proposition 2.7 and Theorem 1.63. □

Proposition 2.26. *Every solution to* (V) *is also a convexity solution to* (V).

Proof. This follows from the fact that, by Proposition 1.60, $\operatorname{Inf} f[\bar{X}] = \operatorname{Inf} f[S]$ implies $\operatorname{Inf} \operatorname{co} f[\bar{X}] = \operatorname{Inf} \operatorname{co} f[S]$. □

Example 2.27. Let $X = \mathbb{R}^2$, $Y = \mathbb{R}^2$ partially ordered by the cone $\mathbb{R}^2_+$ and $Z = \overline{Y}$. Then, $(Z, \leq)$ is a complete lattice. Let

$$S = \left\{x \in \mathbb{R}^2 \middle|\ x_1 > 0,\ x_2 > 0,\ 2x_1 + x_2 > 2,\ x_1 + 2x_2 > 2\right\} \cup \left\{(0,2)^T, (2,0)^T, \left(\frac{2}{3}, \frac{2}{3}\right)^T\right\}$$

and let f be the identity map. Then, $\bar{X} = \left\{(0,2)^T, (2,0)^T, \left(\frac{2}{3}, \frac{2}{3}\right)^T\right\}$ is a convexity solution but not a solution to (V), see Figure 2.6.

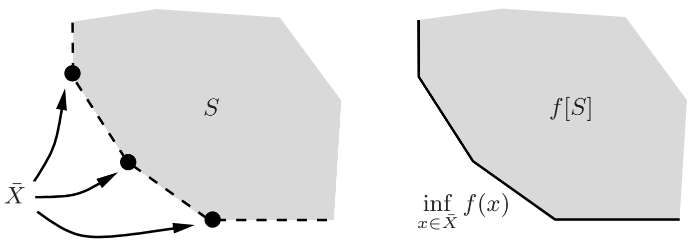

Fig. 2.6 Illustration of Example 2.27. The set $\bar{X}$ is a convexity solution but not a solution. The infimum on the right refers to the complete lattice $\mathcal{I}_{\text{co}}$. It coincides with $\inf_{x \in S} f(x)$.

This example looks somewhat artificial. Convexity solutions will play a role in Section 2.5, where we introduce *mild solutions* by relaxing the condition $f[\bar{X}] = \operatorname{Min} f[S]$. *Mild convexity solutions* will naturally occur in linear vector optimization problems.

2.3 Semicontinuity concepts

Lower semicontinuity of the objective function is typically required as an assumption for the existence of minimal solutions. This section provides a summary of different notions of lower semicontinuity for functions with values in Z, where Z is a partially ordered set and sometimes even a complete lattice. We are mainly interested in the general case without any a priori topology on Z. However, we also consider the special case where $Z = \overline{Y} := Y \cup \{\pm\infty\}$ is the extension of a partially ordered topological vector space Y. In particular, we examine $\overline{Y}$- and $\mathcal{F}$-valued functions. Note that $(\mathcal{F}, \supseteq)$ is isomorphic and isotone to $(\mathcal{I}, \preccurlyeq)$, see Proposition 1.52.

If $f : X \to Z = \overline{Y}$ is a function from a topological space X into the extended real numbers $\overline{\mathbb{R}}$, i.e., $Y = \mathbb{R}$ and $Z = \overline{\mathbb{R}}$, then the following five properties are equivalent characterizations of lower semicontinuity of f.

(a) For all $z \in Z$ the level sets $L_f(z) := \{x \in X |\ f(x) \le z\}$ are closed.
(b) For all $y \in Y$ the level sets $L_f(y)$ are closed.
(c) For all $\bar{x} \in X$,

$$f(\bar{x}) \le \sup_{U \in \mathcal{U}(\bar{x})} \inf_{x \in U} f(x) =: \liminf_{x \to \bar{x}} f(x)$$

holds true, where $\mathcal{U}(\bar{x})$ is a neighborhood base of $\bar{x}$ (that is, for every neighborhood U of $\bar{x}$, there exists some $\bar{U} \in \mathcal{U}(\bar{x})$ such that $\bar{U} \subseteq U$).
(d) For every $\bar{x} \in X$, every $\bar{y} \in Y$ with $\bar{y} \le f(\bar{x})$ and every neighborhood V of $\bar{y}$ there is some neighborhood U of $\bar{x}$ such that

$$\forall x \in U,\ \exists y \in V : f(x) \ge y.$$

(e) The epigraph of f, $\operatorname{epi} f := \{(x, y) \in X \times Y |\ f(x) \le y\}$, is closed.

For more general instances of Z these five properties do not coincide any longer. If Z is merely a complete lattice without any additional structure, then only the properties (a) and (c) are applicable.

Definition 2.28. Let X be a topological space, and let $(Y, \le)$ and $(Z, \le)$ be partially ordered sets. A function $f : X \to Z$ is called *level closed* if property (a) holds. $f : X \to \overline{Y}$ is called *weakly level closed* if property (b) holds. In case Z is a complete lattice, a function $f : X \to Z$ is called *lattice lower semi-continuous (lattice-l.s.c.)* if property (c) holds. In case Y is a partially ordered topological space $f : X \to \overline{Y}$ is called *topologically l.s.c.* if property (d) holds and *epi-closed* if (e) holds.

In the following we investigate the relationships between these properties. First we clarify the connection between the two notions that do not require further structural assumptions for the image space Z in addition to the lattice property.

Proposition 2.29. *Let X be a topological space and $(Z, \le)$ a complete lattice. If a function $f : X \to Z$ is lattice-l.s.c., then it is level closed.*

Proof. Assume that f is lattice-l.s.c. but not level closed, i.e., there is some $\bar{z} \in Z$ such that $L_f(\bar{z})$ is not closed. Then there is some $\bar{x} \in X$ with $\bar{x} \notin L_f(\bar{z})$ such that for all $U \in \mathcal{U}(\bar{x})$ there exists some $x \in U$ with $x \in L_f(\bar{z})$. This implies

$$\sup_{U \in \mathcal{U}(\bar{x})} \inf_{x \in U} f(x) \leq \bar{z}.$$

Since f is lattice-l.s.c., we conclude $f(\bar{x}) \leq \bar{z}$. But this means $\bar{x} \in L_f(\bar{z})$, a contradiction. □

The converse is generally not true as the following example shows.

Example 2.30. Let $X = \mathbb{R}$ and $Z = \overline{\mathbb{R}^2}$, where $\mathbb{R}^2$ is partially ordered by $\mathbb{R}^2_+$. The function $f : X \to Z$ defined by

$$f(x) = \begin{cases} (1,0)^T & \text{if } x \geq 0 \\ (0, -1/x)^T & \text{if } x < 0 \end{cases}$$

is level closed since

$$L_f(y) = \{x \in X |\ f(x) \leq y\}$$
$$= \begin{cases} [0, +\infty) & \text{if } y_2 = 0,\ y_1 \geq 1 \\ (-\infty, -1/y_2] \cup [0, +\infty) & \text{if } y_2 > 0,\ y_1 \geq 1 \\ (-\infty, -1/y_2] & \text{if } y_2 > 0,\ 0 \leq y_1 < 1 \\ \mathbb{R} & \text{if } y = +\infty \\ \emptyset & \text{otherwise.} \end{cases}$$

But f is not lattice-l.s.c. at $\bar{x} = 0$. Indeed, if we take the set of open ε-intervals as a neighborhood base of $\bar{x} = 0$, i.e., $\mathcal{U}(0) = \{(-\varepsilon, +\varepsilon) |\ \varepsilon > 0\}$, we obtain for every $U = (-\varepsilon, +\varepsilon) \in \mathcal{U}(0)$,

$$\inf_{x \in U} f(x) = \inf \left\{(0, 1/\varepsilon)^T, (1,0)^T\right\} = (0,0)^T.$$

We conclude

$$\sup_{U \in \mathcal{U}(0)} \inf_{x \in U} f(x) = (0,0)^T \not\geq (1,0)^T = f(0)$$

i.e., the condition of f being lattice-l.s.c. at $\bar{x} = 0$ is violated.

The following relations between epi-closedness, weak level closedness and level closedness follow immediately from the definitions.

Proposition 2.31. *Let X be a topological space, $(Y, \leq)$ a partially ordered topological space and $f : X \to \overline{Y}$. The following statements hold.*

(i) If f is epi-closed, then it is weakly level closed.
(ii) If f is level closed, then it is weakly level closed.

The converse implications are only true under additional assumptions.

Proposition 2.32. *Let X be a topological space.*

(i) *Assume that $(Y, \leq_C)$ is a topological vector space ordered by a pointed closed convex cone C with nonempty interior. If a function $f : X \to \overline{Y}$ is weakly level closed, then it is epi-closed.*
(ii) *Assume that $(Y, \leq)$ is a partially ordered set having no least element. If a function $f : X \to \overline{Y}$ is weakly level closed, then it is level closed.*

Proof. (i) Assume that f is weakly level closed, i.e., for every $\bar{x} \in X$ and $y \in Y$ we have

$$[\forall U \in \mathcal{U}(\bar{x}),\ \exists x \in U : f(x) \leq_C y] \quad \Longrightarrow \quad f(\bar{x}) \leq_C y. \tag{2.5}$$

In order to prove that f is epi-closed we assume that $(\bar{x}, \bar{y}) \in \operatorname{cl}(\operatorname{epi} f)$, i.e.,

$$\forall U \in \mathcal{U}(\bar{x}),\ \forall V \in \mathcal{V},\ \exists x \in U,\ \exists y \in V :\ f(x) \leq_C \bar{y} + y, \tag{2.6}$$

where $\mathcal{V}$ denotes a neighborhood base of 0 in Y. We have to show that $f(\bar{x}) \leq_C \bar{y}$. Since for every $c \in \operatorname{int} C$ there is some $V \in \mathcal{V}$ with $V \subseteq c - C$ we obtain from (2.6),

$$\forall c \in \operatorname{int} C,\ \forall U \in \mathcal{U}(\bar{x}),\ \exists x \in U :\ f(x) \leq_C \bar{y} + c.$$

Now, (2.5) implies that $f(\bar{x}) \leq_C \bar{y} + c$ holds for all $c \in \operatorname{int} C$. Therefore, we have $f(\bar{x}) \leq_C \bar{y}$ as C is closed.

(ii) It remains to show that $L_f(+\infty)$ and $L_f(-\infty)$ are closed. $L_f(+\infty) = X$ is closed by definition. Since Y has no least element, for $z \in \overline{Y}$ we have $z = -\infty$ if and only if $z \leq y$ for all $y \in Y$. Hence

$$L_f(-\infty) = \bigcap_{y \in Y} L_f(y)$$

is a closed set as well. □

In general, there is no inclusion between the sets of lattice-l.s.c., topologically l.s.c. and epi-closed functions (Gerritse, 1997, appendix). Some of the inclusions are valid under additional assumptions (Penot and Théra, 1982; Gerritse, 1997; Ait Mansour *et al.*, 2007). In this context we only mention the following two results.

Proposition 2.33. *Let X be a topological space and let $(Y, \leq)$ be a partially ordered topological space that has no greatest element. If the ordering of Y is closed (i.e., the set $G := \{(z, y) \in Y \times Y \mid z \leq y\}$ is closed) then every topologically l.s.c. function $f : X \to \overline{Y}$ is epi-closed.*

Proof. In order to prove that $\operatorname{epi} f$ is closed we take a pair $(\bar{x}, \bar{y}) \in (X \times Y) \setminus (\operatorname{epi} f)$ and show that there are neighborhoods $U \in \mathcal{U}(\bar{x})$ and $W \in \mathcal{V}(\bar{y})$ such that $(U \times W) \cap (\operatorname{epi} f) = \emptyset$.

If $(\bar{x},\bar{y}) \in (X \times Y) \setminus (\operatorname{epi} f)$, then $f(\bar{x}) \neq -\infty$ and $(f(\bar{x}),\bar{y}) \notin G$. Let $\hat{y} \in Y$ be chosen such that $\hat{y} \leq f(\bar{x})$ and $(\hat{y},\bar{y}) \notin G$. Such an element $\hat{y}$ always exists. Indeed, if $f(\bar{x}) \in Y$, we can use $\hat{y} = f(\bar{x})$. On the other hand, assuming that no such $\hat{y}$ exists in the case $f(\bar{x}) = +\infty$, we obtain that $\bar{y}$ is the greatest element of Y, a contradiction.

Since $(\hat{y},\bar{y}) \notin G$ and G is closed there exist neighborhoods V of $\hat{y}$ and W of $\bar{y}$ such that $(V \times W) \cap G = \emptyset$. Since f is topologically l.s.c. there exists a neighborhood $U \in \mathcal{U}(\bar{x})$ such that

$$\forall x \in U,\ \exists y \in V:\ y \leq f(x). \tag{2.7}$$

This implies $(U \times W) \cap (\operatorname{epi} f) = \emptyset$. Otherwise there would exist $\hat{x} \in U, \hat{y} \in W$ with $f(\hat{x}) \leq \hat{y}$. By (2.7) there would exist $y \in V$ with $y \leq f(\hat{x})$. Hence we obtain $y \leq \hat{y}$, which contradicts $(V \times W) \cap G = \emptyset$. □

Let us summarize the connections between the different notions. If Y is a partially ordered topological space with a closed ordering that has no greatest element such that $\overline{Y}$ is a complete lattice, then for functions $f : X \to \overline{Y}$ the concept of weak level closedness is the weakest one. It is equivalent to level closedness if Y has no least element. Moreover, epi-closedness, level closedness and weak level closedness coincide and the first two concepts are stronger than the last three.

We next study the relationship between the different concepts for functions with values in the complete lattice $(Z, \leq) = (\mathcal{F}, \supseteq)$, where $\mathcal{F} := \mathcal{F}_C(Y)$ is the space of upper closed subsets of a partially ordered topological vector space Y with an ordering cone C such that $\emptyset \neq \operatorname{int} C \neq Y$. Moreover, the ordering cone C is supposed to be closed. By Proposition 1.52, the corresponding results for the space $\mathcal{I}$ follow immediately.

In Propositions 2.34 and 2.35 the assumption $\emptyset \neq \operatorname{int} C \neq Y$ could be relaxed so that C is only required to be proper. In this case we would need a new definition of the upper closure, because our definition involves the interior of C. For this purpose the condition in Proposition 1.40 could be used.

We only consider the notions of (weakly) level closedness and lattice-semicontinuity in the case of $\mathcal{F}$-valued functions. A topology for $\mathcal{F}$ is not considered, but we investigate the connections to semi-continuity notions based on the topology of the underlying topological vector space Y. If we identify a function $f : X \to \mathcal{F}$ with a corresponding multivalued map $f : X \rightrightarrows Y$, the set

$$\operatorname{gr} f := \{(x,y) \in X \times Y |\ y \in f(x)\},$$

is called the *graph* of f.

Proposition 2.34. *A function $f : X \to \mathcal{F}$ is lattice-l.s.c. if and only if* $\operatorname{gr} f$ *is closed.*

Proof. We have

$$\sup_{U \in \mathcal{U}(\bar{x})} \inf_{x \in U} f(x) = \bigcap_{U \in \mathcal{U}(\bar{x})} \operatorname{cl} \bigcup_{x \in U} f(x).$$

It follows that

$$\bar{y} \in \sup_{U \in \mathcal{U}(\bar{x})} \inf_{x \in U} f(x)$$

is equivalent to

$$\forall U \in \mathcal{U}(\bar{x}),\ \forall V \in \mathcal{V}(\bar{y}),\ \exists x \in U,\ \exists y \in V:\ y \in f(x), \tag{2.8}$$

where $\mathcal{V}(\bar{y})$ denotes a neighborhood base of $\bar{y}$ in Y. Consequently f is lattice-l.s.c. if and only if for all $(\bar{x}, \bar{y})$ satisfying (2.8) one has $\bar{y} \in f(\bar{x})$. But, this is equivalent to $\operatorname{gr} f$ being closed. □

Proposition 2.35. *A function $f : X \to \mathcal{F}$ is level closed if and only if for all $y \in Y$ the sets $\{x \in X |\ y \in f(x)\}$ are closed.*

Proof. If f is level closed then the sets $L_f(\operatorname{Cl}_+\{y\})$ are closed for all $y \in Y$. We have $y \in f(x)$ if and only if $\operatorname{Cl}_+\{y\} \subseteq \operatorname{Cl}_+ f(x) = f(x)$. Thus the "only if"-part follows. The "if"-part follows from

$$L_f(A) = \bigcap_{y \in A} \{x \in X |\ y \in f(x)\}$$

and the fact that the intersection of closed sets is closed. □

Corollary 2.36. *Let $f : X \to \overline{Y}$ be an extended vector-valued function and $\tilde{f} : X \to \mathcal{F}$ its $\mathcal{F}$-valued extension, defined by $\tilde{f}(x) := \operatorname{Cl}_+\{f(x)\}$. Then $\tilde{f}$ is level closed if and only if f is weakly level closed.*

Proof. By Proposition 2.35, $\tilde{f}$ is level closed if and only if for all $y \in Y$ the sets $\{x \in X |\ y \in \tilde{f}(x)\}$ are closed. Similarly to Proposition 2.17, we have $y \in \tilde{f}(x)$ if and only if $y \geq_C f(x)$, where we use that C is closed. Thus the statement follows. □

By Proposition 2.29, every lattice-l.s.c. function is also level closed. For functions with values in $\mathcal{F}$ the converse implication also holds. As seen in Example 2.30 this is generally not true.

Proposition 2.37. *A function $f : X \to \mathcal{F}$ is lattice-l.s.c. if and only if it is level closed.*

Proof. Assume that f is level closed. We show that $\operatorname{gr} f$ is closed. By Proposition 2.34, this implies that f is lattice-l.s.c.. Assume that $(\bar{x}, \bar{y}) \in X \times Y$ is given such that for all $U \in \mathcal{U}(\bar{x}), V \in \mathcal{V}(\bar{y})$ there exist $x \in U, y \in V$ with $y \in f(x)$. We have to show that $\bar{y} \in f(\bar{x})$.

Take $z \in \{\bar{y}\} + \operatorname{int} C$ arbitrarily. Then there exists some neighborhood $\tilde{V} \in \mathcal{V}(\bar{y})$ such that $y \leq_C z$, i.e., $z \in \operatorname{Cl}_+\{y\}$ holds for all $y \in \tilde{V}$. Thus, for all $U \in \mathcal{U}(\bar{x})$ there exist some $x \in U$ and some $y \in \tilde{V}$ with $y \in f(x)$, hence $z \in \mathrm{Cl}_+ \{y\} \subseteq f(x)$. By Proposition 2.35 we get

$$\bar{x} \in \mathrm{cl}\, \{x \in X \mid z \in f(x)\} = \{x \in X \mid z \in f(x)\}\,.$$

Thus we have $\bar{y} + \mathrm{int}\, C \subseteq f(\bar{x})$ and consequently $\bar{y} \in \mathrm{cl}\,(\bar{y} + \mathrm{int}\, C) \subseteq f(\bar{x})$. □

We next formulate a sufficient condition for the domination property of the general optimization problem $(\mathcal{L})$. As in the classical Weierstrass theorem, the assumptions are lower semicontinuity of f and compactness of the feasible set. The appropriate semicontinuity condition for the function f in the general case is level closedness.

Proposition 2.38. *Let X be a compact topological space, $(Z, \leq)$ be a partially ordered set and $f : X \to Z$ a level closed function. Then the domination property holds, i.e., for every $x \in X$ there exists a minimal element $y \in f[X]$ with $y \leq f(x)$.*

Proof. We have to show that for every $x \in X$ the set $\{y \in f[X] \mid y \leq f(x)\} = f[L_f(f(x))]$ has minimal elements. Because of Zorn's lemma it suffices to show that every chain in $f[L_f(f(x))]$ has a lower bound in $f[L_f(f(x))]$. Since every lower bound (in $f[X]$) of a subset W of $f[L_f(f(x))]$ is obviously in $f[L_f(f(x))]$, it is sufficient to prove that every chain in $f[X]$ has a lower bound.

Let W be a chain in $f[X]$. A subset W of $f[X]$ has a lower bound in $f[X]$ if and only if the set

$$\{x \in X \mid \forall w \in W : f(x) \leq w\} = \bigcap_{w \in W} L_f(w)$$

is nonempty. If B is a finite subset of W then $\bigcap_{b \in B} L_f(b)$ is nonempty since every finite chain in $f[X]$ has a least element and hence a lower bound. Since all the sets $L_f(w)$ are closed, X being compact implies that $\bigcap_{w \in W} L_f(w)$ is nonempty, too. Hence W has a lower bound. □

For special cases of the complete lattice Z, the semicontinuity assumption in the latter result can be replaced by other concepts. For the case $(Z, \leq) = (\mathcal{I}, \preccurlyeq)$ this is pointed out in the next section. As a consequence we obtain the existence of solutions based on a variety of different semicontinuity notions. This matches the situation in scalar optimization.

2.4 A vectorial Weierstrass theorem

The results of the previous section can be applied to the vector optimization problems (V) and its lattice extension $(\mathcal{V})$ in order to obtain conditions for the existence of solutions.

Let X be a topological space and $S \subseteq X$. Moreover, let $\overline{Y}$ be an extended partially ordered topological vector space, let the ordering cone C of Y be closed and let $\emptyset \neq \operatorname{int} C \neq Y$. We consider the vector optimization problem (V) as introduced in Section 2.2 as well as its lattice extension $(\mathcal{V})$. The semicontinuity concept required for the existence result can be characterized in terms of the objective function $f : X \to \overline{Y}$ of (V) and in terms of the objective function $\bar{f} : X \to \mathcal{I}$ of the lattice extension $(\mathcal{V})$ of (V), where $\bar{f}$ is defined by f as

$$\bar{f}(x) := \operatorname{Inf}\{f(x)\}\,. \tag{2.9}$$

Theorem 2.39. *For a function $f : X \to \overline{Y}$ and the corresponding function $\bar{f} : X \to \mathcal{I}$ according to* (2.9)*, the following statements are equivalent:*

(i) *f is epi-closed, i.e., the epigraph of f is closed;*
(ii) *f is level closed, i.e., f has closed level sets for all levels in $\overline{Y}$;*
(iii) *f is weakly level closed, i.e., f has closed level sets for all levels in Y;*
(iv) *$\bar{f}$ is level closed, i.e., $\bar{f}$ has closed level sets for all levels in $\mathcal{I}$;*
(v) *$\bar{f}$ is lattice-l.s.c., i.e., for all $\bar{x} \in X$ one has $\bar{f}(\bar{x}) \preccurlyeq \liminf_{x \to \bar{x}} \bar{f}(x)$.*

Proof. The equivalence of (i), (ii) and (iii) follows directly from Proposition 2.31 and Proposition 2.32. The equivalence of (iii), (iv) and (v) follows from Corollary 2.36, Proposition 2.37, the fact (see Proposition 1.52) that a function $g : X \to \mathcal{F}$ is level closed (lattice l.s.c.) if and only if $j \circ g : X \to \mathcal{I}$ is level closed (lattice l.s.c.) and the fact that for the $\mathcal{I}$-valued extension $\bar{f}$ and the $\mathcal{F}$-valued extension $\tilde{f}$ of a function $f : X \to \overline{Y}$, $\bar{f} = j \circ \tilde{f}$ holds true. □

Applying Proposition 2.38 we can formulate the following existence result for a solution to a vector optimization problem. The result is a vectorial analogue of the famous Weierstrass theorem.

Theorem 2.40. *If one of the equivalent characterizations of lower semicontinuity in the preceding theorem is satisfied for the objective function $f : X \to \overline{Y}$ of* (V) *and if S is a compact subset of X, then there exists a solution to* (V)*.*

Proof. This is a direct consequence of Proposition 2.15, Proposition 2.38 and Theorem 2.39. □

It is remarkable that $\bar{f} : X \to \mathcal{I}$ being lattice-l.s.c. is an adequate semicontinuity assumption for a vectorial Weierstrass existence result. The condition that $f : X \to \overline{Y}$ is lattice-l.s.c. is usually (if it is well-defined at all) too strong and not satisfiable.

2.5 Mild solutions

For a solution $\bar{X}$ to the complete-lattice-valued optimization problem $(\mathcal{L})$ as defined in Section 2.1, the condition $f[\bar{X}] = \operatorname{Min} f[S]$ is part of the definition. This requirement can be by several reasons too strong. Relaxing this condition, we obtain an alternative solution concept.

Definition 2.41. A nonempty set $\hat{X}$ with $f[\hat{X}] \subseteq \operatorname{Min} f[S]$ is called a *mild solution* to $(\mathcal{L})$ if the infimum of the canonical extension F over 2^S is attained at $\hat{X}$.

The idea of a mild solution can be explained as follows. A mild solution $\hat{X}$ is allowed to be a smaller set than a solution. However, as the attainment of the infimum is required, the set $\hat{X}$ cannot become arbitrarily small. This ensures that $\hat{X}$ contains a sufficient amount of information. Of course, every solution to $(\mathcal{L})$ is also a mild solution to $(\mathcal{L})$. But a mild solution can be a proper subset of a solution.

Theorem 2.42. *If a mild solution to* $(\mathcal{L})$ *exists, then there exists a solution to* $(\mathcal{L})$.

Proof. Let $\hat{X}$ be a mild solution to $(\mathcal{L})$. Set $\bar{X} := \operatorname{Eff}(\mathcal{L})$, then $S \supseteq \bar{X} \supseteq \hat{X} \neq \emptyset$. Since $\inf_{x \in S} f(x) = \inf_{x \in \hat{X}} f(x)$, we get $\inf_{x \in S} f(x) = \inf_{x \in \bar{X}} f(x)$. Thus $\bar{X}$ is a solution to $(\mathcal{L})$. □

We now consider the vector optimization problem (V) as defined in Section 2.2.

Definition 2.43. A set $\hat{X}$ is called *mild solution* to the vector optimization problem (V) if it is a mild solution to its lattice extension $(\mathcal{V})$.

For the special case of a vector optimization problem, we have the following characterization of a mild solution.

Theorem 2.44. *Assume that a solution to* (V) *exists. A set* $\hat{X} \subseteq S$ *is a mild solution to* (V) *if and only if*

$$f[\hat{X}] \subseteq \operatorname{Min} f[S] \subseteq \operatorname{Inf} f[\hat{X}]. \tag{2.10}$$

Proof. If $\{x \in \hat{X} |\ f(x) = -\infty\} \neq \emptyset$, then

$$\{-\infty\} = \operatorname{Min} f[S] = \operatorname{Inf} f[S] = \operatorname{Inf} f[\hat{X}].$$

Hence $\hat{X}$ is a mild solution if and only if $f[\hat{X}] = \{-\infty\}$.

In case that $f(x) = +\infty$ for all $x \in S$ we have

$$\{+\infty\} = f[\hat{X}] = \operatorname{Min} f[S] = \operatorname{Inf} f[S] = \operatorname{Inf} f[\hat{X}]$$

for every nonempty subset $\hat{X} \subseteq S$. Therefore, every nonempty subset $\hat{X} \subseteq S$ is a mild solution.

We can assume that $f[S] \subseteq Y$ because otherwise we have

$$\operatorname{Min} f[S] = \operatorname{Min}(f[S] \setminus \{+\infty\}) \quad \text{and} \quad \operatorname{Inf} f[S] = \operatorname{Inf}(f[S] \setminus \{+\infty\}).$$

If $\hat{X}$ is a mild solution to (V), we have

$$\emptyset \neq \hat{X} \subseteq S \quad \wedge \quad f[\hat{X}] \subseteq \operatorname{Min} f[S] \quad \wedge \quad \operatorname{Inf} f[\hat{X}] = \operatorname{Inf} f[S].$$

It remains to show $\operatorname{Min} f[S] \subseteq \operatorname{Inf} f[\hat{X}]$. Let $y \in \operatorname{Min} f[S]$, i.e.,

$$y \in f[S] \qquad \wedge \qquad y \notin f[S] + C \setminus \{0\}.$$

It follows

$$\{y\} + \operatorname{int} C \subseteq f[S] + \operatorname{int} C \qquad \wedge \qquad y \notin f[S] + \operatorname{int} C.$$

We have $\emptyset \neq \operatorname{Cl}_+ f[S] \neq Y$. By Corollary 1.48 (ii) we get $y \in \operatorname{Inf} f[S] = \operatorname{Inf} f[\hat{X}]$.

Let $\bar{X}$ be a solution to (V) and let (2.10) be satisfied. It follows $f[\hat{X}] \subseteq f[\bar{X}] \subseteq \operatorname{Inf} f[\hat{X}]$. From Corollary 1.49 (i), we get $\operatorname{Cl}_+ f[\bar{X}] = \operatorname{Cl}_+ f[\hat{X}]$. Proposition 1.52 yields $\operatorname{Inf} f[\bar{X}] = \operatorname{Inf} f[\hat{X}]$. Hence $\hat{X}$ is a mild solution to (V). □

We next focus on a relationship to properly efficient solutions (e.g. Luc, 1988; Göpfert *et al.*, 2003; Jahn, 2004). The famous theorem by Arrow *et al.* (1953) and related results state that, under certain assumptions, the set of properly minimal vectors is a dense subset of the set of minimal vectors. In the literature, there are many density results for different types of proper efficiency (e.g. Borwein, 1980; Jahn, 1988; Ferro, 1999; Fu, 1996; Göpfert *et al.*, 2003). The following theorem shows that the set of proper efficient solutions is just an instance of a mild solution, whenever (under certain assumptions) a corresponding density result holds.

Theorem 2.45. *Assume that a solution to* (V) *exists. Let* $\hat{X} \subseteq S$ *be a set such that* $f[\hat{X}] \subseteq Y$ *and*

$$f[\hat{X}] \subseteq \operatorname{Min} f[S] \subseteq \operatorname{cl} f[\hat{X}]. \tag{2.11}$$

Then $\hat{X}$ *is a mild solution to* (V).

Proof. Let $\bar{X}$ be a solution to (V). Then $\operatorname{Min} f[S]$ is nonempty, hence $\operatorname{cl} f[\hat{X}]$ is nonempty and thus $\hat{X}$ is nonempty, too. We have

$$f[\hat{X}] \subseteq \operatorname{Min} f[S] = f[\bar{X}] \subseteq \operatorname{cl} f[\hat{X}].$$

Using Corollary 1.49 (i) and the fact $\operatorname{Cl}_+ \operatorname{cl} f[\hat{X}] = \operatorname{Cl}_+ f[\hat{X}]$, we get $\operatorname{Cl}_+ f[\bar{X}] = \operatorname{Cl}_+ f[\hat{X}]$. Proposition 1.52 yields $\operatorname{Inf} f[\bar{X}] = \operatorname{Inf} f[\hat{X}]$. Hence $\hat{X}$ is a mild solution to (V). □

In general, (2.11) does not hold for a mild solution $\hat{X}$ to (V).

Example 2.46. Let $X = \mathbb{R}^2$, $Y = \mathbb{R}^2$ partially ordered by $C = \mathbb{R}^2_+$, f the identity map and

$$S = \left\{x \in \mathbb{R}^2 \middle|\ x_1 > 0,\ x_2 > 0,\ x_1 + x_2 \geq 1\right\} \cup \left\{x \in \mathbb{R}^2 \middle|\ x_1 = 0, x_2 \geq 2\right\}.$$

Then $\hat{X} := \{\lambda\,(0,1)^T + (1-\lambda)\,(1,0)^T \mid \lambda \in (0,1)\}$ is a mild solution. But $(0,2)^T \in \operatorname{Min} f[S] \setminus \operatorname{cl} f[\hat{X}]$, hence (2.11) is violated, see Figure 2.7.

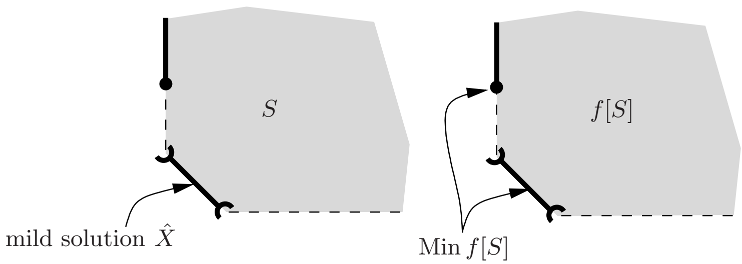

Fig. 2.7 Illustration of Example 2.46. The mild solution $\hat{X}$ does not satisfy the density condition (2.11).

If we additionally assume that $f[S]+C$ is closed and Y is a finite dimensional space, say $Y = \mathbb{R}^q$, we obtain that a mild solution satisfies (2.11). For instance, if S is a polyhedral convex set, C is polyhedral and f linear (see Chapter 4), then $f[S]+C$ is closed (Rockafellar, 1972, Theorem 19.3). Also, the assumptions of the Weierstrass existence result, Theorem 2.40, imply that $f[S]+C$ is closed (this follows from epi f being closed and S compact).

Theorem 2.47. *Let $Y = \mathbb{R}^q$. If $\hat{X}$ is a mild solution to* (V), *$f[S] \subseteq \mathbb{R}^q$ and $f[S]+C$ is closed, then*

$$f[\hat{X}] \subseteq \operatorname{Min} f[S] \subseteq \operatorname{cl} f[\hat{X}].$$

Proof. It remains to show the second inclusion. Let $y \in \operatorname{Min} f[S]$, i.e.,

$$y \in f[S] \subseteq \operatorname{cl}\,(f[S]+C) = \operatorname{Cl}_+ f[S].$$

and (take into account that the cone C is pointed and convex and $f[S]+C$ is closed)

$$\begin{aligned} y \notin f[S]+C\setminus\{0\} &= (f[S]+C)+C\setminus\{0\} \\ &= \operatorname{cl}\,(f[S]+C)+C\setminus\{0\} = \operatorname{Cl}_+ f[S]+C\setminus\{0\}. \end{aligned}$$

This yields $y \in \operatorname{Min}\operatorname{Cl}_+ f[S]$. As $\hat{X}$ is a mild solution, we have $\operatorname{Inf} f[\hat{X}] = \operatorname{Inf} f[S]$. Proposition 1.52 implies $\operatorname{Cl}_+ f[\hat{X}] = \operatorname{Cl}_+ f[S]$. Thus we have $y \in \operatorname{Min}\operatorname{Cl}_+ f[\hat{X}]$.

It remains to show that $\operatorname{Min}\operatorname{Cl}_+ f[\hat{X}] \subseteq \operatorname{cl} f[\hat{X}]$. Assuming the contrary, there exists some $y \in \operatorname{Cl}_+ f[\hat{X}] = \operatorname{cl}\,(f[\hat{X}]+C)$ such that $y \notin \operatorname{cl} f[\hat{X}]$ and

$$\big(y - C\setminus\{0\}\big) \cap \operatorname{cl}\,(f[\hat{X}]+C) = \emptyset. \tag{2.12}$$

Let (b_n) and (c_n) be sequences, respectively, in $f[\hat{X}]$ and C such that $b_n + c_n \to y$. There is no subsequence of c_n that converges to 0, because otherwise we get the contradiction $y \in \operatorname{cl} f[\hat{X}]$. Hence there exists $n_0 \in \mathbb{N}$ and $\alpha > 0$ such that $\|c_n\| \geq \alpha$ for all $n \geq n_0$. There is a subsequence $(c_n)_{n \in M}$ (M an infinite subset of $\{n \in \mathbb{N} |\; n \geq n_0\}$) such that

$$\tilde{c}_n := \frac{\alpha c_n}{\|c_n\|} \xrightarrow{M} \tilde{c} \in C \setminus \{0\}.$$

It follows

$$b_n + \left(1 - \frac{\alpha}{\|c_n\|}\right) c_n = b_n + c_n - \tilde{c}_n \xrightarrow{M} y - \tilde{c}.$$

We obtain $y - \tilde{c} \in \operatorname{cl}(f[\hat{X}] + C)$ which contradicts (2.12). □

In Section 2.2 we introduced convexity solutions to (V). To this end the complete lattice $\mathcal{I}$ is replaced by the complete lattice $\mathcal{I}_{\text{co}}$. We proceed in the same way and introduce *mild convexity solutions* to (V).

Definition 2.48. A nonempty set $\hat{X} \subseteq X$ is called a *mild convexity solution* or *mild* $\mathcal{I}_{\text{co}}$*-solution* to the vector optimization problem (V) if $\hat{X}$ is a mild solution to the corresponding convex lattice extension ($\mathcal{V}_{\text{co}}$).

Parallel to Theorem 2.25, mild convexity solutions can be characterized in terms of the vectorial objective function f.

Theorem 2.49. *A set $\hat{X} \subseteq X$ is a mild convexity solution to the vector optimization problem* (V) *if and only if the following three conditions are satisfied:*

(i) $\hat{X} \subseteq S$,
(ii) $f[\hat{X}] \subseteq \operatorname{Min} f[S]$,
(iii) $\operatorname{Inf} \operatorname{co} f[\hat{X}] = \operatorname{Inf} \operatorname{co} f[S]$.

Proof. This follows in the same way as Theorem 2.25. □

Corollary 2.50. *Every convexity solution to* (V) *is also a mild convexity solution to* (V).

Proof. This follows from Theorem 2.25 and Theorem 2.49. □

Corollary 2.51. *Every mild solution to* (V) *is also a mild convexity solution to* (V).

Proof. This follows from the fact that, by Proposition 1.60, $\operatorname{Inf} f[\hat{X}] = \operatorname{Inf} f[S]$ implies $\operatorname{Inf} \operatorname{co} f[\hat{X}] = \operatorname{Inf} \operatorname{co} f[S]$. □

The different solution concepts to (V) are compared in Figure 2.8.

The next example illustrates a mild convexity solution to a linear vector optimization problem. An essential advantage of mild convexity solutions is

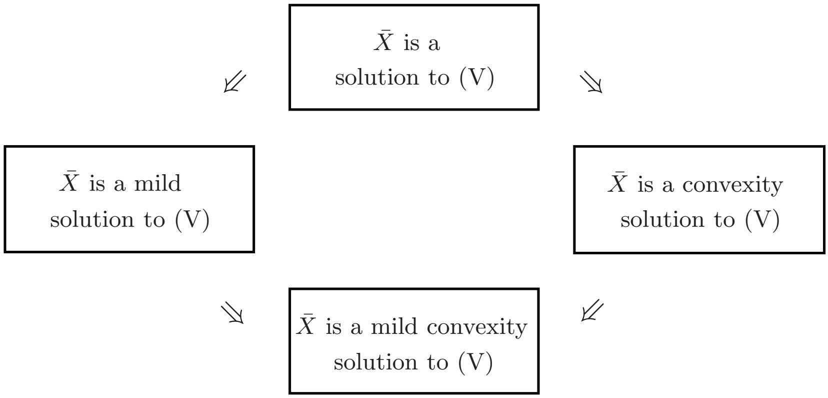

Fig. 2.8 Connections between different solution concepts to (V)

that finite sets sometimes are sufficient. In Chapter 4 we consider a modification of this concept in order to ensure that a “solution” to a linear vector optimization problem can always be a finite set. To this end we have to involve directions of the feasible set.

Example 2.52. Let $X = \mathbb{R}^2$, $Y = \mathbb{R}^2$ partially ordered by the cone $\mathbb{R}^2_+$. Consider Problem (V) with

$$S = \left\{x \in \mathbb{R}^2 |\ x_1 \geq 0,\ x_2 \geq 0,\ 2x_1 + x_2 \geq 2,\ x_1 + 2x_2 \geq 2\right\}$$

and let f be the identity map. Then

$$\hat{X} := \left\{ (0,2)^T, (2,0)^T, \left(\frac{2}{3}, \frac{2}{3}\right)^T \right\}$$

is a mild convexity solution to (V), see Figure 2.9.

2.6 Maximization problems and saddle points

Saddle points play a crucial role in duality theory. The goal of this section is to introduce saddle points in the framework of complete-lattice-valued optimization problems. As a consequence we obtain a corresponding saddle point notion for vector optimization problems, which differs from those in the literature. It is necessary to consider minimization and maximization problems simultaneously and it should be initially clarified how the solution concepts apply in this case.

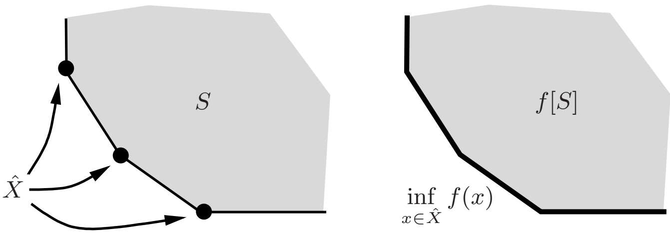

Fig. 2.9 Illustration of Example 2.52. On the left we see a mild convexity solution $\hat{X}$. On the right, the infimum of f over $\hat{X}$ with respect to the complete lattice $\mathcal{I}_{\text{co}}$ is shown. It coincides with $\inf_{x\in S} f(x)$.

Let V be a nonempty set, $T \subseteq V$ and let $(Z, \leq)$ be a complete lattice. Parallel to the minimization problem $(\mathcal{L})$ introduced in Section 2.1, we consider the complete-lattice-valued maximization problem

$$\text{maximize } g : V \to Z \text{ with respect to } \leq \text{ over } T. \qquad (\mathcal{L}_{\max})$$

The canonical extension of the function $g : V \to Z$ in the complete-lattice-valued maximization problem $(\mathcal{L}_{\max})$ is the function

$$G : 2^V \to Z, \quad G(B) := \sup_{v\in B} g(v).$$

The set of *maximal elements* of a set $B \subseteq Z$ is defined by

$$\operatorname{Max} B := \{z \in B |\ (y \in B \wedge y \geq z) \Rightarrow y = z\}.$$

A solution to $(\mathcal{L}_{\max})$ can now be defined in the same way as for Problem $(\mathcal{L})$ in Definition 2.8.

Definition 2.53. A nonempty set $\bar{V}$ with $g[\bar{V}] = \operatorname{Max} g[T]$ is called a *solution* to $(\mathcal{L}_{\max})$ if the supremum of the canonical extension G over 2^T is attained in $\bar{V}$.

In terms of g a solution can be characterized as follows.

Corollary 2.54. *A nonempty set $\bar{V}$ is a solution to $(\mathcal{L}_{\max})$ if and only if the following conditions hold:*

(i) $\bar{V} \subseteq T$,

(ii) $g[\bar{V}] = \operatorname{Max} g[T]$,

(iii) $\sup_{v\in\bar{V}} g(v) = \sup_{v\in T} g(v)$.

Proof. This follows from an analogous result to Proposition 2.7. □

Let X also be a nonempty set. We consider a function $l : X \times V \to Z$ depending on two variables, where we minimize with respect to the first variable and we maximize with respect to the second one. It turns out to be useful to distinguish between two types of canonical extensions for a function l depending on two variables. The function

$$L_l : 2^X \times 2^V \to Z, \quad L_l(\bar{X}, \bar{V}) := \sup_{v \in \bar{V}} \inf_{x \in \bar{X}} l(x, v)$$

is called the *lower canonical extension* of $l : X \times V \to Z$, and

$$L_u : 2^X \times 2^V \to Z, \quad L_u(\bar{X}, \bar{V}) := \inf_{x \in \bar{X}} \sup_{v \in \bar{V}} l(x, v)$$

is called the *upper canonical extension* of $l : X \times V \to Z$. This notion can be motivated by the fact that for all $(\bar{X}, \bar{V}) \in 2^X \times 2^V$ one has

$$L_l(\bar{X}, \bar{V}) \leq L_u(\bar{X}, \bar{V}),$$

which is an easy consequence of Z being a complete lattice.

Denoting by $+\infty$ and $-\infty$, respectively, the largest and the smallest element in Z, we set

$$S := \left\{ x \in X \,\middle|\, \sup_{v \in V} l(x, v) \neq +\infty \right\}$$

and

$$T := \left\{ v \in V \,\middle|\, \inf_{x \in X} l(x, v) \neq -\infty \right\}.$$

Let $p : X \to Z$ and $d : V \to Z$ be two functions such that

$$\forall x \in S : \; p(x) = \sup_{v \in V} l(x, v),$$

$$\forall v \in T : \; d(u) = \inf_{x \in X} l(x, v).$$

We assign to $l : X \times V \to Z$ the pair of dual optimization problems

$$\text{minimize } p : X \to Z \text{ with respect to } \leq \text{ over } S \subseteq X, \tag{2.13}$$

$$\text{maximize } d : V \to Z \text{ with respect to } \leq \text{ over } T \subseteq V. \tag{2.14}$$

Problem (2.13) corresponds to minimize $l : X \times V \to Z$ with respect to the first variable and likewise, Problem (2.14) corresponds to maximize $l : X \times V \to Z$ with respect to the second variable. Note that weak duality relation always holds, that is

$$\inf_{x \in S} p(x) = \inf_{x \in X} \sup_{v \in V} l(x, v) \leq \sup_{v \in V} \inf_{x \in X} l(x, v) = \sup_{v \in T} d(v).$$

According to our solution concept we propose the following notion of a saddle point for complete-lattice-valued problems.

Definition 2.55. Let X, V be two nonempty sets, $(Z, \leq)$ a complete lattice and let a function $l : X \times V \to Z$ be given. An element $(\bar{X}, \bar{V}) \in 2^S \times 2^T$, where $\bar{X} \neq \emptyset$ and $\bar{V} \neq \emptyset$, is called a *saddle point* of l if the following conditions are satisfied:

(i) $p[\bar{X}] = \operatorname{Min} p[S]$,
(ii) $d[\bar{V}] = \operatorname{Max} d[T]$,
(iii) $\forall A \in 2^X, \forall B \in 2^V : L_u(\bar{X}, B) \leq L_u(\bar{X}, \bar{V}) = L_l(\bar{X}, \bar{V}) \leq L_l(A, \bar{V})$.

Condition (iii) in the latter definition is a generalization of the well-known saddle point condition for an extended real-valued function, i.e., $(\bar{x}, \bar{v}) \in X \times V$ with $l(\bar{x}, \bar{v}) \in \mathbb{R}$ is a saddle point of $l : X \times V \to \overline{\mathbb{R}}$ if

$$\forall a \in X, \forall b \in V : \; l(\bar{x}, b) \leq l(\bar{x}, \bar{v}) \leq l(a, \bar{v}). \tag{2.15}$$

Note that in the extended real-valued case, $(\bar{x}, \bar{v}) \in S \times T$ implies $l(\bar{x}, \bar{v}) \in \mathbb{R}$. Vice versa, (2.15) and $l(\bar{x}, \bar{v}) \in \mathbb{R}$ implies $(\bar{x}, \bar{v}) \in S \times T$. Note further that condition (2.15) implies

$$\operatorname{Min} p[S] = \{p(\bar{x})\} \qquad \text{and} \qquad \operatorname{Max} d[T] = \{d(\bar{v})\}.$$

Consequently, conditions like (i) and (ii) of Definition 2.55 do not occur in the scalar case.

In our general setting, $(\bar{X}, \bar{V}) \in 2^S \times 2^T$ implies the following two conditions:

$$\forall a \in \bar{X} : \; L_u(\{a\}, \bar{V}) \neq +\infty \tag{2.16}$$

$$\forall b \in \bar{V} : \; L_l(\bar{X}, \{b\}) \neq -\infty. \tag{2.17}$$

Vice versa, if (iii) in Definition 2.55 holds, (2.16) $\wedge$ (2.17) implies $(\bar{X}, \bar{V}) \in 2^S \times 2^T$.

The following equivalent characterization of condition (iii) in Definition 2.55 is useful.

Lemma 2.56. *For nonempty sets $\bar{X} \subseteq X$ and $\bar{V} \subseteq V$, statement (iii) in Definition 2.55 is equivalent to*

$$\sup_{v \in \bar{V}} d(v) = \inf_{x \in \bar{X}} p(x). \tag{2.18}$$

Proof. From (iii) in Definition 2.55, we get

$$L_u(\bar{X}, V) \leq L_l(X, \bar{V})$$

and hence

$$\inf_{x \in \bar{X}} \sup_{v \in V} l(x, v) \leq \sup_{v \in \bar{V}} \inf_{x \in X} l(x, v).$$

Moreover, we have

$$\sup_{v\in\bar V}\inf_{x\in X} l(x,v) \le \inf_{x\in X}\sup_{v\in\bar V} l(x,v) \le \inf_{x\in\bar X}\sup_{v\in V} l(x,v).$$

This means that (2.18) is obtained from (iii) in Definition 2.55.

Now, let (2.18) be satisfied. It follows that

$$\forall A\in 2^X,\ \forall B\in 2^V:\quad \inf_{x\in\bar X}\sup_{v\in B} l(x,v) \le \sup_{v\in\bar V}\inf_{x\in A} l(x,v).$$

In particular, this implies

$$\forall A\in 2^X:\quad L_u(\bar X,\bar V) \le L_l(A,\bar V),$$

$$\forall B\in 2^V:\quad L_u(\bar X,B) \le L_l(\bar X,\bar V),$$

$$L_u(\bar X,\bar V) \le L_l(\bar X,\bar V).$$

Moreover, we have

$$L_l(\bar X,\bar V) \le L_u(\bar X,\bar V).$$

The last four statements imply statement (iii) in Definition 2.55. □

We are now able to relate saddle points to solutions of (2.13) and (2.14).

Theorem 2.57. *The following statements are equivalent:*

(i) $\bar X$ *is a solution to* (2.13), $\bar V$ *is a solution to* (2.14) *and*

$$\sup_{v\in T} d(v) = \inf_{x\in S} p(x);$$

(ii) $(\bar X,\bar V)$ *is a saddle point of* l.

Proof. Condition (i) can be equivalently expressed as

(a) $\operatorname{Min} p[S] = p[\bar X],\quad \emptyset \ne \bar X,\quad \bar X\subseteq S,$

(b) $\inf_{x\in\bar X} p(x) = \inf_{x\in S} p(x),$

(c) $\operatorname{Max} d[T] = d[\bar V],\quad \emptyset \ne \bar V,\quad \bar V\subseteq T,$

(d) $\sup_{v\in\bar V} d(v) = \sup_{v\in T} d(v),$

(e) $\sup_{v\in T} d(v) = \inf_{x\in S} p(x).$

In view of Lemma 2.56 it remains to show that (b)∧(d)∧(e) is equivalent to (2.18) in the present situation. Of course, (b)∧(d)∧(e) implies (2.18). On the other hand, since $\bar X\subseteq S$ and $\bar V\subseteq T$, (2.18) implies that

$$\inf_{x\in\bar X} p(x) = \sup_{v\in\bar V} d(v) \le \sup_{v\in T} d(v) \le \inf_{x\in S} p(x) \le \inf_{x\in\bar X} p(x).$$

The last expression holds with equality. This yields (b)∧(d)∧(e). □

We next focus on the special case $(Z, \leq) = (\overline{\mathbb{R}}, \leq)$ and show that an ordinary saddle point is obtained.

Theorem 2.58. *For $l : X \times V \to \overline{\mathbb{R}}$ the following is equivalent.*

(i) *$(\bar{X}, \bar{V})$ is a saddle point of l in the sense of Definition 2.55.*
(ii) *Every $(\bar{x}, \bar{v}) \in \bar{X} \times \bar{V}$ is a saddle point of l in the classic sense, that is $(\bar{x}, \bar{v}) \in X \times V$ with $l(\bar{x}, \bar{v}) \in \mathbb{R}$ such that (2.15) holds.*

Proof. As discussed above, $(\bar{x}, \bar{v}) \in S \times T$ corresponds to $l(\bar{x}, \bar{v}) \in \mathbb{R}$ in the present situation. By Theorem 2.57 and Theorem 2.13, (i) is equivalent to

$$\forall \bar{x} \in \bar{X},\ \forall \bar{v} \in \bar{V} : \quad p(\bar{x}) = \inf_{x \in S} p(x) = \sup_{v \in T} d(v) = d(\bar{v}). \tag{2.19}$$

From the definition of p and d we get $p(\bar{x}) \geq l(\bar{x}, \bar{v}) \geq d(\bar{v})$ and (2.19) yields $p(\bar{x}) = l(\bar{x}, \bar{v}) = d(\bar{v})$ for all $\bar{x} \in \bar{X}$ and all $\bar{v} \in \bar{V}$. This implies (ii).

On the other hand, (ii) implies that for all $\bar{x} \in \bar{X}$ and all $\bar{v} \in \bar{V}$ one has

$$\begin{aligned} \inf_{x \in S} p(x) \leq p(\bar{x}) &= \sup_{b \in V} l(\bar{x}, b) \leq l(\bar{x}, \bar{v}) \\ &\leq \inf_{a \in X} l(a, \bar{v}) = d(\bar{v}) \leq \sup_{v \in T} d(v). \end{aligned}$$

Weak duality yields equality. This implies (2.19). □

Similarly to mild solutions we can define mild saddle points by relaxing the conditions (i) and (ii) in Definition 2.55.

Definition 2.59. Let X, V be two nonempty sets, $(Z, \leq)$ a complete lattice and let a function $l : X \times V \to Z$ be given. An element $(\hat{X}, \hat{V}) \in 2^S \times 2^T$, where $\hat{X} \neq \emptyset$ and $\hat{V} \neq \emptyset$, is called a *mild saddle point* of l if the following conditions are satisfied:

(i) $p[\hat{X}] \subseteq \operatorname{Min} p[S]$,
(ii) $d[\hat{V}] \subseteq \operatorname{Max} d[T]$,
(iii) $\forall A \in 2^X, \forall B \in 2^V : L_u(\hat{X}, B) \leq L_u(\hat{X}, \hat{V}) = L_l(\hat{X}, \hat{V}) \leq L_l(A, \hat{V})$.

A corresponding characterization follows immediately.

Theorem 2.60. *The following statements are equivalent:*

(i) *$\hat{X}$ is a mild solution to (2.13), $\hat{V}$ is a mild solution to (2.14) and*

$$\sup_{v \in T} d(v) = \inf_{x \in S} p(x);$$

(ii) *$(\hat{X}, \hat{V})$ is a mild saddle point of l.*

Proof. Similarly to the proof of Theorem 2.57. □

The notion of a saddle point introduced in this section can be used for arbitrary $\mathcal{I}$-valued problems. In case of a vector optimization problem we consider its lattice extension which yields an $\mathcal{I}$-valued problem. We obtain an $\mathcal{I}$-valued Lagrangian and an $\mathcal{I}$-valued dual problem. Thus, the saddle point notions introduced in this section easily apply to vector optimization.

2.7 Notes on the literature

In the framework of a mathematical optimization theory, the notion of an efficient element seems to be first used by Koopmans (1951), compare (Stadler, 1979), but the ideas can be traced back to the early works by Pareto and Edgeworth. Modifications of efficient solutions, such as weakly or properly efficient solutions, are commonly considered in the literature (Luc, 1988; Jahn, 1986, 2004; Ehrgott, 2000; Boţ *et al.*, 2009). The idea to compute a subset of the efficient solutions in order to present it to a decision maker is standard in the literature on vector optimization. Nevertheless, there is no unique and precise specification of such a subset, which is understood as a solution concept.

The solution concept for complete-lattice-valued problems in Section 2.1 and its application to vector optimization in Section 2.2 including the notion of a mild solution first appeared in (Heyde and Löhne, 2010). It should be mentioned that these ideas arose from several discussions about solution concepts for set-valued optimization problems between Andreas H. Hamel and the mentioned authors. The notion of (mild) convexity solutions and all the related results are new in this book.

Section 2.3 is a collection of results on semicontinuity concepts for set-valued maps which can be found similarly in the literature. The results and proofs in the presented form are taken from Heyde and Löhne (2010) and are due to the first author. Definition 2.28 follows the articles by Gerritse (1997) and Ait Mansour *et al.* (2007). Note that in (Ait Mansour *et al.*, 2007) the term "level closed" is used for property (b). We call a function level closed if all level sets are closed and we speak about weak level closedness if the weaker property (b) holds. The notions of lattice- and topological semicontinuity are introduced in (Gerritse, 1997). The definition of lattice semicontinuity coincides with that of Gerritse (1997). The definition of topological semicontinuity differs slightly from that in (Gerritse, 1997) since we do not require a topological structure on the whole set $\overline{Y}$. It coincides, however, with the concept denoted simply by lower semicontinuity in (Ait Mansour *et al.*, 2007). Note also that Gerritse (1997) deals with upper rather than lower semicontinuity. Proposition 2.33 is slightly different from (Penot and Théra, 1982, Proposition 1.3.a) but the proof follows essentially the lines of the one in (Penot and Théra, 1982). Note further that it was shown by Liu and Luo (1991, Theorem 3.6.) that every level closed function $f : X \to Z$

is lattice-l.s.c. if and only if Z is a completely distributive lattice (compare Proposition 2.29).

The existence result in Section 2.4, the notion of mild solutions as well as all related results in Section 2.5 are due to Heyde and Löhne (2010). There are other existence results in the literature; partially they are related to the domination property (e.g. Jahn, 1986, 2004; Luc, 1988; Sonntag and Zălinescu, 2000).

Saddle points for complete-lattice-valued problems as well as all related concepts and results in Section 2.6 seem to be new and arose from discussions with Andreas H. Hamel. In the literature (see e.g. Rödder, 1977; Luc, 1988; Tanaka, 1990, 1994; Li and Wang, 1994; Tan *et al.*, 1996; Li and Chen, 1997; Ehrgott and Wiecek, 2005b; Adán and Novo, 2005) there are other notions of saddle points for vector optimization problems which are not based on the structure of a complete lattice.

Chapter 3
Duality

The dual of a scalar optimization problem is again a scalar optimization problem. This simple fact seems to be the reason that the dual of a vector optimization problem is commonly expected to be a vector optimization problem. But the following is also true: The dual of a scalar optimization problem is a hyperplane-valued optimization problem. This is due to the simple fact that every hyperplane in $\mathbb{R}$ consists of a single real number. It is also true that in scalar optimization the values of the dual objective function are self-infimal sets, because the space of self-infimal sets in $\overline{\mathbb{R}}$ is just the family of all singleton sets in $\overline{\mathbb{R}}$, which can be identified with $\overline{\mathbb{R}}$. This demonstrates that hyperplane-valued or $\mathcal{I}$-valued dual problems are by no means unnatural.

Duality for vector optimization problems can be developed on the basis of the complete lattice $\mathcal{I}$. In this way a vectorial counterpart to the scalar duality theory with a high degree of analogy can be established. In contrast to many other approaches to duality theory of vector optimization problems in the literature (see e.g. the recent book by Boţ *et al.*, 2009) the most results are formulated completely parallel to their scalar versions. The infimum and supremum are consequently used and the dual variables are kept to be linear functionals rather than operators. The dual problems are $\mathcal{I}$-valued and can typically not be derived from vector optimization problems, at least not in the way as the primal $\mathcal{I}$-valued problem is derived from a given vector optimization problem (but compare Section 4.5).

The content of this chapter can be understand as a demonstration how a scalar duality theory can be transformed into a vectorial framework. Out of a variety of possibilities, we only consider conjugate duality and an instance of Lagrange duality with set-valued constraints as well as a related saddle point theorem.

Two types of $\mathcal{I}$-valued dual problems are studied. The main feature of the type I dual problem is a simple variable space, in fact, the same as for the corresponding scalar dual problem. Type II dual problems have slightly more complex dual variables, but therefore, the values of the dual objective function have a simpler structure, in fact, they are essentially hyperplanes. In

A. Löhne, *Vector Optimization with Infimum and Supremum*, Vector Optimization,
DOI 10.1007/978-3-642-18351-5_4, © Springer-Verlag Berlin Heidelberg 2011

this sense, the type II dual objective functions are not more complicated than the (original) vector-valued primal objective functions. Dual problems of type II also have the advantage that, as in scalar duality theory, the existence of a solution to the dual problem can be shown as a part of the strong duality theorem.

3.1 A general duality concept applied to vector optimization

We consider a complete-lattice-valued optimization problem and introduce a general duality concept, where vector optimization problems constitute a special case.

Let $p : X \to Z$, where X is an arbitrary nonempty set and $(Z, \leq)$ is a complete lattice. For a nonempty subset $S \subseteq X$, we consider the optimization problem

$$\text{minimize } p : X \to Z \text{ with respect to } \leq \text{ over } S, \tag{P}$$

which is called the *primal problem*. Simultaneously, we consider a dual optimization problem. Let $d : V \to Z$, where V is an arbitrary nonempty set and $T \subseteq V$ is a nonempty subset, called the *dual feasible set*. We consider the *dual problem*

$$\text{maximize } d : V \to Z \text{ with respect to } \leq \text{ over } T. \tag{D}$$

Definition 3.1. We say that *weak duality* holds for the pair of problems (P) and (D) if we have the implication

$$(x \in S \;\wedge\; v \in T) \implies d(v) \leq p(x).$$

Definition 3.2. We say that *strong duality* holds for the pair of problems (P) and (D) if we have

$$\sup_{v \in T} d(v) = \inf_{x \in S} p(x).$$

It is clear that weak duality can equivalently be characterized by the inequality

$$\sup_{v \in T} d(v) \leq \inf_{x \in S} p(x).$$

Hence strong duality implies weak duality.

We next show that duality is invariant with respect to the canonical extension, which we introduced in Section 2.1 and in Section 2.6. We consider the extended problems

$$\text{minimize } P : 2^X \to Z \text{ with respect to } \leq \text{ over } 2^S, \tag{$\mathcal{P}$}$$

$$\text{maximize } D : 2^V \to Z \text{ with respect to } \leq \text{ over } 2^T, \qquad (\mathcal{D})$$

where we set

$$P : 2^X \to Z, \quad P(A) := \inf_{x \in A} p(x),$$

$$D : 2^V \to Z, \quad D(B) := \sup_{v \in B} d(v).$$

Proposition 3.3. *Weak (strong) duality holds for the pair of problems* (P) *and* (D) *if and only if weak (strong) duality holds for the pair of their canonical extensions* ($\mathcal{P}$) *and* ($\mathcal{D}$)*; that is*

$$\sup_{v \in T} d(v) \leq \inf_{x \in S} p(x) \quad \Longleftrightarrow \quad \sup_{B \in 2^T} D(B) \leq \inf_{A \in 2^S} P(A),$$

$$\sup_{v \in T} d(v) = \inf_{x \in S} p(x) \quad \Longleftrightarrow \quad \sup_{B \in 2^T} D(B) = \inf_{A \in 2^S} P(A).$$

Proof. Since

$$\sup_{B \in 2^T} D(B) = \sup_{B \in 2^T} \sup_{v \in B} d(v) = \sup_{v \in T} d(v)$$

and

$$\inf_{A \in 2^S} P(A) = \inf_{A \in 2^S} \inf_{x \in A} p(x) = \inf_{x \in S} p(x),$$

the statement is obvious. □

We now apply the general duality principle to vector optimization problems. Let X be a nonempty set and $\overline{Y}$ an extended partially ordered topological vector space. Assume that the ordering cone C of Y is closed and $\emptyset \neq \operatorname{int} C \neq Y$. Note that C is automatically convex and closed since we suppose a partial ordering.

Let $p : X \to \overline{Y}$ a given vectorial objective function and $S \subseteq X$ a given feasible set. The primal problem is considered to be the vector optimization problem

$$\text{minimize } p : X \to \overline{Y} \text{ with respect to } \leq_C \text{ over } S. \qquad \text{(VOP)}$$

We assign to (VOP) a corresponding $\mathcal{I}$-valued problem, i.e., a problem of type (P), where the complete lattice $(Z, \leq) = (\mathcal{I}, \preccurlyeq)$ is used. As already seen in Section 2.2, we obtain a closely related complete-lattice-valued problem, even if $\overline{Y}$ is not a complete lattice with respect to the ordering relation generated by C.

We set

$$p_{\mathcal{I}} : X \to \mathcal{I}, \quad p_{\mathcal{I}}(x) := \operatorname{Inf}\{p(x)\}$$

and assign to (VOP) the problem

$$\text{minimize } p_{\mathcal{I}} : X \to \mathcal{I} \text{ with respect to } \preccurlyeq \text{ over } S. \qquad (\mathrm{P}_{\mathcal{I}})$$

Based on the properties of the complete lattice $\mathcal{I}$, a dual problem can be derived for special classes of problems such that weak and strong duality can be shown. We consider a set V and a feasible subset $T \subseteq V$, a dual objective function $d_{\mathcal{I}} : V \to \mathcal{I}$ and the dual problem

$$\text{maximize } d_{\mathcal{I}} : V \to \mathcal{I} \text{ with respect to } \preccurlyeq \text{ over } T. \tag{$\mathrm{D}_{\mathcal{I}}$}$$

The optimal values of ($\mathrm{P}_{\mathcal{I}}$) and ($\mathrm{D}_{\mathcal{I}}$) are defined, respectively, by

$$\bar{p}_{\mathcal{I}} := \inf_{x \in S} p_{\mathcal{I}}(x) \qquad \text{and} \qquad \bar{d}_{\mathcal{I}} := \sup_{v \in T} d_{\mathcal{I}}(v).$$

In the case that duality assertions can be shown for a class of problems, the dual problem ($\mathrm{D}_{\mathcal{I}}$) can usually not be re-interpreted as a vector optimization problem, at least not in the way as ($\mathrm{P}_{\mathcal{I}}$) is obtained from (VOP) (but compare Section 3.5 and Section 4.5). Therefore, ($\mathrm{D}_{\mathcal{I}}$) itself is considered to be the dual problem to (VOP). It is possible to obtain set-valued dual problems with an easy structure; for certain classes of problems we get a hyperplane-valued dual objective function.

Subsequently we consider special classes of problems and derive weak and strong duality statements between ($\mathrm{P}_{\mathcal{I}}$) and ($\mathrm{D}_{\mathcal{I}}$). A duality theory for $\mathcal{I}$-valued problems is developed which applies to the lattice extension ($\mathrm{P}_{\mathcal{I}}$)of a vector optimization problem (VOP). However, the duality theory also works for $\mathcal{I}$-valued problems which do not arise from a vector optimization problem.

3.2 Conjugate duality

As a first instance of duality, conjugate duality (also called Fenchel duality) is studied in this section.

Throughout this section let $\overline{Y}$ be an extended partially ordered locally convex space with an ordering cone $C \subseteq Y$ such that $\emptyset \neq \operatorname{int} C \neq Y$ and let Y^* be the topological dual space. We set $\mathcal{I} := \mathcal{I}_C\left(\overline{Y}\right)$. In general, C is not required to be closed. However, if the lattice extension of a vector optimization problem is considered, C being closed ensures that the ordering in $\mathcal{I}$ is an extension of the ordering in the vector space Y, compare Proposition 2.17.

Furthermore, we assume X and U to be locally convex Hausdorff spaces. We denote by X^* and U^* their topological dual spaces and by $\langle X, X^* \rangle$ and $\langle U, U^* \rangle$ the corresponding dual pairs, see Definition 1.38. Let $B\colon X \to U$ be a linear continuous function. The map $B^*\colon U^* \to X^*$, defined by

$$\forall u^* \in U^*,\ \forall x \in X:\ \langle u^*, Bx \rangle_U = \langle B^* u^*, x \rangle_X\,,$$

is called the *adjoined function* of B. As shown, for instance, in (Göpfert, 1973, p. 75f), the adjoined map is well-defined, linear and continuous (where X^* and U^* are equipped with their weak* topologies).

Definition 3.4. The conjugate of a function $f : X \to \mathcal{I}$ (with respect to some fixed $c \in Y$) is defined by

$$f_c^* : X^* \to \mathcal{I}, \qquad f_c^*(x^*) := \sup_{x \in X} \big(\langle x^*, x \rangle \{c\} - f(x) \big).$$

We know that if f is an $\mathcal{I}$-valued function, so is $-f$. This follows from the fact

$$A \in \mathcal{I} \quad \iff \quad -A \in \mathcal{I},$$

which has been shown in Corollary 1.55. So the term $\langle x^*, x \rangle \{c\} - f(x)$ stands for a shift of $-f(x) \in \mathcal{I}$. This means that the minus sign has to be interpreted in the sense of Minkowski-addition. As a result we have $\langle x^*, x \rangle \{c\} - f(x) \in \mathcal{I}$ and the supremum in the definition of the conjugate is well-defined.

3.2.1 Conjugate duality of type I

Let $f : X \to \mathcal{I}$, $g : U \to \mathcal{I}$ and let $B : X \to U$ be a continuous linear map. The primal problem ($\mathrm{P}_{\mathcal{I}}$) in Section 3.1 is now considered for the objective function

$$p : X \to \mathcal{I}, \qquad p(x) := f(x) \oplus g(Bx).$$

For the dual problem ($\mathrm{D}_{\mathcal{I}}$), we specify the objective function as

$$d_c : U^* \to \mathcal{I}, \qquad d_c(u^*) := -f_c^*(B^* u^*) \oplus -g_c^*(-u^*). \tag{3.1}$$

If we set $S = X$ and $T = U^*$, the problems ($\mathrm{P}_{\mathcal{I}}$) and ($\mathrm{D}_{\mathcal{I}}$) turn into the following problems; the conjugate (or Fenchel) primal problem

$$\text{minimize } p : X \to \mathcal{I} \text{ with respect to } \preccurlyeq \text{ over } X \tag{$\mathrm{P_F}$}$$

and the type I dual problem associated to ($\mathrm{P_F}$)

$$\text{maximize } d_c : U^* \to \mathcal{I} \text{ with respect to } \preccurlyeq \text{ over } U^*. \tag{$\mathrm{D_F^I}$}$$

The *optimal values* of ($\mathrm{P_F}$) and ($\mathrm{D_F^I}$) are denoted, respectively, by

$$\bar{p} := \inf_{x \in X} p(x) \in \mathcal{I} \qquad \text{and} \qquad \bar{d}_c := \sup_{u^* \in U^*} d_c(u^*) \in \mathcal{I}.$$

Continuity of an $\mathcal{I}$-valued function is defined with respect the uniform topology in $\mathcal{I}$ introduced in Section 1.8. Subsequently, we set

$$B_c := \{y \in C^\circ |\, y^*(c) = -1\}\,, \tag{3.2}$$

where C° is the polar cone as defined in (1.10). If $c \in \operatorname{int} C$, then B_c is a *base* of C°, that is, each nonzero element y^* of C° has a unique representation by an element b^* of the convex set B_c by $y^* = \lambda b^*$ for some $\lambda > 0$ (Peressini, 1967, p. 25). For a comparison of different types of bases of cones the reader is referred to Göpfert *et al.* (2003).

We continue with a conjugate duality theorem. The formulation is exactly the same as in the scalar case. This high degree of analogy is one of the essential advantages of using the supremum and infimum in vector optimization.

Theorem 3.5. *The problems* $(\mathrm{P_F})$ *and* $(\mathrm{D^I_F})$ *(with arbitrary* $c \in Y$*) satisfy the weak duality inequality, that is,* $\bar{d}_c \preccurlyeq \bar{p}$*. Furthermore, let* f *and* g *be proper convex functions,* $c \in \operatorname{int} C$*, and let the following constraint qualification be satisfied:*

$$\textit{There exists } u \in \operatorname{dom} g \cap B(\operatorname{dom} f) \textit{ such that } g \textit{ is continuous at } u. \tag{3.3}$$

Then strong duality holds, that is, $\bar{d}_c = \bar{p}$.

Proof. For all $u^* \in U^*$, $x \in X$ and $u \in U$,

$$-f^*(B^*u^*) \oplus -g^*(-u^*) \preccurlyeq (f(x) - \langle B^*u^*, x\rangle) \oplus (g(u) + \langle u^*, u\rangle).$$

Setting $u := Bx$ and recollecting that $\langle B^*u^*, x\rangle = \langle u^*, Bx\rangle$, we get $d_c(u^*) \preccurlyeq p(x)$ for all $u^* \in U^*$ and all $x \in X$. Taking the supremum over $u^* \in U^*$ and the infimum over $x \in X$, we obtain the weak duality inequality $\bar{d}_c \preccurlyeq \bar{p}$.

If $\bar{p} = \{-\infty\}$, strong duality follows from weak duality. Note further that $\operatorname{dom} p$ is nonempty, hence $\bar{p} \neq \{+\infty\}$. Therefore, it remains to prove strong duality for the case $\bar{p} \in \mathcal{I} \setminus \{\{-\infty\}, \{+\infty\}\}$.

We use the scalarization functional $\varphi_A : C^\circ \setminus \{0\} \to \overline{\mathbb{R}}$ $(A \in \mathcal{I})$ as introduced in Section 1.7. As $p : X \to \mathcal{I}$ is convex, Proposition 1.64 implies that $\bar{p} \in \mathcal{I}_{\mathrm{co}}$.

By Corollary 1.68, (the concave function) $\varphi_{\bar{p}}$ is proper, in particular $\operatorname{dom} \varphi_{\bar{p}} \neq \emptyset$. As $c \in \operatorname{int} C$, for every $y^* \in C^\circ \setminus \{0\}$ we have $y^*(c) < 0$. We fix some $y^* \in \operatorname{dom} \varphi_{\bar{p}} \cap B_c$ and consider the extended real-valued functions $\xi : X \to \overline{\mathbb{R}}$ and $\eta : U \to \overline{\mathbb{R}}$ being defined, respectively, by $\xi(x) := \varphi\big(y^* |\, f(x)\big)$ and $\eta(u) := \varphi\big(y^* |\, g(u)\big)$. It follows

$$\varphi(y^* |\, \bar{p}) = \inf_{x \in X} \big(\xi(x) + \eta(Bx)\big). \tag{3.4}$$

By Corollary 1.69, ξ and η are convex, $\operatorname{dom} f = \operatorname{dom} \xi$ and $\operatorname{dom} g = \operatorname{dom} \eta$. As $y^* \in \operatorname{dom} \varphi_{\bar{p}}$ (that is $\varphi_{\bar{p}}(y^*) > -\infty$ as $\varphi_{\bar{p}} : C^\circ \setminus \{0\} \to \overline{\mathbb{R}}$ is concave), ξ and η are proper. By Theorem 1.75, the constraint qualification (3.3) implies a corresponding condition for the scalar problem;

$$\text{There exists } u \in \operatorname{dom} \eta \cap B(\operatorname{dom} \xi) \text{ such that } \eta \text{ is continuous at } u.$$

A scalar duality result (see e.g. Göpfert, 1973, Chapter 3, Theorems 6 and 7) yields

$$\varphi(y^*|\,\bar{p}) = \sup_{u^* \in U^*} \left(-\xi^*(B^* u^*) - \eta^*(-u^*)\right).$$

Moreover, if $\varphi(y^*|\,\bar{p})$ is finite, the supremum is attained, that is

$$\exists \bar{u}^* \in U^* : \qquad \varphi(y^*|\,\bar{p}) = -\xi^*(B^*\bar{u}^*) - \eta^*(-\bar{u}^*). \tag{3.5}$$

Furthermore, it is true that

$$\forall t \in \mathbb{R}: \;\; \varphi\big(y^*|t \cdot \{c\}\big) = -y^*(t \cdot c) = t. \tag{3.6}$$

We have

$$\begin{aligned}
\varphi\,(y^*|\,\bar{p}) &= -\xi^*(B^*\bar{u}^*) - \eta^*(-\bar{u}^*) \\
&= \inf_{x \in X}\big(-\langle B^*\bar{u}^*, x\rangle + \xi(x)\big) + \inf_{u \in U}\big(\langle \bar{u}^*, u\rangle + \eta(u)\big) \\
&\overset{(3.6)}{=} \inf_{x \in X}\Big(\varphi\big(y^*\big|-\langle B^*\bar{u}^*, x\rangle \cdot \{c\}\big) + \varphi\big(y^*\big|f(x)\big)\Big) \\
&\quad + \inf_{u \in U}\Big(\varphi\big(y^*\big|\langle \bar{u}^*, u\rangle \cdot \{c\}\big) + \varphi\big(y^*\big|g(u)\big)\Big) \\
&= \varphi\left(y^*\,\middle|\, \inf_{x \in X}\big(-\langle B^*\bar{u}^*, x\rangle\{c\} + f(x)\big)\right. \\
&\quad \left. \oplus \; \inf_{u \in U}\big(\langle \bar{u}^*, u\rangle\{c\} + g(u)\big)\right) \\
&= \varphi\left(y^*\,\middle|- \sup_{x \in X}\big(\langle B^*\bar{u}^*, x\rangle\{c\} - f(x)\big)\right. \\
&\quad \left. \oplus - \sup_{u \in U}\big(\langle -\bar{u}^*, u\rangle\{c\} - g(u)\big)\right) \\
&= \varphi\big(y^*\big| - f_c^*(B^*\bar{u}^*) \oplus -g_c^*(-\bar{u}^*)\big) = \varphi\,(y^*|d_c(\bar{u}^*)).
\end{aligned}$$

We deduce that

$$\forall y^* \in \operatorname{dom}\varphi_{\bar{p}} \cap B_c,\; \exists \bar{u}^* \in U^* : \quad \varphi(y^*|d_c(\bar{u}^*)) = \varphi(y^*|\,\bar{p}\,). \tag{3.7}$$

For every $A \in \mathcal{I}$ and $\alpha > 0$, we have $\varphi(\alpha \cdot y^*|\,A\,) = -\alpha\varphi(y^*|\,A\,)$. We conclude from (3.7) that $\varphi(y^*|\,\bar{d}_c\,) \geq \varphi(y^*|\,\bar{p}\,)$ for all $y^* \in C^\circ \setminus \{0\}$. As $\bar{p} \in \mathcal{I}_{\rm co}$, Theorem 1.65 (iv) yields $\bar{d}_c \succcurlyeq \bar{p}$. By the weak duality inequality we obtain $\bar{d}_c = \bar{p}$. $\square$

We know from the scalar optimization theory that strong duality results usually consist of two statements. The first one is the equality of the primal and dual optimal values and the second one is the dual attainment, that is, if the primal optimal value is finite, then a solution to the dual problem exists (see e.g. Borwein and Lewis, 2000, Theorem 3.3.5). In the current framework we cannot answer the question whether the dual attainment holds or not.

Open Problem 3.6 *Let the assumptions of Theorem 3.5 be satisfied. Does a solution to the dual problem* $(\mathrm{D}^{\mathrm{I}}_{\mathrm{F}})$ *exists, whenever* $\bar{p} \neq \{-\infty\}$*?*

However, a surrogate result can be shown. Usually, the dual optimal value can be expressed as

$$\bar{d}_c = \sup_{u^* \in U^*} d_c(u^*) = \operatorname{Sup} \bigcup_{u^* \in U^*} d_c(u^*).$$

It is shown in the following result that the supremal set can be replaced by the set of weakly maximal elements. Of course this result is rather unsatisfactory, because it is not compatible with the solution concepts introduced in Chapter 2. However, the dual attainment can be obtained for type II dual problems as considered in the next section.

Theorem 3.7. *Let all the assumptions of Theorem 3.5 be satisfied and let* $\bar{p} \neq \{-\infty\}$*, then*

$$\bar{d}_c = \operatorname{wMax} \bigcup_{u^* \in U^*} d_c(u^*).$$

Proof. We have $\bar{p} \notin \{\{-\infty\}, \{+\infty\}\}$ and hence $\emptyset \neq \operatorname{Cl}_+ \bar{p} \neq Y$. Let

$$\bar{y} \in \operatorname{Sup} \bigcup_{u^* \in U^*} d_c(u^*) = \bar{d}_c = \bar{p}.$$

By Proposition 1.64 we have $\bar{p} \in \mathcal{I}_{\mathrm{co}}$. We get $\bar{y} \notin \bar{p} + \operatorname{int} C$ and the set $\bar{p} + \operatorname{int} C$ is convex. By the separation theorem 1.34, there exists some $\bar{y}^* \in Y^* \setminus \{0\}$ such that

$$\bar{y}^*(\bar{y}) \geq \sup_{y \in \bar{p} + \operatorname{int} C} \bar{y}^*(y).$$

Assuming that $y^* \notin C^\circ$, we get a contradiction as the supremum becomes $+\infty$. Thus we have $y^* \in C^\circ \setminus \{0\}$ and hence

$$\bar{y}^*(\bar{y}) \geq \sigma_{\bar{p} + \operatorname{int} C}(\bar{y}^*) = \sigma_{\operatorname{Cl}_+ \bar{p}}(\bar{y}^*) = -\varphi_{\bar{p}}(\bar{y}^*).$$

Without loss of generality we can assume that $\bar{y}^*(c) = -1$. By (3.7) there exists some $\bar{u}^* \in U^*$ such that

$$\bar{y}^*(\bar{y}) \geq -\varphi_{\bar{p}}(\bar{y}^*) = -\varphi_{d_c(\bar{u}^*)}(\bar{y}^*).$$

Assuming that $\bar{y} \in d_c(\bar{u}^*) + \operatorname{int} C = d_c(\bar{u}^*) + \operatorname{int} C + \operatorname{int} C$, we obtain

$$\forall y^* \in C^\circ \setminus \{0\}: \quad y^*(\bar{y}) < \sigma_{d_c(\bar{u}^*) + \operatorname{int} C}(y^*) = -\varphi_{d_c(\bar{u}^*)}(y^*),$$

a contradiction. It follows $\bar{y} \notin d_c(\bar{u}^*) + \operatorname{int} C$.

On the other hand, $\{\bar{y}\} + \operatorname{int} C \subseteq \bar{p} + \operatorname{int} C \subseteq d_c(\bar{u}^*) + \operatorname{int} C$. We get $\bar{y} \in \operatorname{Inf} d_c(\bar{u}^*) = d_c(\bar{u}^*) \subseteq \bigcup_{u^* \in U^*} d_c(u^*)$. Together we have $\bar{y} \in \operatorname{wMax} \bigcup_{u^* \in U^*} d_c(u^*)$. □

3.2.2 Duality result of type II and dual attainment

A modified dual problem is now introduced. The advantage is that the dual objective function has an easier structure, in fact, it is essentially hyperplane-valued. Another benefit is that the strong duality result now includes the dual attainment. The larger variable space (pre-image space of the objective function) is the downside of the type II duality.

It is shown in this section that both types of dual problems are closely related. Duality results of the one type can be easily obtained from the results of the other type[1].

Let us start with some auxiliary results. Consider the function

$$h : Y^* \to \mathcal{I}_{\mathrm{co}}\,, \qquad h(y^*) := \operatorname{Inf}\{y \in Y |\; y^*(y) \leq 0\}\,,$$

the values of which are either hyperplanes or $\{-\infty\}$.

Proposition 3.8. *The following statement is true:*

$$h(y^*) = \begin{cases} \{y \in Y |\; y^*(y) = 0\} & \text{if} \quad y^* \in C^\circ \setminus \{0\} \\ \{-\infty\} & \text{otherwise.} \end{cases}$$

Proof. First, let $y^* \in C^\circ \setminus \{0\}$. Then we have

$$\operatorname{Cl}_+ \{y \in Y |\; y^*(y) \leq 0\} \;=\; \{y \in Y |\; y^*(y) \leq 0\}$$

and

$$\{y \in Y |\; y^*(y) \leq 0\} + \operatorname{int} C \;=\; \{y \in Y |\; y^*(y) < 0\}\,.$$

According to Corollary 1.48 (v), we have $y \in h(y^*)$ if and only if

$$y \in \operatorname{Cl}_+ \{y \in Y |\; y^*(y) \leq 0\} \quad \wedge \quad y \notin \operatorname{Cl}_+ \{y \in Y |\; y^*(y) \leq 0\} + \operatorname{int} C.$$

This is equivalent to $y \in \{y \in Y |\; y^*(y) = 0\}$.

Secondly, let $y^* \notin C^\circ \setminus \{0\}$. There exists some $y \in -\operatorname{int} C$ such that $y^*(y) \leq 0$. Indeed, if $y^* \neq 0$, we get some $z \in -C$ such that $y^*(z) < 0$. The convex combination $\lambda w + (1-\lambda) z$ for some $w \in -\operatorname{int} C$ provides the desired vector y if $\lambda \in (0,1)$ is chosen sufficiently small. The case $y^* = 0$ is obvious.

We conclude that

[1] Roughly speaking, the results are equivalent.

$$Y \subseteq \{\mu y \mid \mu > 0\} + \operatorname{int} C \subseteq \{y \in Y \mid \; y^*(y) \leq 0\} + \operatorname{int} C \subseteq Y.$$

This implies $\operatorname{Inf}\{y \in Y \mid y^*(y) \leq 0\} = \{-\infty\}$. □

Also the sum of some $A \in \mathcal{I}$ and $h(y^*)$ gives either a hyperplane or $\{-\infty\}$.

Proposition 3.9. *For every $A \in \mathcal{I} \setminus \{\{+\infty\}\}$ we have*

$$A \oplus h(y^*) = \begin{cases} \{y \in Y \mid y^*(y) + \varphi_A(y^*) = 0\} & \textit{if} \quad y^* \in \operatorname{dom} \varphi_A \\ \{-\infty\} & \textit{otherwise.} \end{cases}$$

Proof. Let $y^* \in \operatorname{dom} \varphi_A \subseteq C^\circ \setminus \{0\}$. Theorem 1.65 (i) yields $A \neq \{-\infty\}$. For $z^* \in C^\circ \setminus \{0\}$, we have

$$\varphi_{h(y^*)}(z^*) = \begin{cases} 0 & \text{if } z^* = \alpha y^* \text{ for some } \alpha > 0 \\ -\infty & \text{otherwise.} \end{cases}$$

The first case is obvious. In the second case we have $z^* \notin \operatorname{cone}\{y^*\}$. The set $\operatorname{cone}\{y^*\} := \{\lambda y^* \mid \lambda \geq 0\}$ is a convex w^*-closed set, since (Y^*, w^*) is Hausdorff. The separation theorem 1.37 yields some nonzero w^*-continuous linear functional ϕ such that

$$\phi(z^*) > \sup_{\lambda \geq 0} \lambda \phi(y^*) = 0.$$

There is some $y \in Y$ such that $y^*(y) = \phi(y^*)$ for all $y^* \in Y^*$ (e.g. Kelley *et al.*, 1963, Theorem 17.6). We have $y^*(y) = \phi(y^*) \leq 0$ and thus $y \in \operatorname{Cl}_+ h(y^*)$. From $z^*(y) > 0$, we obtain $\varphi_{h(y^*)}(z^*) = -\infty$.

As $y^* \in \operatorname{dom} \varphi_A$, we have $y^* \in \operatorname{dom} \varphi_{A \oplus h(y^*)}$. We deduce that $A \oplus h(y^*) \notin \{\{-\infty\}, \{+\infty\}\}$ and equivalently $\emptyset \neq \operatorname{Cl}_+(A + h(y^*)) \neq Y$. By Corollary 1.48 (v) we have

$$\bar{y} \in A \oplus h(y^*) \iff \left[\bar{y} \in \operatorname{Cl}_+(A + h(y^*)) \;\wedge\; \bar{y} \notin \operatorname{Cl}_+(A + h(y^*)) + \operatorname{int} C\right]. \tag{3.8}$$

Using Proposition 3.8, we get

$$\operatorname{Cl}_+(A + h(y^*)) = \operatorname{cl}\big(A + h(y^*) + C\big) = \operatorname{cl}\big(A + \{y \mid y^*(y) \leq 0\}\big),$$

which shows that $A + h(y^*)$ is convex. This implies $A \oplus h(y^*) \in \mathcal{I}_{\text{co}}$. Using Theorem 1.65 (iii), (iv) we obtain

$$\bar{y} \in \operatorname{Cl}_+(A + h(y^*)) \iff A \oplus h(y^*) \preccurlyeq \operatorname{Inf}\{\bar{y}\} \iff \varphi_{A \oplus h(y^*)} \leq \varphi_{\operatorname{Inf}\{\bar{y}\}}.$$

Taking into account that

$$\varphi_{A \oplus h(y^*)}(z^*) = \begin{cases} \varphi_A(z^*) & \text{if } z^* = \alpha y^* \text{ for some } \alpha > 0 \\ -\infty & \text{otherwise,} \end{cases}$$

we get the equivalence

$$\varphi_{A\oplus h(y^*)} \le \varphi_{\mathrm{Inf}\{\bar y\}} \quad\iff\quad y^*(\bar y)+\varphi_A(y^*)\le 0.$$

Summarizing these results, we obtain

$$\bar y \in \mathrm{Cl}_+(A+h(y^*)) \quad\iff\quad y^*(\bar y)+\varphi_A(y^*)\le 0. \tag{3.9}$$

As an easy consequence of (3.9) we have

$$\bar y \in \mathrm{Cl}_+(A+h(y^*))+\operatorname{int} C \quad\iff\quad y^*(\bar y)+\varphi_A(y^*)< 0. \tag{3.10}$$

From (3.8), (3.9) and (3.10) we conclude

$$\bar y \in A\oplus h(y^*) \quad\iff\quad y^*(\bar y)+\varphi_A(y^*)= 0.$$

Let $y^* \in Y^*\setminus \operatorname{dom}\varphi_A$. Then we have $\varphi_A(y^*) = -\infty$. It follows that $\varphi_{A\oplus h(y^*)} \equiv -\infty$. Since $A\oplus h(y^*) \in \mathcal{I}_{\mathrm{co}}$, Theorem 1.65 (i) yields $A\oplus h(y^*) = \{-\infty\}$. □

An element $A \in \mathcal{I}_{\mathrm{co}}$ can be re-obtained from the sum $A\oplus h(y^*)$ by taking the supremum over all y^*.

Proposition 3.10. *If $A \in \mathcal{I}_{\mathrm{co}}$, then*

$$\sup_{y^*\in Y^*}(A\oplus h(y^*)) = \sup_{y^*\in\operatorname{dom}\varphi_A}(A\oplus h(y^*)) = A.$$

Proof. If $A=\{+\infty\}$, we have $\operatorname{dom}\varphi_A \neq \emptyset$ and the statement is obvious. If $A=\{-\infty\}$ we have $\operatorname{dom}\varphi_A = \emptyset$. Since $\sup\emptyset = \{-\infty\}$, the second equality follows directly and the first one follows from Proposition 3.9.

We now assume that $A \in \mathcal{I}_{\mathrm{co}}\setminus\{\{-\infty\},\{+\infty\}\}$; by Corollary 1.68 this means that φ_A is proper. By Proposition 3.9 the first equality is obvious. For $y^*\in\operatorname{dom}\varphi_A \subseteq C^\circ\setminus\{0\}$ we have $h(y^*)=\{y\in Y \mid y^*(y)=0\}$ and thus $0\in h(y^*)$. It follows $h(y^*) \preccurlyeq \operatorname{Inf}\{0\}$. We get $A\oplus h(y^*)\preccurlyeq A$ and hence $\sup_{y^*\in\operatorname{dom}\varphi_A}\big(A\oplus h(y^*)\big)\preccurlyeq A$.

Suppose now that $A \preccurlyeq \sup_{y^*\in\operatorname{dom}\varphi_A}\big(A\oplus h(y^*)\big)$ is not true. As $A\in\mathcal{I}_{\mathrm{co}}$, Theorem 1.65 (iv) implies that

$$\varphi(\,\cdot\mid A) \not\le \varphi\left(\cdot\,\middle|\, \sup_{y^*\in\operatorname{dom}\varphi_A}(A\oplus h(y^*))\right).$$

Hence there exists some $\bar y^*\in\operatorname{dom}\varphi_A$ such that

$$\begin{aligned}\varphi(\bar y^*\mid A) &> \varphi\left(y^*\,\middle|\, \sup_{y^*\in\operatorname{dom}\varphi_A}(A\oplus h(y^*))\right) \ge \varphi\big(\bar y^*\,\big|\, A\oplus h(\bar y^*)\big)\\ &= \varphi(\bar y^*\mid A)\oplus\varphi(\bar y^*\mid h(\bar y^*)) = \varphi(\bar y^*\mid A),\end{aligned}$$

which is a contradiction. □

The supremum in the preceding result is sufficient to be taken over a base B_c of the polar cone C° as defined in (3.2). This is pointed out in the following result.

Corollary 3.11. *Let* $c \in \operatorname{int} C$ *and* $A \in \mathcal{I}_{\mathrm{co}}$*, then*

$$\sup_{y^* \in B_c \cap \operatorname{dom} \varphi_A} (A \oplus h(y^*)) = A.$$

Proof. We have $\varphi_A(\alpha y^*) = \alpha \varphi_A(y^*)$ and $h(y^*) = h(\alpha y^*)$ for all $\alpha > 0$. Thus, the statement follows from Proposition 3.10. □

The type II dual problem is now introduced. It is derived from the type I dual problem ($\mathrm{D}^{\mathrm{I}}_{\mathrm{F}}$). For $c \in \operatorname{int} C$ we consider the dual objective function

$$d : U^* \times C^\circ \setminus \{0\} \to \mathcal{I}, \qquad d(u^*, y^*) := d_c\left(\frac{-u^*}{y^*(c)}\right) \oplus h(y^*), \tag{3.11}$$

and the dual problem

$$\text{maximize } d : U^* \times C^\circ \setminus \{0\} \to \mathcal{I} \text{ w. r. t. } \preccurlyeq \text{ over } U^* \times C^\circ \setminus \{0\}. \tag{$\mathrm{D}^{\mathrm{II}}_{\mathrm{F}}$}$$

We will show in Theorem 3.14 below that the new dual objective function d is independent of the choice of the parameter $c \in \operatorname{int} C$. Note further that it is sufficient for many reasons to consider d on the set $U^* \times B_c$. Then, the definition has the simpler form

$$d : U^* \times B_c \to \mathcal{I}, \qquad d(u^*, y^*) := d_c(u^*) \oplus h(y^*).$$

The *optimal value* of Problem ($\mathrm{D}^{\mathrm{II}}_{\mathrm{F}}$) is defined as

$$\bar{d} := \sup_{(u^*, y^*) \in U^* \times C^\circ \setminus \{0\}} d(u^*, y^*).$$

Let us show some properties of the dual objective d. In the following proposition we state that a value of d is either a hyperplane or $\{-\infty\}$. From Theorem 3.14 below, we deduce that the expression of the hyperplane does not depend on the choice of $c \in \operatorname{int} C$.

Proposition 3.12. *Suppose that the primal optimal value* $\bar{p}$ *in Problem* (P_{F}) *is not* $\{+\infty\}$*. For the function* d *as defined in* (3.11) *we have*

$$d(u^*, y^*) = \begin{cases} \left\{ y \in Y \,\middle|\, y^*(y) = -\varphi\left(y^* \middle| d_c\left(\frac{-u^*}{y^*(c)}\right)\right)\right\} & \textit{if } \varphi\left(y^* \middle| d_c\left(\frac{-u^*}{y^*(c)}\right)\right) \neq -\infty \\ \{-\infty\} & \textit{otherwise.} \end{cases}$$

Proof. By the weak duality we have $d_c(v^*) \neq \{+\infty\}$ for all $v^* \in U^*$. Proposition 3.9 yields the result. □

Let us combine formula (3.11) with the expression of the dual objective function of Problem ($\mathrm{D}^{\mathrm{I}}_{\mathrm{F}}$) in (3.1). We have $h(y^*) = h(y^*) \oplus h(y^*)$ for all $y^* \in Y^*$. If $y^*(c) \neq 0$, we get

$$d(u^*, y^*) = \left(-f_c^* \left(\frac{-B^* u^*}{y^*(c)} \right) \oplus h(y^*) \right) \oplus \left(-g_c^* \left(\frac{u^*}{y^*(c)} \right) \oplus h(y^*) \right).$$

It is natural to define a *conjugate* of type II by

$$-f^* : X^* \times C^\circ \setminus \{0\} \to \mathcal{I}, \quad -f^*(x^*, y^*) := -f_c^* \left(\frac{-x^*}{y^*(c)} \right) \oplus h(y^*). \quad (3.12)$$

Again we can work with a base B_c of the polar cone C°. This yields the alternative and simpler definition

$$-f^* : X^* \times B_c \to \mathcal{I}, \quad -f^*(x^*, y^*) := -f_c^* \left(x^* \right) \oplus h(y^*).$$

As a consequence of (3.12), the dual objective $d : U^* \times Y^* \to \mathcal{I}$ can be expressed as

$$d(u^*, y^*) = -f^*(B^* u^*, y^*) \oplus -g^*(-u^*, y^*).$$

We next show that the conjugate defined by (3.12) coincides with a conjugate introduced by Hamel (2009a) in a more general set-valued framework. Since we work in the space $\mathcal{I}$ instead of the space $\mathcal{F}$ of upper closed sets like in Hamel (2009a), we have to suppose additional assumptions to the ordering cone C. For instance, we assume that $\operatorname{int} C \neq \emptyset$. The advantage of our approach, based on the complete lattice $\mathcal{I}$, is a closer connection to vector optimization. The following notion has been introduced and used by Hamel (2009a). For $x^* \in X^*$, $y^* \in C^\circ$, let

$$S_{(x^*,y^*)}(x) := \{ y \in Y \colon \langle x^*, x \rangle + y^*(y) \leq 0 \}.$$

This map can be understood as a generalization of a linear function which maps from X into the conlinear space $\mathcal{F}$. We adopt this concept to our $\mathcal{I}$-valued framework. To this end the map S is modified in the following way. For $x^* \in X^*$, $y^* \in Y^*$, we consider the expression

$$M_{(x^*,y^*)}(x) := \operatorname{Inf} \{ y \in Y \colon \langle x^*, x \rangle + y^*(y) \leq 0 \}.$$

Proposition 3.13. *Let $c \in Y$ and $y^* \in Y^*$ be given such that $y^*(c) = -1$, then*

$$M_{(x^*,y^*)}(x) = \langle x^*, x \rangle \{c\} + h(y^*).$$

Proof. Clearly, we have

$$\{y \in Y \colon \langle x^*, x\rangle + y^*(y) \leq 0\} = \langle x^*, x\rangle \{c\} + \{y \in Y |\ y^*(y) \leq 0\}.$$

It follows

$$\begin{aligned} M_{(x^*,y^*)}(x) &= \operatorname{Inf}\big(\langle x^*, x\rangle \{c\} + \{y \in Y |\ y^*(y) \leq 0\}\big) \\ &= \langle x^*, x\rangle \{c\} + \operatorname{Inf} \{y \in Y |\ y^*(y) \leq 0\}, \end{aligned}$$

which completes the proof. □

The independence of the type II conjugate from the parameter c is shown next.

Theorem 3.14. *Let $c \in \operatorname{int} C$. The type II conjugate defined by* (3.12) *can be expressed as*

$$-f^* : X^* \times C^\circ \setminus \{0\} \to \mathcal{I}, \quad -f^*(x^*, y^*) = \inf_{x \in X} \big(f(x) \oplus M_{(x^*,y^*)}(-x)\big).$$

In particular, $-f^*$ does not depend on the choice of $c \in \operatorname{int} C$.

Proof. As $c \in \operatorname{int} C$ and $y^* \in C^\circ \setminus \{0\}$, we have $y^*(c) < 0$. Using the preceding result, we conclude

$$\begin{aligned} -f^*(x^*, y^*) &\overset{(3.12)}{=} -f_c^*\left(\tfrac{-x^*}{y^*(c)}\right) \oplus h(y^*) \\ &= \inf_{x \in X} \left(f(x) - \langle x^*, x\rangle \left\{\frac{-c}{y^*(c)}\right\}\right) \oplus h(y^*) \\ &= \inf_{x \in X} \left(f(x) \oplus \left(\langle x^*, -x\rangle \left\{\frac{-c}{y^*(c)}\right\} + h(y^*)\right)\right) \\ &\overset{\text{Pr. 3.13}}{=} \inf_{x \in X} \big(f(x) \oplus M_{(x^*,y^*)}(-x)\big), \end{aligned}$$

which completes the proof. □

Theorem 3.14 shows that our conjugate is closely related to the conjugate introduced by Hamel (2009a). The connection becomes clear by the fact that, under our assumptions, $(\mathcal{I}, \preccurlyeq)$ (our choice) is isomorphic and isotone to $(\mathcal{F}, \supseteq)$ (Hamel's choice).

Theorem 3.15. *Let $f : X \to \mathcal{I}$ and $g : U \to \mathcal{I}$ be convex and let $c \in \operatorname{int} C$. Then both dual problems* $(\mathrm{D}^{\mathrm{I}}_{\mathrm{F}})$ *and* $(\mathrm{D}^{\mathrm{II}}_{\mathrm{F}})$ *have the same optimal values, that is, $\bar{d}_c = \bar{d}$.*

Proof. We show that

$$d_c(u^*) = \sup_{y^* \in B_c} d(u^*, y^*) = \sup_{y^* \in C^\circ \setminus \{0\}} d(u^*, y^*). \tag{3.13}$$

Indeed, the statement follows from Corollary 3.11 and Proposition 3.10 if we can verify that $d_c(u^*) \in \mathcal{I}_{\rm co}$. The type I dual objective d_c is defined by $d_c(u^*) = -f^*(B^*u^*) \oplus -g^*(-u^*)$. But $-f^*(B^*u^*) = \inf_{x\in X}\big(f(x) - \langle B^*u^*, x\rangle \{c\}\big)$, where $\big(f(\,\cdot\,) - \langle B^*u^*, \cdot\,\rangle \{c\}\big) : X \to \mathcal{I}$ is convex. By Proposition 1.64, we have $-f^*(B^*u^*) \in \mathcal{I}_{\rm co}$. Likewise we get $-g^*(-u^*) \in \mathcal{I}_{\rm co}$. As $\mathcal{I}_{\rm co}$ is a conlinear space, we obtain $d_c(u^*) \in \mathcal{I}_{\rm co}$.

From (3.13) we get

$$\begin{aligned} \bar d_c = \sup_{u^*\in U^*} d_c(u^*) &= \sup_{u^*\in U^*} \sup_{y^*\in C^\circ\setminus\{0\}} d(u^*, y^*) \\ &= \sup_{(u^*, y^*)\in U^*\times C^\circ\setminus\{0\}} d(u^*, y^*) = \bar d, \end{aligned}$$

which completes the proof. □

From the preceding theorem we immediately obtain a duality theorem which is closely related to (Hamel, 2009b, Theorems 1 and 2). Vice versa, we see that Hamel's result in the form of the following theorem implies the duality theorem 3.5.

Theorem 3.16. *The problems* $(\mathrm{P_F})$ *and* $(\mathrm{D_F^{II}})$ *(with arbitrary $c \in Y$) satisfy the weak duality inequality $\bar d \preccurlyeq \bar p$. Furthermore, let f and g be proper convex functions, $c \in \operatorname{int} C$, and let the constraint qualification* (3.3) *be satisfied. Then strong duality between* $(\mathrm{P_F})$ *and* $(\mathrm{D_F^{II}})$ *holds; that is, $\bar d = \bar p$.*

Proof. Follows from Theorems 3.5 and 3.15.

The advantage of the type II duality theorem is that the dual problem can be shown to have a solution; that is, we have dual attainment in $(\mathrm{D_F^{II}})$.

Theorem 3.17. *Let all the assumptions of Theorem 3.16 be satisfied and let $\bar p \neq \{-\infty\}$. Then, the dual problem* $(\mathrm{D_F^{II}})$ *has a solution.*

Proof. Let $c \in \operatorname{int} C$ and consider the set-valued map $Q : B_c \rightrightarrows U^*$, defined by

$$Q(y^*) := \{u^* \in U^* |\; \varphi(y^*|\bar p) = \varphi(y^*|d_c(u^*)) \in \mathbb{R}\}\,.$$

By (3.7) we have

$$\forall y^* \in \operatorname{dom} \varphi_{\bar p} \cap B_c : \quad Q(y^*) \neq \emptyset.$$

Note further that $\operatorname{dom} p \neq \emptyset$ and thus $\bar p \neq \{+\infty\}$. Moreover, we have $\bar p \neq \{-\infty\}$ as well as $\bar p \in \mathcal{I}_{\rm co}$. By Corollary 1.68, $\varphi_{\bar p}$ is proper; hence $\operatorname{dom} \varphi_{\bar p} \neq \emptyset$. We show that the nonempty set

$$\bar V := \operatorname{gr} Q^{-1} = \big\{(u^*, y^*) \in U^* \times B_c \big|\; u^* \in Q(y^*)\big\}$$

is a solution to the dual problem $(\mathrm{D_F^{II}})$, where we have to verify the conditions (i) to (iii) of Corollary 2.54.

(i) Of course, $\bar{V}$ is a subset of the feasible set $T := U^* \times C^\circ \setminus \{0\}$.

Before we show (ii), we consider some auxiliary assertions. Let some $(\bar{u}^*, \bar{y}^*) \in T$ with $d(\bar{u}^*, \bar{y}^*) \neq \{-\infty\}$ be given and let $(u^*, y^*) \in T$ such that $d(\bar{u}^*, \bar{y}^*) \preccurlyeq d(u^*, y^*)$. By the weak duality, we have $d(u^*, y^*) \neq \{+\infty\}$. Proposition 3.9 yields

$$\bar{y}^* \in \operatorname{dom} \varphi_{d_c(\bar{u}^*)} \quad \wedge \quad d(\bar{u}^*, \bar{y}^*) = \{y \in Y |\ \bar{y}^*(y) + \varphi(\bar{y}^* |\ d_c(\bar{u}^*)) = 0\}$$

and

$$y^* \in \operatorname{dom} \varphi_{d_c(u^*)} \quad \wedge \quad d(u^*, y^*) = \{y \in Y |\ y^*(y) + \varphi(y^* |\ d_c(u^*)) = 0\}\,.$$

It follows that $\{-\infty\} \neq d(\bar{u}^*, \bar{y}^*) \preccurlyeq d(u^*, y^*)$ is equivalent to (note that $d(u^*, y^*) \neq \{+\infty\}$)

$$\bar{y}^* = y^* \in \operatorname{dom} \varphi_{d_c(\bar{u}^*)} \qquad \wedge \qquad \varphi(\bar{y}^* |\ d_c(\bar{u}^*)) \leq \varphi(y^* |\ d_c(u^*)).$$

Therefore, $\{-\infty\} \neq d(\bar{u}^*, \bar{y}^*) \in \operatorname{Max} d[T]$ is equivalent to

$$\bar{y}^* \in \operatorname{dom} \varphi_{d_c(\bar{u}^*)} \quad \wedge \quad \forall u^* \in U^* :\ \varphi(\bar{y}^* | d_c(u^*)) \leq \varphi(\bar{y}^* | d_c(\bar{u}^*)). \tag{3.14}$$

(ii) Let $A \in d[\bar{V}]$. There exits some $(\bar{u}^*, \bar{y}^*) \in U^* \times \big(\operatorname{dom} \varphi_{\bar{p}} \cap B_c\big)$ such that $\varphi(\bar{y}^* | \bar{p}) = \varphi(\bar{y}^* | d_c(\bar{u}^*))$ and $A = d(\bar{u}^*, \bar{y}^*) = d_c(\bar{u}^*) \oplus h(\bar{y}^*)$. As $\varphi(\bar{y}^* |\ d_c(\bar{u}^*)) \neq -\infty$, we obtain $d_c(\bar{u}^*) \neq \{-\infty\}$ and hence $d(\bar{u}^*, \bar{y}^*) \neq \{-\infty\}$. Moreover, weak duality implies

$$\forall u^* \in U^* : \varphi(\bar{y}^* | d_c(u^*)) \leq \varphi(\bar{y}^* | \bar{p}) = \varphi(\bar{y}^* | d_c(\bar{u}^*)).$$

Hence $A = d(\bar{u}^*, \bar{y}^*) \in \operatorname{Max} d[T]$.

Conversely, let $A \in \operatorname{Max} d[T]$. If A $= \{-\infty\}$, we get $d(u^*, y^*) = \{-\infty\}$ for all $(u^*, y^*) \in T$ and by the strong duality this contradicts the assumption $\bar{p} \neq \{-\infty\}$. Otherwise there exists some $(\bar{u}^*, \bar{y}^*) \in T$ with $A = d(\bar{u}^*, \bar{y}^*)$ such that (3.14) is satisfied. By the strong duality it follows that $\varphi(\bar{y}^* | d_c(\bar{u}^*)) = \varphi(\bar{y}^* | \bar{p})$. Without loss of generality we can assume $\bar{y}^*(c) = -1$. Consequently, we have $A \in d[\bar{V}]$.

(iii) It remains to show

$$\sup_{(u^*,\, y^*) \in \bar{V}} d(u^*, y^*) = \sup_{(u^*,\, y^*) \in T} d(u^*, y^*).$$

Let $(u^*, y^*) \in T$. If $y^* \notin \operatorname{dom} \varphi_{d_c(u^*)}$, we have $d(u^*, y^*) = \{-\infty\}$. Otherwise, we get $y^* \in \operatorname{dom} \varphi_{\bar{p}}$ and without loss of generality we can assume that $y^*(c) = -1$. By (3.7) there exists some $\bar{u}^* \in U^*$ such that

$$\varphi_{d_c(u^*)}(y^*) \leq \varphi_{d_c(\bar{u}^*)}(y^*) = \varphi_{\bar{p}}(y^*).$$

By Proposition 3.9 we get $d(u^*, y^*) \preccurlyeq d(\bar{u}^*, y^*)$, where $(\bar{u}^*, y^*) \in \bar{V}$. This implies

$$\sup_{(u^*, y^*)\in \bar{V}} d(u^*, y^*) \succcurlyeq \sup_{(u^*, y^*)\in T} d(u^*, y^*).$$

The opposite inequality is obvious. □

3.2.3 The finite dimensional and the polyhedral case

As known from the scalar theory, the constraint qualification can be relaxed in the finite dimensional case. If we suppose finite dimensional pre-image spaces $X = \mathbb{R}^n$ and $U = \mathbb{R}^m$ we get the following duality theorem. We denote by $\operatorname{ri} A$ the relative interior of a convex set A.

Theorem 3.18. *The problems* $(\mathrm{P_F})$ *and* $(\mathrm{D_F^I})$ *(with arbitrary $c \in Y$) satisfy the weak duality inequality $\bar{d}_c \preccurlyeq \bar{p}$. Furthermore, let f and g be proper convex functions, $c \in \operatorname{int} C$, and let*

$$0 \in \operatorname{ri}(\operatorname{dom} g - B(\operatorname{dom} f)). \tag{3.15}$$

Then strong duality holds, that is, $\bar{d}_c = \bar{p}$.

Proof. The proof is the same as the proof of Theorem 3.5 (in particular ξ and η are defined there) but using a finite dimensional scalar result, for instance (Borwein and Lewis, 2000, Theorem 3.3.5). The constraint qualification $0 \in \operatorname{ri}(\operatorname{dom} \eta - B(\operatorname{dom} \xi))$ of the scalar duality theorem is easily obtained as we have $\operatorname{dom} f = \operatorname{dom} \xi$ and $\operatorname{dom} g = \operatorname{dom} \eta$ □

In the scalar duality theory, the constraint qualification (3.15) can be further weakened if the objective function is polyhedral. In this case it suffices to assume

$$\operatorname{dom} g \cap B(\operatorname{dom} f) \neq \emptyset \tag{3.16}$$

(compare e.g. Borwein and Lewis, 2000, Corollary 5.1.9). This is important for linear problems. Let us consider the special case $X = \mathbb{R}^n$, $U = \mathbb{R}^m$, $Y = \mathbb{R}^q$, $P \in \mathbb{R}^{q\times n}$,

$$f : \mathbb{R}^n \to \mathcal{I}, \quad f(x) = \operatorname{Inf}\{Px\}$$

and

$$g : \mathbb{R}^m \to \mathcal{I}, \quad g(u) := \begin{cases} \operatorname{Inf}\{0\} & \text{if } u \geq b \\ \{+\infty\} & \text{otherwise.} \end{cases}$$

The constraint qualification (3.15) in Theorem 3.18 can be weakened to (3.16), which can be expressed as

$$\exists x \in \mathbb{R}^n : \; Bx \geq 0.$$

This follows by similar considerations as in the proof of Theorem 3.18, but using an adapted scalar result.

3.3 Lagrange duality

In this section, $\mathcal{I}$-valued optimization problems with set-valued constraints are studied. As shown in Section 2.2, vector optimization problems can be regarded as a subclass of $\mathcal{I}$-valued problems. Like in the previous section, duality results are derived from corresponding scalar results. Of course, other variants of scalar Lagrange duality could serve as a template. This section can be understood as a further demonstration, how vectorial duality results can be derived from corresponding scalar results.

Throughout this section, let $\overline{Y}$ be an extended partially ordered locally convex space with an ordering cone $C \subseteq Y$ such that $\emptyset \neq \operatorname{int} C \neq Y$. The topological dual space of Y is denoted by Y^*. We set $\mathcal{I} := \mathcal{I}_C\left(\overline{Y}\right)$. Let X be a linear space, U a Hausdorff locally convex space with topological dual space U^*. We denote by $\langle U, U^* \rangle$ the corresponding dual pair, see Definition 1.38.

Let $f : X \to \mathcal{I}$, let $g : X \rightrightarrows U$ be a set-valued map and let $D \subseteq U$ be a nonempty closed convex cone. The primal problem is given as

$$\text{minimize } f : X \to \mathcal{I} \text{ w.r.t. } \preccurlyeq \text{ over } S := \{x \in X |\; g(x) \cap -D \neq \emptyset\}. \qquad (\mathrm{P_L})$$

The *optimal value* of $(\mathrm{P_L})$ is

$$\bar{p} := \inf_{x \in S} f(x).$$

The set-valued map g is said to be *D-convex* (see e.g. Jahn, 2003) if

$$\forall x_1, x_2 \in X,\; \forall \lambda \in [0,1]:\; g\big(\lambda x_1 + (1-\lambda)x_2\big) + D \supseteq \lambda g(x_1) + (1-\lambda) g(x_2).$$

Of course, the set-valued map g can be understood as a function from X into 2^U. The power set 2^U equipped with the usual Minkowski operations provides a conlinear space. The conlinear space is quasi-ordered (i.e., the ordering is reflexive and transitive) by

$$A \leq B \quad :\Longleftrightarrow \quad A + D \supseteq B + D.$$

Therefore, the notion of D-convexity can be interpreted as convexity of a function with values in this quasi-ordered conlinear space. For the origin of the mentioned quasi-ordering the reader is referred to Hamel (2005).

3.3.1 The scalar case

Initially, a scalar Lagrange duality result with set-valued constraints is provided. The scalar result is used to prove strong duality for $\mathcal{I}$-valued problems. Let us consider the scalar case of Problem ($\mathrm{P_L}$), i.e., let the objective function be $f : X \to \overline{\mathbb{R}}$, where $\overline{\mathbb{R}}$ is equipped with the inf-addition (see Section 1.3). Commonly, a scalar optimization problem is shortly denoted by

$$\hat{p} := \inf_{x \in S} f(x). \tag{3.17}$$

The *Lagrangian* is defined (compare Oettli, 1982) by

$$L : X \times U^* \to \overline{\mathbb{R}}, \qquad L(x, u^*) = f(x) + \inf_{u \in g(x)+D} \langle u^*, u \rangle . \tag{3.18}$$

The dual objective is defined as

$$\phi : U^* \to \overline{\mathbb{R}}, \quad \phi(u^*) := \inf_{x \in X} L(x, u^*)$$

and the dual problem is

$$\hat{d} := \sup_{u^* \in U^*} \phi(u^*). \tag{3.19}$$

Of course, weak duality holds, i.e., $\hat{d} \leq \hat{p}$. Under convexity assumptions and some constraint qualification, we get the following strong duality assertion, which we prove in a common way (e.g. Borwein and Lewis, 2000, Proposition 4.3.5).

Theorem 3.19. *Let $f : X \to \overline{\mathbb{R}}$ be convex, let $g : X \rightrightarrows U$ be D-convex and let*

$$g(\operatorname{dom} f) \cap -\operatorname{int} D \neq \emptyset. \tag{3.20}$$

Then, we have strong duality between (3.17) *and* (3.19)*; that is, $\hat{d} = \hat{p}$. If $\hat{p}$ is finite, then there exists a solution to the dual problem* (3.19)*.*

Proof. The value function is defined by

$$v : U \to \overline{\mathbb{R}} : \qquad v(u) := \inf \left\{ f(x) \middle| \; x \in X : \; g(x) \cap (\{u\} - D) \neq \emptyset \right\}.$$

As f is convex and g is D-convex, v is convex. Moreover, we have $v(0) = \hat{p}$. If $\hat{p} = -\infty$, we obtain $\hat{d} = \hat{p}$ from the weak duality. Therefore, let $\hat{p} > -\infty$. For the conjugate $v^* : U^* \to \overline{\mathbb{R}}$ of v, we have

$$\begin{aligned}
-v^*(-u^*) &= \inf\{\langle u^*,u\rangle + v(u) \mid u\in U\}\\
&= \inf\{\langle u^*,u\rangle + f(x) \mid u\in U,\ x\in X:\ g(x)\cap(\{u\}-D)\neq\emptyset\}\\
&= \inf\{\langle u^*,u\rangle + f(x) \mid x\in X,\ u\in g(x)+D\}\\
&= \inf_{x\in X} L(x,u^*) = \phi(u^*).
\end{aligned}$$

It follows $v^{**}(0)=\hat{d}$. We next show that v is lower semi-continuous at 0 (even continuous). Indeed, by (3.20) there is some $\bar{x}\in\operatorname{dom} f$ and some $\bar{z}\in -\operatorname{int} D$ such that $\bar{z}\in g(\bar{x})$. There exists some neighborhood U of 0 such that $\{\bar{z}\}-U\subseteq -\operatorname{int} D$. It follows that $f(\bar{x})$ is an upper bound of v on U. This implies that v is continuous at 0 (see e.g. Ekeland and Temam, 1976, Lemma 2.1).

We have $v(0)=(\operatorname{lsc} v)(0)=(\operatorname{cl} v)(0)$ (compare Rockafellar, 1974, Theorem 4). By the biconjugation theorem, see e.g. (Ekeland and Temam, 1976, Proposition 4.1) or (Zălinescu, 2002, Theorem 2.3.4), we have $\operatorname{cl} v = v^{**}$. This yields $\hat{p}=v(0)=v^{**}(0)=\hat{d}$.

If $\hat{p}$ is finite, there exists $\bar{u}^*\in\partial v(0)$ (see e.g. Ekeland and Temam, 1976, Proposition 5.2). It follows $v(0)+v^*(\bar{u}^*)=\langle\bar{u}^*,0\rangle$ and hence $\hat{d}=\phi(-\bar{u}^*)$. Thus, $-\bar{u}^*$ solves the dual problem. □

Typically, in Lagrange duality it is shown that the primal problem is reobtained from the Lagrangian. To this end we need the additional assumption that $g(x)+D$ is closed and convex for every $x\in X$. If g is D-convex, like in the strong duality theorem, we have convexity of the values, but in general not the closedness. As D is assumed to be closed, the sum $g(x)+D$ is closed whenever $g(x)$ is compact. Of course, the important case of single-valued maps g is also covered.

Proposition 3.20. *Let $f: X\to\overline{\mathbb{R}}$ be a proper function and let the set $g(x)+D$ be closed and convex for every $x\in X$. Then*

$$\sup_{u^*\in U^*} L(x,u^*) = \begin{cases} f(x) & \text{if } x\in S\\ +\infty & \text{otherwise.}\end{cases}$$

Proof. Note first that $x\in S$ is equivalent to $0\in g(x)+D$. If $x\in S$ we get

$$\begin{aligned}
\sup_{u^*\in U^*} L(x,u^*) &= \sup_{u^*\in U^*}\left(f(x)+\inf_{u\in g(x)+D}\langle u^*,u\rangle\right)\\
&\le \sup_{u^*\in U^*}\left(f(x)+\langle u^*,0\rangle\right) = f(x).
\end{aligned}$$

On the other hand

$$\sup_{u^*\in U^*} L(x,u^*) \ge L(x,0) = f(x),$$

i.e., we have equality.

Assuming $x \notin S$, we obtain $0 \notin g(x) + D$. As the latter set is closed and convex, there exists by Theorem 1.37 (separation theorem) some $\bar{u}^* \in U^*$ such that $\inf_{u \in g(x)+D} \langle \bar{u}^*, u \rangle > 0$. If we consider multiples $u_n^* := n \cdot \bar{u}^* \in U^*$, the latter expression tends to $+\infty$ for $n \to +\infty$. As f is supposed to be proper, we have $f(x) \neq -\infty$ for every x. Hence we get $L(x, u_n^*) \to +\infty$, which proves the statement. □

Note that the assumptions in the last proposition are only used for the proof in the case $x \notin S$. They cannot be dropped as the following examples show.

Example 3.21. Let $f : \mathbb{R}^2 \to \overline{\mathbb{R}}$ be a proper function such that $0 \in \operatorname{dom} f$, let $U = \mathbb{R}^2$, $D = \mathbb{R}^2_+$ and $g(x) = \{x\} + A$, where

$$A := \left\{ a \in \mathbb{R}^2 |\ a_1 > 0 \wedge a_1 a_2 \le -1 \right\}.$$

Note that the sets $g(x)$ and D are closed for all x, but the sum $g(x) + D$ is not. We have $\sup_{u^* \in \mathbb{R}^2} L(0, u^*) = f(0) \in \mathbb{R}$, but $g(0) \cap -D = \emptyset$, i.e., $0 \notin S$.

Example 3.22. Let $f : \mathbb{R} \to \overline{\mathbb{R}}$ be a function such that $f(1) = -\infty$, let $U = \mathbb{R}$, $D = \mathbb{R}_+$ and $g(x) = \{x\}$. We have $\sup_{u^* \in \mathbb{R}} L(1, u^*) = -\infty$, but $g(1) \cap -D = \emptyset$, i.e., $1 \notin S$.

3.3.2 Lagrange duality of type I

An $\mathcal{I}$-valued version of Theorem 3.19 is now considered. The Lagrangian of Problem ($\mathrm{P_L}$) (with respect to $c \in Y$) is defined by

$$L_c : X \times U^* \to \mathcal{I}, \quad L_c(x, u^*) = f(x) \oplus \inf_{u \in g(x)+D} \big(\langle u^*, u \rangle \{c\} + \operatorname{bd} C \big). \quad (3.21)$$

Recall that $\operatorname{bd} C = \operatorname{Inf}\{0\}$ plays the role of the zero element in the space $\mathcal{I}$. It is used in (3.21) to transform the vector $\langle u^*, u \rangle \{c\}$ into an element of $\mathcal{I}$. This ensures the infimum being well-defined. Note that the vector c and the zero element $\operatorname{bd} C$ are the only structural differences to the Lagrangian (3.18) in the scalar case. In the special case $Y = \mathbb{R}$, $C = \mathbb{R}_+$, $c = 1$, the Lagrangian coincides with the Lagrangian (3.18) of the scalar problem (3.17).

The vector $c \in Y$ can be arbitrarily chosen for the moment. For the most assertions, however, we have to assume $c \in \operatorname{int} C$. For every choice of $c \in \operatorname{int} C$ we may have a different Lagrangian and a different corresponding dual problem, but the same duality results hold for all these problems.

The scalar counterpart of the following result is well known, compare Proposition 3.20.

Proposition 3.23. *For every $x \in S$,*

$$\sup_{u^* \in U^*} L_c(x, u^*) = f(x).$$

Proof. Note that $x \in S$ is equivalent to $0 \in g(x) + D$. It follows

$$\begin{aligned}\sup_{u^* \in U^*} L_c(x, u^*) &= \sup_{u^* \in U^*} \left(f(x) \oplus \inf_{u \in g(x)+D} \big(\langle u^*, u \rangle \{c\} + \operatorname{bd} C \big) \right) \\ &\preccurlyeq \sup_{u^* \in U^*} \left(f(x) \oplus \left(\langle u^*, 0 \rangle \{c\} + \operatorname{bd} C \right) \right) = f(x).\end{aligned}$$

From

$$\sup_{u^* \in U^*} L_c(x, u^*) \geq L_c(x, 0) = f(x)$$

we get equality. □

We are also interested in the case where $x \notin S$. Likewise to the scalar case in Proposition 3.20 we need some additional assumptions.

Proposition 3.24. *Let $f : X \to \mathcal{I}$ be a proper function, let the set $g(x) + D$ be closed and convex for every $x \in X$ and let $c \in \operatorname{int} C$. Then*

$$\sup_{u^* \in U^*} L_c(x, u^*) = \begin{cases} f(x) & \text{if } x \in S \\ \{+\infty\} & \text{otherwise.} \end{cases}$$

Proof. The first case has already been shown in Proposition 3.23.

Let $x \notin S$. For all $u^* \in U^*$, we have

$$A := \sup_{u^* \in U^*} L_c(x, u^*) \succcurlyeq f(x) \oplus \inf_{u \in g(x)+D} \left(\langle u^*, u \rangle \{c\} + \operatorname{bd} C \right).$$

From Theorem 1.65 (iii), (v) and (vii) we get

$$\varphi_A \geq \varphi_{f(x)} + \inf_{u \in g(x)+D} \varphi_{\{\langle u^*, u \rangle \{c\} + \operatorname{bd} C\}}.$$

Let $\bar{y}^* \in \operatorname{dom} \varphi_{f(x)} \cap B_c$. Then,

$$\varphi_{\{\langle u^*, u \rangle \{c\} + \operatorname{bd} C\}}(\bar{y}^*) = \langle u^*, u \rangle.$$

As shown in the proof of Proposition 3.20 (using a separation theorem), in case of $x \notin S$ there exists a sequence (u_n^*) in U^* such that $\inf_{u \in g(x)+D} \langle u_n^*, u \rangle$ tends to $+\infty$. It follows that $\varphi_A(\bar{y}^*) = +\infty$. By Theorem 1.65 (ii) we get $A = \{+\infty\}$. □

We next define the dual problem. The dual objective function (with respect to $c \in Y$) is defined by

$$\phi_c : U^* \to \mathcal{I}, \qquad \phi_c(u^*) := \inf_{x \in X} L_c(x, u^*).$$

The dual problem (with respect to $c \in Y$) associated to $(\mathrm{P_L})$ is

$$\text{maximize } \phi : U^* \to \mathcal{I} \text{ with respect to } \preccurlyeq \text{ over } T \subseteq U^*, \tag{$\mathrm{D}^{\mathrm{I}}_{\mathrm{L}}$}$$

where T is subset of U^*, such that $\{u^* \in U^* | \ \phi(u^*) \neq \{-\infty\}\} \subseteq T$. The set T is called the dual feasible set. There are important special cases, where the set T can be determined explicitly. In the linear case, for instance, a description by inequalities is possible. In the present framework we do not loose generality by setting $T = U^*$. The *dual optimal value* is denoted by

$$\bar{d}_c := \sup_{u^* \in U^*} \phi_c(u^*). \tag{3.22}$$

Theorem 3.25 (weak duality). *Let $c \in \operatorname{int} C$. Then the problems* $(\mathrm{P}_{\mathrm{L}})$ *and* $(\mathrm{D}^{\mathrm{I}}_{\mathrm{L}})$ *satisfy the weak duality inequality $\bar{d}_c \preccurlyeq \bar{p}$.*

Proof. Since $\mathcal{I}$ is a complete lattice, we immediately have

$$\sup_{u^* \in U^*} \inf_{x \in X} L_c(x, u^*) \preccurlyeq \inf_{x \in X} \sup_{u^* \in U^*} L_c(x, u^*) \tag{3.23}$$

(even if L_c would be replaced by an arbitrary function from $X \times U^*$ into $\mathcal{I}$). By Proposition 3.23 we know that $\inf_{x\in X} \sup_{u^* \in U^*} L_c(x, u^*) \preccurlyeq \bar{p}$ in case of $c \in \operatorname{int} C$. □

The main result of this section is the following Lagrange duality result of type I.

Theorem 3.26 (strong duality). *Suppose that f is convex and g is D-convex. Let*

$$g(\operatorname{dom} f) \cap -\operatorname{int} D \neq \emptyset, \tag{3.24}$$

and $c \in \operatorname{int} C$. Then strong duality between $(\mathrm{P}_{\mathrm{L}})$ *and* $(\mathrm{D}^{\mathrm{I}}_{\mathrm{L}})$ *holds; that is, $\bar{p} = \bar{d}_c$.*

Proof. If $\bar{p} = \{-\infty\}$, strong duality follows from weak duality. Note further that $\operatorname{dom} f$ is nonempty, hence $\bar{p} \neq \{+\infty\}$. Therefore, it remains to prove strong duality for the case $\bar{p} \in \mathcal{I} \setminus \{\{-\infty\}, \{+\infty\}\}$. We use the scalarization functional $\varphi_A : C^\circ \setminus \{0\} \to \overline{\mathbb{R}}$ $(A \in \mathcal{I})$ as introduced in Section 1.7. As $f : X \to \mathcal{I}$ is convex and S is a convex set (as g is D-convex), Proposition 1.64 implies $\bar{p} \in \mathcal{I}_{\mathrm{co}}$. By Corollary 1.68, $\varphi_{\bar{p}}$ is proper, in particular $\operatorname{dom} \varphi_{\bar{p}} \neq \emptyset$.

For $y^* \in B_c$ we have

$$\begin{aligned}
&\varphi\Big(y^* \Big| \inf_{u \in g(x)+D} \big(\langle u^*, u\rangle \{c\} + \operatorname{bd} C\big)\Big) \\
&\overset{\text{Th. 1.65 (vii)}}{=} \inf_{u \in g(x)+D} \varphi\big(y^* \big| \ \langle u^*, u\rangle \{c\} + \operatorname{bd} C\big) \\
&= \inf_{u \in g(x)+D} -\sigma\big(y^* \big| \ \langle u^*, u\rangle \{c\} + C\big) \\
&= \inf_{u \in g(x)+D} \langle u^*, u\rangle .
\end{aligned}$$

Let $y^* \in \operatorname{dom} \varphi_{\bar{p}} \cap B_c$. By Theorem 3.19 there exists some $\bar{u}^*$ (a solution to the scalar dual problem) such that

$$\begin{aligned}
\varphi(y^*|\bar{p}) &= \varphi\left(y^*\middle| \inf_{g(x)\cap -D\neq\emptyset} f(x)\right)\\
&\overset{\text{Th. 1.65 (vii)}}{=} \inf_{g(x)\cap -D\neq\emptyset} \varphi\big(y^*\big|\, f(x)\big)\\
&\overset{\text{Th. 3.19}}{=} \inf_{x\in X}\left(\varphi\big(y^*\big|\, f(x)\big) + \inf_{u\in g(x)+D} \langle \bar{u}^*, u\rangle\right)\\
&= \inf_{x\in X}\left(\varphi\big(y^*\big|\, f(x)\big)\right.\\
&\qquad \left. + \varphi\left(y^*\middle| \inf_{u\in g(x)+D} \left(\langle \bar{u}^*, u\rangle \{c\} + \operatorname{bd} C\right)\right)\right)\\
&\overset{\text{Th. 1.65 (v), (vii)}}{=} \varphi\big(y^*\big|\, \phi_c(\bar{u}^*)\big).
\end{aligned}$$

Together we have

$$\forall y^* \in \operatorname{dom} \varphi_{\bar{p}} \cap B_c,\ \exists \bar{u}^* \in U^* : \quad \varphi(y^*|\phi_c(\bar{u}^*)) = \varphi(y^*|\,\bar{p}\,). \tag{3.25}$$

For every $A \in \mathcal{I}$ and $\alpha > 0$ it is true that $\varphi(\alpha \cdot y^*|\,A\,) = -\alpha\varphi(y^*|\,A\,)$. We conclude from (3.25) that $\varphi(y^*|\,\bar{d}_c\,) \geq \varphi(y^*|\,\bar{p}\,)$ for all $y^* \in C^\circ \setminus \{0\}$. As $\bar{p} \in \mathcal{I}_{\text{co}}$, Theorem 1.65 (iv) yields $\bar{d}_c \succcurlyeq \bar{p}$. By the weak duality inequality we obtain $\bar{d}_c = \bar{p}$. □

Note that strong duality implies that (3.23) is satisfied with equality.

As for type I conjugate duality, the existence of a solution of the dual problem remains an open problem. Similarly to Theorem 3.7 we obtain the following result.

Theorem 3.27. *Let all the assumptions of Theorem 3.26 be satisfied and let $\bar{p} \neq \{-\infty\}$, then*

$$\bar{d}_c = \operatorname{wMax} \bigcup_{u^*\in U^*} \phi_c(u^*).$$

Proof. The proof is the same as the proof of Theorem 3.7 but using (3.25) instead of (3.7). □

3.3.3 Lagrange duality of type II

A type II variant, compare Section 3.2.2, can also be proven for Lagrange duality. As in the case of conjugate duality, we obtain the existence of a solution to the dual problem, whereas it remains open whether a solution to the type I dual problem always exists.

Based on the original definition in (3.21), we redefine the Lagrangian of $(\mathrm{P_L})$ by

$$L : X \times U^* \times C^\circ \setminus \{0\} \to \mathcal{I}, \quad L(x, u^*, y^*) := L_c\left(x, \frac{-u^*}{y^*(c)}\right) \oplus h(y^*), \quad (3.26)$$

where we assume $c \in \operatorname{int} C$. Note that it is sufficient for many reasons to consider the following simplified Lagrangian

$$L : X \times U^* \times B_c \to \mathcal{I}, \qquad L(x, u^*, y^*) := L_c(x, u^*) \oplus h(y^*).$$

The Lagrangian is independent of the choice of the parameter $c \in \operatorname{int} C$ and can be expressed using the map

$$M_{(u^*,y^*)}(u) := \operatorname{Inf}\{y \in Y\colon\ \langle u^*, u\rangle + y^*(y) \le 0\},$$

which we already introduced in Section 3.2.2.

Theorem 3.28. *Let $c \in \operatorname{int} C$. The Lagrangian L as defined in* (3.26) *can be expressed as*

$$L(x, u^*, y^*) = f(x) \oplus \inf_{u \in g(x)+D} M_{(u^*,y^*)}(u). \qquad (3.27)$$

In particular, L is independent of the choice of $c \in \operatorname{int} C$.

Proof. We have $c \in \operatorname{int} C$ and $y^* \in C^\circ \setminus \{0\}$, hence $y^*(c) < 0$. We obtain

$$\begin{aligned}
L(x, u^*, y^*) &\overset{(3.26)}{=} L_c\left(x, \tfrac{-u^*}{y^*(c)}\right) \oplus h(y^*) \\
&\overset{(3.21)}{=} f(x) \oplus \inf_{u \in g(x)+D} \left(\langle u^*, u\rangle \left\{\frac{-c}{y^*(c)}\right\} + \operatorname{bd} C\right) \oplus h(y^*) \\
&\overset{\text{Pr. 1.56 (i)}}{=} f(x) \oplus \inf_{u \in g(x)+D} \left(\langle u^*, u\rangle \left\{\frac{-c}{y^*(c)}\right\} + h(y^*)\right) \\
&\overset{\text{Pr. 3.13}}{=} f(x) \oplus \inf_{u \in g(x)+D} M_{(u^*,y^*)}(u),
\end{aligned}$$

which completes the proof. □

For feasible x, $f(x)$ is re-obtained from the Lagrangian. In contrast to the corresponding result for the type I Lagrangian, we need $f(x) \in \mathcal{I}_{\mathrm{co}}$ in the next result because we use Proposition 3.10 in the proof. This assumption cannot be dropped, because the special case $g \equiv \{0\}$, $D = \{0\}$ implies the statement of Proposition 3.10 and the assumption cannot be dropped there.

Proposition 3.29. *Let $f : X \to \mathcal{I}_{\mathrm{co}}$ and $x \in S$, then*

$$\sup_{(u^*,y^*) \in U^* \times C^\circ \setminus \{0\}} L(x, u^*, y^*) = f(x).$$

Proof. We have

$$\begin{aligned}
&\sup_{(u^*,y^*)\in U^*\times C^\circ\setminus\{0\}} L(x,u^*,y^*)\\
&\qquad\overset{(3.26)}{=} \sup_{y^*\in C^\circ\setminus\{0\}} \sup_{u^*\in U^*} \left(h(y^*)\oplus L_c\left(x, \frac{-u^*}{y^*(c)}\right)\right)\\
&\qquad\overset{\text{Pr. 1.56 (ii)}}{\preccurlyeq} \sup_{y^*\in C^\circ\setminus\{0\}} \left(h(y^*)\oplus \sup_{u^*\in U^*} L_c\left(x, \frac{-u^*}{y^*(c)}\right)\right)\\
&\qquad\overset{\text{Pr. 3.23}}{=} \sup_{y^*\in C^\circ\setminus\{0\}} \left(h(y^*)\oplus f(x)\right) \overset{\text{Pr. 3.10}}{=} f(x),
\end{aligned}$$

Conversely, we have

$$\begin{aligned}
\sup_{(u^*,y^*)\in U^*\times C^\circ\setminus\{0\}} L(x,u^*,y^*) &\succcurlyeq \sup_{y^*\in C^\circ\setminus\{0\}} L(x,0,y^*)\\
&= \sup_{y^*\in C^\circ\setminus\{0\}} \left(h(y^*)\oplus f(x)\right)\\
&\overset{\text{Pr. 3.10}}{=} f(x).
\end{aligned}$$

Together we obtain the desired equality. □

Likewise to Proposition 3.24 (type I) and Proposition 3.20 (scalar problem) we now consider the case $x \notin S$. Again we need an assumption to g.

Proposition 3.30. *Let $f : X \to \mathcal{I}_{\mathrm{co}}$ be a proper function, let the set $g(x)+D$ be closed and convex for every $x \in X$ and let $c \in \operatorname{int} C$. Then*

$$\sup_{(u^*,y^*)\in U^*\times C^\circ\setminus\{0\}} L(x,u^*,y^*) = \begin{cases} f(x) & \text{if } x\in S\\ \{+\infty\} & \text{otherwise.}\end{cases}$$

Proof. The first case has been shown in Proposition 3.29.

Let $x \notin S$. We set

$$A := \sup_{(u^*,y^*)\in U^*\times C^\circ\setminus\{0\}} L(x,u^*,y^*).$$

For all $u^* \in U^*$ and all $y^* \in C^\circ \setminus \{0\}$, we have

$$A \succcurlyeq f(x) \oplus h(y^*) \oplus \inf_{u\in g(x)+D} \left(\langle u^*, u\rangle \{c\} + \operatorname{bd} C\right).$$

From Theorem 1.65 (iii), (v) and (vii) we get

$$\varphi_A \geq \varphi_{f(x)} + \varphi_{h(y^*)} + \inf_{u\in g(x)+D} \varphi_{\{\langle u^*,u\rangle\{c\}+\operatorname{bd} C\}}.$$

Let $\bar{y}^* \in \operatorname{dom} \varphi_{f(x)} \cap B_c$. Then

$$\varphi_{\{\langle u^*,u\rangle\{c\}+\mathrm{bd}\,C\}}(\bar{y}^*) = \langle u^*, u\rangle \quad \text{and} \quad \varphi_{h(\bar{y}^*)}(\bar{y}^*) = 0.$$

As shown in the proof of Proposition 3.20 (using a separation theorem), in case of $x \notin S$ there exists a sequence (u_n^*) in U^* such that $\inf_{u\in g(x)+D} \langle u_n^*, u\rangle$ tends to $+\infty$. It follows that $\varphi_A(\bar{y}^*) = +\infty$. By Theorem 1.65 (ii) we get $A = \{+\infty\}$. □

The dual objective function of type II is defined by

$$\phi : U^* \times C^\circ \setminus \{0\} \to \mathcal{I}, \quad \phi(u^*, y^*) := \inf_{x\in X} L(x, u^*, y^*).$$

By Proposition 1.56 (i), the dual objective function can be equivalently expressed as

$$\phi : U^* \times C^\circ \setminus \{0\} \to \mathcal{I}, \quad \phi(u^*, y^*) = \phi_c\left(\frac{-u^*}{y^*(c)}\right) \oplus h(y^*). \tag{3.28}$$

If the type I dual objective function ϕ_c is $\mathcal{I}_{\mathrm{co}}$-valued, it can be re-obtained from ϕ using Corollary 3.11 as

$$\phi_c(u^*) = \sup_{y^*\in B_c} \phi(u^*, y^*). \tag{3.29}$$

The type II dual problem associated to $(\mathrm{P_L})$ is defined as

$$\text{maximize } \phi : U^* \times C^\circ \setminus \{0\} \to \mathcal{I} \text{ w.r.t. } \preccurlyeq \text{ over } U^* \times C^\circ \setminus \{0\}. \tag{$\mathrm{D_L^{II}}$}$$

The *optimal value* of $(\mathrm{D_L^{II}})$ is denoted by

$$\bar{d} := \sup_{(u^*,y^*)\in U^*\times C^\circ\setminus\{0\}} \phi(u^*, y^*). \tag{3.30}$$

The relationship to the type I dual problem is now pointed out.

Theorem 3.31. *Let $f : X \to \mathcal{I}$ be convex, let $g : X \to U$ be D-convex and $c \in \operatorname{int} C$. Then both dual problems $(\mathrm{D_L^I})$ and $(\mathrm{D_L^{II}})$ have the same optimal values; that is, $\bar{d}_c = \bar{d}$.*

Proof. It is a straightforward exercise to show that the convexity assumptions to f and g imply that $L(\cdot, u^*)$ is a convex function for every $u^* \in U^*$. By Proposition 1.64, we get $\phi_c(u^*) \in \mathcal{I}_{\mathrm{co}}$ for all $u^* \in U^*$. From (3.29) we obtain

$$\begin{aligned}\bar{d}_c = \sup_{u^*\in U^*} \phi_c(u^*) &= \sup_{u^*\in U^*} \sup_{y^*\in C^\circ\setminus\{0\}} \phi(u^*, y^*)\\ &= \sup_{(u^*, y^*)\in U^*\times C^\circ\setminus\{0\}} \phi(u^*, y^*) = \bar{d},\end{aligned}$$

which completes the proof. □

Now we immediately obtain a type II duality theorem from the corresponding type I result (Theorem 3.26). Conversely, we see that the following result implies the type I result.

Theorem 3.32 (strong duality of type II). *Let f be convex, let g be D-convex and*

$$g(\operatorname{dom} f) \cap -\operatorname{int} D \neq \emptyset.$$

Then, strong duality holds between $(\mathrm{P_L})$ *and* $(\mathrm{D_L^{II}})$*; that is,* $\bar{p} = \bar{d}$*.*

Proof. Follows from Theorems 3.26 and 3.31.

As for the conjugate duality, we can show that the Lagrange dual problem of type II has a solution, whenever the primal optimal value is finite.

Theorem 3.33. *Let all the assumptions of Theorem 3.32 be satisfied and let* $\bar{p} \neq \{-\infty\}$*. Then the dual problem* $(\mathrm{D_L^{II}})$ *has a solution.*

Proof. The proof is almost the same as the proof of Theorem 3.17 but using (3.25) instead of (3.7). □

3.4 Existence of saddle points

In the scalar theory the existence of a saddle point implies the existence of a solution to the primal and the dual problem. Under convexity assumptions and if a constraint qualification holds, the converse is also true: The existence of a solution to the primal problem implies the existence of a saddle point of the Lagrangian. We show in this section that analogous results are true for vector optimization problems. To this end the lattice extension of the vector optimization problem and the Lagrangian of type II has to be considered.

In Section 2.6, saddle points of a complete-lattice-valued function have been studied. The notion of a saddle point for the two types of $\mathcal{I}$-valued Lagrangians, as introduced in the preceding sections, is immediately obtained. The primal problem (2.13) in Section 2.6 is defined by the Lagrange-type function $l : X \times V \to Z$, which is now replaced by the Lagrangians of type I or II. The objective function in Section 2.6 is defined by

$$\forall x \in S : \; p(x) = \sup_{v \in V} l(x, v).$$

Moreover, the primal feasible set is introduced by the Lagrange-type function $l : X \times V \to Z$ as

$$S := \left\{ x \in X \,\middle|\, \sup_{v \in V} l(x, v) \neq +\infty \right\}.$$

In order to can use the concepts and results of Section 2.6, it is important to ensure that the statements of Propositions 3.24 and 3.30 hold. This is realized by the following assumptions: Let $f : X \to \mathcal{I}$ be proper and let the sets $g(x) + D$ be closed and convex for each $x \in X$. In case of the type II dual problem, we assume additionally that $f : X \to \mathcal{I}_{\rm co}$.

It is clear from Theorem 2.57 that the existence of a saddle point $(\bar X, \bar U^*) \in 2^X \times 2^{U^*}$ of the Lagrangian (3.21) implies that $\bar X$ is a solution to $(\mathrm{P_L})$. Similarly the existence of a saddle point $(\bar X, \bar V^*) \in 2^X \times 2^{U^* \times C^\circ \setminus \{0\}}$ of the Lagrangian (3.26) implies that $\bar X$ is a solution to $(\mathrm{P_L})$.

Under convexity assumptions and a constraint qualification, the opposite implication holds for the type II Lagrangian in (3.26).

Theorem 3.34. *Let $\bar X$ be a solution to* $(\mathrm{P_L})$*, where f is convex, g is D-convex and*

$$g(\operatorname{dom} f) \cap -\operatorname{int} D \neq \emptyset.$$

Then there exists $\bar V^* \subseteq U^* \times C^\circ \setminus \{0\}$ such that $(\bar X, \bar V^*)$ is a saddle point of the Lagrangian (3.26)*.*

Proof. By Theorem 3.33 there exists a solution $\bar V$ to $(\mathrm{D_L^{II}})$. Theorem 3.32 yields the equality $\inf_{x \in S} f(x) = \sup_{(u^*,y^*) \in T} \phi(u^*, y^*)$. The result now follows from Theorems 2.57. □

It remains an open problem whether or not a corresponding result is valid for the type I Lagrangian (3.21). Of course this problem is equivalent to the open problem 3.6.

3.5 Connections to classic results

In the literature one can find duality results with a vector-valued dual objective function (e.g. Boţ *et al.*, 2009, Chapter 4). We demonstrate in this section, how results of this type can be obtained from the $\mathcal{I}$-valued duality theory.

Let X be a set and let $\overline{Y}$ be an extended partially ordered topological vector space, let the ordering cone C of Y be closed and let $\emptyset \neq \operatorname{int} C \neq Y$. Let $p : X \to \overline{Y}$ be the objective function and $S \subseteq X$ be the feasible set of a vector optimization problem

$$\underset{x \in S}{\operatorname{wMin}}\, p(x), \tag{P_Y}$$

where we assume that $p[S] \subseteq Y$. One is interested in finding weakly efficient solutions to (P_Y). A vector $\bar x \in S$ is called a *weakly efficient solution* to (P_Y) if $p(\bar x) \in \operatorname{wMin} p[S]$.

We consider the lattice extension $(\mathrm{P}_{\mathcal{I}})$ with the objective function

$$p_{\mathcal{I}} : X \to \mathcal{I}, \quad p_{\mathcal{I}}(x) := \operatorname{Inf}\{p(x)\}$$

as well as the dual problem $(\mathrm{D}_{\mathcal{I}})$ as defined in Section 3.1. Let $d_{\mathcal{I}} : V \to \mathcal{I}$ be the objective function and $T \subseteq V$ the feasible set of Problem $(\mathrm{D}_{\mathcal{I}})$. For simplicity we assume that $d_{\mathcal{I}}[T] \subseteq \mathcal{I} \setminus \{\{-\infty\}, \{+\infty\}\}$. We consider the dual vector optimization problem

$$\operatorname*{wMax}_{(v,y)\in\mathcal{T}} d(v,y), \tag{D_Y}$$

where we set

$$d : V \times \overline{Y} \to \overline{Y}, \;\; d(v,y) := y \quad \text{ and } \quad \mathcal{T} := \{(v,y) \in T \times Y \mid y \in d_{\mathcal{I}}(v)\}.$$

A vector $(\bar{v},\bar{y}) \in \mathcal{T}$ is called a *weakly efficient solution* to (D_Y) if $d(\bar{v},\bar{y}) \in \operatorname{wMax} d[\mathcal{T}]$. Note that we have

$$d_{\mathcal{I}}(T) = \bigcup_{v\in T} d_{\mathcal{I}}(v) = \{y \mid (y,v) \in \mathcal{T}\} = d[\mathcal{T}]. \tag{3.31}$$

In order to formulate a weak duality assertion, we write $y^1 <_C y^2$, whenever $y^2 - y^1 \in \operatorname{int} C$.

Theorem 3.35 (weak duality). *The following statements are equivalent:*

(i) *Weak duality between* $(\mathrm{P}_{\mathcal{I}})$ *and* $(\mathrm{D}_{\mathcal{I}})$ *holds, that is, if* $\bar{x} \in S$ *and* $\bar{v} \in T$, *then* $d_{\mathcal{I}}(\bar{v}) \preccurlyeq p_{\mathcal{I}}(\bar{x})$;
(ii) *Weak duality between* (P_Y) *and* (D_Y) *holds, that is, there is no* $x \in X$ *and no* $(v,y) \in \mathcal{T}$ *such that* $p(x) <_C d(v,y)$.

Proof. Let (i) be satisfied and let $\bar{x} \in S$ and $(\bar{v},\bar{y}) \in \mathcal{T}$ be given. We get $\bar{y} \in d_{\mathcal{I}}(\bar{v})$ and $d_{\mathcal{I}}(\bar{v}) \preccurlyeq p_{\mathcal{I}}(\bar{x})$. By assumption we have $d_{\mathcal{I}}(\bar{v}) \in \mathcal{I} \setminus \{\{-\infty\}, \{+\infty\}\}$ and $p[S] \subseteq Y$ implies $p_{\mathcal{I}}(\bar{x}) \in \mathcal{I} \setminus \{\{-\infty\}, \{+\infty\}\}$. We get $\emptyset \neq \operatorname{Cl}_+ p_{\mathcal{I}}(\bar{x}) \subseteq \operatorname{Cl}_+ d_{\mathcal{I}}(\bar{v}) \neq Y$ and hence $p(\bar{x}) \in \operatorname{Cl}_+ d_{\mathcal{I}}(\bar{v})$. Corollary 1.48 (ix) and (x) yield $p(\bar{x}) \notin d_{\mathcal{I}}(\bar{v}) - \operatorname{int} C$. Thus we get $p(\bar{x}) \notin \{\bar{y}\} - \operatorname{int} C$. It follows that $p(\bar{x}) \not<_C d(\bar{v},\bar{y})$, i.e., (ii) holds.

Let (ii) be satisfied and let $\bar{x} \in S$ and $\bar{v} \in T$ be given. By assumption we have $d_{\mathcal{I}}(\bar{v}) \in \mathcal{I} \setminus \{\{-\infty\}, \{+\infty\}\}$. Hence $d_{\mathcal{I}}(\bar{v})$ is a nonempty subset of Y. For all $\bar{y} \in d_{\mathcal{I}}(\bar{v})$ we have $p(\bar{x}) \not<_C d(\bar{v},\bar{y}) = \bar{y}$. We get $p(\bar{x}) \notin d_{\mathcal{I}}(\bar{v}) - \operatorname{int} C$. Corollary 1.48 (ix) and (xi) yield $p(\bar{x}) \in \operatorname{Cl}_+ d_{\mathcal{I}}(\bar{v})$. Hence $d_{\mathcal{I}}(\bar{v}) \preccurlyeq p_{\mathcal{I}}(\bar{x})$. □

In Theorems 3.7 and 3.27 we have shown that the dual optimal value $\bar{d}_{\mathcal{I}}$ for both of the type I dual problems $(\mathrm{D}^{\mathrm{I}}_{\mathrm{F}})$ and $(\mathrm{D}^{\mathrm{I}}_{\mathrm{L}})$ can be expressed as

$$\bar{d}_{\mathcal{I}} := \sup_{v\in T} d_{\mathcal{I}}(v) = \operatorname{wMax} d_{\mathcal{I}}(T). \tag{3.32}$$

As a consequence, strong duality between $(\mathrm{P}_{\mathcal{I}})$ and $(\mathrm{D}_{\mathcal{I}})$ entails a classical scheme of strong duality.

Theorem 3.36 (strong duality). *Assume that strong duality holds between* $(\mathrm{P}_{\mathcal{I}})$ *and* $(\mathrm{D}_{\mathcal{I}})$*, that is,* $\bar{p}_{\mathcal{I}} = \bar{d}_{\mathcal{I}}$*, and let* (3.32) *be satisfied. Then, strong duality between* (P_Y) *and* (D_Y) *holds; that is, if* $\bar{x}$ *is a weakly efficient solution to* (P_Y)*, then there exists a weakly efficient solution* $(\bar{v}, \bar{y})$ *to* (D_Y) *such that* $p(\bar{x}) = d(\bar{v}, \bar{y})$*.*

Proof. Let $\bar{y} = p(\bar{x}) \in \operatorname{wMin} p[S]$. We get

$$\begin{aligned}\operatorname{wMin} p[S] \subseteq \operatorname{Inf} p[S] &= \operatorname{Inf} \bigcup_{x \in S} \operatorname{Inf}\{p(x)\} = \inf_{x \in S} p_{\mathcal{I}}(x) \\ &= \bar{p}_{\mathcal{I}} = \bar{d}_{\mathcal{I}} \overset{(3.32)}{=} \operatorname{wMax} d_{\mathcal{I}}(T) \overset{(3.31)}{=} \operatorname{wMax} d[\mathcal{T}].\end{aligned}$$

It follows $\bar{y} \in \operatorname{wMax} d[\mathcal{T}] \subseteq d[\mathcal{T}]$. Hence there exists $\bar{v} \in T$ such that $(\bar{v}, \bar{y}) \in \mathcal{T}$ and $p(\bar{x}) = \bar{y} = d(\bar{v}, \bar{y})$. □

Under the common (but restrictive) assumption that $p[S] + C$ is closed, we get also the so-called converse duality.

Theorem 3.37 (converse strong duality). *Assume that strong duality holds between* $(\mathrm{P}_{\mathcal{I}})$ *and* $(\mathrm{D}_{\mathcal{I}})$ *and let* $p[S] + C$ *be closed. Then, converse strong duality between* (P_Y) *and* (D_Y) *holds; that is, if* $(\bar{v}, \bar{y})$ *is a weakly efficient solution to* (D_Y)*, then* $\bar{y} \in \operatorname{wMin}(p[S] + C)$*.*

Proof. Let $\bar{y} = d(\bar{v}, \bar{y}) \in \operatorname{wMax} d[\mathcal{T}]$. As $p[S] + C$ is closed, we get

$$\operatorname{Inf} p[S] = \operatorname{wMin} \operatorname{cl}(P[S] + C) = \operatorname{wMin}(p[S] + C).$$

It follows

$$\begin{aligned}\operatorname{wMax} d[\mathcal{T}] &\overset{(3.31)}{=} \operatorname{wMax} d_{\mathcal{I}}(T) \subseteq \operatorname{Sup} d_{\mathcal{I}}(T) = \bar{d}_{\mathcal{I}} \\ &= \bar{p}_{\mathcal{I}} = \operatorname{Inf} p[S] = \operatorname{wMin}(p[S] + C),\end{aligned}$$

which completes the proof. □

The opposite direction of the statements in the last two theorems can be shown when $p[S] + C$ is assumed to be closed.

Theorem 3.38. *Assume that strong duality and converse strong duality hold between* (P_Y) *and* (D_Y)*. Further let* (3.32) *be satisfied and let* $p[S] = \operatorname{cl}(p[S] + C)$*. Then strong duality between* $(\mathrm{P}_{\mathcal{I}})$ *and* $(\mathrm{D}_{\mathcal{I}})$ *holds.*

Proof. By the assumption $p[S] = \operatorname{cl}(p[S] + C)$, we get

$$\operatorname{wMin} p[S] = \operatorname{wMin} \operatorname{cl}(p[S] + C) = \operatorname{Inf} p[S] = \bar{p}_{\mathcal{I}}$$

and (3.32) yields

$$\operatorname{wMax} d[\mathcal{T}] = \operatorname{wMax} d_{\mathcal{I}}(T) = \bar{d}_{\mathcal{I}}.$$

Strong duality between (P_Y) and (D_Y) yields $\mathrm{wMin}\, p[S] \subseteq \mathrm{wMax}\, d[\mathcal{T}]$. Converse strong duality between (P_Y) and (D_Y) yields $\mathrm{wMax}\, d[\mathcal{T}] \subseteq \mathrm{wMin}(p[S]+C)$.

We have $p[S] \subseteq p[S]+C \subseteq \mathrm{cl}\,(p[S]+C) \subseteq p[S]$ and hence $p[S] = p[S]+C$. Together we obtain $\bar{p}_{\mathcal{I}} = \bar{d}_{\mathcal{I}}$. □

We observe that duality between (P_Y) and (D_Y) involves the existence of weakly minimal elements. In the scalar duality theory (and likewise in the $\mathcal{I}$-valued theory) we obtain the existence of a solution to the dual problem as a result. But a solution to the primal problem is not required to exist in order to get duality assertions.

3.6 Notes on the literature

A general classification of vectorial duality can be found in the recent book by Boţ *et al.* (2009) on duality in vector optimization. The authors distinguish between duality via scalarization, Wolfe and Mond-Weir duality concepts and duality based on vector conjugacy. Some results of this chapter are related to results of the first and third class of problems in Boţ *et al.* (2009), but our philosophy is a different one. Among all other approaches to duality in vector optimization the one by Tanino (1992) seems to be the closest. This paper is followed, for instance, by (Song, 1997, 1998; Chen and Li, 2009; Li *et al.*, 2009). The paper by Tanino (1992) is partially based on several earlier works (Nieuwenhuis, 1980; Kawasaki, 1981, 1982; Sawaragi *et al.*, 1985; Tanino, 1988).

The conjugate in Definition 3.4 can be seen as a combination of the k-conjugate introduced by Tanino and Sawaragi (1980) and the conjugate considered by Tanino (1992), see also Chapter 7 in (Boţ *et al.*, 2009). One can sometimes observe similarities between the mentioned results from the literature and the results of this chapter. The difference is, however, that the infimality concept used by Nieuwenhuis (1980); Tanino (1992) and others was not considered to be an infimum in a complete lattice.

Another approach to vectorial duality is to suppose the existence of the supremum of every bounded subset of the vector space (least upper bound property) (e.g. Zowe, 1975; Zălinescu, 1983; Pallaschke and Rolewicz, 1997). This approach seems to be only of theoretical interest (compare e.g. Jahn, 2004, p. 107: *"The notion of strong minimality is very restrictive and is often not applicable in practice."*). Duality on a very abstract level, but with connections to our approach has been studied by Martínez-Legaz and Singer (1994).

The general duality concept in Section 3.1 seems to be new, in particular, in view of its application in vector optimization. The canonical extension has been introduced in (Heyde and Löhne, 2010).

The conjugate duality of type I has been published in a finite dimensional setting in (Löhne and Tammer, 2007). Similar proof techniques have been already used in (Löhne, 2005b). The type II conjugate duality is due to Hamel (2009a,b); Schrage (2009), where the complete lattice $\mathcal{F}$ is used and the assumptions to the cone C are therefore weaker. The existence of a solution to the dual problem has been proven by Hamel (2009a,b) with respect to another but related solution concept. In particular, the generalized affine map $M_{(x^*,y^*)}(x)$, compare Proposition 3.13, has been introduced and studied by Hamel (2009a). A new aspect in our exposition is that Y does not need to be Hausdorff.

Lagrange duality and saddle points in vector optimization and set-valued optimization have been investigated by Tanaka (1990); Martein (1990); Tammer (1991); Li and Wang (1994); Tanaka (1994); Tan *et al.* (1996); Li and Chen (1997); Götz and Jahn (1999); Adán and Novo (2005); Ehrgott and Wiecek (2005b); Ha (2005) and many others. The set-valued constraints in Section 3.3 were used by several authors, among them Oettli (1982), Borwein (1977); Corley (1987); Luc (1988); Götz and Jahn (1999). In (Löhne, 2005a) similar Lagrange duality results in a set-valued framework can be found. The saddle point theorem in Section 3.4 is a new result based on the definition of a saddle point due to Andreas H. Hamel and the author. It seems to be the first ever saddle point theorem in vector optimization, which is based on the notion of infimum and supremum.

The classical duality scheme which is considered in Section 3.5 is due to Jahn (1983) and is discussed in (Boţ *et al.*, 2009, Section 4.3.4), where specific classes of problems are considered.

Part II
Linear Problems

Chapter 4
Solution concepts and duality

Linear vector optimization problems provide an important subclass of convex vector optimization problems that is of special importance. On the one hand, they occur in a wide range of applications in economics, engineering, finance and other fields. On the other hand, in contrast to the general case, the linear structure allows to use special methods and yields specific results. It is an important feature of linear vector optimization problems that one can deal with only polyhedral sets.

Even though several results of this chapter are already covered by the convex theory in the first part of this book, the following exposition of the linear case is self-contained. The concepts already introduced in the first part are recalled and all the results are proven directly. Nevertheless, at several places the relationship to results and concepts of Part I is discussed.

We consider the problem to minimize q linear objective maps $P_i : \mathbb{R}^n \to \mathbb{R}$ under linear constraints. The objective functions P_i can be regarded as the rows of a $(q \times n)$-matrix P. The feasible set is given by

$$S := \{x \in \mathbb{R}^n |\; Bx \geq b\},$$

where $B \in \mathbb{R}^{m \times n}$ and $b \in \mathbb{R}^m$. The problem can be expressed as

$$\text{minimize } P : \mathbb{R}^n \to \mathbb{R}^q \text{ with respect to } \leq \text{ over } S, \tag{P}$$

where $\leq$ stands for the corresponding componentwise ordering relation in $\mathbb{R}^q$ and $\mathbb{R}^m$.

This chapter is organized as follows. We start with scalarization methods in Section 4.1. A solution concept for (P) is introduced in Section 4.2. In particular, we point out that the solution concepts introduced in Chapter 2 are meaningful and useful in the linear case. Motivated by the ideas of Chapter 3, we present in Section 4.3 a duality theory for linear vector optimization problems. In Section 4.4, the theory is related to the type II duality results for the convex case, which have been studied in Chapter 3. We shall demonstrate that the concepts and results for the linear case can be reformulated using

A. Löhne, *Vector Optimization with Infimum and Supremum*, Vector Optimization,
DOI 10.1007/978-3-642-18351-5_5,

infimum and supremum. In Section 4.5 it is shown that the set-valued dual problem has a vector-valued counterpart, which is called the geometric dual problem. Homogeneous problems shall be studied in Section 4.6. We continue in Section 4.7 with a criterion to identify those faces of the polyhedral image set of a linear vector optimization problem that consist of only minimal elements.

4.1 Scalarization

One of the most important techniques in vector optimization is scalarization. In this section we point out two basic scalarization methods. Moreover, we will study solutions of scalarized problems. Several basic techniques, which will be frequently used throughout this chapter, are provided. The scalarization methods are also motivated from an applicational point of view. It turns out that it is beneficial to consider both scalarization methods simultaneously. This can be understood as a kind of duality.

4.1.1 Basic methods

Two fundamental scalarization methods are introduced in this subsection. Subsequently, both methods will play a crucial role. By the way we will recall several duality statements for scalar linear programs.

For the first method, we create a new objective map based on the given q linear objective functions $x \mapsto P_i x$. To this end we assign a nonnegative weight $w_i \in \mathbb{R}_+$ to each of the objective maps. The weighted sum is considered to be our new objective map

$$\sum_{i=1}^{q} w_i P_i x = w^T P x.$$

This method is called *weighted sum scalarization*. For each vector $w \in \mathbb{R}^q,\ w \geq 0$, we obtain a scalar linear program

$$\min w^T P x \quad \text{ subject to } \ Bx \geq b. \tag{P$_1(w)$}$$

The weights can be normalized so that we can assume $\sum_{i=1}^{q} w_i = 1$, which can be written as $e^T w = 1$ for $e^T = (1, \ldots, 1)$. If one weight w_i is zero, then the corresponding objective map has no influence. Therefore, it can be useful to assume $w_i > 0$ in certain situations.

The dual problem of $(\mathrm{P}_1(w))$ is obtained as

$$\max b^T u \quad \text{subject to} \quad \begin{cases} B^T u = P^T w \\ u \geq 0. \end{cases} \tag{$\mathrm{D}_1(w)$}$$

The scalar duality theory turns out to be a principal tool in the following sections. Thus let us recall the basic facts. If x is feasible for $\mathrm{P}_1(w)$ and u is feasible for $\mathrm{D}_1(w)$, then we have $b^T u \leq w^T P x$ (weak duality). Moreover, solutions to both problems exist in this case. If a solution to $\mathrm{P}_1(w)$ exists, then there also exists a solution to $\mathrm{D}_1(w)$ and vice versa. The optimal values coincide in this case (strong duality). The same is true for other pairs of dual linear programs.

Note that the term *solution* always means *optimal solution* in this book. The term *feasible solution*, which is often used in the literature on linear programming, is not used here. Instead, we speak about *feasible points (or feasible elements, or feasible vectors)*.

We next consider a second scalarization method. The q linear objective functions are associated to a single reference variable. The i-th objective function is restrained from being larger than a common reference variable z plus a fixed real number y_i, that is,

$$\begin{aligned} P_1 x &\leq y_1 + z \\ P_2 x &\leq y_2 + z \\ &\cdots \\ P_q x &\leq y_q + z. \end{aligned}$$

The reference variable z becomes the objective function that has to be minimized. This standard method is called *scalarization by a reference variable.* Setting $e = (1, \ldots, 1)^T$, we get for each vector $y \in \mathbb{R}^q$ the scalar linear program

$$\min z \quad \text{subject to} \quad \begin{cases} Bx \geq b \\ Px \leq y + z \cdot e. \end{cases} \tag{$\mathrm{P}_2(y)$}$$

The dual program is easily obtained as

$$\max b^T u - y^T w \quad \text{subject to} \quad \begin{cases} B^T u - P^T w = 0 \\ e^T w = 1 \\ (u, w) \geq 0. \end{cases} \tag{$\mathrm{D}_2(y)$}$$

Both pairs of dual problems $(\mathrm{P}_1(w))$, $(\mathrm{D}_1(w))$ and $(\mathrm{P}_2(y))$, $(\mathrm{D}_2(y))$ will play a crucial role in the following. In the subsequent results and proofs, both pairs often occur together. In some sense, the pairs can therefore considered to be dual to each other.

4.1.2 Solutions of scalarized problems

We recall in this section the classical notions of weakly efficient solutions and efficient solutions. It is shown that they correspond to solutions of scalarized problems. Even though the term "solution" is used, we distinguish these concepts from solutions to vector optimization problems, which shall be introduced in the subsequent sections.

For vectors $\bar{y}, \hat{y} \in \mathbb{R}^q$, we write $\bar{y} < \hat{y}$ for the componentwise strict order, which is equivalent to $\hat{y} - \bar{y} \in \operatorname{int} \mathbb{R}^q_+$. Recall that a feasible vector $\bar{x} \in S$ is called *weakly efficient solution* to (P) if there is no feasible vector $x \in S$ with $Px < P\bar{x}$. The set of weakly efficient solutions to (P) is denoted by wEff (P).

The following result shows that a weakly efficient solution to (P) can be characterized by solutions to both types of scalarized problems $(\mathrm{P}_1(w))$ and $(\mathrm{P}_2(y))$ for suitable parameters w and y. The notation $w \gneq 0$ stands for $[w \geq 0 \ \wedge \ w \neq 0]$ or equivalently $w \in \mathbb{R}^q_+ \setminus \{0\}$.

Theorem 4.1. *For the linear vector optimization problem* (P) *the following statements are equivalent.*

(i) *There is a weight vector* $w \gneq 0$ *such that* $\bar{x}$ *solves* $(\mathrm{P}_1(w))$.
(ii) *There is a vector* $y \in \mathbb{R}^q$ *such that* $(\bar{x}, 0)$ *solves* $(\mathrm{P}_2(y))$.
(iii) $\bar{x}$ *is a weakly efficient solution to* (P).

Proof. (i) $\Rightarrow$ (iii). There is no $x \in S$ with $w^T Px < w^T P\bar{x}$. Hence there is no $x \in S$ with $Px < P\bar{x}$ because otherwise $w \gneq 0$ would imply $w^T Px < w^T P\bar{x}$.

(iii) $\Rightarrow$ (ii). We set $y := P\bar{x}$. Then $(\bar{x}, 0)$ is feasible for $(\mathrm{P}_2(y))$. Suppose that $(\bar{x}, 0)$ is not optimal for $(\mathrm{P}_2(y))$. Then there is some feasible (x, z) for $(\mathrm{P}_2(y))$ such that $z < 0$. This implies $Px \leq y + ze = P\bar{x} + ze < P\bar{x}$, i.e., (iii) does not hold.

(ii) $\Rightarrow$ (i). Let $(\bar{x}, 0)$ be a solution to $(\mathrm{P}_2(y))$. From the scalar duality theorem we obtain a solution $(\bar{u}, \bar{w})$ of $(\mathrm{D}_2(y))$ and the optimal values of both problems coincide, that is, $b^T \bar{u} - y^T \bar{w} = 0$. It can be seen that $\bar{u}$ is a solution to $(\mathrm{D}_1(\bar{w}))$. Applying again the scalar duality theorem we get a solution $\hat{x}$ to $(\mathrm{P}_1(\bar{w}))$. The optimal values coincide, that is, $b^T \bar{u} = \bar{w}^T P\hat{x}$. It follows $\bar{w}^T P\hat{x} = y^T \bar{w}$. Since $(\bar{x}, 0)$ is feasible for $(\mathrm{P}_2(y))$, we have $P\bar{x} \leq y$ and hence $\bar{w}^T P\bar{x} \leq \bar{w}^T y = y^T \bar{w}$. It follows that $\bar{w}^T P\bar{x} \leq \bar{w}^T P\hat{x}$. Thus, $\bar{x}$ solves $(\mathrm{P}_1(\bar{w}))$. □

Recall that

$$\operatorname{wMin} A = \left\{ y \in A \,|\, (\{y\} - \operatorname{int} \mathbb{R}^q_+) \cap A = \emptyset \right\}$$

is the set of *weakly minimal vectors* of a set $A \subseteq \mathbb{R}^q$ with respect to $\mathbb{R}^q_+$. The set of weakly maximal vectors of A is $\operatorname{wMax} A = -\operatorname{wMin}(-A)$. For a function $f : X \to Y$ and a set $A \subseteq X$, we use the notation

$$f[A] := \{ f(x) \in \mathbb{R}^q \,|\, x \in A \}.$$

The set $P[S]$ is called *image* of Problem (P). It can easily be seen that

$$P[\text{wEff}\,(\text{P})] = \text{wMin}\, P[S].$$

We proceed with similar considerations for efficient solutions, which we already introduced in a more general framework. A feasible vector $\bar{x} \in S$ is an *efficient solution* to (P) if and only if there is no feasible vector x with $Px \lneq P\bar{x}$. The set of efficient solutions to (P) is denoted by Eff (P).

It is possible to give a characterization of efficient solutions similar to the characterization of weakly efficient solutions in Theorem 4.1. To this end we consider the following dual pair of scalar problems:

$$\min e^T Px \quad \text{subject to} \quad \begin{cases} Bx \geq b \\ Px \leq y, \end{cases} \tag{$\text{P}_3(y)$}$$

$$\max b^T u - y^T v \quad \text{subject to} \quad \begin{cases} B^T u - P^T v = P^T e \\ (u, v) \geq 0. \end{cases} \tag{$\text{D}_3(y)$}$$

Again, the weighted sum scalarization ($\text{P}_1(w)$) is involved in the characterization, but in contrast to Theorem 4.1 non-zero weights are not allowed.

Theorem 4.2. *For the linear vector optimization problem* (P) *the following statements are equivalent:*

(i) *There exists a weight vector* $w > 0$ *such that* $\bar{x}$ *solves* ($\text{P}_1(w)$)*;*
(ii) *There exists some* $y \in \mathbb{R}^q$ *such that* $\bar{x}$ *solves* ($\text{P}_3(y)$)*;*
(iii) $\bar{x}$ *is an efficient solution to* (P)*.*

Proof. (i) $\Rightarrow$ (iii). There is no $x \in S$ such that $w^T Px < w^T P\bar{x}$. Hence, there is no $x \in S$ with $Px \lneq P\bar{x}$. Otherwise, $w > 0$ would imply $w^T Px < w^T P\bar{x}$, which contradicts (i).

(iii) $\Rightarrow$ (ii). We set $y := P\bar{x}$. Then $\bar{x}$ is feasible for ($\text{P}_3(y)$). Assuming that $\bar{x}$ is not a solution to ($\text{P}_3(y)$), we get a feasible x with $e^T Px < e^T P\bar{x}$. Taking into account $Px \leq y \leq P\bar{x}$, we obtain $Px \lneq P\bar{x}$. This is a contradiction as (iii) is not satisfied.

(ii) $\Rightarrow$ (i). There exists some $y \in \mathbb{R}^q$ such that $\bar{x}$ solves ($\text{P}_3(y)$). From a scalar duality result we get the existence of a solution $(\bar{u}, \bar{v})$ to ($\text{D}_3(y)$). Moreover, the optimal values of both problems coincide, that is, $b^T \bar{u} - y^T \bar{v} = e^T P\bar{x}$. It can be shown that $\bar{u}$ is an optimal solution to ($\text{D}_1(\bar{w})$), where $\bar{w} := \bar{v} + e > 0$. Applying again the duality theorem, we get the existence of a solution $\hat{x}$ to ($\text{P}_1(\bar{w})$). As the optimal values of both problems coincide, we get $b^T \bar{u} = (\bar{v} + e)^T P\hat{x}$. It follows $(\bar{v} + e)^T P\hat{x} = \bar{v}^T y + e^T P\bar{x}$. Since $\bar{x}$ is feasible for ($\text{P}_3(y)$), we get $P\bar{x} \leq y$ and hence $\bar{v}^T P\bar{x} \leq \bar{v}^T y$. We obtain $\bar{w}^T P\bar{x} \leq \bar{w}^T P\hat{x}$. Thus $\bar{x}$ solves ($\text{P}_1(\bar{w})$), too. □

Recall that

$$\text{Min}\, A = \left\{ y \in A \middle| \; \left(\{y\} - \mathbb{R}^q_+ \setminus \{0\} \right) \cap A = \emptyset \right\}$$

is the set of *minimal vectors* of a set $A \subseteq \mathbb{R}^q$ with respect to $\mathbb{R}^q_+$. The set of *maximal vectors* of A is $\operatorname{Max} A := -\operatorname{Min}(-A)$. From the definitions we immediately get

$$P[\operatorname{Eff}(\mathrm{P})] = \operatorname{Min} P[S].$$

The following statement is a special result for the linear case.

Corollary 4.3. *Let* Eff (P) *be nonempty. Then*

$$\operatorname{Min} P[S] + \mathbb{R}^q_+ = P[S] + \mathbb{R}^q_+.$$

Proof. The inclusion $\subseteq$ is obvious. To show the opposite inclusion let $y \in P[S] + \mathbb{R}^q_+$. Choose some $\bar{x} \in \operatorname{Eff}(\mathrm{P})$. By Theorem 4.2, there exists some $\bar{y} \in \mathbb{R}^q$ such that $\bar{x}$ solves $(\mathrm{P}_3(\bar{y}))$. From the scalar duality theory, we conclude that $(\mathrm{D}_3(\bar{y}))$ has a solution. It follows that $(\mathrm{D}_3(y))$ has a feasible point. Since $y \in P[S] + \mathbb{R}^q_+$, there exists $x \in S$ such that $y \geq Px$. This means that x is feasible for $(\mathrm{P}_3(y))$. From the scalar duality theory, we deduce that an optimal solution $\hat{x}$ to $(\mathrm{P}_3(y))$ exists. We have $y \geq P\hat{x}$ and Theorem 4.2 yields $P\hat{x} \in \operatorname{Min} P[S]$. Hence $y \in \operatorname{Min} P[S] + \mathbb{R}^q_+$. □

4.2 Solution concept for the primal problem

A solution concept for the primal linear vector optimization problem (P) is now introduced. This concept is related (but slightly different) to those introduced in Chapter 2. Taking into account the polyhedral structure of the problem, a solution is envisioned to be a finite subset of the feasible set. In order to realize this, not only vectors but also directions have to be considered. A solution $\bar{S} \subseteq S$ is intended to be a subset of the efficient solutions. This condition can be written as

$$P[\bar{S}] \subseteq \operatorname{Min} P[S].$$

On the other hand we expect that enough information is contained in a solution $\bar{S}$ in order to characterize the set

$$\mathcal{P} := P[S] + \mathbb{R}^q_+,$$

which is called the *upper image* of (P). Note that in scalar optimization, the upper image is completely determined by a single minimal solution but this is typically not the case in vector optimization. Moreover, bear in mind that the latter requirement refers to the attainment of the infimum in Chapter 2.

Finitely many vectors are not adequate to fix a *polyhedral set*, which is defined to be the intersection of finitely many halfspaces[1]. However, a finite

[1] This is commonly called a polyhedral *convex* set. Since we only consider convex sets, we use the shorter term, likewise for related concepts.

characterization is possible by directions and vectors. Given a nonempty convex set $A \subseteq \mathbb{R}^n$, we say that $y \in \mathbb{R}^n \setminus \{0\}$ is a *direction* (of recession) in A if the half-line $L(x,y) := \{x + \lambda y \mid \lambda \geq 0\}$ belongs to A for all vectors $x \in A$. For closed convex sets it is sufficient to require the existence of some $x \in A$ such that $L(x,y) \subseteq A$ in order to obtain a direction (Rockafellar, 1972, Theorem 8.3). The set of all directions in $A \subseteq \mathbb{R}^n$ together with $0 \in \mathbb{R}^n$ is called the *recession cone* (or asymptotic cone). It is denoted by A_∞. A direction y in A is called *extreme* if there are no directions $\bar{y}, \hat{y}$ in A with $\bar{y} \neq \alpha \hat{y}$ for all $\alpha > 0$ such that $y = \bar{y} + \hat{y}$.

A vector $\hat{x} \in \mathbb{R}^n \setminus \{0\}$ is a direction of the feasible set S of Problem (P), called a *feasible direction* if and only if $\hat{x}$ is a nonzero solution of the homogeneous system of inequalities, that is,

$$0 \neq \hat{x} \in S^h := \{x \in \mathbb{R}^n \mid Bx \geq 0\}.$$

It is known (Rockafellar, 1972, Theorem 19.1) that every polyhedral set A can be expressed as

$$S = Q + K,$$

where Q is a polytope (that is a bounded polyhedral set) and K a polyhedral cone having the same directions as A. Recall that each polytope is the convex hull of finitely many vectors and each polyhedral cone is the conical hull of finitely many directions. The *conical hull* of directions $k^1, \dots, k^s \in \mathbb{R}^n \setminus \{0\}$ is the set

$$\operatorname{cone}\left\{k^1, \dots, k^s\right\} := \left\{\sum_{i=1}^{s} \mu_i k^i \,\middle|\, \mu_1, \dots, \mu_s \geq 0\right\}.$$

Furthermore, we set $\operatorname{cone} \emptyset := \{0\}$. If $A \subseteq \mathbb{R}^n$ is a nonempty polyhedron, there are $x^1, \dots, x^r \in \mathbb{R}^n$ $(r \geq 1)$ and $k^1, \dots, k^s \in \mathbb{R}^n \setminus \{0\}$ $(s \geq 0)$ such that

$$A = \operatorname{co}\left\{x^1, \dots, x^r\right\} + \operatorname{cone}\left\{k^1, \dots, k^s\right\}.$$

In case of $s = 0$, no directions occur. This is exactly the case when the set A is bounded (see e.g. Rockafellar, 1972, Theorem 8.4). The recession cone of A can be expressed as

$$A_\infty = \operatorname{cone}\left\{k^1, \dots, k^s\right\}.$$

As a consequence, for a linear map $L : \mathbb{R}^n \to \mathbb{R}^p$ we have

$$L\left(A_\infty\right) = L(A)_\infty, \tag{4.1}$$

and two nonempty polyhedral sets A^1, A^2 satisfy

$$A^1_\infty + A^2_\infty = \left(A^1 + A^2\right)_\infty. \tag{4.2}$$

The feasible set $S \subseteq \mathbb{R}^n$ can be represented by a finite set $\bar{S} \subseteq S$ and a finite set $\bar{S}^h \subseteq S^h \setminus \{0\}$ as

$$S = \operatorname{co} \bar{S} + \operatorname{cone} \bar{S}^h.$$

This implies the following finite description of the upper image of (P),

$$\mathcal{P} = P[S] + \mathbb{R}^q_+ = \operatorname{co} P[\bar{S}] + \operatorname{cone} P[\bar{S}^h] + \mathbb{R}^q_+.$$

We next propose a solution concept for the linear vector optimization problem (P), which is based on the above ideas.

Definition 4.4. A nonempty set $\bar{S} \subseteq \mathbb{R}^n$ together with a (possibly empty) set $\bar{S}^h \subseteq \mathbb{R}^n \setminus \{0\}$ is called a *finitely generated solution* to the linear vector optimization problem (P) if

(i) $\bar{S}$ is a finite subset of S,
(ii) $\bar{S}^h$ is a finite subset of S^h,
(iii) $P[\bar{S}] \subseteq \operatorname{Min} P[S]$,
(iv) $P[\bar{S}^h] \subseteq \operatorname{Min} P[S^h]$,
(v) $P[S] \subseteq \operatorname{co} P[\bar{S}] + \operatorname{cone} P[\bar{S}^h] + \mathbb{R}^q_+$.

Note that, if $\bar{S}^h = \emptyset$, the conditions (ii) and (iv) disappear (as they are obviously satisfied) and (v) reduces to $P[S] \subseteq \operatorname{co} P[\bar{S}] + \mathbb{R}^q_+$.

If the there is no risk of confusion, in particular, in the context of linear vector optimization, a finitely generated solution can also be called a *solution* to (P). In the general context of vector optimization we have to distinguish solutions (as defined in Section 2.2) from finitely generated solutions. The relationship to the solution concepts from Chapter 2 is pointed out below.

If the feasible set S in the linear vector optimization problem (P) is replaced by S^h, we obtain the homogeneous problem

$$\text{minimize } P : \mathbb{R}^n \to \mathbb{R}^q \text{ with respect to } \leq \text{ over } S^h. \tag{P^h}$$

The upper image of (P^h) is defined as

$$\mathcal{P}^h := P[S^h] + \mathbb{R}^q_+.$$

We also consider the corresponding scalarized problem

$$\min w^T P x \quad \text{subject to } Bx \geq 0 \tag{$\mathrm{P}^h_1(w)$}$$

as well as its dual problem

$$\max 0^T u \quad \text{subject to } \begin{cases} B^T u = P^T w \\ u \geq 0. \end{cases} \tag{$\mathrm{D}^h_1(w)$}$$

A nonzero efficient solution to (P^h) is called an *efficient direction* to (P). This notion is motivated by the following result.

Proposition 4.5. *The following statements are equivalent:*

(i) $S \neq \emptyset$ *and* $k \in \mathrm{Eff}\,(\mathrm{P}^h)$,
(ii) $\exists x \in S, \forall \lambda \geq 0 : x + \lambda k \in \mathrm{Eff}\,(\mathrm{P})$.

Proof. Let (i) be satisfied. By Theorem 4.2, there exists $w > 0$ such that k solves $(\mathrm{P}_1^h(w))$. Hence there exists a solution to the dual problem $(\mathrm{D}_1^h(w))$. It follows that the feasible set of $(\mathrm{D}_1(w))$ is nonempty. Since $S \neq \emptyset$, $(\mathrm{P}_1(w))$ has a solution x. We have $Pk = 0$ (by duality), hence $x + \lambda k$ is a solution to $(\mathrm{P}_1(w))$ for all $\lambda \geq 0$. Theorem 4.2 implies that $x + \lambda k$ is an efficient solution to (P).

Let (ii) be satisfied. We have $B(x + \lambda k) \geq b$ for all $\lambda \geq 0$ and hence

$$\forall \lambda > 0 : \quad B\left(\frac{x}{\lambda} + k\right) \geq \frac{b}{\lambda}.$$

Taking the limit for $\lambda \to \infty$, we get $k \in S^h$. Let $d \in S^h$ such that $Pd \leq Pk$. We have $P(x+d) \leq P(x+k)$. Since $P(x+k) \in \mathrm{Min}\, P[S]$, we get $P(x+d) = P(x+k)$ and hence $Pd = Pk$, i.e., $k \in \mathrm{Eff}\,(\mathrm{P}^h)$. □

We next show that extreme directions of the upper image $\mathcal{P}^h$ of (P) are minimal except the unit vectors, which are the extreme directions of the ordering cone.

Lemma 4.6. *Let* $\mathrm{Min}\, \mathcal{P}^h \neq \emptyset$. *If* $k \in \mathbb{R}^q \setminus \{0\}$ *is an extreme direction of* $\mathcal{P}^h$ *such that* $k \notin \mathbb{R}^q_+$, *then* $k \in \mathrm{Min}\, \mathcal{P}^h$.

Proof. Assume that k is not minimal in $\mathcal{P}^h$. There exists some

$$z \in \big(\{k\} - \mathbb{R}^q_+ \setminus \{0\}\big) \cap \mathcal{P}^h.$$

We conclude that $y := k + (k - z) \in \mathcal{P}^h + \mathbb{R}^q_+ = \mathcal{P}^h$ and $z = k - (k - z) \in \mathcal{P}^h$. Since $\frac{1}{2}y + \frac{1}{2}z = k$ and k is an extreme direction of $\mathcal{P}^h$ there exists $\alpha > 0$ such that $y = \alpha z$. We obtain

$$(\alpha - 1)k = (1 + \alpha)(k - z) \in \mathbb{R}^q_+ \setminus \{0\}.$$

Thus $\alpha \neq 1$. Since $k \notin \mathbb{R}^q_+$ we get $\alpha < 1$ and hence $k \in -\mathbb{R}^q_+ \setminus \{0\}$. This means that $\mathcal{P}^h$ has a direction that belongs to $-\mathbb{R}^q_+ \setminus \{0\}$. Hence $\mathcal{P}^h$ has no minimal element, which contradicts the assumption $\mathrm{Min}\, \mathcal{P}^h \neq \emptyset$. □

In the present setting, the solution concepts introduced in Chapter 2 can be characterized by weaker conditions. The reader who is interested in a self-contained theory of linear vector optimization can consider Proposition 4.7 as a definition and can continue with Theorem 4.11.

Proposition 4.7. *A nonempty set* $\bar{X} \subseteq \mathbb{R}^n$ *is a solution to* (P) *(in the sense of Definition 2.20) if and only if*

(i) $\bar{X} \subseteq S$,
(ii) $P[\bar{X}] = \operatorname{Min} P[S]$,

Proof. In view of Theorem 2.21 we have to show that (i) and (ii) imply $\operatorname{Inf} P[\bar{X}] = \operatorname{Inf} P[S]$. By Corollary 4.3, we have $\operatorname{Min} P[S] + \mathbb{R}^q_+ = P[S] + \mathbb{R}^q_+$. From (ii) we get $P[\bar{X}] + \mathbb{R}^q_+ = P[S] + \mathbb{R}^q_+$. This implies $\operatorname{Cl}_+ P[\bar{X}] = \operatorname{Cl}_+ P[S]$ and hence $\operatorname{Inf} P[\bar{X}] = \operatorname{Inf} P[S]$. □

Proposition 4.8. *A nonempty set $\bar{X} \subseteq \mathbb{R}^n$ is a mild solution to* (P) *if and only if*

(i) $\bar{X} \subseteq S$,
(ii) $P[\bar{X}] \subseteq \operatorname{Min} P[S]$,
(iii) $P[S] \subseteq \operatorname{cl}\left(P[\bar{X}] + \mathbb{R}^q_+\right)$.

Proof. The proof is similar to the proof of Proposition 4.7 (even easier). Note that (iii) cannot be omitted, because (ii) does not hold with equality. □

Proposition 4.9. *For a nonempty set $\bar{S} \subseteq \mathbb{R}^n$ the following statements are equivalent:*

(i) $\bar{X}$ *is a convexity solution to* (P),
(ii) $\bar{X}$ *is a solution to* (P).

Proof. This follows from the fact that $P[\bar{X}] + \mathbb{R}^q_+ = \operatorname{Min} P[S] + \mathbb{R}^q_+$ is convex, which is a consequence of Corollary 4.3. □

Proposition 4.10. *A nonempty set $\bar{S} \subseteq \mathbb{R}^n$ is a mild convexity solution to* (P) *if and only if*

(i) $\bar{X} \subseteq S$,
(ii) $P[\bar{X}] \subseteq \operatorname{Min} P[S]$,
(iii) $P[S] \subseteq \operatorname{cl}\operatorname{co}\left(P[\bar{X}] + \mathbb{R}^q_+\right)$.

Proof. The proof is similar to the proof of Proposition 4.7 (even easier). But (iii) cannot be omitted, because (ii) does not hold with equality. □

We proceed with existence results with respect to the solution concepts introduced in the second chapter.

Theorem 4.11. *If* $\operatorname{Eff}(\mathrm{P}) \neq \emptyset$, *then there exists a solution (in the sense of Definition 2.20) to the linear vector optimization problem* (P).

Proof. The nonempty set $\bar{X} := \operatorname{Eff}(\mathrm{P})$ satisfies (i) and (ii) of Proposition 4.7 and is therefore a solution. □

Corollary 4.12. *If* $\operatorname{Eff}(\mathrm{P}) \neq \emptyset$, *then there exist a mild solution, a convexity solution and a mild convexity solution to* (P).

Proof. This is immediate by the above characterizations of the solution concepts and Corollary 4.3. □

We next show the existence of a finitely generated solution to the linear vector optimization problem (P).

Theorem 4.13. *Let* $\operatorname{vert} S \neq \emptyset$ *and* $\operatorname{Eff}(\mathrm{P}) \neq \emptyset$ *and set*

$$\bar{S} := \operatorname{vert} S \cap \operatorname{Eff}(\mathrm{P}),$$

$$\bar{S}^h := \operatorname{extdir} S \cap \operatorname{Eff}(\mathrm{P}^h),$$

where $\operatorname{extdir} S$ *denotes the set of extreme directions of* S*. Then* $(\bar{S}, \bar{S}^h)$ *is a finitely generated solution to* (P)*.*

Proof. Conditions (i) to (iv) of Definition 4.4 are obviously satisfied. It remains to show (v) and $\bar{S} \neq \emptyset$.

Set $\bar{X} := \operatorname{Eff}(\mathrm{P})$ which is nonempty by assumption. Since S is a polyhedron, we obtain that $P[S] + \mathbb{R}^q_+$ is a polyhedron (see e.g. Rockafellar, 1972, Theorem 19.3). By Corollary 4.3, we have $\operatorname{Min} P[S] + \mathbb{R}^q_+ = P[S] + \mathbb{R}^q_+$. Using $P[\bar{X}] = \operatorname{Min} P[S]$, we obtain

$$P[\bar{X}] + \mathbb{R}^q_+ = P[S] + \mathbb{R}^q_+.$$

Let $y \in P[S]$. There exists some $\bar{x} \in \bar{X} = \operatorname{Eff}(\mathrm{P})$ such that $y \in \{P\bar{x}\} + \mathbb{R}^q_+$. By Theorem 4.2, there exists some $w > 0$ such that $\bar{x}$ solves $(\mathrm{P}_1(w))$. Let

$$F := \left\{x \in S \,|\, w^T P\bar{x} = w^T Px\right\}$$

be the solution set of $(\mathrm{P}_1(w))$. Of course, $F \subseteq S$ is a polyhedron. As $\operatorname{vert} S \neq \emptyset$, S contains no lines (follows from Rockafellar, 1972, Theorem 8.3). Hence F contains no lines and by (Rockafellar, 1972, Theorem 18.5) we know that F can be expressed as

$$F = \operatorname{co} \operatorname{vert} F + \operatorname{cone} \operatorname{extdir} F.$$

In particular, we get $\operatorname{vert} F \neq \emptyset$. We next show that

$$\operatorname{vert} F \subseteq \operatorname{vert} S \cap \operatorname{Eff}(\mathrm{P}).$$

Clearly, we have $\operatorname{vert} F \subseteq \operatorname{Eff}(\mathrm{P})$. To show that $\operatorname{vert} F \subseteq \operatorname{vert} S$, let $x \in \operatorname{vert} F \subseteq S$, let $x^1, x^2 \in S$ and $\lambda \in (0,1)$ such that $x = \lambda x^1 + (1-\lambda)x^2$. Assuming that $x^1 \notin F$, we get $w^T Px^1 > w^T P\bar{x}$. Since $x \in F$, we have $w^T P\bar{x} = w^T Px$ and hence $w^T Px^2 < w^T P\bar{x}$. This contradicts $\bar{x}$ being a solution to $(\mathrm{P}_1(w))$. We conclude that $x_1 \in F$. Likewise we obtain $x^2 \in F$. Since $x \in \operatorname{vert} F$, we obtain $x^1 = x = x^2$. This means that $x \in \operatorname{vert} S$. In particular, we see that $\bar{S}$ is nonempty.

Let us show that

$$\operatorname{extdir} F \subseteq \operatorname{extdir} S \cap \operatorname{Eff}(\mathrm{P}^h).$$

By Proposition 4.5, every direction of $F \subseteq \operatorname{Eff}(\mathrm{P})$ belongs to $\operatorname{Eff}(\mathrm{P}^h)$. Let $\hat{x} \in \operatorname{extdir} F \subseteq S_\infty$ and let $\hat{x}^1, \hat{x}^2$ be directions of S such that $\hat{x} = \hat{x}^1 + \hat{x}^2$. Let $x \in F$ be fixed. Assuming that $\hat{x}^1$ is not a direction of F, we obtain some $\mu > 0$ such that $w^T P(x + \mu\hat{x}^1) > w^T P\bar{x}$. Since $\hat{x}$ is a direction of F have $w^T P\bar{x} = w^T P(x + \mu\hat{x})$ and hence $w^T P(x + \mu\hat{x}^2) < w^T P\bar{x}$. Since $x + \mu\hat{x}^2 \in S$, this contradicts $\bar{x}$ being a solution to $(\mathrm{P}_1(w))$. We conclude that $\hat{x}_1$ and likewise $\hat{x}_2$ are directions of F. Since $\hat{x} \in \operatorname{extdir} F$ we obtain some $\alpha > 0$ such that $\hat{x}^1 = \alpha\hat{x}^2$. This means $x \in \operatorname{extdir} S$.

We now conclude that

$$\begin{aligned} y \in \{P\bar{x}\} + \mathbb{R}^q_+ &\subseteq P[F] + \mathbb{R}^q_+ \\ &\subseteq \operatorname{co} P[\operatorname{vert} F] + \operatorname{cone}[\operatorname{extdir} F] + \mathbb{R}^q_+ \\ &\subseteq \operatorname{co} P[\bar{S}] + \operatorname{cone} P[\bar{S}^h] + \mathbb{R}^q_+. \end{aligned}$$

Thus, (v) of Definition 4.4 holds. □

4.3 Set-valued duality

A set-valued dual problem is now assigned to the linear vector optimization problem (P) and duality results are established. We choose a hyperplane-valued dual objective function. In order to determine a hyperplane in $\mathbb{R}^q$, we need the same amount of information as for a vector in $\mathbb{R}^q$, in fact, q real numbers. From this point of view hyperplane-valued functions are not more complicated than vector-valued functions.

We consider the following dual objective map

$$D : \mathbb{R}^m \times \mathbb{R}^q_+ \setminus \{0\} \rightrightarrows \mathbb{R}^q, \quad D(u, w) := \left\{ y \in \mathbb{R}^q |\; w^T y = b^T u \right\}.$$

The following notation is used for a subset T of $\mathbb{R}^m \times \mathbb{R}^q_+ \setminus \{0\}$:

$$D[T] := \{D(u, w) |\; (u, w) \in T\}.$$

Note the important fact that $D[T]$ is a collection of hyperplanes (and not the union). The union is denoted by

$$D(T) := \bigcup_{(u,w) \in T} D(u, w).$$

We use the following ordering relation in $2^{\mathbb{R}^q}$:

$$A^1 \preccurlyeq A^2 \quad :\Longleftrightarrow \quad A^2 \subseteq A^1 + \mathbb{R}^q_+ \quad \Longleftrightarrow \quad A^2 + \mathbb{R}^q_+ \subseteq A^1 + \mathbb{R}^q_+.$$

Note that the ordering $\preccurlyeq$ is an extension of the usual vector ordering in $\mathbb{R}^q$ to the power set, i.e., for singleton sets it reduces to the usual vector ordering

$\leq$ with respect to the cone $\mathbb{R}^q_+$. The following lemma points out the typical usage of the ordering.

Lemma 4.14. *Let be given two hyperplanes* $\bar{H} := \left\{y \in \mathbb{R}^q |\ \bar{w}^T y = \bar{\gamma}\right\}$ $(\bar{w} \neq 0)$ *and* $\hat{H} := \left\{y \in \mathbb{R}^q |\ \hat{w}^T y = \hat{\gamma}\right\}$ $(\hat{w} \neq 0)$. *If* $\bar{w}, \hat{w} \geq 0$ *such that* $e^T \bar{w} = e^T \hat{w} = 1$, *then*

$$\bar{H} \preccurlyeq \hat{H} \quad \iff \quad (\bar{w} = \hat{w} \ \wedge \ \bar{\gamma} \leq \hat{\gamma}) .$$

Proof. Let $\bar{H} \preccurlyeq \hat{H}$, i.e., $\hat{H} + \mathbb{R}^q_+ \subseteq \bar{H} + \mathbb{R}^q_+$. From (4.2) and the fact that $A \subseteq B$ implies $A_\infty \subseteq B_\infty$ for closed convex sets (follows from Rockafellar, 1972, Theorem 8.3), we deduce $\hat{H}_\infty + \mathbb{R}^q_+ \subseteq \bar{H}_\infty + \mathbb{R}^q_+$. Using (Rockafellar, 1972, Corollary 16.4.2), we obtain $(\hat{H}_\infty)^\circ \cap \mathbb{R}^q_- \supseteq (\bar{H}_\infty)^\circ \cap \mathbb{R}^q_-$. As $\bar{w}, \hat{w} \geq 0$, we get $\mathbb{R}_- \hat{w} \supseteq \mathbb{R}_- \bar{w}$. Since $e^T \bar{w} = e^T \hat{w} = 1$, we conclude $\hat{w} = \bar{w}$. The inequality $\hat{\gamma} \geq \bar{\gamma}$ is now immediate.

The opposite implication is obvious. □

Note that, in the situation of the preceding lemma, the ordering relation $\preccurlyeq$ coincides with the ordering $\preccurlyeq$ of the complete lattice $\mathcal{I}$, introduced in Section 1.5.

We consider the following dual problem to (P):

$$\text{maximize } D : \mathbb{R}^m \times \mathbb{R}^q_+ \setminus \{0\} \rightrightarrows \mathbb{R}^q \text{ with respect to } \preccurlyeq \text{ over } T, \qquad \text{(D)}$$

where

$$T := \left\{(u, w) \in \mathbb{R}^m \times \mathbb{R}^q |\ (u, w) \geq 0,\ e^T w = 1,\ B^T u = P^T w\right\} .$$

Let us define a solution concept for the dual problem (D). The upper image $\mathcal{P}$ of the primal problem (P) is a polyhedral set in $\mathbb{R}^q$. We expect the closure of the complement of a polyhedral set to be the dual counterpart. Such a set is also fully determined by a finite number of hyperplanes so that a solution is envisioned to be a finite subset of the feasible set. This set is intended to consist of efficient solutions, which are defined as follows.

Definition 4.15. A feasible vector $\bar{v} \in T$ is called an *efficient solution* to (D) if there is no feasible vector $v \in T$ such that $D(\bar{v}) \preccurlyeq D(v)$ and $D(\bar{v}) \neq D(v)$. The set of efficient solutions to (D) is denoted by Eff (D). Further we write $\operatorname{Max} D[T] := D[\operatorname{Eff}(D)]$ for the set of *maximal elements* in $D[T]$ with respect to $\preccurlyeq$.

Note that this definition is in accordance with Definition 2.1. One should bear in mind that an efficient solution is not regarded to be a solution to (D). It is rather connected to a solution of a scalarized problem. This is shown in Theorem 4.19 below.

The set

$$\mathcal{D} := D(T) - \mathbb{R}^q_+$$

is called the *lower image* of Problem (D). We expect that a solution $\bar{T} \subseteq T$, based on a suitable solution concept, determines the lower image, that is,

$$\mathcal{D} = D(T) - \mathbb{R}^q_+ = D(\bar{T}) - \mathbb{R}^q_+.$$

Following these ideas we introduce a solution concept for (D).

Definition 4.16. A nonempty set $\bar{T} \subseteq \mathbb{R}^{m+q}$ is called a *finitely generated solution* to (D) if

(i) $\bar{T}$ is a finite subset of T,
(ii) $D[\bar{T}] \subseteq \operatorname{Max} D[T]$,
(iii) $D(T) \subseteq D(\bar{T}) - \mathbb{R}^q_+$.

If the context of linear vector optimization is clear, we can simply speak about a *solution* to (D). The relationship to the solution concepts introduced in the second chapter can be found in Section 4.4. To this end it is necessary to work with a suitable complete lattice and to indicate the corresponding infimum and supremum.

Subsequently, we use the following notation, where $\mathcal{P}$ denotes the upper image of (P) as introduced in Section 4.2:

$$F(u,w) := D(u,w) \cap \mathcal{P}.$$

The next statement is a kind of weak duality. The relationship to the scalar case can be seen by the lattice theoretical interpretation in Theorem 4.35 below.

Lemma 4.17. *If* $(u,w) \in T$ *and* $y \in \mathcal{P}$*, then* $w^T y \geq b^T u$*. Moreover, we have*

$$\mathcal{D} \cap \left(\mathcal{P} + \operatorname{int} \mathbb{R}^q_+\right) = \emptyset.$$

Proof. Since $y \in \mathcal{P}$, there is some $x \in S$ such that $y \geq Px$. Hence $(x, 0)$ is feasible for $(\mathrm{P}_2(y))$. Duality between $(\mathrm{P}_2(y))$ and $(\mathrm{D}_2(y))$ implies $b^T u - y^T w \leq 0$.

Assume that there exist $(u,w) \in T$, $z \in D(u,w)$, $d \in \mathbb{R}^q_+$, $c \in \operatorname{int} \mathbb{R}^q_+$ and $y \in \mathcal{P}$ such that $z - d = y + c$. We get $b^T u = w^T z = w^T y + w^T(c+d) > w^T y$, which contradicts the first statement. □

The counterpart to weakly minimal elements of $\mathcal{P}$ are weakly maximal elements of the set $D(T)$.

Lemma 4.18. *If* $(\bar{u},\bar{w}) \in T$ *and* $\bar{y} \in F(\bar{u},\bar{w})$*, then* $\bar{y} \in \operatorname{wMax} D(T)$*.*

Proof. Let $(\bar{u},\bar{w}) \in T$ and $\bar{y} \in F(\bar{u},\bar{w})$. Then, $\bar{y} \in D(T) \cap \mathcal{P}$. We show that $\left(\{\bar{y}\} + \operatorname{int} \mathbb{R}^q_+\right) \cap D(u,w) = \emptyset$ for all $(u,w) \in T$. Assuming the contrary, we obtain $(u,w) \in T$ and $y \in D(u,w)$ such that $y > \bar{y}$. Since $w \gneq 0$, this implies $w^T y > w^T \bar{y} \geq b^T u = w^T y$, a contradiction. □

The following theorem provides different characterizations of efficient solutions to (D).

Theorem 4.19. *Let* $(\bar{u}, \bar{w}) \in T$*. Then the following statements are equivalent:*

(i) $(\bar{u}, \bar{w})$ *is an efficient solution to* (D),
(ii) $\bar{u}$ *solves* $(\mathrm{D}_1(\bar{w}))$,
(iii) *there exists* $\bar{x} \in S$ *such that* $\bar{w}^T P \bar{x} = b^T \bar{u}$,
(iv) $F(\bar{u}, \bar{w})$ *is nonempty,*
(v) $D(\bar{u}, \bar{w}) \cap \mathrm{wMax}\, D(T) \neq \emptyset$.

Proof. (i) $\Rightarrow$ (ii). Let u be feasible for $\mathrm{D}_1(\bar{w})$ and let $b^T \bar{u} \leq b^T u$. Then we have $D(\bar{u}, \bar{w}) \preccurlyeq D(u, \bar{w})$. From (i) we get $b^T \bar{u} = b^T u$, i.e., $\bar{u}$ solves $\mathrm{D}_1(\bar{w})$.

(ii) $\Rightarrow$ (iii). If $\bar{u}$ solves $(\mathrm{D}_1(\bar{w}))$ then by duality between the linear programs $(\mathrm{P}_1(\bar{w}))$ and $(\mathrm{D}_1(\bar{w}))$ there is some $\bar{x} \in S$ such that $\bar{w}^T P \bar{x} = b^T \bar{u}$.

(iii) $\Rightarrow$ (iv). The condition $\bar{w}^T P \bar{x} = b^T \bar{u}$ can be written as $P\bar{x} \in D(\bar{u}, \bar{w})$. Since $P\bar{x} \in \mathcal{P}$, we get $P\bar{x} \in F(\bar{u}, \bar{w})$.

(iv) $\Rightarrow$ (v). This follows from Lemma 4.18.

(v) $\Rightarrow$ (i). Assume that (i) is not true, i.e., there exists $(u, w) \in T$ such that $D(\bar{u}, \bar{w}) \preccurlyeq D(u, w)$ and $D(\bar{u}, \bar{w}) \neq D(u, w)$. This implies $\bar{w} = w$ and $b^T \bar{u} < b^T u$. But for each $y \in D(\bar{u}, \bar{w})$, we get

$$y + e(b^T u - b^T \bar{u}) \in \left(\{y\} + \mathrm{int}\, \mathbb{R}^q_+\right) \cap D(u, \bar{w}),$$

which contradicts (v). □

We continue with a strong duality theorem in the sense that the set of weakly minimal vectors of the upper image $\mathcal{P}$ of (P) and the set of weakly maximal vectors of the set $D(T)$ coincide.

Theorem 4.20. *The following statements are equivalent:*

(i) $\bar{y} \in \mathrm{wMin}\, \mathcal{P}$,
(ii) *there is some* $\bar{x} \in \mathbb{R}^n$ *such that* $(\bar{x}, 0)$ *solves* $(\mathrm{P}_2(\bar{y}))$,
(iii) *there is some* $(\bar{u}, \bar{w}) \in T$ *with* $b^T \bar{u} = \bar{y}^T \bar{w}$ *solving* $(\mathrm{D}_2(\bar{y}))$,
(iv) $\bar{y} \in \mathrm{wMax}\, D(T)$.

Proof. (ii) $\Rightarrow$ (i). If $(\bar{x}, 0)$ solves $(\mathrm{P}_2(\bar{y}))$, then $\bar{x} \in S$ and $P\bar{x} \leq \bar{y}$. Hence $\bar{y} \in \mathcal{P}$. Assume that there is some $y \in \mathcal{P}$ (i.e. there is some $x \in S$ with $Px \leq y$) with $y < \bar{y}$. Then there is some $z < 0$ such that $y \leq \bar{y} + ez$. This implies $Px - ez \leq y - ez \leq \bar{y}$, i.e., (x, z) is feasible for $(\mathrm{P}_2(\bar{y}))$ and $z < 0$ contradicts the optimality of $(\bar{x}, 0)$.

(i) $\Rightarrow$ (ii). If $\bar{y} \in \mathrm{wMin}\, \mathcal{P}$, then there exists some $\bar{x} \in S$ with $P\bar{x} \leq \bar{y}$, i.e., $(\bar{x}, 0)$ is feasible for $(\mathrm{P}_2(\bar{y}))$. Assume that there is some $(x, z) \in \mathbb{R}^{n+1}$ with $z < 0$ being feasible for $(\mathrm{P}_2(\bar{y}))$. Let $y := \bar{y} + ze$. Then $y < \bar{y}$ and $Px \leq \bar{y} + ez = y$, i.e., $y \in \mathcal{P}$. This contradicts the weak minimality of $\bar{y}$.

(ii) $\Leftrightarrow$ (iii). By duality of $(\mathrm{P}_2(\bar{y}))$ and $(\mathrm{D}_2(\bar{y}))$.

(iii) ⇔ (iv). We have $\bar y \in \operatorname{wMax} D(T)$ if and only if

$$\bar y \in D(T) \quad \text{and} \quad \bar y \notin D(T) - \operatorname{int} \mathbb{R}^q_+. \tag{4.3}$$

Condition (4.3) is equivalent to

$$\exists (\bar u, \bar w) \in T : \bar y^T \bar w = b^T \bar u \quad \text{and} \quad \forall (u, w) \in T : \bar y^T w \geq b^T u. \tag{4.4}$$

Since (iii) is equivalent to (4.4), the statement follows. □

A more symmetric variant of the preceding result can be formed, where the set $D(T)$ is replaced by the lower image $\mathcal{D}$ of (D). This is a consequence of the following lemma.

Lemma 4.21. *We have* $\operatorname{wMax} \mathcal{D} = \operatorname{wMax} D(T)$*.*

Proof. We know that

$$y \in \operatorname{wMax} D(T) \quad \Longleftrightarrow \quad \left[y \in D(T) \;\wedge\; y \notin D(T) - \operatorname{int} \mathbb{R}^q_+ \right]$$

and

$$y \in \operatorname{wMax} \left(D(T) - \mathbb{R}^q_+ \right) \quad \Longleftrightarrow \quad \left[y \in D(T) - \mathbb{R}^q_+ \;\wedge\; y \notin D(T) - \operatorname{int} \mathbb{R}^q_+ \right].$$

Thus it remains to show that

$$\left[y \in D(T) - \mathbb{R}^q_+ \;\wedge\; y \notin D(T) - \operatorname{int} \mathbb{R}^q_+ \right] \quad \Longrightarrow \quad y \in D(T).$$

Indeed, $y \notin D(T) - \operatorname{int} \mathbb{R}^q_+$ implies $y^T w \geq b^T u$ for all $(u, w) \in T$ and $y \in D(T) - \mathbb{R}^q_+$ implies the existence of some $(\bar u, \bar w) \in T$ with $y^T \bar w \leq b^T \bar u$. Thus we obtain $y^T \bar w = b^T \bar u$, i.e., $y \in D(T)$. □

We next prove several statements showing the relationship between proper faces (in particular facets) of $\mathcal{P}$ and efficient solutions to (D). Let us recall some facts concerning the facial structure of polyhedral sets. Let $A \subseteq \mathbb{R}^q$ be a convex set. A convex subset $F \subseteq A$ is called a *face* of A if

$$(\bar y, \hat y \in A \quad \wedge \quad \lambda \in (0,1) \quad \wedge \quad \lambda \bar y + (1 - \lambda) \hat y \in F) \quad \Longrightarrow \quad \bar y, \hat y \in F.$$

A face F of A is called *proper* if $\emptyset \neq F \neq A$. A set $E \subseteq A$ is called an *exposed face* of A if there are $w \in \mathbb{R}^q \setminus \{0\}$ and $\gamma \in \mathbb{R}$ such that $A \subseteq \left\{ y \in \mathbb{R}^q | \; w^T y \geq \gamma \right\}$ and $E = \left\{ y \in \mathbb{R}^q | \; w^T y = \gamma \right\} \cap A$. The proper $(r-1)$-dimensional faces of an r-dimensional polyhedral set A are called *facets* of A. A point $y \in A$ is called a *vertex* of A if $\{y\}$ is a face of A. We denote by $\operatorname{ri} A$ the *relative interior* of a convex set A, that is, the interior if A is regarded to be a subset of its affine hull (see e.g. Rockafellar, 1972). The *relative boundary* is the set

$$\operatorname{rbd} A := \operatorname{cl} A \setminus \operatorname{ri} A.$$

Theorem 4.22. *Let A be a polyhedral set in $\mathbb{R}^q$. Then A has a finite number of faces, each of which is exposed and a polyhedral set. Every proper face of A is the intersection of those facets of A that contain it, and* $\operatorname{rbd} A$ *is the union of all the facets of A. If A has a nonempty face of dimension s, then A has faces of all dimensions from s to* $\dim A$.

Proof. See (Webster, 1994, Theorem 3.2.2).

If $\mathcal{P} \neq \emptyset$, then $\mathcal{P}$ is a q-dimensional polyhedral set. Hence the facets of $\mathcal{P}$ are the $(q-1)$-dimensional faces of $\mathcal{P}$, i.e., the maximal (w.r.t. inclusion) proper faces. A subset $F \subseteq \mathcal{P}$ is a proper face if and only if it is a proper exposed face, i.e., F is a proper face if and only if there is a supporting hyperplane H to $\mathcal{P}$ such that $F = H \cap \mathcal{P}$. We call a hyperplane $H := \left\{y \in \mathbb{R}^q \middle|\, w^T y = \gamma\right\}$ $(w \neq 0)$ *supporting* to $\mathcal{P}$ if $w^T y \geq \gamma$ for all $y \in \mathcal{P}$ and there is some $\bar{y} \in \mathcal{P}$ such that $w^T \bar{y} = \gamma$.

Lemma 4.23. *If $H = \left\{y \in \mathbb{R}^q \middle|\, w^T y = \gamma\right\}$ is a supporting hyperplane to $\mathcal{P}$, then $w \gneq 0$.*

Proof. We have $w \neq 0$ since otherwise H is not a hyperplane. If H is a supporting hyperplane to $\mathcal{P}$, then there exists some $\bar{y} \in \mathcal{P}$ such that $w^T \bar{y} = \gamma$ and $w^T y \geq \gamma$ for all $y \in \mathcal{P}$. By definition of $\mathcal{P}$, we have $\bar{y} + z \in \mathcal{P}$ for all $z \in \mathbb{R}^q_+$, hence $w^T z \geq 0$ for all $z \in \mathbb{R}^q_+$. This implies $w \geq 0$. □

The next result states that proper faces of $\mathcal{P}$ are generated by efficient solutions to (D).

Lemma 4.24. *A set $F \subseteq \mathcal{P}$ is a proper face of $\mathcal{P}$ if and only if there is an efficient solution $(u, w) \in T$ to* (D) *such that $F = F(u, w)$.*

Proof. (i) If $(u, w) \in T$ is an efficient solution to (D), then (by Theorem 4.19) there is some $\bar{x} \in S$ such that $P\bar{x} \in D(u, w)$. Hence $P\bar{x} \in F(u, w)$. Moreover, if $y \in \mathcal{P}$ then $w^T y \geq b^T u$, by Lemma 4.17. Consequently, $D(u, w)$ is a supporting hyperplane to $\mathcal{P}$ and $F(u, w)$ is a proper face of $\mathcal{P}$.

(ii) If F is a proper face of $\mathcal{P}$, then there is some $w \in \mathbb{R}^q \setminus \{0\}$ and some $\gamma \in \mathbb{R}$ such that $H := \left\{y \in \mathbb{R}^q \middle|\, w^T y = \gamma\right\}$ is a supporting hyperplane to $\mathcal{P}$ and $F = H \cap \mathcal{P}$. By Lemma 4.23, we have $w \geq 0$. Since $w \neq 0$, we obtain $e^T w > 0$. Without loss of generality we can assume that $e^T w = 1$. Since H is a supporting hyperplane, we have $w^T y \geq \gamma$ for all $y \in \mathcal{P}$ and $w^T \bar{y} = \gamma$ for some $\bar{y} \in \mathcal{P}$. Hence there is some $\bar{x} \in S$ such that $w^T P\bar{x} = w^T \bar{y} = \gamma$, i.e.,

$$\gamma = w^T P\bar{x} = \min\left\{w^T P x \middle|\, x \in S\right\}.$$

By duality between $(\mathrm{P}_1(w))$ and $(\mathrm{D}_1(w))$, Problem $(\mathrm{D}_1(w))$ has a solution u such that $b^T u = \gamma = w^T P\bar{x}$. Thus $(u, w) \in T$ is an efficient solution to (D) by Theorem 4.19, and $D(u, w) = H$. Hence $F = F(u, w)$. □

As a result of strong duality in Theorem 4.20, we obtain that proper faces of $\mathcal{P}$ consist of only weakly minimal elements. Those faces are called *weakly minimal faces.*

Corollary 4.25. *Every proper face of* $\mathcal{P}$ *is weakly minimal.*

Proof. Let F be a proper face of $\mathcal{P}$. By the preceding lemma there is an efficient solution $(u, w) \in T$ to (D) such that $F = F(u, w)$. Let $y \in F = F(u, w)$. Then $y \in \mathcal{P}$ (implying the existence of $x \in S$ such that $Px \leq y$, i.e., $(x, 0)$ is feasible for $(\mathrm{P}_2(y))$) and $b^T u = w^T y$. Duality between $(\mathrm{P}_2(y))$ and $(\mathrm{D}_2(y))$ implies that (u, w) is optimal for $(\mathrm{D}_2(y))$ and $(x, 0)$ is optimal for $(\mathrm{P}_2(y))$. By Theorem 4.20, y is weakly minimal. □

The next result has already been shown in a more general context in Section 1.4, see Corollary 1.44.

Corollary 4.26. $\operatorname{wMin} \mathcal{P} \neq \emptyset$ *if and only if* $\emptyset \neq \mathcal{P} \neq \mathbb{R}^q$.

Proof. This is a direct consequence of Corollary 4.25, Theorem 4.22 and the fact that a nonempty set in $A \subseteq \mathbb{R}^q$ has a nonempty boundary if and only if $A \neq \mathbb{R}^q$. □

The following lemma shows that the set of facets of $\mathcal{P}$ is completely determined by those efficient solutions to (D) that are vertices of the feasible set T.

Lemma 4.27. *If* F *is a facet of* $\mathcal{P}$ *then there is an efficient solution* $(\bar{u}, \bar{w})$ *to* (D) *that is a vertex of* T *such that* $F = F(\bar{u}, \bar{w})$.

Proof. Let

$$\bar{T} := \{(u, w) \in T |\; F(u, w) = F\}.$$

By Theorem 4.19, all points of $\bar{T}$ are efficient solutions to (D) because F is nonempty as a facet of $\mathcal{P}$. Let $y \in \operatorname{ri} F$ be arbitrary. Since F is a $(q-1)$-dimensional face we have $(u, w) \in \bar{T}$ if and only if $(u, w) \in T$ and $y \in D(u, w)$, i.e., $b^T u = y^T w$. Hence $\bar{T} = T \cap D_y$ where

$$D_y := \left\{(u, w) \in \mathbb{R}^m \times \mathbb{R}^q |\; y^T w - b^T u = 0\right\}.$$

Since $y \in \operatorname{wMin} \mathcal{P}$ by Corollary 4.25, Theorem 4.20 implies that D_y is a supporting hyperplane to T, hence $\bar{T}$ is a nonempty face of T. Since $\bar{T} \subseteq T \subseteq \mathbb{R}^{m+q}_+$ contains no lines, there is a vertex $(\bar{u}, \bar{w})$ of $\bar{T}$ (Rockafellar, 1972, Corollary 18.5.3). Hence $(\bar{u}, \bar{w})$ is also a vertex of T. □

Now we can extend the strong duality result in Theorem 4.20. We relate the duality statement to the solution concepts introduced in this chapter. A finitely generated solution of the dual problem is shown to exist. In particular, it is indicated that this solution consists of those vertices of the dual feasible set T that correspond to facets of the upper image $\mathcal{P}$ of the primal problem (P). We denote by $\operatorname{bd} \mathcal{P}$ the boundary of $\mathcal{P}$ and by $\operatorname{vert} T$ the set of all vertices of T.

Theorem 4.28. *For the set* $\bar{T} := \mathrm{Eff}\,(\mathrm{D}) \cap \mathrm{vert}\, T$ *we have the following chain of equalities:*

$$\mathrm{wMin}\,\mathcal{P} = \mathrm{bd}\,\mathcal{P} = F(\bar{T}) = \mathrm{wMax}\, D(\bar{T}) = \mathrm{wMax}\, D(T) = \mathrm{wMax}\,\mathcal{D}. \qquad (4.5)$$

If $S \neq \emptyset$ *and* $T \neq \emptyset$, *then* $\bar{T}$ *is a finitely generated solution to* (D), *in particular, the sets in* (4.5) *are nonempty.*

If $(\bar{S}, \bar{S}^h)$ *is a finitely generated solution to* (P) *and* $\bar{T}$ *is a finitely generated solution to* (D), *then*

$$\mathrm{wMin}\,\mathcal{P} = \mathrm{wMin}\left(\mathrm{co}\, P[\bar{S}] + \mathrm{cone}\, P[\bar{S}^h] + \mathbb{R}^q_+\right) = \mathrm{wMax}\, D(\bar{T}) = \mathrm{wMax}\,\mathcal{D}.$$

Proof. We set

$$\mathcal{F} := \{F \subseteq \mathcal{P} |\ F \text{ is a proper face of } \mathcal{P}\}$$

and

$$\mathcal{G} := \{F \subseteq \mathcal{P} |\ F \text{ is a facet of } \mathcal{P}\}\,.$$

Theorem 4.22, Lemma 4.27, Lemma 4.24 and Corollary 4.25 imply the following chain of inclusions:

$$\begin{aligned}\mathrm{bd}\,\mathcal{P} = \bigcup_{F \in \mathcal{G}} F \subseteq \bigcup_{(u,w)\in \bar{T}} F(u,w) \subseteq \bigcup_{(u,w)\in \mathrm{Eff}\,(\mathrm{D})} F(u,w) \\ = \bigcup_{F \in \mathcal{F}} F \subseteq \mathrm{wMin}\,\mathcal{P} \subseteq \mathrm{bd}\,\mathcal{P}.\end{aligned}$$

Hence the first two equalities of (4.5) hold.

The equalities $\mathrm{wMin}\,\mathcal{P} = \mathrm{wMax}\, D(T) = \mathrm{wMax}\,\mathcal{D}$ have already been proven in Theorem 4.20 and in Lemma 4.21. Thus it remains to show that $F(\bar{T}) = \mathrm{wMax}\, D(\bar{T})$. If $y \in F(\bar{T})$ then there exists some $(\bar{u}, \bar{w}) \in \bar{T} \subseteq T$ such that $y \in F(\bar{u}, \bar{w}) = D(\bar{u}, \bar{w}) \cap \mathcal{P}$, i.e., $y \in D(\bar{T})$. From Lemma 4.18 we obtain $y \in \mathrm{wMax}\, D(T)$ and hence $(\{y\} + \mathrm{int}\,\mathbb{R}^q_+) \cap D(T) = \emptyset$. It follows $(\{y\} + \mathrm{int}\,\mathbb{R}^q_+) \cap D(\bar{T}) = \emptyset$. Together we obtain $y \in \mathrm{wMax}\, D(\bar{T})$.

On the other hand, if $y \in \mathrm{wMax}\, D(\bar{T})$, then $y \in D(\bar{T})$ and $y \notin D(\bar{T}) - \mathrm{int}\,\mathbb{R}^q_+$. This is equivalent to

$$\exists (\bar{u}, \bar{w}) \in \bar{T} : y^T \bar{w} = b^T \bar{u} \qquad (4.6)$$

and

$$\forall (u, w) \in \bar{T} : y^T w \geq b^T u. \qquad (4.7)$$

By Theorem 4.19, $\bar{u}$ solves $(\mathrm{D}_1(\bar{w}))$, hence $S \neq \emptyset$ by duality between $(\mathrm{P}_1(\bar{w}))$ and $(\mathrm{D}_1(\bar{w}))$. Thus the feasible set of $(\mathrm{P}_2(y))$ is nonempty as well. Since $(\bar{u}, \bar{w}) \in T$, i.e., $T \neq \emptyset$, Problem $(\mathrm{D}_2(y))$ has an optimal solution $(\hat{u}, \hat{w})$ that is a vertex of T. Optimality of $(\hat{u}, \hat{w})$ for $(\mathrm{D}_2(y))$ implies optimality of $\hat{u}$ for $(\mathrm{D}_1(\hat{w}))$, hence $(\hat{u}, \hat{w}) \in \bar{T}$ by Theorem 4.19. Now, (4.7) implies that $y^T \hat{w} \geq b^T \hat{u}$. Moreover, optimality of $(\hat{u}, \hat{w})$ for $(\mathrm{D}_2(y))$ implies $b^T \hat{u} - y^T \hat{w} \geq$

$b^T\bar{u} - y^T\bar{w} = 0$, i.e., $y^T\hat{w} = b^T\hat{u}$. Consequently, we have $y \in D(\hat{u}, \hat{w})$ and $y \in \operatorname{wMin} \mathcal{P} \subseteq \mathcal{P}$ by Theorem 4.20, i.e., $y \in F(\bar{T})$.

Let us prove the second assertion. Since $S \neq \emptyset$ and $T \neq \emptyset$, we deduce that $\emptyset \neq \mathcal{P} \neq \mathbb{R}^q$ (e.g. by Lemma 4.17). Hence $\mathcal{P}$ has a facet and by Lemma 4.27, we get $\bar{T} \neq \emptyset$. We conclude that the sets in (4.5) are nonempty. Conditions (i) and (ii) of Definition 4.16 are obviously satisfied. It remains to show that $D(T) \subseteq D(\bar{T}) - \mathbb{R}^q_+$. Let $y \in D(T)$. There exists some $(u, w) \in T$ such that $y \in D(u, w)$, in particular, (u, w) is a feasible vector for $(\mathrm{D}_2(y))$. As S is assumed to be nonempty, there is a feasible vector for $(\mathrm{P}_2(y))$. Hence there is an optimal solution $(\bar{x}, \bar{z})$ to $(\mathrm{P}_2(y))$, where $\bar{z} \geq 0$. It follows that $(\bar{x}, 0)$ is an optimal solution to $(\mathrm{P}_2(y + \bar{z} \cdot e))$, i.e., $y + \bar{z} \cdot e \in \operatorname{wMax} D(T)$, by Theorem 4.20. Hence

$$y \in \operatorname{wMax} D(T) - \mathbb{R}^q_+ = \operatorname{wMax} D(\bar{T}) - \mathbb{R}^q_+ \subseteq D(\bar{T}) - \mathbb{R}^q_+.$$

For the third statement we still have to show the first equality. Taking into account $P[S] \supseteq \operatorname{co} P[\bar{S}] + \operatorname{cone} P[\bar{S}^h]$ and condition (v) in Definition 4.4, we obtain $\mathcal{P} = \operatorname{co} P[\bar{S}] + \operatorname{cone} P[\bar{S}^h] + \mathbb{R}^q_+$, which implies the desired equality. □

4.4 Lattice theoretical interpretation of duality

The more general convex case of the vectorial duality theory was developed on the basis of the complete lattice $\mathcal{I}$ of self-infimal sets as introduced in Section 1.5. The goal of this section is a corresponding approach to the linear case. To this end we reformulate the above results using the infimum and supremum in the complete lattice $(\mathcal{I}, \preccurlyeq)$. We also relate the notion of a finitely generated solution of (D) to the solution concepts developed in Chapter 2 for general complete-lattice-valued problems. This section is based on several concepts and results of the first part of this book. The section can be skipped if the reader is interested in a self-contained theory for linear problems.

Subsequently, we use the complete lattice $\mathcal{I} = \mathcal{I}_C\left(\overline{Y}\right)$ for the sets $Y = \mathbb{R}^q$ and $C = \mathbb{R}^q_+$, compare Section 1.5. Based on the primal problem (P), we consider the $\mathcal{I}$-valued objective function

$$p : \mathbb{R}^n \to \mathcal{I}, \qquad p(x) := \operatorname{Inf}\{Px\} = \{Px\} + \operatorname{bd} \mathbb{R}^q_+.$$

Of course, the ordering in $\mathcal{I}$ is just an extension of the vector ordering. Thus, for all $x^1, x^2 \in \mathbb{R}^n$, we have

$$Px^1 \leq Px^2 \quad \iff \quad p(x^1) \preccurlyeq p(x^2).$$

This yields

$$\operatorname{Eff}(\mathrm{P}) = \operatorname{Eff}(\mathrm{P}').$$

By Theorem 1.54, we obtain

$$\inf_{x\in S} p(x) = \operatorname{Inf} \bigcup_{x\in S} \operatorname{Inf}\{Px\} = \operatorname{Inf} \bigcup_{x\in S} \{Px\} = \operatorname{Inf} P[S],$$

where the second equality is follows easily using Corollaries 1.48 and 1.49. It follows from Definition 1.45 and Proposition 1.40 that $\operatorname{wMin}\mathcal{P} = \operatorname{Inf} P[S]$ if and only if $S \neq \emptyset$ and $\mathcal{P} \neq \mathbb{R}^q$ and hence

$$\inf_{x\in S} p(x) = \begin{cases} \{+\infty\} & \text{if} \quad S = \emptyset \\ \{-\infty\} & \text{if} \quad \mathcal{P} = \mathbb{R}^q \\ \operatorname{wMin}\mathcal{P} & \text{otherwise.} \end{cases}$$

Note that, by Corollary 4.26, $\operatorname{wMin}\mathcal{P} \neq \emptyset$ if and only if $\emptyset \neq \mathcal{P} \neq \mathbb{R}^q$. This means, if the set $\operatorname{wMin}\mathcal{P}$ is nonempty, it coincides with $\inf_{x\in S} p(x)$. Otherwise if $\operatorname{wMin}\mathcal{P}$ is empty, we distinguish between two cases: $\inf_{x\in S} p(x) = \{+\infty\}$ if $S = \emptyset$ and $\inf_{x\in S} p(x) = \{+\infty\}$ if $S \neq \emptyset$. Thus, (P) is closely related to the following $\mathcal{I}$-valued problem:

$$\text{minimize } p : \mathbb{R}^n \to \mathcal{I} \text{ with respect to } \preccurlyeq \text{ over } S. \tag{P$'$}$$

Moreover, it is easy to see that $\bar{x} \in S$ is an efficient solution to (P) if and only if it is an efficient solution to (P$'$), that is

$$\big[x \in S,\ p(x) \preccurlyeq p(\bar{x})\big] \quad \Longrightarrow \quad p(x) = p(\bar{x}).$$

Let us introduce a solution concept for the $\mathcal{I}$-valued problem (P$'$) taking into account its linear (or polyhedral) structure.

Definition 4.29. A nonempty set $\bar{S} \subseteq \mathbb{R}^n$ together with a (possibly empty) set $\bar{S}^h \subseteq \mathbb{R}^n \setminus \{0\}$ is said to be *finitely generated solution* to (P$'$) if

(i) $\bar{S}$ is a finite subset of S,
(ii) $\bar{S}^h$ is a finite subset of S^h,
(iii) $p[\bar{S}] \subseteq \operatorname{Min} p[S]$,
(iv) $p[\bar{S}^h] \subseteq \operatorname{Min} p[S^h]$,
(v) $\inf_{x\in S} p(x) = \inf_{x\in \operatorname{co}\bar{S}} p(x) \oplus \inf_{x\in \operatorname{cone}\bar{S}^h} p(x)$.

We show that the problems (P) and (P$'$) are equivalent in the sense that they have the same solutions.

Proposition 4.30. *Consider* (P) *and* (P$'$) *based on the same data. The following statements are equivalent:*

(i) $(\bar{S}, \bar{S}^h)$ *is a finitely generated solution to* (P),
(ii) $(\bar{S}, \bar{S}^h)$ *is a finitely generated solution to* (P$'$).

Proof. Let us denote the five conditions of Definition 4.4 by (i)°, ... , (v)°, whereas the five condition of Definition 4.29 are denoted by (i)′, ... , (v)′. It is clear that (i)° to (iv)° are satisfied if and only if (i)′ to (iv)′ hold.

Note first that (i)° and (ii)° imply that (v)° is equivalent to

$$P[S] + \mathbb{R}^q_+ = \operatorname{co} P[\bar{S}] + \operatorname{cone} P[\bar{S}^h] + \mathbb{R}^q_+.$$

By Proposition 1.40 and the fact that the sum of polyhedral sets is closed, this is equivalent to

$$\operatorname{Cl}_+ P[S] = \operatorname{Cl}_+ \left(\operatorname{co} P[\bar{S}] + \operatorname{cone} P[\bar{S}^h]\right). \tag{4.8}$$

We have

$$\operatorname{co} P[\bar{S}] = P[\operatorname{co} \bar{S}] \quad \text{and} \quad \operatorname{cone} P[\bar{S}^h] = P[\operatorname{cone} \bar{S}^h].$$

Using Corollary 1.49, we obtain that (4.8) holds if and only if

$$\operatorname{Inf} P[S] = \operatorname{Inf} \left(\operatorname{Inf} P[\operatorname{co} \bar{S}] + \operatorname{Inf} P[\bar{S}^h]\right). \tag{4.9}$$

For arbitrary nonempty sets $A \subseteq \mathbb{R}^n$, we have

$$\operatorname{Inf} P[A] = \operatorname{Inf} \bigcup_{x \in A} \{Px\} = \operatorname{Inf} \bigcup_{x \in A} \operatorname{Inf} \{Px\} = \operatorname{Inf} p(A).$$

Theorem 1.54 implies that

$$\operatorname{Inf} p(A) = \inf_{x \in A} p(x).$$

Hence, (4.9) is equivalent to (v)′. □

Remark 4.31. Note that in Definition 4.29, (v) can be replaced by

(v′) $\inf_{x \in S} p(x) = \inf_{x \in \bar{S}} p(x) \oplus \inf_{x \in \bar{S}^h} p(x)$

if the infimum is taken with respect to the complete lattice $\mathcal{I}_{\text{co}}$ as introduced in Section 1.6.

We next reformulate the dual problem (D) from Section 4.3 by using the supremum in $\mathcal{I}$. We start with two lemmas.

Lemma 4.32. *The set $D(T) - \mathbb{R}^q_+$ is closed.*

Proof. Let $\left(y^i\right)_{i \in \mathbb{N}}$ be a sequence in $D(T) - \mathbb{R}^q_+$ converging to $\bar{y} \in \mathbb{R}^q$. For each i there exists some $(u^i, w^i) \in T$ such that $y^i \in D(u^i, w^i) - \mathbb{R}^q_+$. Hence ${y^i}^T w^i \leq b^T u^i$. We have to show that there exists some $(\bar{u}, \bar{w}) \in T$ with $\bar{y}^T \bar{w} \leq b^T \bar{u}$.

Assuming the contrary, we get $\bar{y}^T w - b^T u > 0$ for all $(u, w) \in T$. Since T is polyhedral, there is some $\gamma > 0$ such that $\bar{y}^T w - b^T u \geq \gamma$ for all $(u, w) \in T$.

Let us denote by $\|\cdot\|_\infty$ and $\|\cdot\|_1$ the maximum norm and the sum norm in $\mathbb{R}^q$, respectively. Take $j \in \mathbb{N}$ such that $\|y^j - \bar{y}\|_\infty < \gamma$, then

$$(\bar{y} - y^j)^T w^j \leq \|y^j - \bar{y}\|_\infty \|w^j\|_1 < \gamma$$

hence

$$\bar{y}^T w^j - b^T u^j < {y^j}^T w^j + \gamma - b^T u^j \leq \gamma,$$

a contradiction. □

We now define an $\mathcal{I}$-valued dual objective function as

$$d : \mathbb{R}^m \times \mathbb{R}^q \setminus \{0\} \to \mathcal{I}, \qquad d(u,w) := \operatorname{Inf} \left\{ y \in \mathbb{R}^q |\; w^T y \geq b^T u \right\}.$$

This function is closely related to the hyperplane-valued objective map $D : \mathbb{R}^m \times \mathbb{R}^q \setminus \{0\} \rightrightarrows \mathbb{R}^q$ in Problem (D).

Lemma 4.33. *The dual objective function $d : \mathbb{R}^m \times \mathbb{R}^q \setminus \{0\} \to \mathcal{I}$ can be expressed as*

$$d(u,w) = \begin{cases} D(u,w) & \text{if} \quad w \gneq 0 \\ \{-\infty\} & \text{if} \quad w \not\gneq 0. \end{cases}$$

Proof. Since $w \neq 0$, we have

$$d(u,w) = \frac{b^T u}{w^T w} w + \operatorname{Inf} \left\{ y \in \mathbb{R}^q |\; w^T y \geq 0 \right\}.$$

The result now follows from Proposition 3.8. □

The following dual problem of (P′) is considered:

$$\text{maximize } d : \mathbb{R}^m \times \mathbb{R}^q \setminus \{0\} \to \mathcal{I} \text{ with respect to } \preccurlyeq \text{ over } T. \tag{D′}$$

For arbitrary $(u^1, w^1), (u^2, w^2) \in T$ we have

$$D(u^1,w^1) \preccurlyeq D(u^2,w^2) \quad \Longleftrightarrow \quad d(u^1,w^1) \preccurlyeq d(u^2,w^2),$$

where $\preccurlyeq$ on the left stands for the ordering introduced in Section 4.3 and $\preccurlyeq$ on the right is the ordering of the complete lattice $\mathcal{I}$. Both ordering relations coincide on the set $\{D(u,w)|\; (u,w) \in T\}$. Taking into account Lemma 4.33, we get

$$\operatorname{Eff}(\mathrm{D}) = \operatorname{Eff}(\mathrm{D}'). \tag{4.10}$$

The next lemma points out a further relationship between (D) and (D′).

Lemma 4.34. *The optimal value of* (D′) *can be expressed as*

$$\sup_{(u,w)\in T} d(u,w) = \begin{cases} \{-\infty\} & \text{if} \quad T = \emptyset \\ \{+\infty\} & \text{if} \quad D(T) - \mathbb{R}^q_+ = \mathbb{R}^q \\ \operatorname{wMax} D(T) & \text{otherwise.} \end{cases}$$

Proof. (i) If $T = \emptyset$, we have $\sup_{(u,w)\in T} d(u,w) = \{-\infty\}$, by definition.

(ii) If $D(T) - \mathbb{R}^q_+ = \mathbb{R}^q$, then $\operatorname{Cl}_- D(T) \supseteq D(T) - \mathbb{R}^q_+ = \mathbb{R}^q$. From the definition of the supremal set, which is analogous to Definition 1.45 (see also (1.11)), and from Theorem 1.54, we get $\sup_{(u,w)\in T} d(u,w) = \{+\infty\}$.

(iii) Let $T \neq \emptyset$ and $D(T) - \mathbb{R}^q_+ \neq \mathbb{R}^q$. By Lemma 4.32 we have $\emptyset \neq \operatorname{Cl}_- D(T) = D(T) - \mathbb{R}^q_+ \neq \mathbb{R}^q$. Lemma 4.21 yields that

$$\operatorname{Sup} D(T) = \operatorname{wMax} \operatorname{Cl}_- D(T) = \operatorname{wMax}(D(T) - \mathbb{R}^q_+) = \operatorname{wMax} D(T).$$

By Theorem 1.54 and Lemma 4.33, we have

$$\sup_{(u,w)\in T} d(u,w) = \operatorname{Sup} D(T).$$

Together, we obtain $\sup_{(u,w)\in T} d(u,w) = \operatorname{wMax} D(T)$. □

We continue with a reformulation of the duality results for the linear problem (P) and its dual problem (D). We obtain duality results for (P′) and (D′). Let us begin with weak duality.

Theorem 4.35 (weak duality). *Let $x \in S$ and $(u,w) \in T$, then*

$$d(u,w) \preccurlyeq p(x).$$

Proof. For all $y \in p(x) = \{Px\} + \operatorname{bd} \mathbb{R}^q_+ \subseteq \mathcal{P}$, Lemma 4.17 yields $y^T w \geq b^T u$, hence $p(x) \subseteq d(u,w) + \mathbb{R}^q_+$. This implies $d(u,w) \preccurlyeq p(x)$. □

The next result shows strong duality between (P′) and (D′). The distinction of the three cases is well-known from scalar linear programming.

Theorem 4.36 (strong duality). *Let at least one of the sets S and T be nonempty. Then strong duality holds between* (P′) *and* (D′)*, that is,*

$$\bar{d} := \sup_{(u,w)\in T} d(u,w) = \inf_{x\in S} p(x) =: \bar{p}.$$

Moreover, the following statements are true:

(i) *If $S \neq \emptyset$ and $T \neq \emptyset$, then $\{-\infty\} \neq \bar{d} = \bar{p} \neq \{+\infty\}$ and*

$$\bar{d} = \operatorname{wMax} D(T) = \operatorname{wMin} \mathcal{P} = \bar{p} \neq \emptyset.$$

(ii) *If $S = \emptyset$ and $T \neq \emptyset$, then $\bar{d} = \bar{p} = \{+\infty\}$.*
(iii) *If $S \neq \emptyset$ and $T = \emptyset$, then $\bar{d} = \bar{p} = \{-\infty\}$.*

Proof. By the weak duality we have

$$\sup_{(u,w)\in T} d(u,w) \preccurlyeq \inf_{x\in S} p(x).$$

(i) If $S \neq \emptyset$ and $T \neq \emptyset$, neither $\sup_{(u,w)\in T} d(u,w)$ nor $\inf_{x\in S} p(x)$ can be $\{-\infty\}$ or $\{+\infty\}$ (by weak duality). Hence, Theorem 4.20 and Lemma 4.34 imply

$$\sup_{(u,w)\in T} d(u,w) = \operatorname{wMax} D(T) = \operatorname{wMin} \mathcal{P} = \inf_{x\in S} p(x).$$

(ii) If $S = \emptyset$ and $T \neq \emptyset$, we have $\inf_{x\in S} p(x) = \{+\infty\}$. Theorem 4.20 implies that

$$\operatorname{wMax} D(T) = \operatorname{wMin} \mathcal{P} = \emptyset.$$

Moreover, we have $D(T) - \mathbb{R}^q_+ = \mathbb{R}^q$. Assuming the contrary, there exists $y \notin D(T) - \mathbb{R}^q_+$, i.e.,

$$\forall (u,w) \in T: \quad b^T u - y^T w < 0. \tag{4.11}$$

Problem $(\mathrm{D}_2(y))$ is feasible but $(\mathrm{P}_2(y))$ is not. From the scalar duality theory we conclude that $(\mathrm{D}_2(y))$ is unbounded, which contradicts (4.11). Lemma 4.34 yields $\sup_{(u,w)\in T} d(u,w) = \{+\infty\}$.

(iii) If $S \neq \emptyset$ and $T = \emptyset$, we have $\sup_{(u,w)\in T} d(u,w) = \{-\infty\}$. Theorem 4.20 implies

$$\operatorname{wMin} \mathcal{P} = \operatorname{wMax} D(T) = \emptyset.$$

Since $S \neq \emptyset$, Corollary 4.26 implies $\mathcal{P} = \mathbb{R}^q$. Hence $\inf_{x\in S} p(x) = \{-\infty\}$. □

Let us define finitely generated solutions to (D′). Note that condition (iii) corresponds to the attainment of the supremum as discussed in Chapter 2.

Definition 4.37. A nonempty set $\bar{T}$ is called a *finitely generated solution* to (D′) if

(i) $\bar{T}$ is a finite subset of T,

(ii) $d[\bar{T}] \subseteq \operatorname{Max} d[T]$,

(iii) $\sup_{(u,w)\in T} d(u,w) = \sup_{(u,w)\in \bar{T}} d(u,w)$.

Proposition 4.38. *Consider the problems* (D) *and* (D′) *based on the same data. The following statements are equivalent:*

(i) $\bar{T}$ *is a finitely generated solution to* (D),

(ii) $\bar{T}$ *is a finitely generated solution to* (D′).

Proof. Let us denote the three conditions of Definition 4.16 by (i)°, (ii)° and (iii)°, whereas the three condition of Definition 4.37 are denoted by (i)′, (ii)′ and (iii)′.

From Lemma 4.33 and (4.10) we deduce that (i)° and (ii)° are equivalent to (i)′ and (ii)′.

In the following steps we assume that (i)° and (i)′ hold. Then, condition (iii)° is equivalent to

$$d(T) - \mathbb{R}^q_+ = d(\bar{T}) - \mathbb{R}^q_+. \tag{4.12}$$

By Lemma 4.32, we know that $d(T) - \mathbb{R}^q_+$ is closed. Since $\bar{T}$ is finite, the set $d(\bar{T}) - \mathbb{R}^q_+$ is closed, too. Hence (4.12) is equivalent to

$$\text{cl}\,\left(d(T) - \mathbb{R}^q_+\right) = \text{cl}\,\left(d(\bar{T}) - \mathbb{R}^q_+\right). \tag{4.13}$$

By an analogous statements to Proposition 1.40 and by Lemma 4.33, (4.13) is equivalent to

$$\text{Cl}_{\,-}D(T) = \text{Cl}_{\,-}D(\bar{T}).$$

A counterpart of Corollary 1.48 (vi) (see also (1.11)) yields that the last statements holds if and only if

$$\text{Sup}\,D(T) = \text{Sup}\,D(\bar{T}).$$

By Theorem 1.54, this is equivalent to (iii)$'$. □

The existence of a finitely generated solution to (D$'$) (in the case where $S \neq \emptyset$ and $T \neq \emptyset$) follows from Theorem 4.28 and the latter result.

In Chapter 2 we introduced several solution concepts for the $\mathcal{I}$-valued problem (D$'$). Let us point out a characterization of a solution to (D$'$), as defined in Definition 2.53, because this notion has a simple structure in the present setting.

Proposition 4.39. *A nonempty set $\bar{T} \subseteq \mathbb{R}^m \times \mathbb{R}^q$ is a solution to* (D$'$) *(in the sense of Definition 2.53) if and only if*

(i) $\bar{T} \subseteq T$,
(ii) $d[\bar{T}] = \text{Max}\, d[T]$.

Proof. It remains to show that (i) and (ii) imply that the supremum of the canonical extension of $d : \mathbb{R}^m \times \mathbb{R}^q \to \mathcal{I}$ over T is attained in $\bar{T}$. In terms of d this means

$$\sup_{(u,w)\in T} d(u,w) = \sup_{(u,w)\in \bar{T}} d(u,w).$$

From (i) and (ii) we get $d[\bar{T}] = d[\text{Eff}\,(\text{D}')]$ and hence

$$\sup_{(u,w)\in \bar{T}} d(u,w) = \sup_{(u,w)\in \text{Eff}\,(\text{D}')} d(u,w).$$

Since Eff (D) = Eff (D$'$) $\neq \emptyset$, Theorem 4.19 implies $S \neq 0$. Furthermore, we have

$$\begin{aligned}\sup_{(u,w)\in \text{Eff}\,(\text{D}')\cap \text{vert}\, T} d(u,w) &\preccurlyeq \sup_{(u,w)\in \text{Eff}\,(\text{D}')} d(u,w)\\ &= \sup_{(u,w)\in \bar{T}} d(u,w) \preccurlyeq \sup_{(u,w)\in T} d(u,w).\end{aligned}$$

By Theorem 4.28, $\tilde{T} := \text{Eff}\,(\text{D}) \cap \text{vert}\, T$ is a finitely generated solution to (D). From Proposition 4.38, we deduce that $\tilde{T}$ is a finitely generated solution to

(D′). Using condition (iii) in Definition 4.37, we obtain that the last statement holds with equality. □

Propositions 4.7 and 4.39 show that, in the linear case, Eff (P′) and Eff (D′) always provide solutions to (P′) and (D′), respectively, whenever these sets are nonempty. This means that the attainment of infimum is automatically satisfied in the linear case. This fact is well-known from scalar linear programming.

4.5 Geometric duality

The dual problem (D) of the linear vector optimization problem (P) is based on a hyperplane-valued objective function. We already mentioned that a hyperplane in $\mathbb{R}^q$ is a kind of dual object to a vector in $\mathbb{R}^q$ carrying the same amount of information in the sense that both objects are determined by q real numbers. It is natural to ask in which sense the dual objective is linear. A more general concern is whether the hyperplane-valued dual problem is equivalent to a vector optimization problem. This question is the subject of this section.

The idea of such a connection between vector-valued and hyperplane-valued problems is taken from the classical theory of polytopes. It is well-known from the theory of convex polytopes (see e.g. Grünbaum, 2003) that two polytopes $\mathcal{P}$ and $\mathcal{P}^*$ in $\mathbb{R}^q$ are said to be dual to each other provided there exists a one-to-one mapping Ψ between the set of all faces of $\mathcal{P}$ and the set of all faces of $\mathcal{P}^*$ such that Ψ is inclusion reversing, i.e., faces F^1 and F^2 of $\mathcal{P}$ satisfy $F^1 \subseteq F^2$ if and only if the faces $\Psi(F^1)$ and $\Psi(F^2)$ satisfy $\Psi(F^1) \supseteq \Psi(F^2)$ (see Figure 4.1).

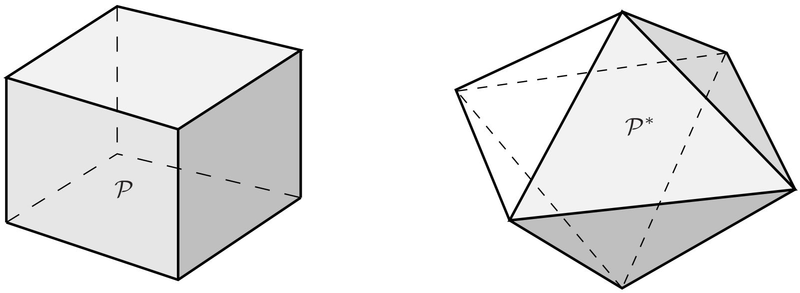

Fig. 4.1 Example of a pair of dual polytopes in $\mathbb{R}^3$.

Instead of speaking about strong duality if the optimal values of a pair of dual optimization problems are equal, we deal with a duality relation between the polyhedral image set of the primal problem and the polyhedral image of the dual problem. This relation is similar to duality of polytopes. Denoting by $\mathcal{P}$ and $\mathcal{D}^*$ the (slightly modified) images of the objective functions of our given problem (P) and a suitable dual problem (D*), respectively, we show that there is an inclusion reversing one-to-one map Ψ between the set of all K-maximal proper faces of $\mathcal{D}^*$ (i.e. proper faces of $\mathcal{D}^*$ that only contain K-maximal elements of $\mathcal{D}^*$) and the set of all weakly minimal proper faces of $\mathcal{P}$ (i.e. proper faces of $\mathcal{P}$ that only contain weakly minimal elements of $\mathcal{P}$), where K is an appropriate ordering cone for the dual problem. With the aid of such a map Ψ we can compute the weakly minimal faces of $\mathcal{P}$ whenever we know the K-maximal faces of $\mathcal{D}^*$ and vice versa. In particular, we are given by Ψ a one-to-one correspondence between weakly minimal vertices (facets) of $\mathcal{P}$ and K-maximal facets (vertices) of $\mathcal{D}^*$.

In the following we consider two special ordering cones, in fact, the cone $\mathbb{R}^q_+$ for the primal problem and the cone

$$K := \mathbb{R}_+ \cdot (0,0,\dots,0,1)^T = \{y \in \mathbb{R}^q |\; y_1 = \dots = y_{q-1} = 0, y_q \geq 0\}$$

for the dual problem. We write

$$\operatorname{Max}_K A := \{y \in A |\; (\{y\} + K \setminus \{0\}) \cap A = \emptyset\}$$

for the set of maximal vectors of a set $A \subseteq \mathbb{R}^q$ with respect to the cone K. The counterpart in the primal problem is the set

$$\operatorname{wMin} A = \left\{y \in A |\; (\{y\} - \operatorname{int} \mathbb{R}^q_+) \cap A = \emptyset\right\}$$

of weakly minimal vectors of A. Note that an index is only used for cones apart from $\mathbb{R}^q_+$.

Using the relative interior of the ordering cones $\mathbb{R}^q_+$ and K, we can subsume the notions $\operatorname{wMin} A$ and $\operatorname{Max}_K A$ under the common notation

$$\operatorname{Min}_C A := \{y \in A |\; (\{y\} - \operatorname{ri} C) \cap A = \emptyset\} \quad \text{and} \quad \operatorname{Max}_C A := \operatorname{Min}_{(-C)} A.$$

In case of $C = K$ we have $\operatorname{ri} K = K \setminus \{0\}$ and for the choice $C = \mathbb{R}^q_+$ we have $\operatorname{ri} C = \operatorname{int} \mathbb{R}^q_+$, which yields the concepts introduced above.

We next assign to the linear vector optimization problem (P) a *geometric dual problem*. In contrast to the dual problem (D) from Section 4.3 with a hyperplane-valued objective function, we now consider a linear vector-valued dual objective function

$$D^* : \mathbb{R}^m \times \mathbb{R}^q \to \mathbb{R}^q, \quad D^*(u,w) := \left(w_1, ..., w_{q-1}, b^T u\right)^T.$$

The geometric dual problem to (P) is

$$\text{maximize } D^* : \mathbb{R}^m \times \mathbb{R}^q \to \mathbb{R}^q \text{ with respect to } \leq_K \text{ over } T. \tag{D*}$$

where we use the same feasible set T as in Problem (D), that is,

$$T := \left\{(u,w) \in \mathbb{R}^m \times \mathbb{R}^q |\ (u,w) \geq 0,\ B^T u = P^T w,\ e^T w = 1\right\},$$

where $e := (1,\dots,1)^T$. Even though both problems (P) and (D*) are vectorial linear programs, they are not symmetric. This feature, however, is shared by all attempts to duality for vector optimization problems. The special choice of the cone K for the dual problem reflects a parametric character of the dual problem. In fact, a point $D^*(\bar{u},\bar{w})$ is a K-maximal point of $D^*[T]$ if and only if for fixed $\bar{w}$, the point $\bar{u}$ maximizes $b^T u$ over the set $\{u \in \mathbb{R}^m |\ (u,\bar{w}) \in T\}$.

We aim to show a duality relation between the sets

$$\begin{aligned} \mathcal{P} &:= P[S] + \mathbb{R}^q_+ = \{y \in \mathbb{R}^q | \exists x \in S : y \in \{Px\} + \mathbb{R}^q_+\} \quad \text{and} \\ \mathcal{D}^* &:= D^*[T] - K = \{y^* \in \mathbb{R}^q | \exists (u,w) \in T : y^* \in \{D^*(u,w)\} - K\}. \end{aligned}$$

To this end we construct an inclusion reversing one-to-one map Ψ between the K-maximal proper faces of $\mathcal{D}^*$ and the weakly minimal proper faces of $\mathcal{P}$.

We consider the coupling function $\varphi : \mathbb{R}^q \times \mathbb{R}^q \to \mathbb{R}$, defined by

$$\varphi(y,y^*) := \sum_{i=1}^{q-1} y_i y_i^* + y_q \left(1 - \sum_{i=1}^{q-1} y_i^*\right) - y_q^*.$$

Note that $\varphi(\,\cdot\,,y^*)$ and $\varphi(y,\,\cdot\,)$ are affine functions. Choosing the values of the primal and dual objective function as arguments, we obtain

$$\varphi(Px, D^*(u,w)) = w^T Px - b^T u. \tag{4.14}$$

Throughout, we use the notation

$$w(y^*) := \left(y_1^*,\dots,y_{q-1}^*, 1 - \sum_{i=1}^{q-1} y_i^*\right)^T \tag{4.15}$$

and

$$w^*(y) := \left(y_1 - y_q, \dots, y_{q-1} - y_q, -1\right)^T. \tag{4.16}$$

We can write

$$\varphi(y,y^*) = w(y^*)^T y - y_q^* = w^*(y)^T y^* + y_q.$$

The following assertion can be understood as a weak duality result.

Theorem 4.40. *The following implication is true:*

$$[y \in \mathcal{P} \wedge y^* \in \mathcal{D}^*] \implies \varphi(y,y^*) \geq 0.$$

Proof. Let $y \in \mathcal{P}$. There exists some $x \in S$ such that $y \geq Px$. Let $y^* \in \mathcal{D}^*$. We obtain some $(u, w) \in T$ such that $y_1^* = w_1, \dots, y_{q-1}^* = w_{q-1}$ and $y_q^* \leq b^T u$. Using the weak duality between $(\mathrm{P}_1(w))$ and $(\mathrm{D}_1(w))$, and the fact $w \geq 0$, we get

$$w^T y \geq w^T P x \geq b^T u \geq y_q^*.$$

Taking into account that $e^T w = 1$, we obtain

$$\varphi(y, y^*) = w^T y - y_q^* \geq 0,$$

which completes the proof. □

The next statement is a first strong duality result.

Theorem 4.41. *Let the feasible sets S of* (P) *and T of* (D*) *be nonempty. Then*

$$[\forall y^* \in \mathcal{D}^* : \varphi(y, y^*) \geq 0] \implies y \in \mathcal{P},$$
$$[\forall y \in \mathcal{P} : \varphi(y, y^*) \geq 0] \implies y^* \in \mathcal{D}^*.$$

If we set

$$\Phi^*(A) := \{y^* \in \mathbb{R}^q |\ \forall y \in A :\ \varphi(y, y^*) \geq 0\},$$

where $A \subseteq \mathbb{R}^q$ is an arbitrary set, and

$$\Phi(A^*) := \{y \in \mathbb{R}^q |\ \forall y^* \in A^* :\ \varphi(y, y^*) \geq 0\},$$

where $A^* \subseteq \mathbb{R}^q$, the following statements hold:

$$A \subseteq \mathcal{P} \iff \Phi^*(A) \supseteq \mathcal{D}^*,$$

$$A^* \subseteq \mathcal{D}^* \iff \Phi(A^*) \supseteq \mathcal{P}.$$

Proof. Let $y \in \mathbb{R}^q$ such that $\varphi(y, y^*) \geq 0$ for all $y^* \in \mathcal{D}^*$. It follows that $(\mathrm{D}_2(y))$ is feasible and bounded. Hence there exist solutions (u, w) to $(\mathrm{D}_2(y))$ and (x, z) to $(\mathrm{P}_2(y))$. We have $y^* := D^*(u, w) \in \mathcal{D}^*$. From $\varphi(y, y^*) \geq 0$ we conclude $w^T y \geq b^T u$. Strong duality between $(\mathrm{P}_2(y))$ and $(\mathrm{D}_2(y))$ yields $z \leq 0$ and hence $Px \leq y + ze \leq y$, i.e., $y \in \mathcal{P}$. This proves the first implication.

To prove the second one, let $y^* \in \mathbb{R}^q$ be given such that $\varphi(y, y^*) \geq 0$ for all $y \in \mathcal{P}$. Consequently, $(\mathrm{P}_1(w(y^*)))$ is feasible and bounded. There exist solutions x to $(\mathrm{P}_1(w(y^*)))$ and u to $(\mathrm{D}_1(w(y^*)))$. We have $y := Px \in \mathcal{P}$. From $\varphi(y, y^*) \geq 0$ we conclude $w(y^*)^T P x \geq y_q^*$. Strong duality between $(\mathrm{P}_1(w(y^*)))$ and $(\mathrm{D}_1(w(y^*)))$ implies that $b^T u \geq y_q^*$, whence $y^* \in \mathcal{D}^*$.

The other results are consequences of the first two implications and Theorem 4.40 (weak duality). □

Next we shall show a strong duality result which involves the facial structure of $\mathcal{P}$ and $\mathcal{D}^*$. The coupling function φ is used to define the following two set-valued maps

$$H : \mathbb{R}^q \rightrightarrows \mathbb{R}^q, \quad H(y^*) := \{y \in \mathbb{R}^q |\ \varphi(y, y^*) = 0\},$$

$$H^* : \mathbb{R}^q \rightrightarrows \mathbb{R}^q, \quad H^*(y) := \{y^* \in \mathbb{R}^q |\ \varphi(y, y^*) = 0\}.$$

Of course, $H(y^*)$ and $H^*(y)$ are hyperplanes in $\mathbb{R}^q$ for all $y^*, y \in \mathbb{R}^q$. Using (4.15) and (4.16) we obtain the expressions

$$H(y^*) = \left\{y \in \mathbb{R}^q |\ w(y^*)^T y = y_q^*\right\}$$

and

$$H^*(y) = \left\{y^* \in \mathbb{R}^q |\ w^*(y)^T y^* = -y_q\right\}.$$

Obviously, the set-valued maps H and H^* are injective. Moreover, we have

$$y^* \in H^*(y) \iff y \in H(y^*). \tag{4.17}$$

The dual problem (D^*) is related to the hyperplane-valued dual problem (D) as defined in Section 4.3. Indeed, both problems have the same constraints and, if $e^T w = 1$, we have

$$D(u, w) = H(D^*(u, w)).$$

The map H is used to define the function

$$\Psi : 2^{\mathbb{R}^q} \to 2^{\mathbb{R}^q}, \quad \Psi(F^*) := \bigcap_{y^* \in F^*} H(y^*) \cap \mathcal{P}.$$

The following geometric duality theorem states that Ψ is a duality map between $\mathcal{P}$ and $\mathcal{D}^*$.

Theorem 4.42 (geometric duality). *Ψ is an inclusion reversing one-to-one map between the set of all K-maximal proper faces of $\mathcal{D}^*$ and the set of all weakly minimal proper faces of $\mathcal{P}$. The inverse map is given by*

$$\Psi^{-1}(F) = \bigcap_{y \in F} H^*(y) \cap \mathcal{D}^*.$$

Moreover, if F^* is a K-maximal proper face of $\mathcal{D}^*$, then

$$\dim F^* + \dim \Psi(F^*) = q - 1.$$

The proof of this theorem is given below after preparing it by several lemmas. Prior to this, an important special case is considered. Vertices as well as facets are actually the most important faces from the viewpoint of applications. Therefore, we extract some corresponding conclusions from the above theorem.

Corollary 4.43. *The following statements are equivalent:*

(i) y^* *is a K-maximal vertex of* $\mathcal{D}^*$,

(ii) $H(y^*) \cap \mathcal{P}$ *is a weakly minimal* $(q-1)$*-dimensional facet of* $\mathcal{P}$*.*

Moreover, if F *is a weakly minimal* $(q-1)$*-dimensional facet of* $\mathcal{P}$*, then there is some uniquely defined point* $y^* \in \mathbb{R}^q$ *such that* $F = H(y^*) \cap \mathcal{P}$*.*

Proof. (i) $\Rightarrow$ (ii). Since $H(y^*) \cap \mathcal{P} = \Psi(\{y^*\})$, Theorem 4.42 implies that $H(y^*) \cap \mathcal{P}$ is a weakly minimal proper face of $\mathcal{P}$. Theorem 4.42 also implies that $\dim(H(y^*) \cap \mathcal{P}) = q - 1 - \dim\{y^*\} = q - 1$.

(ii) $\Rightarrow$ (i). Let $H(y^*) \cap \mathcal{P}$ be a weakly minimal $(q-1)$-dimensional facet of $\mathcal{P}$. By Theorem 4.42, $\Psi^{-1}(H(y^*) \cap \mathcal{P})$ is a K-maximal vertex of $\mathcal{D}^*$, denoted by $\bar{y}^*$. It follows that $\Psi \circ \Psi^{-1}(H(y^*) \cap \mathcal{P}) = \Psi(\{\bar{y}^*\})$ and hence $H(y^*) \cap \mathcal{P} = H(\bar{y}^*) \cap \mathcal{P}$ implying $H(y^*) = H(\bar{y}^*)$ as $\dim(H(y^*) \cap \mathcal{P}) = q-1$. The mapping H being injective implies $y^* = \bar{y}^*$.

To show the last statement, let F be a weakly minimal $(q-1)$-dimensional facet of $\mathcal{P}$. Hence $\Psi^{-1}(F)$ consists of a K-maximal vertex of $\mathcal{D}^*$, denoted by y^*. It follows $F = \Psi \circ \Psi^{-1}(F) = \Psi(\{y^*\}) = H(y^*) \cap \mathcal{P}$. By $\dim(H(y^*) \cap \mathcal{P}) = q - 1$ and H being injective, y^* is uniquely defined. □

The next result is the dual counterpart of the last one.

Corollary 4.44. *The following statements are equivalent:*

(i) y *is a weakly minimal vertex of* $\mathcal{P}$*,*
(ii) $H^*(y) \cap \mathcal{D}^*$ *is a* K*-maximal* $(q-1)$*-dimensional facet of* $\mathcal{D}^*$*.*

Moreover, if F^* *is a* K*-maximal* $(q-1)$*-dimensional facet of* $\mathcal{D}^*$*, then there is some uniquely defined point* $y \in \mathbb{R}^q$ *such that* $F^* = H^*(y) \cap \mathcal{D}^*$*.*

Proof. (i) $\Rightarrow$ (ii). Let y be a weakly minimal vertex of $\mathcal{P}$. By Theorem 4.42, the set $F^* := \Psi^{-1}(\{y\}) = H^*(y) \cap \mathcal{D}^*$ is a K-maximal face of $\mathcal{D}^*$. From Theorem 4.42 we also conclude that $\dim F^* = q - 1 - \dim\{y\} = q - 1$. Thus F^* is a facet of $\mathcal{D}^*$.

(ii) $\Rightarrow$ (i). Let $H^*(y) \cap \mathcal{D}^*$ be a K-maximal $(q-1)$-dimensional facet of $\mathcal{D}^*$. By Theorem 4.42, $\Psi(H^*(y) \cap \mathcal{D}^*)$ consists of a weakly minimal vertex of $\mathcal{P}$, denoted by $\bar{y}$. It follows that

$$\Psi^{-1} \circ \Psi(H^*(y) \cap \mathcal{D}^*) = \Psi^{-1}(\{\bar{y}\})$$

and hence $H^*(y) \cap \mathcal{D}^* = H^*(\bar{y}) \cap \mathcal{D}^*$. Since $\dim(H^*(y) \cap \mathcal{D}^*) = q-1$ and H^* is injective, we get $y = \bar{y}$.

To show the last statement, let F^* be a K-maximal $(q-1)$-dimensional facet of $\mathcal{D}^*$. Hence $\Psi(F^*)$ is a minimal vertex of $\mathcal{P}$, denoted by y. It follows that

$$F^* = \Psi^{-1} \circ \Psi(F^*) = \Psi^{-1}(\{y\}) = H^*(y) \cap \mathcal{D}^*.$$

By $\dim(H^*(y) \cap \mathcal{D}^*) = q-1$ and H^* being injective, y is uniquely defined. □

Geometric duality is illustrated by the following two examples.

Example 4.45. Consider Problem (P) with the data

$$P = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \qquad B = \begin{pmatrix} 0 & -1 \\ 2 & 1 \\ 1 & 1 \\ 1 & 2 \end{pmatrix}, \qquad b = \begin{pmatrix} -4 \\ 4 \\ 3 \\ 4 \end{pmatrix}.$$

The set $\mathcal{D}^*$ (see Figure 4.2) can easily be evaluated as

$$\mathcal{D}^* = \text{co}\left\{\left(\tfrac{1}{3}, \tfrac{4}{3}\right)^T, \left(\tfrac{1}{2}, \tfrac{3}{2}\right)^T, \left(\tfrac{2}{3}, \tfrac{4}{3}\right)^T, \left(1, 0\right)^T\right\} - K.$$

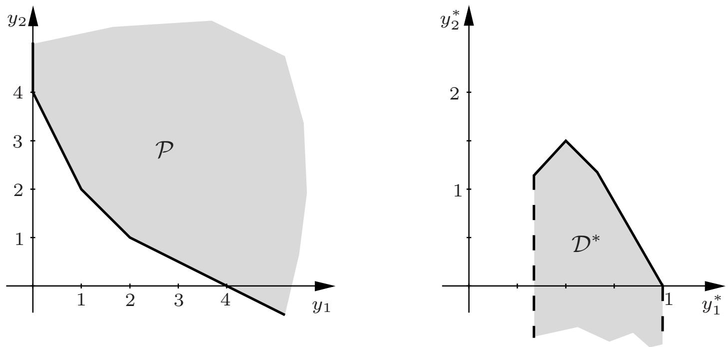

Fig. 4.2 The three weakly minimal vertices of $\mathcal{P}$ correspond to the three K-maximal facets of $\mathcal{D}^*$ and the four weakly minimal facets of $\mathcal{P}$ correspond to the four K-maximal vertices of $\mathcal{D}^*$.

Example 4.46. Consider Problem (P) with the data

$$P = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \qquad B = \begin{pmatrix} 1 & 1 & 1 \\ -1 & -1 & 1 \\ -1 & 1 & -1 \\ 1 & -1 & -1 \end{pmatrix}, \qquad b = \begin{pmatrix} 1 \\ -1 \\ -1 \\ -1 \end{pmatrix}.$$

An easy computation shows that

$$\mathcal{D}^* = \text{co}\left\{(0,0,0)^T, (1,0,0)^T, (0,1,0)^T, \left(\tfrac{1}{3}, \tfrac{1}{3}, \tfrac{1}{3}\right)^T\right\} - K,$$

see Figure 4.3.

We now turn to the proof of the geometric duality theorem. It is based on the two pairs of scalar linear programs ($\text{P}_1(w)$), ($\text{D}_1(w)$) and ($\text{P}_2(y)$), ($\text{D}_2(y)$), as introduced in Section 4.1. Note that the lower image of (D^*) can be expressed as

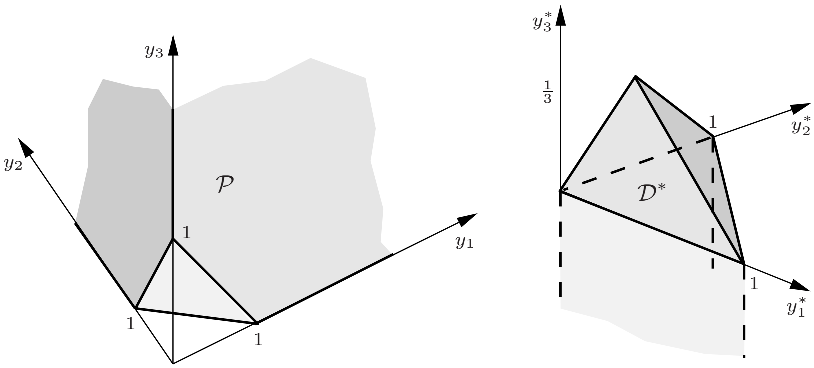

Fig. 4.3 The three weakly minimal vertices of $\mathcal{P}$ correspond to the three K-maximal facets of $\mathcal{D}^*$, the six weakly minimal edges of $\mathcal{P}$ correspond to the six K-maximal edges of $\mathcal{D}^*$ and the four weakly minimal facets of $\mathcal{P}$ correspond to the four K-maximal vertices of $\mathcal{D}^*$.

$$\mathcal{D}^* = \left\{ y^* \in \Delta |\ \exists u \in \mathbb{R}^m :\ u \geq 0,\ B^T u = P^T w(y^*),\ b^T u \geq y_q^* \right\}, \qquad (4.18)$$

where

$$\Delta := \{y^* \in \mathbb{R}^q |\ w(y^*) \geq 0\}.$$

We start with several lemmas.

Lemma 4.47. *Every K-maximal proper face of $\mathcal{D}^*$ contains a vertex.*

Proof. Let F^* be a K-maximal proper face of $\mathcal{D}^*$. It suffices to show that F^* contains no lines (Rockafellar, 1972, Corollary 18.5.3). Assume on the contrary that F^* contains a line, i.e., there are $\bar{y}^* \in F^*$ and $\psi \in \mathbb{R}^q \setminus \{0\}$ such that $\bar{y}^* + \lambda\psi \in F^*$ for all $\lambda \in \mathbb{R}$. For every $y^* \in F^* \subseteq \mathcal{D}^*$, we have $y_1^* \geq 0, \dots, y_{q-1}^* \geq 0$ and hence $\psi_1 = \dots = \psi_{q-1} = 0$. Thus, $\psi \neq 0$ implies $K \subseteq \{\lambda\psi |\ \lambda \in \mathbb{R}\}$. We get $\{\bar{y}^*\} + K \subseteq F^*$, which contradicts the K-maximality of F^*. □

Lemma 4.48. *Consider a hyperplane*

$$H^* := \left\{ y^* \in \mathbb{R}^q |\ {w^*}^T y^* = \gamma \right\}.$$

Then the following statements are equivalent:

(i) *H^* is a supporting hyperplane to $\mathcal{D}^*$ such that $H^* \cap \mathcal{D}^*$ is K-maximal;*
(ii) *H^* is a supporting hyperplane to $D^*[T]$ and $w_q^* < 0$.*

Proof. (i) ⇒ (ii). If H^* is a supporting hyperplane to $\mathcal{D}^*$, then there is some $\bar{y}^* \in \mathcal{D}^*$ such that ${w^*}^T \bar{y}^* = \gamma$ and for $y^* \in \mathcal{D}^*$ we have ${w^*}^T y^* \geq \gamma$. From the definition of $\mathcal{D}^*$ we get $\hat{y}^* := \bar{y}^* - e^q \in \mathcal{D}^*$, where e^q is the q-th unit vector. Hence $w_q^* \leq 0$. Since $w_q^* = 0$ would imply $\hat{y}^* \in H^* \cap \mathcal{D}^*$

and $\bar{y}^* \in (\hat{y}^* + K \setminus \{0\}) \cap \mathcal{D}^*$, contradicting the maximality of $H^* \cap \mathcal{D}^*$, we conclude $w_q^* < 0$. As $\bar{y}^* \in \mathcal{D}^*$, there are $\tilde{y}^* \in D^*[T] \subseteq \mathcal{D}^*$ and $z \geq 0$ such that $\bar{y}^* = \tilde{y}^* - e^q z$. Hence ${w^*}^T \tilde{y}^* = {w^*}^T \bar{y}^* + w_q^* z \leq \gamma$. This implies ${w^*}^T \tilde{y}^* = \gamma$. Therefore, H^* is a supporting hyperplane to $D^*[T]$.

(ii) $\Rightarrow$ (i). If H^* is a supporting hyperplane to $D^*[T]$, then there is some $\bar{y}^* \in D^*[T]$ such that ${w^*}^T \bar{y}^* = \gamma$ and, for all $y^* \in D^*[T]$, one has ${w^*}^T y^* \geq \gamma$. Since $w_q^* < 0$, it follows ${w^*}^T y^* \geq \gamma$ for all $y^* \in D^*[T] - K = \mathcal{D}^*$. From $\bar{y}^* \in \mathcal{D}^*$ and ${w^*}^T \bar{y}^* = \gamma$, we conclude that H^* is a supporting hyperplane to $\mathcal{D}^*$.

In order to show that $H^* \cap \mathcal{D}^*$ is K-maximal, let $\bar{y}^* \in H^* \cap \mathcal{D}^*$ be given. We have ${w^*}^T \bar{y}^* = \gamma$. For every $y^* \in \{\bar{y}^*\} + K \setminus \{0\}$ one has ${w^*}^T y^* < \gamma$, which is a consequence of $w_q^* < 0$. Since ${w^*}^T y^* \geq \gamma$ for all $y^* \in \mathcal{D}^*$, we obtain $(\bar{y}^* + K \setminus \{0\}) \cap \mathcal{D}^* = \emptyset$. □

Weakly minimal points of the upper image $\mathcal{P}$ of (P) refer to K-maximal faces of the lower image $\mathcal{D}^*$ of the geometric dual problem (D*). A weakly minimal point of $\mathcal{P}$ defines a hyperplane $H^*(y)$ that supports $\mathcal{D}^*$ in a K-maximal face. This is shown by the following lemma.

Lemma 4.49. *Let $y \in \mathbb{R}^q$. The following statements are equivalent:*

(i) *y is a weakly minimal point of $\mathcal{P}$,*
(ii) *$H^*(y) \cap \mathcal{D}^*$ is a K-maximal proper face of $\mathcal{D}^*$.*

Moreover, for every K-maximal proper face F^* of $\mathcal{D}^*$ there exists some $y \in \mathbb{R}^q$ such that $F^* = H^*(y) \cap \mathcal{D}^*$.

Proof. Take into account that $\varphi(y, y^*) \geq 0$ holds for all $y^* \in D^*[T]$ (Theorem 4.40). By Theorem 4.20, (i) is equivalent to

(iii) There exists some $(\bar{u}, \bar{w}) \in T$ with $y^T \bar{w} = b^T \bar{u}$ solving (D$_2(y)$).

Because of (4.14), (iii) is equivalent to

(iv) There exists some $\bar{y}^* \in D^*[T]$ such that $\varphi(y, \bar{y}^*) = 0$.

Statement (iv) is equivalent to

(v) $H^*(y)$ is a supporting hyperplane to $D^*[T]$.

Regarding the fact that $H^*(y) = \left\{y^* \in \mathbb{R}^q |\ w^*(y)^T y^* = -y_q\right\}$ with $w^*(y)_q = -1 < 0$, (v) is equivalent to (ii), by Lemma 4.48.

Let F^* be a K-maximal proper face of $\mathcal{D}^*$. Then there exists a supporting hyperplane $H^* := \left\{y^* \in \mathbb{R}^q |\ {w^*}^T y^* = \gamma\right\}$ $(w^* \neq 0)$ to $\mathcal{D}^*$ such that $F^* = H^* \cap \mathcal{D}^*$. By Lemma 4.48, we have $w_q^* < 0$. Setting

$$y := \left(\frac{\gamma - w_1^*}{w_q^*}, \dots, \frac{\gamma - w_{q-1}^*}{w_q^*}, \frac{\gamma}{w_q^*}\right)^T$$

we obtain $H^* = H^*(y)$. Hence $F^* = H^*(y) \cap \mathcal{D}^*$. □

Lemma 4.50. *Consider a hyperplane* $H := \left\{ y \in \mathbb{R}^q \mid w^T y = \gamma \right\}$. *The following statements are equivalent:*

(i) H *is a supporting hyperplane to* $\mathcal{P}$,
(ii) $w \geq 0$ *and* H *is a supporting hyperplane to* $P[S]$.

Proof. (i) $\Rightarrow$ (ii). If H is a supporting hyperplane to $\mathcal{P}$, then there is some $\bar{y} \in \mathcal{P}$ with $w^T \bar{y} = \gamma$ and for all $y \in \mathcal{P}$ one has $w^T y \geq \gamma$. By the definition of $\mathcal{P}$, we have $\bar{y} + \hat{y} \in \mathcal{P}$ for all $\hat{y} \in \mathbb{R}^q_+$, hence $w^T \hat{y} \geq 0$ for all $\hat{y} \in \mathbb{R}^q_+$. This implies $w \geq 0$. Since $\bar{y} \in \mathcal{P}$, there is $\tilde{y} \in P[S] \subseteq \mathcal{P}$ and $\hat{y} \in \mathbb{R}^q_+$ such that $\bar{y} = \tilde{y} + \hat{y}$. Hence $w^T \tilde{y} = w^T \bar{y} - w^T \hat{y} \leq \gamma$. This implies $w^T \tilde{y} = \gamma$. Therefore, H is a supporting hyperplane to $P[S]$.

(ii) $\Rightarrow$ (i). If H is a supporting hyperplane to $P[S]$, there exists some $\bar{y} \in P[S]$ with $w^T \bar{y} = \gamma$ and for all $y \in P[S]$, one has $w^T y \geq \gamma$. Since $w \geq 0$, it follows that $w^T y \geq \gamma$ for all $y \in P[S] + \mathbb{R}^q_+$. By $\bar{y} \in \mathcal{P}$ and $w^T \bar{y} = \gamma$, we conclude that H is a supporting hyperplane to $\mathcal{P}$. □

We continue with the dual counterpart of Lemma 4.49.

Lemma 4.51. *Let* $y^* \in \mathbb{R}^q$. *The following statements are equivalent:*

(i) y^* *is a* K*-maximal point of* $\mathcal{D}^*$,
(ii) $H(y^*) \cap \mathcal{P}$ *is a weakly minimal proper face of* $\mathcal{P}$.

Moreover, for every proper face F *of* $\mathcal{P}$ *there exists some* $y^* \in \mathbb{R}^q$ *such that* $F = H(y^*) \cap \mathcal{P}$.

Proof. Taking into account (4.18), we conclude that (i) is equivalent to

(iii) $w(y^*) \geq 0$ and there exists a solution $\bar{u}$ to $(\mathrm{D}_1(w(y^*)))$ such that $y^*_q = b^T \bar{u}$.

By duality between $(\mathrm{P}_1(w(y^*)))$ and $(\mathrm{D}_1(w(y^*)))$, (iii) is equivalent to

(iv) $w(y^*) \geq 0$ and there exists a solution $\bar{x}$ to $(\mathrm{P}_1(w(y^*)))$ such that $y^*_q = w(y^*)^T P \bar{x}$.

Statement (iv) is equivalent to

(v) $w(y^*) \geq 0$ and $H(y^*)$ is a supporting hyperplane to $P[S]$.

By Lemma 4.50 and Corollary 4.25, (v) is equivalent to (ii).

To show the last conclusion, let F be a proper face of $\mathcal{P}$. Hence there exists a supporting hyperplane $H := \left\{ y \in \mathbb{R}^q \mid w^T y = \gamma \right\}$ $(w \neq 0)$ of $\mathcal{P}$ such that $F = H \cap \mathcal{P}$. By Lemma 4.50, we have $w \geq 0$. Without loss of generality we can assume that $e^T w = 1$. Setting $y^*_i := w_i$ for $i = 1, \dots, q-1$ and $y^*_q := \gamma$, we obtain $H = H(y^*)$. Hence $F = H(y^*) \cap \mathcal{P}$. □

Now we are able to give the proof of the geometric duality theorem.

Proof of Theorem 4.42. (a) We show that, if F^* is a K-maximal proper face of $\mathcal{D}^*$, then $\Psi(F^*)$ is a weakly minimal proper face of $\mathcal{P}$. By Lemma 4.51,

$H(y^*)\cap\mathcal{P}$ is a weakly minimal proper face of $\mathcal{P}$ for each $y^* \in F^*$, hence $\Psi(F^*)$ is a weakly minimal face of $\mathcal{P}$. It remains to show that $\Psi(F^*)$ is nonempty. By Lemma 4.49 there is some $\bar{y} \in \operatorname{wMin}\mathcal{P}$ such that $F^* = H^*(\bar{y})\cap\mathcal{D}^*$. Using (4.17), we get $\bar{y} \in \Psi(F^*)$.

(b) We shall show that $\Psi^*(F) := \bigcap_{y\in F} H^*(y)\cap\mathcal{D}^*$ is a K-maximal proper face of $\mathcal{D}^*$ whenever F is a weakly minimal proper face of $\mathcal{P}$. By Lemma 4.49, $H^*(y)\cap\mathcal{D}^*$ is a K-maximal proper face of $\mathcal{D}^*$ for each $y \in F$. Hence $\Psi^*(F)$ is a K-maximal proper face of $\mathcal{D}^*$ whenever this set is nonempty. Indeed, by Lemma 4.51, there exists some $\bar{y}^* \in \operatorname{Max}_K \mathcal{D}^*$ such that $F = H(\bar{y}^*)\cap\mathcal{P}$. Using (4.17), we obtain $\bar{y}^* \in \Psi^*(F)$.

(c) In order to show that Ψ is a bijection and $\Psi^{-1}(F) = \bigcap_{y\in F} H^*(y)\cap\mathcal{D}^* = \Psi^*(F)$, we have to show the following two statements: (c_1) $\Psi^*(\Psi(F^*)) = F^*$ for all K-maximal proper faces F^* of $\mathcal{D}^*$ and (c_2) $\Psi(\Psi^*(F)) = F$ for all weakly minimal proper faces F of $\mathcal{P}$.

(c_1) First we show that $F^* \subseteq \Psi^*(\Psi(F^*))$. Assuming the contrary, we get some $\bar{y}^* \in F^*$ such that $\bar{y}^* \notin \Psi^*(\Psi(F^*))$. Hence there exists $\bar{y} \in \Psi(F^*)$ such that $\bar{y}^* \notin H^*(\bar{y})\cap\mathcal{D}^*$. This implies $\bar{y}^* \notin H^*(\bar{y})$ since $\bar{y}^* \in \mathcal{D}^*$. It follows $\bar{y} \notin H(\bar{y}^*)$, whence $\bar{y} \notin \Psi(F^*)$, a contradiction. To show the opposite inclusion, let $\bar{y} \in \operatorname{wMin}\mathcal{P}$ such that $F^* = H^*(\bar{y})\cap\mathcal{D}^*$. The existence of such a point $\bar{y}$ is ensured by Lemma 4.49. It follows $\bar{y} \in \Psi(F^*)$. Hence $\Psi^*(\Psi(F^*)) \subseteq H^*(\bar{y})\cap\mathcal{D}^* = F^*$.

(c_2) The proof works analogously using Lemma 4.51 instead of Lemma 4.49.

(d) Obviously, Ψ is inclusion reversing.

(e) It remains to prove that $\dim F^* + \dim\Psi(F^*) = q-1$ for all K-maximal proper faces F^* of $\mathcal{D}^*$. Consider some fixed F^* and set $r := \dim F^*$ and $s := \dim\Psi(F^*)$. By the first part of the proof, $F := \Psi(F^*)$ is a weakly minimal face of $\mathcal{P}$. Hence there exist proper faces $F \subsetneq F^1 \subsetneq F^2 \subsetneq \cdots \subsetneq F^{q-1-s}$ (all of them being weakly minimal by Corollary 4.25) such that $\dim F^{q-1-s} = q-1$. From the properties of Ψ, we conclude that $0 \le \dim\Psi^{-1}(F^{q-1-s}) \le r-(q-1-s)$. Hence $r+s \ge q-1$. Since every K-maximal face of $\mathcal{D}^*$ has a vertex (Lemma 4.47), there are K-maximal faces $F^* \supsetneq F^{*1} \supsetneq F^{*2} \supsetneq \cdots \supsetneq F^{*r}$ such that $\dim F^{*r} = 0$. It follows that $s+r \le \dim\Psi(F^{*r}) \le q-1$. Together we have $s+r = q-1$, which completes the proof. □

We proceed with a conclusion of Theorem 4.42 and Lemma 4.49.

Corollary 4.52. *The following statements are equivalent:*

(i) $\bar{y} \in \mathbb{R}^q$ *belongs to the relative interior of a proper weakly minimal face* F *of* $\mathcal{P}$ *with* $\dim F = r$;
(ii) $H^*(\bar{y})\cap\mathcal{D}^*$ *is a* $(q-r-1)$*-dimensional* K*-maximal proper face of* $\mathcal{D}^*$.

Proof. (i) ⇒ (ii). The geometric duality theorem yields

$$F^* := \Psi^{-1}(F) = \bigcap_{y\in F} H^*(y)\cap\mathcal{D}^* \subseteq H^*(\bar{y})\cap\mathcal{D}^* =: \bar{F}^*,$$

where $\dim F^* = q - r - 1$. By Lemma 4.49, it remains to show that $\dim F^* = \dim \bar{F}^*$. Of course, $\dim F^* \leq \dim \bar{F}^*$. Assume that $\dim F^* < \dim \bar{F}^*$. By Theorem 4.42, we have

$$\bar{F} := \Psi(\bar{F}^*) = \bigcap_{y^* \in \bar{F}^*} H(y^*) \cap \mathcal{P}$$

and $\dim \bar{F} < \dim F = r$. From $\bar{F}^* \subseteq H^*(\bar{y})$ and (4.17), we get $\bar{y} \in \bigcap_{y^* \in \bar{F}^*} H(y^*)$. As $\bar{y} \in \mathcal{P}$, it follows $\bar{y} \in \bar{F}$. Both $\bar{F}$ and F are proper weakly minimal faces of $\mathcal{P}$ such that $F \supseteq \bar{F}$ (as Ψ is inclusion reversing) and $\dim F > \dim \bar{F}$. Thus $\bar{F}$ is a proper face of F. By Theorem 4.22, we get $\bar{F} \subseteq \operatorname{rbd} F$. Whence the contradiction $\bar{y} \in \operatorname{ri} F \cap \operatorname{rbd} F = \emptyset$.

(ii) $\Rightarrow$ (i). Set $F := \Psi(\bar{F}^*)$. According to Lemma 4.49 and Theorem 4.42, it remains to show that $\bar{y} \in \operatorname{ri} F$. Assuming the contrary, we obtain that $\bar{y}$ belongs to the relative interior of a face $\bar{F} \subseteq F$ of $\mathcal{P}$ with $\bar{r} := \dim \bar{F} < r$ (compare Webster, 1994, Theorem 2.6.5). From the first part of the proof we get $\dim \bar{F}^* = q - \bar{r} - 1 > q - r - 1$. This contradicts condition (ii). □

The dual counterpart of the latter result is the following.

Corollary 4.53. *The following statements are equivalent:*

(i) $\bar{y}^* \in \mathbb{R}^q$ *belongs to the relative interior of a proper* K*-maximal face* F^* *of* $\mathcal{D}^*$ *with* $\dim F^* = r$*;*
(ii) $H(\bar{y}^*) \cap \mathcal{P}$ *is a* $(q - r - 1)$*-dimensional weakly minimal proper face of* $\mathcal{P}$*.*

Proof. The proof is completely analogous to the proof of Corollary 4.52. Lemma 4.51 has to be used instead of Lemma 4.49. □

Theorem 4.1 also provides a characterization of weakly minimal vectors of the set $\mathcal{P}$. Let us consider, for fixed $\bar{y} \in \operatorname{wMin} \mathcal{P}$, the set of all $w \in \mathbb{R}^q_+ \setminus \{0\}$ with the property $w^T \bar{y} \leq w^T y$ for all $y \in \mathcal{P}$. This leads to the idea of *weight space decomposition.* We set

$$W(\bar{y}) := \left\{ w \in \mathbb{R}^q_+ \mid e^T w = 1 \text{ and } \forall y \in \mathcal{P} : w^T \bar{y} \leq w^T y \right\}.$$

The set

$$W(\operatorname{wMin} \mathcal{P}) = \bigcup_{\bar{y} \in \operatorname{wMin} \mathcal{P}} W(\bar{y}).$$

is called the *weight space* of (P). It consists of those weights $w \geq 0$ with $e^T w = 1$ that make the weighted sum scalarization ($\mathrm{P}_1(w)$) solvable. If $\mathcal{P}_\infty = \mathbb{R}^q_+$, we have the maximal (w.r.t. inclusion) weight space $W(\operatorname{wMin} \mathcal{P}) = \left\{ w \geq 0 \mid e^T w = 1 \right\}$.

Definition 4.54. A finite family of sets $\left\{ A^1, \ldots, A^r \right\}$ is said to be a *weight space decomposition* of Problem (P) if there exist $\bar{y}^1, \ldots, \bar{y}^r \in \mathcal{P}$ such that $W(\bar{y}^i) = A^i$ for $i = 1, \ldots, r$ and

$$W(\operatorname{wMin}\mathcal{P}) = \bigcup_{i=1}^{r} A^i \quad \wedge \quad \operatorname{ri} A^i \cap \operatorname{ri} A^j = \emptyset \text{ whenever } i \neq j.$$

Geometric duality yields the following result about the weight space decomposition. Recall that $w : \mathbb{R}^q \to \mathbb{R}^q$ is defined in (4.15).

Corollary 4.55. *Let $\{\bar{y}^1, \dots, \bar{y}^r\}$ be the set of weakly minimal vertices of $\mathcal{P}$, and assume this set to be nonempty. Denote by $\{F^{*1}, \dots, F^{*r}\}$ the corresponding K-maximal facets of $\mathcal{D}^*$ according to the geometric duality theorem (in particular Corollary 4.44). Then we have*

$$W(\bar{y}^i) = w[F^{*i}] := \{w(y^*) |\; y^* \in F^{*i}\},$$

and $\{w[F^{*i}] |\; i = 1, \dots, r\}$ is a weight space decomposition of (P)*.*

Proof. It can be shown that $w(\cdot)$ is a one-to-one map from $\operatorname{Max}_K \mathcal{D}^*$ onto $W(\operatorname{wMin}\mathcal{P})$. The inverse map is $w^{-1}(\lambda) = (\lambda_1, \dots, \lambda_{q-1}, y_q^*)^T$, where y_q^* is the optimal value of the linear program $(\mathrm{P}_1(\lambda))$, which is finite for $\lambda \in W(\operatorname{wMin}\mathcal{P})$. Moreover, $w(\cdot)$ is affine on convex subsets of $\operatorname{Max}_K \mathcal{D}^*$, in particular, on each K-maximal facet of $\mathcal{D}^*$.

Let $\bar{y}^i$ be a weakly minimal vertex of $\mathcal{P}$. By Corollary 4.44, $H^*(\bar{y}^i) \cap \mathcal{D}^* = F^{*i}$ is a K-maximal facet of $\mathcal{D}^*$. Note that

$$y^* \in H^*(y) \quad \Longleftrightarrow \quad w(y^*)^T y = y_q^*.$$

Let $\lambda \in W(\bar{y}^i)$. Setting $y^* := w^{-1}(\lambda) \in \operatorname{Max}_K \mathcal{D}^*$, we obtain $w(y^*)^T \bar{y}^i = \lambda^T \bar{y}^i = y_q^*$, which implies that $y^* \in H^*(\bar{y}^i) \cap \mathcal{D}^* = F^{*i}$ and thus $\lambda = w(y^*) \in w[F^{*i}]$.

On the other hand, if $y^* \in F^{*i}$, we have $\bar{y}^i \in H(y^*) \cap \mathcal{P}$ by geometric duality and thus $w(y^*)^T \bar{y}^i = y_q^*$. For every $y \in \mathcal{P}$, we get $y_q^* \leq w(y^*)^T y$ (Theorem 4.40). Together we obtain $w(y^*) \in W(\bar{y}^i)$.

The second statement follows from the fact

$$\operatorname{Max}_K \mathcal{D}^* = \textstyle\bigcup_{i=1}^r F^{*i} \quad \text{and} \quad \operatorname{ri} F^{*i} \cap \operatorname{ri} F^{*j} = \emptyset \quad \text{whenever} \quad i \neq j,$$

compare (Webster, 1994, Corollary 2.6.7). Indeed, using Lemma 4.51, we can show that

$$w\left(\operatorname{Max}_K \mathcal{D}^*\right) = W\left(\operatorname{wMin}\mathcal{P}\right).$$

We conclude that

$$W\left(\operatorname{wMin}\mathcal{P}\right) = w\left[\operatorname{Max}_K \mathcal{D}^*\right] = w\left[\bigcup_{i=1}^{r} F^{*i}\right] = \bigcup_{i=1}^{r} w\left[F^{*i}\right].$$

The second condition in the definition of a weight space decomposition follows directly from the properties of $w(\cdot)$. □

A solution concept for the geometric dual problem (D^*) is now introduced.

Definition 4.56. A nonempty set $\bar{T} \subseteq \mathbb{R}^{m+q}$ is called a *finitely generated solution* to the geometric dual problem (D*) if

(i) $\bar{T}$ is a finite subset of T,
(ii) $D^*[\bar{T}] \subseteq \operatorname{Max}_K D^*[T]$,
(iii) $D^*[T] \subseteq \operatorname{co} D^*[\bar{T}] - K$.

The latter definition is analogous to the definition of a finitely generated solution to (P), but the conditions (ii) and (iv) of Definition 4.4 do not occur here. The reason is that the recession cone of $\mathcal{D}^*$ is always the same, namely $\mathcal{D}^*_\infty = -K$. The recession cone of $\mathcal{P}$, however, depends on the data of the primal problem (P).

We next show that the hyperplane-valued problem (D) and the vector-valued problem (D*) are equivalent in the sense that they have the same finitely generated solution.

Theorem 4.57. *Consider both dual problems* (D) *and* (D*) *of problem* (P). *The following is equivalent:*

(i) $\bar{T}$ *is a finitely generated solution to* (D)*;*
(ii) $\bar{T}$ *is a finitely generated solution to* (D*)*.*

Proof. Let us denote the three conditions in Definition 4.16 by (i)$'$, (ii)$'$ and (iii)$'$ and the conditions in Definition 4.56 by (i)*, (ii)* and (iii)*. Obviously, (i)$'$ is equivalent to (i)*.

(ii)$'$ is equivalent (compare Definition 4.15 and use $\bar{T} \subseteq T$) to

$$\forall (\bar{u}, \bar{w}) \in \bar{T}: \; (\bar{u}, \bar{w}) \in \operatorname{Eff}(\mathrm{D}).$$

By Theorem 4.19 this is equivalent to

$$\forall (\bar{u}, \bar{w}) \in \bar{T}: \; \bar{u} \text{ solves } (\mathrm{D}_1(\bar{w})).$$

Because of (iii) and (i) in Lemma 4.51, the latter assertion is equivalent to

$$\forall (\bar{u}, \bar{w}) \in \bar{T}: \; D^*(\bar{u}, \bar{w}) \text{ is a } K\text{-maximal point of } \mathcal{D}^*.$$

As $\operatorname{Max}_K \mathcal{D}^* = \operatorname{Max}_K D^*[T]$, the last statement is equivalent to (ii)*.

Let us denote the elements of $\bar{T}$ by (u^i, w^i), where $i = 1, \dots, k$. (iii)$'$ is equivalent to

$$\forall (u, w) \in T: \quad D(u, w) - \mathbb{R}^q_+ \subseteq \bigcup_{i=1}^{k} (D(u^i, w^i) - \mathbb{R}^q_+).$$

By De Morgan's law, this is equivalent to

$$\forall (u, w) \in T: \quad \bigcap_{i=1}^{k} \big(D(u^i, w^i) + \operatorname{int} \mathbb{R}^q_+\big) \subseteq D(u, w) + \operatorname{int} \mathbb{R}^q_+,$$

or equivalently,

$$\forall (u,w) \in T : \quad \left([\forall i \in \{1,\dots,k\} :\ y^T w^i > b^T u^i] \implies y^T w > b^T u \right).$$

This is equivalent to

$$\forall (u,w) \in T : \quad \left([\forall i \in \{1,\dots,k\} :\ y^T w^i \geq b^T u^i] \implies y^T w \geq b^T u \right).$$

The last equivalence can easily be shown by using $y + \varepsilon e$ with small $\varepsilon > 0$. By (Rockafellar, 1972, Theorem 22.3), the latter statement holds if and only if

$$\forall (u,w) \in T, \exists \lambda_1,\dots,\lambda_k \geq 0 : \quad w = \sum_{i=1}^{k} \lambda_i w^i \ \wedge\ b^T u \leq \sum_{i=1}^{k} \lambda_i b^T u^i. \tag{4.19}$$

Note that we have

$$\sum_{i=1}^{k} \lambda_i = \sum_{i=1}^{k} \lambda_i e^T w^i = e^T \sum_{i=1}^{k} \lambda_i w^i = e^T w = 1.$$

Setting $\tilde{u} := \sum_{i=1}^{k} \lambda_i u^i$, we have $(\tilde{u}, w) \in \operatorname{co} \bar{T}$ in (4.19), which can be written as

$$\forall (u,w) \in T,\ \exists \tilde{u} \in \mathbb{R}^m : \quad (\tilde{u}, w) \in \operatorname{co} \bar{T} \ \wedge\ b^T u \leq b^T \tilde{u}.$$

Since $\operatorname{co} D^*[\bar{T}] = D^*[\operatorname{co} \bar{T}]$, the latter assertion is equivalent to (iii)*. □

4.6 Homogeneous problems

In Section 4.2 we already introduced the homogeneous problem (P^h). The geometric dual of (P^h) is now investigated. Geometric duality between (P) and (D*) can be extended in such a way that (extreme) directions of $\mathcal{P}$ and faces of $\mathcal{D}^*$ being not K-maximal are taken into account.

The geometric dual problem of (P^h) is

$$\text{maximize } D^{*h} : \mathbb{R}^m \times \mathbb{R}^q \to \mathbb{R}^q \text{ with respect to } \leq_K \text{ over } T, \tag{D^{*h}}$$

where the feasible set T is the same as in the inhomogeneous problem (D*) (see Section 4.5) and the objective function is

$$D^{*h} : \mathbb{R}^m \times \mathbb{R}^q \to \mathbb{R}^q, \quad D^{*h}(u,w) := (w_1, \dots, w_{q-1}, 0)^T.$$

We denote the lower image of Problem (D^{*h}) by

$$\mathcal{D}^{*h} := D^{*h}[T] - K.$$

Observe that the objective function D^{*h} only differs in the last component from the objective function D^* of the inhomogeneous problem (D*) and consequently

$$\mathcal{D}^{*h} + K = \mathcal{D}^* + K. \tag{4.20}$$

An example is given in Figure 4.4. Since $S^h \neq \emptyset$ (because $0 \in S^h$), K-maximal points of $\mathcal{D}^{*h}$ exist, whenever $T \neq \emptyset$. In this case (but obviously also if $T = \emptyset$), $\mathcal{D}^{*h}$ can be expressed as

$$\mathcal{D}^{*h} = \left(\mathcal{D}^{*h} \cap \left\{y^* \in \mathbb{R}^q |\; y^*_q = 0\right\}\right) - K. \tag{4.21}$$

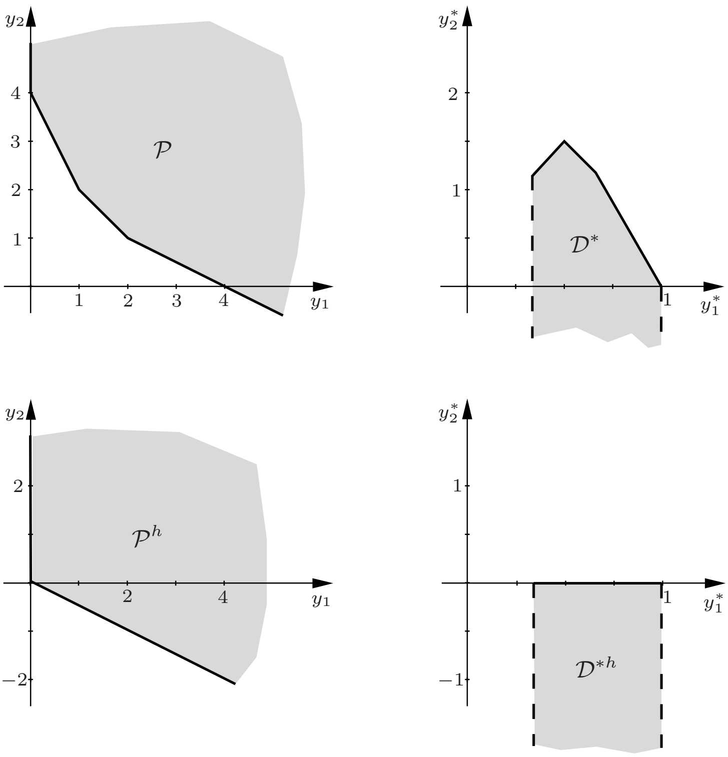

Fig. 4.4 The upper pictures show $\mathcal{P}$ and $\mathcal{D}^*$ from Example 4.45 and the lower pictures show the counterparts of the homogeneous problems.

The following example shows the case where S is empty.

Example 4.58. Consider Problem (P) with the data

$$P = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \qquad B = \begin{pmatrix} 1 & 1 \\ -1 & -1 \\ 0 & 1 \end{pmatrix}, \qquad b = \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}.$$

The feasible set S of (P) is empty, but S^h is nonempty. The upper images of (P) and (P^h) and the lower images of (D^*) and (D^{*h}) are shown in Figure 4.5.

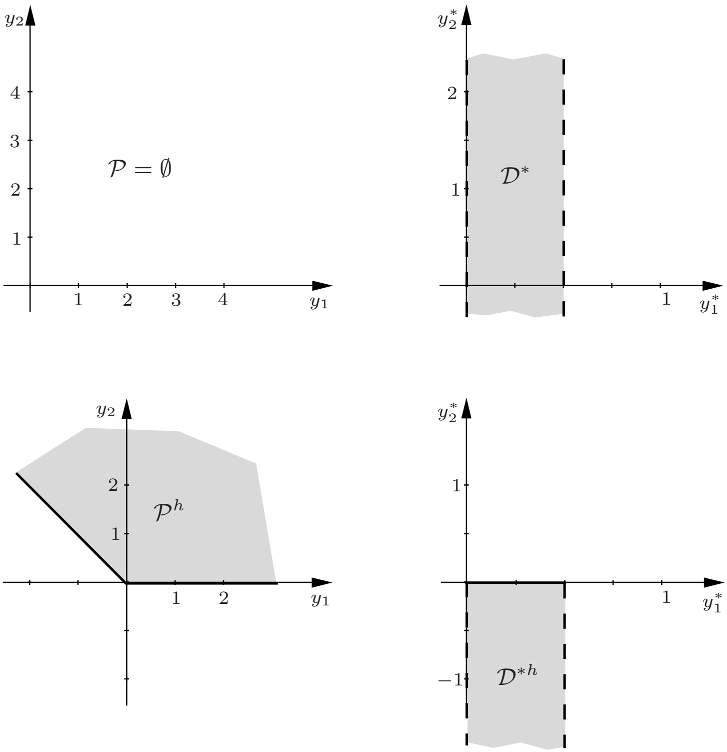

Fig. 4.5 Illustration of Example 4.58. $\mathcal{P}$ is empty and $\mathcal{D}^*$ has no K-maximal elements. But the homogeneous problems have weakly minimal and K-maximal elements, respectively.

The next statement extends the geometric duality theorem. In Theorem 4.42, only K-maximal faces are considered, but the lower image $\mathcal{D}^*$ also contains faces that are not K-maximal. As we will point out below, the facets of $\mathcal{D}^*$ being not K-maximal correspond to the extreme directions of $\mathcal{P}$. We introduce *vertical faces* of $\mathcal{D}^*$ and show that we obtain exactly those faces of $\mathcal{D}^*$ that are not K-maximal.

Definition 4.59. A hyperplane $H = \left\{y \in \mathbb{R}^q |\, \eta^T y = \gamma\right\}$ $(\eta \neq 0)$ is called *vertical* if $\eta_q = 0$. A proper face F of a nonempty polyhedral set $A \subseteq \mathbb{R}^q$ is called *vertical* if every hyperplane that supports A in F is vertical.

Lemma 4.60. *For every nonempty polyhedral set $A \subseteq \mathbb{R}^q$ with $A_\infty = -K$ the following statements are equivalent:*

(i) *F is a vertical face of A;*
(ii) *F is a proper face of A and $F \not\subseteq \operatorname{Max}_K A$.*

Proof. (i) ⇒ (ii). Let $y \in F$. The face F is the intersection of vertical hyperplanes and the set A. Since $A_\infty = -K$, we get $y - e^q \in F$, where e^q is the q-th unit vector. As $y - e^q \notin \operatorname{Max}_K A$, (ii) holds.

(ii) ⇒ (i). Let $H = \left\{y \in \mathbb{R}^q |\, \eta^T y = \gamma\right\}$ be a supporting hyperplane of A containing the face F. Let $y \in F \setminus \operatorname{Max}_K A$. There exists some $\varepsilon > 0$ such that $y + \varepsilon e^q \in A$. Since $A_\infty = -K$, we get $y - \varepsilon e^q \in A$. It follows $\eta^T(y + \varepsilon e^q) \geq \gamma$ and $\eta^T(y - \varepsilon e^q) \geq \gamma$. Hence $\eta_q = (\gamma - \eta^T y)/\varepsilon = 0$ and so H is vertical. As every supporting hyperplane of A containing the face F is vertical, F is a vertical face. □

The homogeneous primal problem (P^h) reflects the asymptotic behavior of the inhomogeneous problem (P). For the upper images of both problems this can be seen as follows.

Lemma 4.61. *Let $S \neq \emptyset$. A vector $y \in \mathbb{R}^q \setminus \{0\}$ is a direction of the upper image $\mathcal{P}$ of Problem* (P) *if and only if y belongs to $\mathcal{P}^h \setminus \{0\}$. Moreover, $\mathcal{P}^h = \mathcal{P}_\infty$.*

Proof. For every $i \in \{1, \dots, m\}$, we have

$$\{x \in \mathbb{R}^n |\, B_i x \geq b\}_\infty = \{x \in \mathbb{R}^n |\, B_i x \geq 0\}.$$

Since $S \neq \emptyset$, we obtain $S^h = S_\infty$ (compare Rockafellar, 1972, Corollary 8.3.3). It follows

$$\begin{aligned} \mathcal{P}^h &= P[S^h] + \mathbb{R}^q_+ = P[S_\infty] + \mathbb{R}^q_+ \\ &\overset{(4.1)}{=} P[S]_\infty + \mathbb{R}^q_+ \overset{(4.2)}{=} (P[S] + \mathbb{R}^q_+)_\infty = \mathcal{P}_\infty, \end{aligned}$$

which yields the desired results. □

A homogeneous counterpart φ^h of the coupling function $\varphi : \mathbb{R}^q \times \mathbb{R}^q \to \mathbb{R}$ (see Section 4.5) is now introduced in order to establish a geometric duality relation between extreme directions of $\mathcal{P}$ and vertical facets of $\mathcal{D}^*$. We set

$$\varphi^h : \mathbb{R}^q \times \mathbb{R}^q \to \mathbb{R}, \quad \varphi^h(y, y^*) := \sum_{i=1}^{q-1} y_i y_i^* + y_q \left(1 - \sum_{i=1}^{q-1} y_i^*\right).$$

Furthermore, we consider the corresponding vertical hyperplanes

$$\begin{aligned} H^{*h}(y) &:= \big\{y^* \in \mathbb{R}^q |\ \varphi^h(y,y^*) = 0\big\} \\ &= \{y^* \in \mathbb{R}^q |\ (y_1 - y_q, \dots, y_{q-1} - y_q, 0) \cdot y^* = -y_q\}\,. \end{aligned}$$

We obtain the following extension of the geometric duality theorem.

Theorem 4.62. *Let $S \neq \emptyset$. The following statements are equivalent:*

(i) $y \in \mathbb{R}^q \setminus \{0\}$ *is an extreme direction of* $\mathcal{P}$*;*
(ii) $H^{*h}(y)$ *supports* $\mathcal{D}^*$ *in a vertical* $(q-1)$*-dimensional facet.*

Proof. From Lemma 4.61, we conclude that $y \in \mathbb{R}^q \setminus \{0\}$ is an extreme direction of $\mathcal{P}$ if and only if it is an extreme direction of $\mathcal{P}^h$ (because $\mathcal{P}$ and $\mathcal{P}^h$ have the same directions). Note further that $H^{*h}(y)$ supports $\mathcal{D}^*$ in a vertical facet if and only if it supports $\mathcal{D}^{*h}$ in a vertical facet. This follows as the dual objective functions D^{*h} and D^* only differ in the last component. Thus it is sufficient to prove the statement for the homogeneous problems (P^h) and (D^{*h}).

Statement (i) holds for (P^h) if and only if the point y belongs to the relative interior of the 1-dimensional face $F = \{\lambda y |\ \lambda \geq 0\}$ of $\mathcal{P}^h$ (in particular, 0 is a vertex of $\mathcal{P}^h$ implying that $\operatorname{int} \mathcal{D}^{*h} \neq \emptyset$). By Corollary 4.52, this is true exactly when $H^*(y)$ supports $\mathcal{D}^{*h}$ in a $(q-2)$-dimensional K-maximal proper face. Because of the special form (4.21) of $\mathcal{D}^{*h}$ (in particular, $\operatorname{int} \mathcal{D}^{*h} \neq \emptyset$) this is equivalent to (ii) being satisfied for (D^{*h}). □

We close this section with some easy consequences of the results from Section 4.5. From the weak duality in Theorem 4.40 we get the next result.

Corollary 4.63. *The following implication is true:*

$$\big(\hat{y} \in \mathcal{P}^h \wedge y^* \in \mathcal{D}^*\big) \implies \varphi^h(\hat{y}, y^*) \geq 0.$$

Proof. If we assign to $y^* \in \mathbb{R}^q$ the vector $y^{*h} := (y_1^*, \dots, y_{q-1}^*, 0)^T$, we get

$$\varphi^h(\hat{y}, y^*) = \varphi(\hat{y}, y^{*h})$$

from the definitions of φ and φ^h. Since $y^{*h} \in \mathcal{D}^{*h}$ (for $y^* \in \mathcal{D}^*$), we obtain $\varphi(\hat{y}, y^{*h}) \geq 0$ from Theorem 4.40 applied to (P^h) and (D^{*h}). □

Corollary 4.64. *Let the feasible sets S and T of, respectively,* (P) *and* (D^*) *be nonempty. Then*

$$\big(\forall y^* \in \mathcal{D}^* : \varphi^h(\hat{y}, y^*) \geq 0\big) \implies \hat{y} \in \mathcal{P}^h.$$

Proof. For $\hat{y}^* \in \mathcal{D}^{*h}$, we have $\hat{y}_q^* \leq 0$ and thus

$$\begin{aligned}\varphi(\hat{y},\hat{y}^*) &= \sum_{i=1}^{q-1} \hat{y}_i \hat{y}_i^* + \hat{y}_q \left(1 - \sum_{i=1}^{q-1} \hat{y}_i^*\right) - \hat{y}_q^* \\ &\geq \sum_{i=1}^{q-1} \hat{y}_i \hat{y}_i^* + \hat{y}_q \left(1 - \sum_{i=1}^{q-1} \hat{y}_i^*\right) \\ &= \varphi^h(\hat{y},\hat{y}^*).\end{aligned}$$

If $\varphi^h(\hat{y}, y^*) \geq 0$ is satisfied for all $y^* \in \mathcal{D}^*$, we also have $\varphi^h(\hat{y},\hat{y}^*) \geq 0$ for all $\hat{y}^* \in \mathcal{D}^{*h}$. Hence $\varphi(\hat{y},\hat{y}^*) \geq 0$ for all $\hat{y}^* \in \mathcal{D}^{*h}$. Theorem 4.41 applied to (P^h) and (D^{*h}) yields $\hat{y} \in \mathcal{P}^h$. □

4.7 Identifying faces of minimal vectors

Geometric duality can help us to decide whether or not a face of $\mathcal{P}$ consists completely of minimal elements. The lower image $\mathcal{D}^*$ of the geometric dual is sometimes easier to analyze, for instance, in the very important 3-dimensional case. The reason is that the recession cone of $\mathcal{D}^*$ has a simple structure, namely $\mathcal{D}^*_\infty = -K$. Therefore, a suitable projection on $\mathbb{R}^{q-1}$ may already contain enough information.

Recalling that every proper face of $\mathcal{P}$ consists of weakly minimal vectors, we get a decision whether or not a face contains weakly minimal elements being not minimal.

We start with a lemma, where we set

$$\Lambda := \left\{ w \in \mathbb{R}^q \,|\, w > 0 \;\wedge\; e^T w = 1 \right\}.$$

Lemma 4.65. *For a nonempty set F the following statements are equivalent:*

(i) *F is a proper face of $\mathcal{P}$ containing only minimal elements,*
(ii) $\exists w \in \Lambda,\ \exists \gamma \in \mathbb{R} :\ \left[\forall y \in \mathcal{P} : w^T y \geq \gamma\right] \;\wedge\; \left[\forall z \in F :\ w^T z = \gamma\right].$

Proof. By Theorem 4.2, $\bar{y} \in \operatorname{Min} \mathcal{P}$ is equivalent to

$$\exists w \in \Lambda,\ \exists \gamma \in \mathbb{R} :\ \left[\forall y \in \mathcal{P} :\ w^T y \geq \gamma\right] \;\wedge\; w^T \bar{y} = \gamma. \tag{4.22}$$

(i) $\Rightarrow$ (ii). Choose some $\bar{y} \in \operatorname{ri} F$. By Corollary 4.25, we have $\bar{y} \in \operatorname{Min} \mathcal{P}$ and thus (4.22) holds. Assume that there is some $\tilde{y} \in F$ such that $w^T \tilde{y} > \gamma$. As $\bar{y} \in \operatorname{ri} F$, by (Rockafellar, 1972, Theorem 6.4), there exists some $\mu > 1$ such that $\hat{y} := (1-\mu)\tilde{y} + \mu \bar{y} \in F$. It follows that $w^T \hat{y} < \gamma$, which is a contradiction.

(ii) $\Rightarrow$ (i). Clearly, F is a proper face of $\mathcal{P}$. From Corollary 4.25 we get $F \subseteq \operatorname{Min} \mathcal{P}$. □

The set

$$\Delta := \left\{ y^* \in \mathbb{R}^q |\ w(y^*) \in \mathbb{R}^q_+ \right\}$$

is a polyhedron with nonempty interior and recession cone $K \cup (-K) = \mathbb{R} \cdot (0,0,\dots,0,1)^T$. We have

$$\operatorname{int} \Delta = \left\{ y^* \in \mathbb{R}^q |\ w(y^*) \in \operatorname{int} \mathbb{R}^q_+ \right\}$$

and

$$\operatorname{bd} \Delta = \left\{ y^* \in \mathbb{R}^q |\ w(y^*) \in \operatorname{bd} \mathbb{R}^q_+ \right\}.$$

The lower image $\mathcal{D}^*$ of the dual problem (D^*) is contained in Δ. The next theorem states that K-maximal faces belonging to $\operatorname{bd} \Delta$ refer to faces of $\mathcal{P}$ that contain elements being not minimal (but only weakly minimal). An example is given in Figure 4.6.

Theorem 4.66. *For every proper face F of $\mathcal{P}$, the following is equivalent:*

(i) $F \subseteq \operatorname{Min} \mathcal{P}$,
(ii) $\Psi^{-1}(F) \cap \operatorname{int} \Delta \neq \emptyset$.

Proof. If F is a proper face of $\mathcal{P}$, then (i) is equivalent to statement (ii) of Lemma 4.65, that is,

$$\exists w \in \Lambda,\ \exists \gamma \in \mathbb{R}:\ \left[\forall y \in \mathcal{P} : w^T y \geq \gamma\right] \ \wedge\ \left[\forall z \in F:\ w^T z = \gamma\right]. \qquad (4.23)$$

(i) $\Rightarrow$ (ii). Setting $y^* := (w_1, w_2, \dots, w_{q-1}, \gamma)^T$, we get $F \subseteq H(y^*) \cap \mathcal{P} =: \bar{F}$, where $\bar{F}$ is a proper face of $\mathcal{P}$. Corollary 4.25 and Lemma 4.51 yield $y^* \in \mathcal{D}^*$. Using (4.17), we get $y^* \in H^*(y)$ for all $y \in F$. Hence $y^* \in \bigcap_{y \in F} H^*(y) \cap \mathcal{D}^* = \Psi^{-1}(F)$. As $w(y^*) > 0$, we have $y^* \in \operatorname{int} \Delta$ and thus (ii) holds.

(ii) $\Rightarrow$ (i). Let $y^* \in \Psi^{-1}(F) \cap \operatorname{int} \Delta \neq \emptyset$, then (4.23) holds for $w := w(y^*) \in \Lambda$ and $\gamma := y^*_q$. □

We proceed with a direct consequence of Theorem 4.66.

Corollary 4.67. *Every vertex of $\mathcal{P}$ is minimal.*

Proof. This follows immediately from Theorem 4.66, but a short direct proof can be given as follows: Let y be a vertex of $\mathcal{P} = P[S] + \mathbb{R}^q_+$ and assume that y is not minimal. Hence, there exists some $z \in \left(\{y\} - \mathbb{R}^q_+ \setminus \{0\}\right) \cap \mathcal{P}$, i.e.,

$$y \in \{z\} + \mathbb{R}^q_+ \setminus \{0\} \subseteq P[S] + \mathbb{R}^q_+ + \left(\mathbb{R}^q_+ \setminus \{0\}\right) = P[S] + \mathbb{R}^q_+ \setminus \{0\}.$$

Therefore, there is some $\bar{x} \in S$ and some $\bar{y} \in \mathbb{R}^q_+ \setminus \{0\}$ such that $y = P\bar{x} + \bar{y} \in \mathcal{P}$. Hence the points $y - \bar{y}$ and $y + \bar{y}$ belong to $\mathcal{P}$ and $y = \frac{1}{2}(y - \bar{y}) + \frac{1}{2}(y + \bar{y})$. This contradicts y being a vertex of $\mathcal{P}$. □

The following result is a dual counterpart of Corollary 4.67.

Proposition 4.68. *Every vertex of $\mathcal{D}^*$ is K-maximal.*

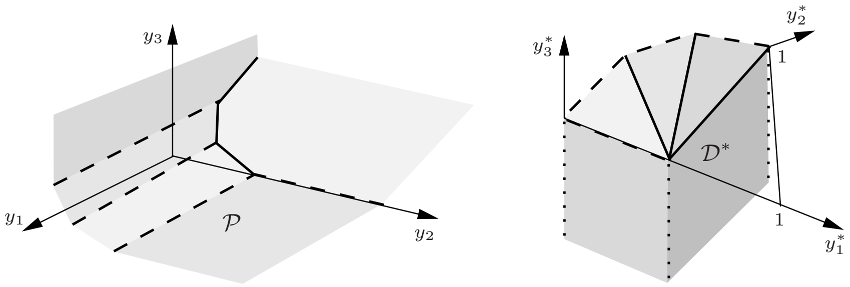

Fig. 4.6 The lower image $\mathcal{D}^*$ of the geometric dual problem on the right has the five vertices $(0,0,0)^T$, $(0,\frac{1}{3},1)^T$, $(0,\frac{2}{3},1)^T$, $(0,1,0)^T$ and $(\frac{1}{2},0,0)^T$. All of them belong to $\operatorname{bd}\Delta$. Therefore, the five facets of the upper image $\mathcal{P}$ of the primal problem are not minimal, i.e., they contain elements being not minimal. The three K-maximal facets of $\mathcal{D}^*$ all have common points with $\operatorname{int}\Delta$ (this is always true for K-maximal facets). This implies that the three vertices of $\mathcal{P}$, namely $(0,3,0)^T$, $(-1,1,1)^T$ and $(-3,0,3)^T$, are minimal (note that vertices of $\mathcal{P}$ are always minimal by Corollary 4.67). Furthermore, we observe that the four dashed K-maximal edges of $\mathcal{D}^*$ belong to $\operatorname{bd}\Delta$. The corresponding edges of $\mathcal{P}$ (also dashed) are therefore not minimal. The other three K-maximal edges of $\mathcal{D}^*$ have common points with $\operatorname{int}\Delta$. The corresponding three edges of $\mathcal{P}$ are therefore minimal. The union of these three edges is the set $\operatorname{Min}\mathcal{P}$.

Proof. This follows from the geometric duality theorem and the fact that every facet of $\mathcal{P}$ consists of weakly minimal elements.

A simple direct proof can be given as follows: Assume there is some vertex $\bar{y}^* \in \mathcal{D}^*$ which is not K-maximal. Then there exists some $y^* \in \left(\{\bar{y}^*\}+K\right)\cap\mathcal{D}^*$ with $y^* \neq \bar{y}^*$. We get $\bar{y}^* = \frac{1}{2}y^* + \frac{1}{2}(\bar{y}^* - (y^* - \bar{y}^*))$, where $y^* \in \mathcal{D}^*$ and $(\bar{y}^* - (y^* - \bar{y}^*)) \in \mathcal{D}^*$ are not equal. This contradicts the fact that $\bar{y}^*$ is a vertex. □

4.8 Notes on the literature

Duality for multiple objective linear programs seems to have its origin in (Gale *et al.*, 1951) followed by (Kornbluth, 1974; Rödder, 1977; Isermann, 1974a, 1978b,a; Brumelle, 1981; Jahn, 1983). More recent expositions are contained in the books by Jahn (1986, 2004), Luc (1988) Göpfert and Nehse (1990), Ehrgott (2000, 2005) and Boţ *et al.* (2009).

While many results from scalar optimization can be generalized to the vectorial case without any restrictions, some difficulties occurred in duality for linear vector optimization, such as a duality gap in the case $b = 0$ (where b is the right-hand side of the inequality constraints). In (Hamel *et al.*, 2004), this duality gap could be closed by a set-valued approach. In (Löhne and Tammer, 2007; Heyde *et al.*, 2009b,a), this set-valued approach has been revisited from a lattice theoretical point of view. The aim of the mentioned papers was to work in an appropriate complete lattice in order to have a duality theory which can be formulated along the lines of the scalar duality theory. Another goal (Heyde *et al.*, 2009a) is to have a "simple" dual problem, which is at least not more complicated than the primal problem. A further aspect is that the dual variables in the mentioned papers are vectors rather than matrices, which is beneficial for an economic interpretation (see Heyde *et al.*, 2009a) and for computational aspects. Note that there are connections between matrices and vectors as dual variables as pointed out, for instance, by (Jahn, 2004, Theorems 2.3 and 7.5). Further insights could be obtained from the vectorial interpretation of the hyperplane-valued dual problem, the geometric duality, which is due to Heyde and Löhne (2008). There are connections to the weight space decomposition, which is a common method in the literature (see e.g. Benson and Sun, 2000).

The results and concepts in Section 4.1 are classic (see e.g. Ehrgott, 2000, 2005). The pair ($\mathrm{P}_2(y)$, ($\mathrm{D}_2(y)$) of dual problems occurred, for instance, in (Isermann, 1974b). Note that these problems are related to a very common scalarization method in vector optimization (e.g. Gerstewitz (Tammer), 1983; Pascoletti and Serafini, 1984; Weidner, 1990), which is related to nonconvex separation theorems (Gerth (Tammer) and Weidner, 1990).

The solution concepts introduced in this chapter are related to those by Heyde and Löhne (2010). Section 4.3 as well as Section 4.4 are based on (Heyde *et al.*, 2009a). Moreover, an application to mathematical finance and a corresponding interpretation of the dual problem can be found there.

The results of Section 4.5 are due to Heyde and Löhne (2008), except some extensions and Corollary 4.55, which is taken from (Ehrgott, Löhne, and Shao, 2007).

Chapter 5
Algorithms

This chapter is devoted to algorithms to compute finitely generated solutions to (P) and (D). There are several methods in the literature that evaluate the set of all efficient vertices (and efficient extreme directions) of the feasible set. For large problems, this can be rather expensive. It is sufficient and often less expensive to compute finitely generated solutions. This means that a subset of the efficient vertices (and efficient extreme directions) is sufficient but the requirement is to get a full description of the upper and lower images, respectively. With the terminology of Chapter 2 this is just the attainment of the infimum or supremum.

If we enter in this discussion from a practical point of view, we observe that the decision maker does not need to know all the efficient solutions. The reason is that the variable space is usually of higher dimension than the objective space. Therefore, it is easier to search for the "best" efficient solution in the objective space. To this end it is sufficient to know the upper or lower images, which are fully determined by finitely generated solutions. Moreover, it is adequate to assume that a decision maker will select a preferred efficient solution based on the objective values rather than variable values.

Benson's outer approximation algorithm, which is studied in this chapter, is a method which is based on these ideas. The algorithm can be understood as a kind of primal-dual method. Geometric duality also provides inner approximations of the the upper image $\mathcal{P}$. Together this can be interpreted as the embedding of the optimal value between the values of the primal and dual objectives at feasible solutions, which is a well-known principle from scalar optimization.

The investigations of this chapter are based on the geometric duality from Section 4.5. We start with a generalization of Benson's original algorithm and continue with a dual variant. Both algorithms require solutions to the homogeneous problems as an input. Under certain boundedness assumptions, the solutions of the homogeneous problems are known. It is shown that each homogeneous problem can be transferred into a problem which satisfies these boundedness assumptions.

A. Löhne, *Vector Optimization with Infimum and Supremum*, Vector Optimization,
DOI 10.1007/978-3-642-18351-5_6,

5.1 Benson's algorithm

Benson (1998b) proposed an outer approximation algorithm in order to *"generate the set of all efficient extreme points in the outcome set"*. We consider a simplified and generalized variant of the algorithm. The main idea is to evaluate the upper image $\mathcal{P} = P[S] + \mathbb{R}^q_+$ of Problem (P) by constructing suitable cutting planes. The algorithm performs an iterative refinement of an outer approximation of $\mathcal{P}$. A sequence $(\mathcal{T}^k)$ of inclusion decreasing polyhedral supersets of $\mathcal{P}$ is constructed. The approximating supersets are evaluated in the sense that one obtains both:

- a description by vertices and directions,
- a description by inequalities.

The algorithm terminates after a finite number of steps. Then, we have $\mathcal{T}^k = \mathcal{P}$ for some k. This means that $\mathcal{P}$ have been evaluated in the same sense, i.e., one has a description by vectors/directions and by inequalities. A description of $\mathcal{P}$ by feasible elements (or directions) of the variable spaces (pre-image spaces) $\mathbb{R}^n$ and $\mathbb{R}^{m+q}$ is nothing but computing finitely generated solutions to (P) and (D*). This means that a suitable subset of S yields the vectors and directions to describe $\mathcal{P}$ and a suitable subset of T yields the inequalities describing $\mathcal{P}$.

Throughout, the feasible set S is assumed to be nonempty. Moreover, we assume that the homogeneous problems (P^h) and (D^{*h}) already have been solved. We also suppose that $\mathcal{D}^{*h}$ has a nonempty interior. This is equivalent to 0 being a vertex of $\mathcal{P}^h$, which is an easy consequence of the geometric duality theorem, see Section 4.5. This is illustrated in Figure 5.1.

An important special case is the one where $\mathcal{P}$ is $\mathbb{R}^q_+$-bounded below, i.e., there exists some $l \in \mathbb{R}^q$ such that $l \leq y$ for all $y \in \mathcal{P}$. In this case, solutions to (P^h) and (D^{*h}) are known and $\mathcal{D}^{*h}$ has always a nonempty interior. The details of this case are discussed in Section 5.3.

The outer approximation of $\mathcal{P}$ is subsequently denoted by $\mathcal{T}$ or by $\mathcal{T}^k$ if we intend to indicate that it was constructed in iteration k. The polyhedral set $\mathcal{T}$ is described in two different ways.

Definition 5.1. Let $\mathcal{T} \subseteq \mathbb{R}^q$ be a polyhedral set with $\mathcal{T}_\infty \supseteq \mathbb{R}^q_+$. A pair of finite sets

$$(\mathcal{T}^p, \hat{\mathcal{T}}^p) = \big(\{y^1, \dots, y^r\}, \{\hat{y}^1, \dots, \hat{y}^s\}\big) \in 2^{\mathbb{R}^q} \times 2^{\mathbb{R}^q \setminus \{0\}},$$

where $\mathcal{T}^p$ is nonempty $(r \geq 1)$ but $\hat{\mathcal{T}}^p$ is allowed to be empty $(s \geq 0)$, is called a *primal representation* of $\mathcal{T}$ if $\mathcal{T}$ is re-obtained from $(\mathcal{T}^p, \hat{\mathcal{T}}^p)$ by

$$\mathcal{T} = \operatorname{co} \mathcal{T}^p + \operatorname{cone} \hat{\mathcal{T}}^p + \mathbb{R}^q_+.$$

A nonempty finite set

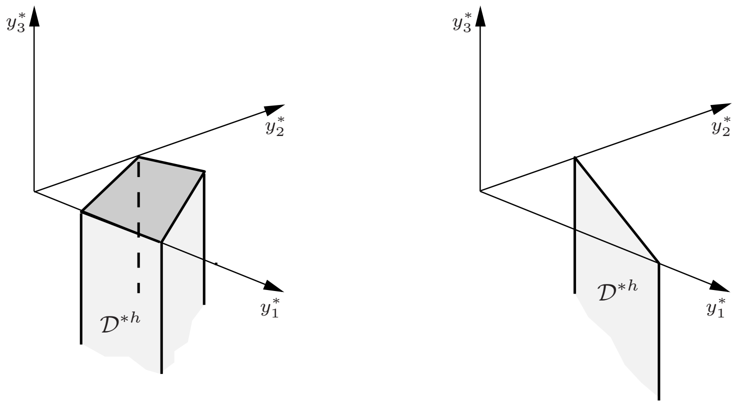

Fig. 5.1 Illustration of the assumption $\operatorname{int} \mathcal{D}^{*h} \neq \emptyset$ for the case $q = 3$. In the example on the left, the assumption is fulfilled. A K-maximal face of dimension $q-1$ exists. By geometric duality, this corresponds to a 0-dimensional face of $\mathcal{P}^h$, that is, a vertex of $\mathcal{P}^h$ exists (of course, this is the origin). On the right, the assumption is not fulfilled. The largest (with respect to inclusion) K-maximal face is of dimension $q - 2$. By geometric duality, the smallest (with respect to inclusion) weakly minimal face of $\mathcal{P}^h$ has dimension 1. As every proper face of $\mathcal{P}^h$ is weakly minimal, this means that $\mathcal{P}^h$ has no vertex.

$$\mathcal{T}^d = \left\{y^{*1}, \ldots, y^{*t}\right\} \quad (t \geq 1)$$

is called a *dual representation* of $\mathcal{T}$ if $\mathcal{T}$ is re-obtained from $\mathcal{T}^d$ by

$$\mathcal{T} = \left\{y \in \mathbb{R}^q |\; \varphi(y, y^{*1}) \geq 0\,, \ldots, \varphi(y, y^{*t}) \geq 0\right\},$$

where $\varphi : \mathbb{R}^q \times \mathbb{R}^q \to \mathbb{R}$ is the coupling function of the geometric duality theorem, see Section 4.5.

Recall that the inequality $\varphi(y, y^*) \geq 0$ can be expressed as $w(y^*)^T y \geq y_q^*$, where

$$w(y^*) = \left(y_1^*, \ldots, y_{q-1}^*, 1 - \sum_{i=1}^{q-1} y_i^*\right)^T.$$

The connections to finitely generated solutions can be easily seen.

Proposition 5.2. *The pair of sets* $(\bar{S}, \bar{S}^h)$ *is a finitely generated solution to* (P) *if and only if*

$$\left(P[\bar{S}], P[\bar{S}^h]\right)$$

is a primal representation of $\mathcal{P}$*,* $\bar{S} \subseteq \operatorname{Eff}(\mathrm{P})$ *and* $\bar{S}^h \subseteq \operatorname{Eff}(\mathrm{P}^h)$*.*

The set $\bar{T}$ *is a finitely generated solution to* (D^*) *if and only if* $D^*[\bar{T}]$ *is a dual representation of* $\mathcal{P}$ *and* $\bar{T} \subseteq \operatorname{Eff}(\mathrm{D}^*)$*.*

Proof. The conditions (i) - (iv) in Definition 4.4 are equivalent to $\bar{S} \subseteq \operatorname{Eff}(\mathrm{P})$, $\bar{S}^h \subseteq \operatorname{Eff}(\mathrm{P}^h)$, and to $\mathcal{T}^p := P[\bar{S}]$ and $\hat{\mathcal{T}}^p := P[\bar{S}^h]$ being finite sets. Because of

$$\operatorname{co} P[\bar{S}] + \operatorname{cone} P[\bar{S}^h] + \mathbb{R}^q_+ = \operatorname{co} \mathcal{T}^p + \operatorname{cone} \hat{\mathcal{T}}^p + \mathbb{R}^q_+,$$

(v) is equivalent to $\mathcal{P} \subseteq \operatorname{co} \mathcal{T}^p + \operatorname{cone} \hat{\mathcal{T}}^p + \mathbb{R}^q_+$. Since $\bar{S} \subseteq S$ and $\bar{S}^h \subseteq S^h$, the last inclusion holds with equality.

Definition 4.56 (i), (ii) is equivalent to $\bar{T} \subseteq \operatorname{Eff}(\mathrm{D}^*)$ and $\mathcal{T}^d := D^*[\bar{T}]$ being a finite set. By $\bar{T} \subseteq T$, condition (iii) in Definition 4.56 is equivalent to $\mathcal{D}^* = \operatorname{co} D^*[\bar{T}] - K$. This is true if and only if the vertices of $\mathcal{D}^*$ are contained in $D^*[\bar{T}]$ (one direction is obvious, for the other one use the definition of a vertex). The statement now follows from the geometric duality theorem. □

The following result shows that the recession cone of $\mathcal{T}$ can be easily obtained from a dual representation by turning to the homogeneous inequalities.

Proposition 5.3. *Let $\mathcal{T} \subseteq \mathbb{R}^q$ be a polyhedron such that $\mathcal{T}_\infty \supseteq \mathbb{R}^q_+$ and let $\mathcal{T}^d$ be a dual representation of $\mathcal{T}$. Then the recession cone of $\mathcal{T}$ is given by*

$$\mathcal{T}_\infty = \left\{ y \in \mathbb{R}^q |\ \forall y^* \in \mathcal{T}^d : w(y^*)^T y \geq 0 \right\}.$$

Proof. We have

$$\mathcal{T} = \bigcap_{y^* \in \mathcal{T}^d} \left\{ y \in \mathbb{R}^q |\ w(y^*)^T y \geq y^*_q \right\} \neq \emptyset.$$

The recession cone of the closed halfspace

$$H(y^*) = \left\{ y \in \mathbb{R}^q |\ w(y^*)^T y \geq y^*_q \right\}$$

is the closed halfspace

$$(H(y^*))_\infty = \left\{ y \in \mathbb{R}^q |\ w(y^*)^T y \geq 0 \right\}.$$

It follows (see e.g. Rockafellar, 1972, Corollary 8.3.3)

$$\mathcal{T}_\infty = \left(\bigcap_{y^* \in \mathcal{T}^d} H(y^*) \right)_\infty = \bigcap_{y^* \in \mathcal{T}^d} (H(y^*))_\infty ,$$

which implies the desired statement. □

The algorithm is now explained geometrically. The steps can be illustrated as a construction of the upper image $\mathcal{P}$ of Problem (P). The line numbers given in the following description refer to Algorithm 1 below:

Input and variables. Besides the data of (P), the input contains a (possibly empty) set $\bar{S}^h$ such that $(\{0\}, \bar{S}^h)$ is a finitely generated solution to (P^h). This means that the extreme directions of $\mathcal{P}$ have to be known. In fact, they are contained in the set $P[\bar{S}^h] \cup \{e^1, \ldots, e^q\}$. The set $\bar{S}^h$ is not changed by the algorithm and is part of the finitely generated solution to (P) computed by the algorithm. This is due to the feature that $\mathcal{P}$ and $\mathcal{P}^h$ have the same extreme directions, which follows from the fact that $\mathcal{P}^h$ is the recession cone of $\mathcal{P}$. Moreover, the input data contain a finitely generated solution $\bar{T}^h$ to (D^{*h}). Geometrically, this means that an inequality description of $\mathcal{P}^h$ is known at the beginning. The most important variables are the following:

$\bar{T}$ ... an array of vectors in $\mathbb{R}^{m+q}$ to construct a finitely generated solution to (D^*),
$\bar{S}$ ... an array of vectors in $\mathbb{R}^n$ to construct the vector part of a finitely generated solution to (P),
$\mathcal{T}^p$ ... an array of vectors in $\mathbb{R}^q$ to store the first part of a primal representation of the current outer approximation,
$\mathcal{T}^d$ ... an array of vectors in $\mathbb{R}^q$ to store a dual representation of the current outer approximation.

1. Initialization. First, an interior point $\hat{p}$ of $\mathcal{P}$ is determined (line 08). To this end a feasible point $x \in S$ is determined and the vector $e = (1, \dots, 1)^T$ is added to Px.

The algorithm constructs a set $\bar{T}$ such that $\mathcal{T}^d := D^*[\bar{T}]$ is a dual representation of an initial outer approximation $\mathcal{T} \supseteq \mathcal{P}$. The set $\bar{T}$ is obtained from the finitely generated solution $\bar{T}^h$ to (D^{*h}) in the following way:

$D^*[\bar{T}^h]$ provides a dual representation of $\mathcal{P}^h$. The halfspaces generated by the corresponding inequalities are shifted appropriately such that they contain $\mathcal{P}$. As a consequence, the desired set $\bar{T}$ is obtained. Note that by Proposition 5.3, $\mathcal{P}^h$ and the first outer approximation $\mathcal{T}$ have an identical recession cone. The same is true for every other outer approximation which is contained in $\mathcal{T}$. As a consequence, the direction part of the primal representation is known for every outer approximation $\mathcal{T}$. It is therefore sufficient to compute the vertices of $\mathcal{T}$ in order to get a primal representation.

2. Iteration. The dual representation of $\mathcal{T}$ is used to get a primal representation. Because of the above remarks on the recession cones, it remains to compute the vertices of the outer approximation $\mathcal{T}$ in each step (line 06).

It is then tested whether the vertices of $\mathcal{T}$ belong to $\mathcal{P}$. To this end, the linear program ($\mathrm{P}_2(t)$) is solved for the vertices t of $\mathcal{T}$ (line 11). The optimal value z of ($\mathrm{P}_2(t)$) is zero if and only if $t \in \mathcal{P}$. If all vertices of an outer approximation $\mathcal{T}$ belong to $\mathcal{P}$, we have $\mathcal{T} = \mathcal{P}$. Finitely generated solutions to (P) and (D^*) can easily be obtained from the primal and dual representations of $\mathcal{P}$ by solving ($\mathrm{P}_2(t)$) and ($\mathrm{D}_2(s)$), respectively (lines 11 and 17). If a vertex $t \in \mathcal{T}$ is detected that does not belong to $\mathcal{P}$, then this vertex is used to construct a better (smaller) outer approximation. Let us explain this for the k-th iteration:

We have a vertex t of $\mathcal{T}^k$ not belonging to $\mathcal{P}$. A point $s \in \operatorname{bd} \mathcal{P}$ on the line between $\hat{p}$ and t is determined. To this end, we solve (see line 15) the linear program

$$\max \alpha \quad \text{subject to} \begin{cases} Bx \geq b \\ \alpha t + (1-\alpha)\hat{p} \geq Px. \end{cases} \tag{R(t)}$$

We next compute a supporting hyperplane of $\mathcal{P}$ that contains the point s. This hyperplane is obtained as $H(D^*(u,w)) = \{y \in \mathbb{R}^q |\ w^T y = b^T u\}$, where (u,w) is a solution to the linear program $(\mathrm{D}_2(s))$ (see line 17) and H is defined in Section 4.5.

Proposition 5.4. *Let $s \in \mathrm{wMin}\,\mathcal{P}$. Then there exists a solution to* $(\mathrm{D}_2(s))$ *and for each solution (u,w) to* $(\mathrm{D}_2(s))$*, $H(D^*(u,w))$ is a supporting hyperplane to $\mathcal{P}$ containing s.*

Proof. By Theorem 4.20 there exists a solution (u,w) to $(\mathrm{D}_2(s))$ such that $b^T u = s^T w$. Of course, the latter equality is also valid for any other optimal solution to $(\mathrm{D}_2(s))$. For arbitrary $s \in \mathcal{P}$, there exists some $x \in S$ such that $s \geq Px$. Hence $(x,0)$ is feasible for $(\mathrm{P}_2(s))$ and duality between $(\mathrm{P}_2(s))$ and $(\mathrm{D}_2(s))$ implies that $y^T w \geq b^T u$ for all $y \in \mathcal{P}$. Hence $H(D^*(u,w)) = \{y \in \mathbb{R}^q |\ y^T w = b^T u\}$ is a supporting hyperplane to $\mathcal{P}$ containing s. $\square$

We now append the solution (u,w) of $(\mathrm{D}_2(s))$ to $\bar{T}$ (line 18). As $D^*[\bar{T}]$ yields a dual representation of the new outer approximation $\mathcal{T}^{k+1}$, appending (u,w) to $\bar{T}$ can be interpreted geometrically as a cut of $\mathcal{T}^k$ by the hyperplane $H(D^*(u,w))$, that is, $\mathcal{T}^{k+1} := \mathcal{T}^k \cap \{y \in \mathbb{R}^q |\ w^T y \geq b^T u\}$.

Output. The output consists of a finitely generated solution $(\bar{S}, \bar{S}^h)$ to (P) and a finitely generated solution $\bar{T}$ to (D^*).

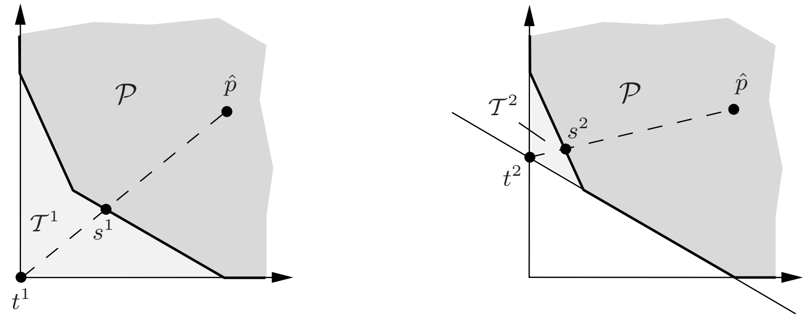

Fig. 5.2 Illustration of Benson's algorithm. In the first step on the left, the vertex t^1 of the outer approximation $\mathcal{T}^1$ does not belong to $\mathcal{P}$. The point $s^1 \in \mathrm{wMin}\,\mathcal{P}$ on the line between t^1 and $\hat{p}$ is computed. The supporting hyperplane of $\mathcal{P}$ containing the point s^1 yields a smaller approximation $\mathcal{T}^2$, which is shown in the figure on the right. Again a vertex t^2 of $\mathcal{T}^2$ not belonging to $\mathcal{P}$ is computed. One proceeds in the same way. If all vertices of the outer approximation $\mathcal{T}^k$ belong to $\mathcal{P}$, the algorithm stops.

The algorithm uses two subroutines. The command *solve*( ) solves a linear program. A solution is returned. For this purpose, a large variety of software is available. The routine *vert*( ) returns the vertices of a polyhedron $\mathcal{T}$, which is given by a dual representation. This means that a finite set of inequalities is known, whose solution set is just $\mathcal{T}$. For this purpose, a method called "Vertex Enumeration by Adjacency Lists" can be used, see Chen *et al.* (1991). Two variants are available. The offline variant has to be used in the first iteration. In the subsequent steps the online variant should be more efficient. The online variant uses the results from the previous steps.

The algorithm is now given in a pseudo code.

Algorithm 1.

Input:
A, b, P (data of Problem (P));
a finitely generated solution $\left(\{0\}, \bar{S}^h\right)$ to (P^h);
a finitely generated solution $\bar{T}^h$ to (D^{*h});
Output:
a finitely generated solution $(\bar{S}, \bar{S}^h)$ to (P);
a finitely generated solution $\bar{T}$ to (D*);

begin
$\hat{p} \leftarrow P\,(\text{solve}(\text{P}_1(0))) + e$;
$\bar{T} \leftarrow \left\{\left(\text{solve}(\text{D}_1(w)), w\right) \middle| \, (u, w) \in \bar{T}^h\right\}$;
repeat
$\mathcal{T}^d \leftarrow \left\{D^*(u, w) \middle| \, (u, w) \in \bar{T}\right\}$;
$\mathcal{T}^p \leftarrow \text{vert}\,(\mathcal{T}^d)$;
$\bar{S} \leftarrow \emptyset$;
for $i = 1$ to $|\mathcal{T}^p|$ do
begin
$t \leftarrow \mathcal{T}^p[i]$;
$(x, z) \leftarrow \text{solve}(\text{P}_2(t))$;
$\bar{S} \leftarrow \bar{S} \cup \{x\}$;
if $z \neq 0$ then
begin
$(x, \alpha) \leftarrow \text{solve}(\text{R}(t))$;
$s \leftarrow \alpha t + (1 - \alpha)\hat{p}$;
$(u, w) \leftarrow \text{solve}(\text{D}_2(s))$;
$\bar{T} \leftarrow \bar{T} \cup \{(u, w)\}$;
break;
end;
end;
until $z = 0$;
end.

Further details of the algorithm are discussed in the proofs of the next two theorems.

Theorem 5.5. *Algorithm 1 works correctly.*

Proof. In line 02, we compute a solution x to $(\mathrm{P}_1(0))$ in order to obtain some $x \in S$ (note that S was assumed to be nonempty). As $e \in \operatorname{int} \mathbb{R}^q_+$ we obtain $\hat{p} = Px + e \in P[S] + \operatorname{int} \mathbb{R}^q_+ = \operatorname{int}(P[S] + \mathbb{R}^q_+) = \operatorname{int} \mathcal{P}$.

Since $\bar{T}^h$ is a finitely generated solution to (D^{*h}), $\bar{T}^h$ is a nonempty subset of the feasible set T. Thus, if $(u, w) \in \bar{T}^h \subseteq T$, u is feasible for $(\mathrm{D}_1(w))$. As the feasible set S of $(\mathrm{P}_1(w))$ is nonempty, in line 03 a solution to $(\mathrm{D}_1(w))$ exists (classical duality theory).

Let us show that $\mathcal{T}$ in line 05, which is given by its dual representation $\mathcal{T}^d$, is an outer approximation of $\mathcal{P}$, i.e., $\mathcal{P} \subseteq \mathcal{T}$. Indeed, if $y \in \mathcal{P}$ and $y^* \in \mathcal{T}^d = D^*[\bar{T}] \subseteq D^*[T]$, weak duality implies $\varphi(y, y^*) \geq 0$ (compare Theorem 4.40). Hence $y \in \mathcal{T}$. Note that for every $y^* \in \mathcal{T}^d$, there is some $y \in \mathcal{P}$ such that $\varphi(y, y^*) = 0$. This is a consequence of u being a solution to $(\mathrm{D}_1(w))$ for every $(u, w) \in \mathcal{T}^d$, compare Theorem 4.19.

The set $\mathcal{T}^p$ in line 06 is nonempty, i.e., there exists a vertex of $\mathcal{T}$. Indeed, we assumed $\operatorname{int} \mathcal{D}^{*h} \neq \emptyset$. As a consequence we obtain that 0 is a vertex of $\mathcal{P}^h$. Geometrically, $\mathcal{T}$ is obtained from $\mathcal{P}^h$ by a shift of the supporting hyperplanes. Thus, $\mathcal{T}$ contains at least one vertex (this follows more precisely by (Rockafellar, 1972, Corollary 18.5.3) and Theorem 5.3).

In line 11 we solve $(\mathrm{P}_2(t))$. Taking $x \in S$ and z sufficiently large, we obtain that (x, z) is feasible for $(\mathrm{P}_2(t))$. In particular, an optimal solution (x, z) to $(\mathrm{P}_2(t))$ always exists. We have either $t \in \operatorname{bd} \mathcal{P} = \operatorname{wMin} \mathcal{P}$ or $t \notin \mathcal{P}$. In the first case we have $z = 0$ by Theorem 4.20 and in the second case we have $z > 0$. This means that the lines 14-20 are performed in the case where $t \notin \mathcal{P}$. Thus, $(\mathrm{R}(t))$ has a solution and we obtain some $s \in \operatorname{bd} \mathcal{P} = \operatorname{wMin} \mathcal{P}$. The linear program $(\mathrm{D}_2(s))$ in line 17 has a solution by Proposition 5.4. In line 19 we break the innermost loop (lines 08-21).

The outer loop (lines 04-22) terminates in the case where for every $t \in \mathcal{T}^p$, $z = 0$ is obtained in line 11. This means $\mathcal{T}^p \subseteq \mathcal{P}$, which implies $\mathcal{T} = \mathcal{P}$. We have $P[\bar{S}] = \mathcal{T}^p$ and $D^*[\bar{T}] = \mathcal{T}^d$.

From Proposition 5.2 we obtain that $(\bar{S}, \bar{S}^h)$ and $\bar{T}$ are finitely generated solutions to (P) and (D^*), respectively. The assumption $\bar{S} \subseteq \operatorname{Eff}(\mathrm{P})$ in Proposition 5.2 follows from the fact that every vertex of $\mathcal{P}$ is minimal (Corollary 4.67). Indeed, if $x \in \bar{S}$, then $(x, 0)$ was obtained as an optimal solution to $(\mathrm{P}_2(t))$ for a vertex t of $\mathcal{P}$ (line 11). We have $Px \leq t + 0 \cdot e$ and the minimality of t implies $Px = t$. The assumption $\bar{S}^h \subseteq \operatorname{Eff}(\mathrm{P}^h)$ follows from the input because $(\{0\}, \bar{S}^h)$ is a finitely generated solution to (P^h). The assumption $\bar{T} \subseteq \operatorname{Eff}(\mathrm{D}^*)$ in Proposition 5.2 follows from Proposition 5.4 and the geometric duality theorem. □

It remains to show that the algorithm terminates after a finite number of steps.

Theorem 5.6. *Algorithm 1 is finite.*

Proof. Since $\hat{p} \in \operatorname{int} \mathcal{P}$, the point $s^k \in \mathcal{P}$ computed in iteration k belongs to $\operatorname{int} \mathcal{T}^k$. We have $\mathcal{T}^{k+1} := \mathcal{T}^k \cap \{y \in \mathbb{R}^q |\ \varphi\big(y, D^*(u^k, w^k)\big) \geq 0\}$, where (u^k, w^k) is the solution to $(\mathrm{D}_2(s^k))$ computed in iteration k. By Proposition 5.4 we know that $F := \{y \in \mathcal{P} |\ \varphi\big(y, D(u^k, w^k)\big) = 0\}$ is a face of $\mathcal{P}$ with $s^k \in F$, where $F \subseteq \operatorname{bd} \mathcal{T}^{k+1}$. For the next iteration, this means that $s^{k+1} \notin F$ (because $s^{k+1} \in \operatorname{int} \mathcal{T}^{k+1}$). Therefore, s^{k+1} belongs to another face of $\mathcal{P}$. Since $\mathcal{P}$ is polyhedral, it has a finite number of faces. This proves that the algorithm is finite. □

5.2 A dual variant of Benson's algorithm

The geometric explanation of Benson's algorithm is based on a construction of the upper image $\mathcal{P}$ of the primal vector optimization problem. By similar ideas, the lower image $\mathcal{D}^*$ of the geometric dual problem can be constructed. The link between $\mathcal{P}$ and $\mathcal{D}^*$ is given by geometric duality. Therefore, a construction of $\mathcal{D}^*$ leads to the same results as a construction of $\mathcal{P}$. A dual algorithm based on this idea is developed in this section.

Note that geometric duality is already involved in the original (primal) algorithm. In the same way the dual algorithm is based on this theory. Again we obtain finitely generated solutions to both problems (P) and (D*).

As in Section 5.1, we assume that the primal feasible set S is nonempty and the homogeneous problems (P^h) and (D^{*h}) have already been solved. Again we suppose that $\mathcal{D}^{*h}$ has a nonempty interior.

The special case where $\mathcal{P}$ is $\mathbb{R}^q_+$-bounded below is also important for the dual algorithm. In this case solutions to (P^h) and (D^{*h}) are known and $\mathcal{D}^{*h}$ has a nonempty interior. The details of the bounded case are discussed in Section 5.3.

We now work with two different descriptions of the lower image $\mathcal{D}^*$ and its outer approximation, which is denoted by $\mathcal{T}$ or by $\mathcal{T}^k$ if the iteration step shall be indicated. Recall that $\Delta := \{y^* \in \mathbb{R}^q |\ w(y^*) \geq 0\}$.

Definition 5.7. Let $\mathcal{T}$ be a polyhedral set with $\mathcal{T} \subseteq \Delta$ and $\mathcal{T}_\infty = -K$. A nonempty finite set

$$\mathcal{T}^p = \{y^{*1}, \dots, y^{*t}\} \in 2^{\mathbb{R}^q}$$

is called a *primal representation* of $\mathcal{T}$ if $\mathcal{T}$ is re-obtained from $\mathcal{T}^p$ by

$$\mathcal{T} = \operatorname{co} \mathcal{T}^p - K.$$

A pair of finite sets

$$\mathcal{T}^d = \big(\{y^1, \dots, y^r\}, \{\hat{y}^1, \dots, \hat{y}^s\}\big) \in 2^{\mathbb{R}^q} \times 2^{\mathbb{R}^q \setminus \{0\}},$$

where $r \geq 1$ and $s \geq 0$ (i.e. the first set is nonempty but the second set can be empty), is called a *dual representation* of $\mathcal{T}$ if $\mathcal{T}$ is re-obtained from $\mathcal{T}^d$

by

$$\mathcal{T} = \{y^* \in \Delta |\, \varphi(y^1, y^*) \geq 0, \dots, \varphi(y^r, y^*) \geq 0,\\ \varphi^h(\hat{y}^1, y^*) \geq 0, \dots, \varphi^h(\hat{y}^s, y^*) \geq 0\},$$

where $\varphi : \mathbb{R}^q \times \mathbb{R}^q \to \mathbb{R}$ is the coupling function of the geometric duality theorem, see Section 4.5, and $\varphi^h : \mathbb{R}^q \times \mathbb{R}^q \to \mathbb{R}$ is the homogeneous variant of the coupling function, see Section 4.6.

The next result shows that a primal (dual) representation of $\mathcal{P}$ is just a dual (primal) representation of $\mathcal{D}^*$. This is a consequence of geometric duality and the fact that we use the term "primal" for a representation by vectors (and directions) and the term "dual" for a representation by inequalities.

Proposition 5.8. *The following statements are equivalent:*

(i) $(\mathcal{A}, \hat{\mathcal{A}})$ *is a primal representation of* $\mathcal{P}$*;*
(ii) $(\mathcal{A}, \hat{\mathcal{A}})$ *is a dual representation of* $\mathcal{D}^*$*.*

Proof. A pair $(\mathcal{A}, \hat{\mathcal{A}})$ is a primal representation of $\mathcal{P}$ if $\mathcal{A} \subseteq \mathbb{R}^q$ is nonempty and finite, and $\hat{\mathcal{A}} \in \mathbb{R}^q \setminus \{0\}$ is finite (possibly empty) such that $\mathcal{P} = \operatorname{co} \mathcal{A} + \operatorname{cone} \hat{\mathcal{A}} + \mathbb{R}^q_+$. Equivalently, all vertices of $\mathcal{P}$ belong to $\mathcal{A}$ and all extreme directions of $\mathcal{P}$ except the unit vectors (the extreme directions of $\mathbb{R}^q_+$) belong to $\hat{\mathcal{A}}$. The vertices of $\mathcal{P}$ correspond to supporting hyperplanes of $\mathcal{D}^*$ in K-maximal facets (Theorem 4.42) and the extreme directions of $\mathcal{P}$ correspond to supporting hyperplanes of $\mathcal{D}^*$ in vertical facets (Theorem 4.62). Taking into account also the duality assertions in Theorems 4.40 and 4.41 and Corollaries 4.63 and 4.64, we obtain the equivalent statement

$$\mathcal{D}^* = \left\{ y^* \in \Delta |\; \forall y \in \mathcal{A} :\; \varphi(y, y^*) \geq 0 \;\;\wedge\;\; \forall \hat{y} \in \hat{\mathcal{A}} :\; \varphi^h(\hat{y}, y^*) \geq 0 \right\}.$$

Note that the inequalities describing Δ correspond to the extreme directions of $\mathbb{R}^q_+$ (the unit vectors). The latter statement means that $(\mathcal{A}, \hat{\mathcal{A}})$ is a dual representation of $\mathcal{D}^*$. □

Proposition 5.9. *The following statements are equivalent:*

(i) $\mathcal{B}$ *is a primal representation of* $\mathcal{D}^*$*;*
(ii) $\mathcal{B}$ *is a dual representation of* $\mathcal{P}$*.*

Proof. A nonempty finite set $\mathcal{B}$ is a primal representation of $\mathcal{D}^*$ if and only if it contains all vertices of $\mathcal{D}^*$. These vertices correspond exactly to the supporting hyperplanes of $\mathcal{D}^*$ in weakly minimal facets by Theorem 4.42. Taking into account also the duality assertions in Theorems 4.40 and 4.41, we get the equivalent statement

$$\mathcal{P} = \{y \in \mathbb{R}^q |\; \forall y^* \in \mathcal{B} :\; \varphi(y, y^*) \geq 0\},$$

which means that $\mathcal{B}$ is a dual representation of $\mathcal{P}$. □

The connections to finitely generated solutions can easily be seen.

Corollary 5.10. *The pair of sets* $(\bar{S}, \bar{S}^h)$ *is a finitely generated solution to* (P) *if and only if*

$$\left(P[\bar{S}], P[\bar{S}^h]\right)$$

is a dual representation of $\mathcal{D}^*$, $\bar{S} \subseteq \mathrm{Eff}\,(\mathrm{P})$ *and* $\bar{S}^h \subseteq \mathrm{Eff}\,(\mathrm{P}^h)$.

The set $\bar{T}$ *is a finitely generated solution to* (D^*) *if and only if* $D^*[\bar{T}]$ *is a primal representation of* $\mathcal{D}^*$ *and* $\bar{T} \subseteq \mathrm{Eff}\,(\mathrm{D}^*)$.

Proof. Follows from Propositions 5.2, 5.8 and 5.9.

Let us show some properties of the first outer approximation $\mathcal{T}$ computed by the dual algorithm.

Proposition 5.11. *Let* $(\{0\}, \bar{S}^h)$ *be a finitely generated solution to* (P^h) *and let* $y \in \mathcal{P}$. *For the set*

$$\mathcal{T} := \left\{ y^* \in \Delta |\; \varphi(y, y^*) \geq 0 \wedge \forall \hat{y} \in P[\bar{S}^h] :\; \varphi^h(\hat{y}, y^*) \geq 0 \right\},$$

the following is true:

(i) $\mathcal{D}^* \subseteq \mathcal{T} \subseteq \Delta$,
(ii) $\mathcal{T}_\infty = -K$,
(iii) $\mathcal{T}$ *and* $\mathcal{D}^*$ *have the same vertical supporting hyperplanes.*

Proof. (i) Let $y^* \in \mathcal{D}^* \subseteq \Delta$. By Corollary 4.63, for every $\hat{y} \in P[\bar{S}^h] \subseteq \mathcal{P}^h$ we have $\varphi^h(\hat{y}, y^*) \geq 0$. From Theorem 4.40 we get $\varphi(y, y^*) \geq 0$. Hence $y^* \in \mathcal{T}$. The second inclusion is obvious.

(ii) By (i) we have $\mathcal{T} \subseteq \Delta$ and hence $\mathcal{T}_\infty \subseteq \Delta_\infty = K \cup (-K)$. Let $y^* \in \mathcal{T}$. Then, $\varphi^h(\hat{y}, y^* + \lambda e^q) \geq 0$ holds for all $\hat{y} \in P[\bar{S}^h]$ and all $\lambda \in \mathbb{R}$. Therefore, $y^* + \lambda e^q$ belongs to $\mathcal{T}$ if and only if

$$\varphi(y, y^* + \lambda e^q) = w(y^*)^T y - y^*_q - \lambda \geq 0.$$

This shows that $y^* + \lambda e^q \in \mathcal{T}$ for all $\lambda \leq 0$ but there exists some $\lambda > 0$ such that $y^* + \lambda e^q \notin \mathcal{T}$. Hence $\mathcal{T}_\infty = -K$.

(iii) This follows from Corollary 5.10. □

In order to construct the set $\mathcal{D}^*$ by outer approximations, we need an interior point of $\mathcal{D}^*$. As shown in the next two propositions, an interior point of $\mathcal{D}^*$ can be obtained from a finitely generated solution to (D^{*h}). Note that the assumption $\operatorname{int} \mathcal{D}^{*h} \neq \emptyset$ is used in both results.

Proposition 5.12. *Let* $\bar{T}^h = \left\{(u^1, w^1), \ldots, (u^k, w^k)\right\}$ *be a finitely generated solution to* (D^{*h}), *then*

$$\eta := \sum_{i=1}^{k} \frac{1}{k} w^i \in \operatorname{int} \left(\mathcal{D}^{*h} + K\right),$$

and $e^T \eta = 1$.

Proof. The vector η is a convex combination with nonzero coefficients of the w-components of the elements of $\bar{T}^h$. Since $D^{*h}[\bar{T}^h]$ contains all vertices of the bounded set $D^{*h}[T]$ (which belongs to $\{y^* \in \mathbb{R}^q |\ y^*_q = 0\}$), $(\eta_1, \ldots, \eta_{q-1}, 0)$ belongs to the relative interior of $D^{*h}[T]$. As $\operatorname{int} \mathcal{D}^{*h}$ is assumed to be nonempty, we have $\dim D^{*h}[T] = q - 1$. It follows that $\eta \in \operatorname{ri} D^{*h}[T] \times \mathbb{R} = \operatorname{int}(\mathcal{D}^{*h} + K)$.

We have $e^T w^i = 1$ for all $i = \{1, \ldots, k\}$ and hence $e^T \eta = 1$. □

An interior point of $\mathcal{D}^*$ is obtained from an interior point of $\operatorname{int}\left(\mathcal{D}^{*h} + K\right)$ by solving a scalar linear program.

Proposition 5.13. *Let* $\eta \in \operatorname{int}\left(\mathcal{D}^{*h} + K\right)$ *such that* $e^T\eta = 1$ *and let* γ *be the optimal value of* $(\mathrm{P}_1(\eta))$*. Then,* $\gamma \in \mathbb{R}$ *and for every* $\mu \in (-\infty, \gamma)$ *we have* $(\eta_1, \ldots, \eta_{q-1}, \mu)^T \in \operatorname{int} \mathcal{D}^*$*.*

Proof. We have $\eta \in \operatorname{int}\left(\mathcal{D}^{*h} + K\right) = \operatorname{int}(D^{*h}[T] + \mathbb{R}\{e^q\})$. Thus there exists $u \in \mathbb{R}^m$ such that $(u, \eta) \in T$. It follows that u is feasible for $(\mathrm{D}_1(\eta))$. We assumed throughout that $S \neq \emptyset$, where S is also the feasible set of $(\mathrm{P}_1(\eta))$. From the duality theory we get $\gamma \in \mathbb{R}$. Moreover, we obtain $s^* := (\eta_1, \ldots, \eta_{q-1}, \gamma)^T \in \mathcal{D}^*$.

For $t^* \in \mathbb{R}^q_+$ we denote by $\gamma(t^*)$ the optimal value of $(\mathrm{P}_1(w(t^*)))$, which is finite on $\mathcal{D}^{*h} + K$ by the same arguments as used above. The function $\gamma : \mathbb{R}^q_+ \to \mathbb{R}$ is concave.

Let $\varepsilon > 0$ be sufficiently small such that the ball $B_\varepsilon(s^*)$ with center s^* and radius ε belongs to $\operatorname{int}(\mathcal{D}^{*h} + K)$. Consequently, $\gamma(\cdot)$ is continuous on $B_\varepsilon(s^*)$, see e.g. (Borwein and Lewis, 2000, Theorem 4.1.3) or (Rockafellar, 1972, Theorem 10.1). It follows that $\gamma(\cdot)$ is bounded on $B_\varepsilon(s^*)$. Let $\rho \in \mathbb{R}$ such that $\gamma(t^*) \geq \rho$ for all $t^* \in B_\varepsilon(s^*)$.

Setting $r^* := (\eta_1, \ldots, \eta_{q-1}, \rho - \varepsilon)^T$, we get $B_\varepsilon(r^*) \subseteq \mathcal{D}^*$. Hence $r^* \in \operatorname{int} \mathcal{D}^*$. Every proper convex combination of $s^* \in \mathcal{D}^*$ and $r^* \in \operatorname{int} \mathcal{D}^*$ belongs to $\operatorname{int} \mathcal{D}^*$, which yields the desired result. □

We proceed with a geometric explanation of the algorithm (compare also Figure 5.3):

Input and variables. The input data are exactly the same as in Algorithm 1: the data of (P), a (possibly empty) set $\bar{S}^h$ such that $(\{0\}, \bar{S}^h)$ is a finitely generated solution to (P^h) and a finitely generated solution $\bar{T}^h$ to (D^{*h}). Also, the following variables (partly with a slightly different meaning) already occurred in the primal algorithm:

$\bar{S}$ ... an array of vectors $\mathbb{R}^n$ to construct the vector part of a finitely generated solution to (P),
$\bar{T}$ ... an array of vectors $\mathbb{R}^{m+q}$ to construct a finitely generated solution to (D^*),
$\mathcal{T}^p$ ... an array of vectors in $\mathbb{R}^q$ to store a primal representation of the current outer approximation,

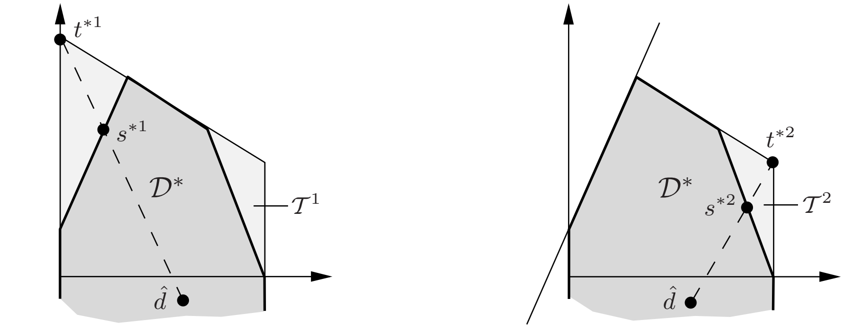

Fig. 5.3 Illustration of the dual variant of Benson's algorithm. In the first step, shown on the left, the vertex t^{*1} of the outer approximation $\mathcal{T}^1$ does not belong to $\mathcal{D}^*$. The point $s^{*1} \in \operatorname{Max}_K \mathcal{D}^*$ on the line between t^{*1} and $\hat{d}$ is computed. A cut with the supporting hyperplane of $\mathcal{D}^*$ containing the point s^{*1} yields a smaller approximation $\mathcal{T}^2$, which is shown in the figure on the right. Again a vertex t^{*2} of $\mathcal{T}^2$ not belonging to $\mathcal{D}^*$ is computed. One proceeds in the same way. If all vertices of the outer approximation $\mathcal{T}^k$ belong to $\mathcal{D}^*$, the algorithm stops.

$\mathcal{T}^d$ ... an array of vectors in $\mathbb{R}^q$ to store the first part of a dual representation of the current outer approximation.

1. Initialization. First, an interior point $\hat{d}$ of $\mathcal{D}^*$ is determined (lines 02 to 04). Simultaneously (lines 03 and 05) the algorithm constructs a singleton set $\bar{S} = \{x\}$ such that the first part of a dual representation $\mathcal{T}^d$ of an initial outer approximation $\mathcal{T} \supseteq \mathcal{D}^*$ is obtained. The second part $\hat{\mathcal{T}}^d$ is obtained from the finitely generated solution $(\{0\}, \bar{S}^h)$ to (P^h). Geometrically speaking, we obtain vertical supporting hyperplanes of $\mathcal{D}^*$ from the extreme directions of $\mathcal{P}^h$. Then we compute some $x \in S$ such that $y = Px$ is a weakly minimal point of $\mathcal{P}$ (line 03), which corresponds by geometric duality to a K-maximal (non-vertical) supporting hyperplane of $\mathcal{D}^*$. Together this yields a dual representation (inequality representation) of an outer approximation $\mathcal{T}$. By Proposition 5.11, we have $\mathcal{D}^* \subseteq \mathcal{T} \subseteq \Delta$, $\mathcal{T}_\infty = -K$ and we know that $\mathcal{T}$ and $\mathcal{D}^*$ have the same vertical supporting hyperplanes. These properties also hold for all subsequent outer approximations. In particular, we see that the set $\hat{\mathcal{T}}^d$ has not to be changed by the algorithm.

2. Iteration. The subroutine vert( ) computes a primal representation of $\mathcal{T}$, i.e., a description by vertices. It is tested whether or not the vertices of $\mathcal{T}$ belong to $\mathcal{D}^*$. To this end, the linear program $(\mathrm{D}_1(w(t^*)))$ is solved for the vertices t^* of $\mathcal{T}$ (line 14). The optimal value $b^T u$ of $(\mathrm{D}_1(w(t^*)))$ equals t^*_q if and only if $t^* \in \mathcal{D}^*$. If all vertices of an outer approximation $\mathcal{T}$ belong to $\mathcal{D}^*$, we have $\mathcal{T} = \mathcal{D}^*$. Finitely generated solutions to (P) and (D^*) can easily be obtained from the primal and dual representations of $\mathcal{D}^*$ by solving $(\mathrm{D}_1(w(t^*)))$ and $(\mathrm{P}_1(w(s^*)))$ (lines 14 and 20), respectively. If a vertex t^* of

$\mathcal{T}$ with $t^* \notin \mathcal{D}^*$ is detected, it is used to construct a better (smaller) outer approximation. This is now explained for the k-th iteration:

We have a vertex t^* of $\mathcal{T}^{k-1}$ not belonging to $\mathcal{D}^*$. A point $s^* \in \operatorname{Max}_K \mathcal{D}^*$ is determined, which is on the line between $\hat{d}$ and t^*. To this end we solve (see line 18) the linear program

$$\max \alpha \quad \text{subject to} \begin{cases} (u,w) \in T \\ \alpha t^* + (1-\alpha)\hat{d} = D^*(u,w). \end{cases} \tag{$\mathrm{R}^*(t^*)$}$$

A supporting hyperplane of $\mathcal{D}^*$ containing s^* is computed as shown in the following proposition.

Proposition 5.14. *Let $s^* \in \operatorname{Max}_K \mathcal{D}$. There exists a solution to* $(\mathrm{P}_1(w(s^*)))$*, and for every solution x to* $(\mathrm{P}_1(w(s^*)))$*, $H^*(Px)$ is a supporting hyperplane of $\mathcal{D}^*$ containing s^*.*

Proof. By $s^* \in \mathcal{D}^*$, the dual problem $(\mathrm{D}_1(w(s^*)))$ is feasible, hence a solution to $(\mathrm{P}_1(w(s^*)))$ exists (note that we assumed $S \neq \emptyset$). By Theorem 4.1 we have $Px \in \operatorname{wMin} \mathcal{P}$ for every solution x to $(\mathrm{P}_1(w(s^*)))$. Lemma 4.49 implies that $H^*(Px)$ is a supporting hyperplane of $\mathcal{D}^*$. Since $s^* \in \operatorname{Max}_K \mathcal{D}^*$, we have $s^*_q = b^T u$, where u is a solution to $(\mathrm{D}_1(w(s^*)))$. Strong duality between $(\mathrm{P}_1(w(s^*)))$ and $(\mathrm{D}_1(w(s^*)))$ implies $s^*_q = w(s^*)^T Px$, which is equivalent to $\varphi(Px, s^*) = 0$. This means $s^* \in H^*(Px)$. □

We now append the solution x of the linear program $(\mathrm{P}_1(w(s^*)))$ to $\bar{S}$ (line 21). Note that $P[\bar{S}]$ yields the first part of a dual representation of the new outer approximation $\mathcal{T}^{k+1}$. Appending x to $\bar{S}$ can be interpreted geometrically as a cut by the hyperplane $H^*(Px)$, that is, $\mathcal{T}^{k+1} = \mathcal{T}^k \cap \{y^* \in \mathbb{R}^q |\ \varphi(Px, y^*) \geq 0\}$.

Output. The output consists of a finitely generated solution $(\bar{S}, \bar{S}^h)$ to (P) and a finitely generated solution $\bar{T}$ to (D^*).

Let us now formulate the *dual outer approximation algorithm* in pseudo code. We use the same subroutines as for the primal algorithm in Section 5.1. A slight difference with respect to the routine *vert*( ) in comparison to the primal algorithm is discussed in the following remark.

Remark 5.15. The routine *vert*( ) returns the vertices of an outer approximation $\mathcal{T}$ of $\mathcal{D}^*$, which is given by a dual representation $\mathcal{T}^d$. This can be realized by "Vertex Enumeration by Adjacency Lists" (Chen *et al.*, 1991). Instead of using the offline variant of this method in the first iteration (compare Section 5.1), the given finitely generated solution $\bar{T}^h$ to (D^{*h}) can be used. The vertices of $\mathcal{T}^1$ are contained in the set

$$\left\{y^* \in \mathbb{R}^q |\ (u,w) \in \bar{T}^h \ \wedge\ y^* = (w_1, \ldots, w_{q-1}, w^T Px)\right\},$$

where x is the solution to $(\mathrm{P}_1(\eta))$ computed in line 03 of Algorithm 2. This follows from Proposition 5.11.

Algorithm 2.

Input:
  A, b, P (data of Problem (P));
  a finitely generated solution $(\{0\}, \bar{S}^h)$ to (P^h);
  a finitely generated solution $\bar{T}^h$ to (D^{*h});
Output:
  a finitely generated solution $(\bar{S}, \bar{S}^h)$ to (P);
  a finitely generated solution $\bar{T}$ to (D^*);

begin
  $(\cdot, \eta) = \sum_{j=1}^{|\bar{T}^h|} \frac{1}{|\bar{T}^h|} \bar{T}^h[j];$
  $x \leftarrow \mathrm{solve}(\mathrm{P}_1(\eta));$
  $\hat{d} \leftarrow \left\{\eta_1, \ldots, \eta_{q-1}, \eta^T P x - 1\right\}^T;$
  $\bar{S} \leftarrow \left\{x\right\};$
  $\hat{\mathcal{T}}^d = \left\{Px \mid x \in \bar{S}^h\right\};$
  repeat
    $\mathcal{T}^d \leftarrow \left\{Px \mid x \in \bar{S}\right\};$
    $\mathcal{T}^p \leftarrow \mathrm{vert}\,((\mathcal{T}^d, \hat{\mathcal{T}}^d));$
    $\bar{T} \leftarrow \emptyset;$
    for $i = 1$ to $|\mathcal{T}^p|$ do
    begin
      $t^* \leftarrow \mathcal{T}^p[i];$
      $u \leftarrow \mathrm{solve}(\mathrm{D}_1(w(t^*)));$
      $\bar{T} \leftarrow \bar{T} \cup \{(u, w(t^*))\};$
      if $t^*_q \neq b^T u$ then
      begin
        $(\cdot, \cdot, \alpha) \leftarrow \mathrm{solve}(\mathrm{R}^*(t^*));$
        $s^* \leftarrow \alpha t^* + (1 - \alpha)\hat{d};$
        $x \leftarrow \mathrm{solve}(\mathrm{P}_1(w(s^*)));$
        $\bar{S} \leftarrow \bar{S} \cup \{x\};$
        break;
      end;
    end;
  until $t^*_q = b^T u;$
end.

Further details of the dual algorithm are discussed in the proofs of the next two theorems.

Theorem 5.16. *Algorithm 2 works correctly.*

Proof. In line 02, we compute a convex combination (with nonzero coefficients) of the w-components of the elements of $\bar{T}^h$. By Proposition 5.12, we obtain $\eta \in \operatorname{int}(D^{*h} + K)$ with $e^T \eta = 1$. By Proposition 5.13, a solution to $(\mathrm{P}_1(\eta))$ exists (line 03) and we have $\hat{d} \in \operatorname{int} \mathcal{D}^*$ (line 04).

By Proposition 5.11, the set $\mathcal{T}$, which is defined by its dual representation $(\mathcal{T}^d, \hat{\mathcal{T}}^d)$ (line 08), is an outer representation of $\mathcal{D}^*$, i.e., $\mathcal{T} \supseteq \mathcal{D}^*$.

The set $\mathcal{T}^p$ in line 09 is nonempty, i.e. there exists a vertex of $\mathcal{T}$. Indeed, by Proposition 5.11 we know that $\mathcal{T}_\infty = -K$. Hence $\mathcal{T}$ contains no lines. By (Rockafellar, 1972, Corollary 18.5.3) a vertex exists.

In line 14 we solve $(\mathrm{D}_1(w(t^*)))$. In the k-th iteration, we have $t^* \in \mathcal{T}^k \subseteq \mathcal{T}^{k-1} \subseteq \cdots \subseteq \mathcal{T}^1$. The first outer approximation $\mathcal{T}^1$ can be expressed like $\mathcal{T}$ in Proposition 5.11, where $y := Px$ and x is the element first stored in $\bar{S}$ (line 05). Hence we have $\varphi(y, t^*) \geq 0$, or equivalently, $w(t^*)^T y \geq t^*_q$. This means that $(\mathrm{P}_1(w(t^*)))$ is bounded and feasible. Therefore, an optimal solution exists. By duality the same is true for the dual problem $(\mathrm{D}_1(w(t^*)))$.

Since t^* is a K-maximal point of an outer approximation of $\mathcal{D}^*$, either $t^* \in \operatorname{Max}_K \mathcal{D}^*$ or $t^* \notin \mathcal{D}^*$ holds. In the first case, we have $t^*_q = b^T u$ by Lemma 4.51 (i), (iii). In the second case, by Theorem 4.41, there exists some $y \in \mathcal{P}$ such that $\varphi(y, y^*) < 0$, or equivalently, $t^*_q > w(t^*)^T y$. Weak duality between $(\mathrm{P}_1(w(t^*)))$ and $(\mathrm{D}_1(w(t^*)))$ implies $t^*_q > b^T u$. This means that the lines 17-23 are performed in the case where $t^* \notin \mathcal{D}^*$. Thus, $(\mathrm{R}^*(t^*))$ has a solution and we obtain some $s^* \in \operatorname{Max}_K \mathcal{D}^*$. The linear program $(\mathrm{P}_1(w(s^*)))$ in line 20 has a solution by Proposition 5.14. In line 22 we break the innermost loop (lines 11-24).

The outer loop (lines 07-25) terminates in the case where for every $t^* \in \mathcal{T}^p$, $t^*_q = b^T u$ is obtained in line 14. This means $\mathcal{T}^p \subseteq \mathcal{D}^*$, which implies $\mathcal{T} = \mathcal{D}^*$. We have $D^*[\bar{T}] = \mathcal{T}^p$ (lines 13-15) and $P[\bar{S}] = \mathcal{T}^d$ (line 08).

From Corollary 5.10 we obtain that $(\bar{S}, \bar{S}^h)$ and $\bar{T}$ are finitely generated solutions to (P) and (D*), respectively. The assumption $\bar{T} \subseteq \operatorname{Eff}(\mathrm{D}^*)$ in Corollary 5.10 follows from the fact that every vertex of $\mathcal{D}^*$ is K-maximal. Indeed, if $(u, w) \in \bar{T}$, then u arose from an optimal solution to $(\mathrm{D}_1(w(t^*)))$ for a vertex t^* of $\mathcal{D}^*$ (line 14) and we have $w = w(t^*)$. We conclude $w_i = t^*_i$ for $i \in \{1, \dots, q-1\}$. Since t^* is K-maximal, we get $t^*_q = b^T u$, by Lemma 4.51 (i), (iii). Hence $t^* = D^*(u, w)$ and $(u, w) \in \operatorname{Eff}(\mathrm{D}^{*h})$.

The assumption $\bar{S} \subseteq \operatorname{Eff}(\mathrm{P})$ in Corollary 5.10 follows from Proposition 5.14 and the geometric duality theorem. The assumption $\bar{S}^h \subseteq \operatorname{Eff}(\mathrm{P}^h)$ follows from the input, where $(\{0\}, \bar{S}^h)$ is a finitely generated solution to (P^h). □

Finally we show that the algorithm terminates after a finite number of steps.

Theorem 5.17. *Algorithm 2 is finite.*

Proof. Since $\hat{d} \in \operatorname{int} \mathcal{D}^*$, the point $s^{*k} \in \mathcal{D}^*$ computed in iteration k belongs to $\operatorname{int} \mathcal{T}^k$. We have $\mathcal{T}^{k+1} := \mathcal{T}^k \cap \left\{y^* \in \mathbb{R}^q \,|\, \varphi(Px^k, y^*) \geq 0\right\}$ and, by

Proposition 5.14, we know that $F^* := \left\{y^* \in \mathcal{D}^* \mid \varphi(Px^k, y^*) = 0\right\}$ is a face of $\mathcal{D}^*$ with $s^{*k} \in F^*$, where $F^* \subseteq \operatorname{bd} \mathcal{T}^{k+1}$. For the next iteration, this means $s^{*k+1} \notin F^*$ (because $s^{*k+1} \in \operatorname{int} \mathcal{T}^{k+1}$). Therefore, s^{*k+1} belongs to another face of $\mathcal{D}^*$. Since $\mathcal{D}^*$ is polyhedral, it has a finite number of faces. Hence the algorithm is finite. □

5.3 Solving bounded problems

Boundedness of the primal problem (P) plays an important role because finitely generated solutions to the homogeneous problems (P^h) and (D^{*h}) can easily be obtained in this case. This means that Algorithms 1 and 2 can be used directly to solve (P) and (D^*).

As we shall show in the next section, the homogeneous problem (P^h) can be transformed into a bounded problem (P^η). Thus, both algorithms (the primal and the dual) can be used to solve an arbitrary homogeneous problem (P^h). In either case, the geometric dual problem (D^{*h}) is solved simultaneously. In a second step, the finitely generated solutions to the homogeneous problems (P^h) and (D^{*h}) can be used to solve the original problem (P), again by one of the Algorithms 1 or 2. In either case, the dual problem (D^*) is solved simultaneously.

Definition 5.18. Problem (P) is said to be *bounded* if there exists some $y \in \mathbb{R}^q$ such that

$$\forall x \in S: \; y \le Px.$$

Of course, the condition $\mathcal{P} \subseteq \{y\} + \mathbb{R}^q_+$ is equivalent to (P) being bounded. This property is commonly called $\mathcal{P}$ being $\mathbb{R}^q_+$-bounded below. If the feasible set S is bounded, then (P) is bounded, but the inverse implication is obviously not true.

Let us relate boundedness of (P) to a corresponding condition for the geometric dual problem (D^*). Recall that $\Delta := \{y^* \in \mathbb{R}^q \mid w(y^*) \ge 0\}$.

Proposition 5.19. *The following statements are equivalent:*

(i) *Problem* (P) *is bounded,*
(ii) $\exists \gamma \in \mathbb{R}: \; \left\{y^* \in \mathbb{R}^q \mid y^*_q \le \gamma\right\} \cap \Delta \subseteq \mathcal{D}^*.$

Proof. Let $y \in \mathbb{R}^q$ such that $\mathcal{P} \subseteq \{y\} + \mathbb{R}^q_+$. Setting $\gamma := \min\{y_i \mid i = 1, \dots, q\}$, we obtain $\mathcal{P} \subseteq \{\gamma e\} + \mathbb{R}^q_+$, where $e = (1, \dots, 1)^T$. Thus (i) is equivalent to the existence of some $\gamma \in \mathbb{R}$ such that $\mathcal{P} \subseteq \{\gamma e\} + \mathbb{R}^q_+ =: \mathcal{T}$

The set $\mathcal{T}$ can be expressed as

$$\mathcal{T} = \left\{y \in \mathbb{R}^q \mid (e^1)^T y \ge \gamma, \dots, (e^q)^T y \ge \gamma\right\},$$

where e^i denotes the i-th unit vector. By the last statement in Theorem 4.41, $\mathcal{P} \subseteq \mathcal{T}$ is equivalent to

$$\mathcal{T}^* := \left\{ \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \\ 0 \\ \gamma \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ \vdots \\ 0 \\ 0 \\ \gamma \end{pmatrix}, \dots, \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 1 \\ 0 \\ \gamma \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ 1 \\ \gamma \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ 0 \\ \gamma \end{pmatrix} \right\} \subseteq \mathcal{D}^*.$$

Equivalently, we have $\operatorname{co}\mathcal{T}^* \subseteq \mathcal{D}^*$. But, $\operatorname{co}\mathcal{T}^* = \left\{ y^* \in \mathbb{R}^q |\; y^*_q \leq \gamma \right\} \cap \Delta$, which completes the proof. □

The homogeneous problem (P^h) of a bounded problem (P) can easily be solved.

Theorem 5.20. *Let* (P) *be bounded and denote by* (P^h) *and* (D^{*h}) *the corresponding homogeneous problems. Then,*

(i) $(\{0\}, \emptyset)$ *is a finitely generated solution to* (P^h),
(ii) $\left\{ (u^1, e^1), \dots, (u^q, e^q) \right\}$ *is a finitely generated solution to* (D^{*h}), *where* u^i *is a feasible point of* $(\mathrm{D}_1(e^i))$ *for* $i \in \{1, \dots, q\}$.

Proof. (i) Note first that the conditions (i), (ii) and (iv) of Definition 4.4 are obviously satisfied for $\bar{S} = \{0\}$ and $\bar{S}^h = \emptyset$, respectively. To show (iii), we observe that 0 is feasible for (P^h). Furthermore, we have $P[S^h] \subseteq \mathcal{P}^h = \mathbb{R}^q_+$ and hence $(\{0\} - \mathbb{R}^q_+ \setminus \{0\}) \cap P[S^h] = \emptyset$. Condition (v) of Definition 4.4 reduces to $P[S^h] \subseteq \operatorname{co} P[\{0\}] + \mathbb{R}^q_+ = \mathbb{R}^q_+$, which is obviously satisfied (recall the convention $\operatorname{cone}\emptyset = \{0\}$).

(ii) Let $\bar{T} := \left\{ (u^1, e^1), \dots, (u^q, e^q) \right\}$. Then $\bar{T}$ is a finite subset of T. We have

$$D^{*h}[\bar{T}] = \left\{ e^1, e^2, \dots, e^{q-1}, 0 \right\}.$$

We conclude that $D^{*h}[\bar{T}] \subseteq \operatorname{Max}_K D^{*h}[T]$ and $D^{*h}[T] \subseteq \operatorname{co} D^{*h}[\bar{T}] - K$. This means that all conditions of Definition 4.56 are satisfied. □

Remark 5.21. Note that the u-components of a finitely generated solution to (D^{*h}) are sometimes not needed. For instance, in Algorithms 1 and 2 only the w-components are used. It is therefore not necessary to determine feasible solutions to $(\mathrm{D}_1(e^i))$ $(i \in \{1, \dots, q\})$ as proposed in the latter theorem. This means that no computational effort is necessary to obtain the input data in Algorithms 1 and 2 for bounded problems (P).

5.4 Solving the homogeneous problem

The aim of this section is to transform an arbitrary homogeneous problem (P^h) into a bounded problem such that a finitely generated solution to the bounded problem yields a finitely generated solution to (P^h).

Note that the homogeneous problem (P^h) is always feasible because $0 \in S^h = \{x \in \mathbb{R}^n |\, Bx \geq 0\}$. Throughout this section, the lower image $\mathcal{D}^{*h}$ of (D^{*h}) is assumed to have a nonempty interior. Note that

$$\operatorname{int} \mathcal{D}^{*h} \neq \emptyset \quad \Longleftrightarrow \quad \operatorname{int} (\mathcal{D}^{*h} + K) \neq \emptyset.$$

In Section 5.5 we provide an algorithm that either evaluates an interior point of $\mathcal{D}^{*h}$ or states that $\operatorname{int} \mathcal{D}^{*h}$ is empty. The latter case is equivalent to $\mathcal{P}^h$ having no vertex. This follows from (4.21) and the geometric duality theorem.

Let us start with an auxiliary assertion.

Lemma 5.22. *Let $y^* \in \operatorname{int} (\mathcal{D}^{*h} + K)$ such that $y^*_q = 1$. The set*

$$\mathcal{B} := H(y^*) \cap \mathcal{P}^h$$

is a bounded base of the polyhedral cone $\mathcal{P}^h$. In particular, $\mathcal{B}$ is a $(q-1)$-dimensional polytope.

Proof. The point $y^{*h} := (y^*_1, \dots, y^*_{q-1}, 0)^T$ is obtained from y^* by setting the last component to zero. We get

$$H(y^{*h}) = \left\{ y \in \mathbb{R}^q |\, w(y^*)^T y = 0 \right\}.$$

By assumption we have $y^* \in \operatorname{int} (\mathcal{D}^{*h} + K)$. Hence, y^{*h} belongs to the relative interior of the only K-maximal facet $F^* := D^{*h}[T] = \mathcal{D}^{*h} \cap \left\{ y^* \in \mathbb{R}^q |\, y^*_q = 0 \right\}$ of $\mathcal{D}^{*h}$ (compare (4.21)). By Corollary 4.53, $H(y^{*h}) \cap \mathcal{P}^h$ consists of exactly one vertex of $\mathcal{P}^h$. But, the only vertex of the cone $\mathcal{P}^h$ can be 0. This implies

$$H(y^{*h}) \cap \mathcal{P}^h = \{0\}. \tag{5.1}$$

In order to show that $\mathcal{B}$ is a base, let $y \in \mathcal{P}^h \setminus \{0\}$. By Theorem 4.40 we have $\varphi(y, y^{*h}) \geq 0$. This can be written as $w(y^*)^T y \geq 0$. It follows the strict inequality $w(y^*)^T y > 0$, since otherwise, we get $y \in H(y^{*h}) \cap \mathcal{P}^h$, which contradicts (5.1). Hence there exists some $\mu > 0$ such that $w(y^*)^T (\mu y) = 1$ and thus $\mu y \in \mathcal{B}$.

Let us show that $\mathcal{B}$ is bounded. Assuming the contrary, we obtain some $y \neq 0$ such that $y \in \mathcal{B}_\infty$. From (Rockafellar, 1972, Theorem 8.3 and Corollary 8.3.3) it follows that $y \in (H(y^{*h}) \cap \mathcal{P})_\infty$, which contradicts (5.1).

Since $H(y^*)$ is a hyperplane in $\mathbb{R}^q$, we obtain $\dim \mathcal{B} \leq q-1$. Moreover, we have $\mathbb{R}^q_+ \subseteq \mathcal{P}^h$ and hence $\mathcal{B} \cap \mathbb{R}^q_+$ is a base of $\mathbb{R}^q_+$. Thus we obtain $\dim \mathcal{B} \geq q-1$. □

Let $\eta \in \operatorname{int} (\mathcal{D}^{*h} + K)$ such that $e^T \eta = 1$. Note that this implies $\eta > 0$. We consider the problem

$$\text{minimize } P : \mathbb{R}^n \to \mathbb{R}^q \text{ with respect to } \leq \text{ over } S^\eta, \tag{P^η}$$

where

$$S^\eta := \left\{x \in \mathbb{R}^n \middle|\ Bx \geq 0,\ \eta^T Px \leq 1\right\}.$$

We write

$$\mathcal{P}^\eta := P[S^\eta] + \mathbb{R}^q_+$$

for the upper image of (P^η). We have

$$y \in \mathcal{P}^\eta \iff \exists x \in \mathbb{R}^n:\ y \geq Px,\ Bx \geq 0,\ \eta^T Px \leq 1.$$

Setting $y^* := (\eta_1, \dots, \eta_{q-1}, 1)^T$, we obtain

$$H(y^*) = \left\{y \in \mathbb{R}^q \middle|\ \eta^T y = 1\right\}.$$

Hence, the upper image $\mathcal{P}^\eta$ can be expressed by the set $\mathcal{B}$, defined in Lemma 5.22, as

$$\mathcal{P}^\eta = [0,1]\mathcal{B} + \mathbb{R}^q_+, \tag{5.2}$$

where $[0,1]\mathcal{B} := \{\mu y|\ 0 \leq \mu \leq 1,\ y \in \mathcal{B}\}$. Note that y is vertex of $\mathcal{B}$ if and only if it is an extreme direction of $\mathcal{P}^h$ with $\eta^T y = 1$. Let us show some properties of (P^η).

Theorem 5.23. *Let $\eta \in \operatorname{int}(\mathcal{D}^{*h} + K)$ such that $e^T\eta = 1$. Then*

(i) *Problem (P^η) is bounded;*
(ii) *If a vector $y \in \mathbb{R}^q \setminus \{0\}$ is a vertex of $\mathcal{P}^\eta$, then it is an extreme direction of $\mathcal{P}^h$ such that $\eta^T y = 1$;*
(iii) *If a vector $y \in \mathbb{R}^q \setminus \{0\}$ is an extreme direction of $\mathcal{P}^h$ with $\eta^T y = 1$, then y is either a vertex of $\mathcal{P}^\eta$ or $y = \frac{1}{\eta_i} e^i$ for some $i \in \{1, \dots, q\}$.*

Proof. (i) This follows from (5.2) and Lemma 5.22.

(ii) Let $y \neq 0$ be a vertex of $\mathcal{P}^\eta = [0,1]\mathcal{B} + \mathbb{R}^q_+$. It remains to show that y is a vertex of $\mathcal{B}$.

We first show that $y \in \mathcal{B}$. Assume that $y \notin [0,1]\mathcal{B}$. There exist $v \in [0,1]\mathcal{B}$ and $c \in \mathbb{R}^q_+ \setminus \{0\}$ such that $y = v + c$. We obtain

$$y = \frac{1}{2}v + \frac{1}{2}(v + 2c).$$

Since $v, v + 2c \in \mathcal{P}^\eta$ and $v \neq y \neq v + 2c$, this contradicts y being a vertex of $\mathcal{P}^\eta$. Hence $y \in [0,1]\mathcal{B}$.

By assumption we have $y \neq 0$. Thus, y belongs to $(0,1]\mathcal{B}$. Assume that $y \notin \mathcal{B}$. There exist $\mu \in (0,1)$ and $v \in \mathcal{B}$ such that $y = \mu v$. Choosing some $\varepsilon > 0$ such that $0 < \mu - \varepsilon < \mu + \varepsilon < 1$, we obtain

$$y = \frac{1}{2}(\mu - \varepsilon)v + \frac{1}{2}(\mu + \varepsilon)v, \qquad (\mu \pm \varepsilon)v \in [0,1]\mathcal{B} \subseteq \mathcal{P}^\eta,$$

This contradicts y being a vertex of $\mathcal{P}^\eta$.

We have shown that a given vertex $y \neq 0$ of $\mathcal{P}^\eta$ belongs to $\mathcal{B} \subseteq \mathcal{P}^\eta$. Using the definition of a vertex, we conclude that y is a vertex of $\mathcal{B}$.

(iii) Let y be an extreme direction of $\mathcal{P}^h$ such that $\eta^T y = 1$, i.e., y is a vertex of $\mathcal{B}$.

We first show that y is a vertex of $[0,1]\mathcal{B}$. Let $v^1, v^2 \in [0,1]\mathcal{B}$ and $\lambda \in (0,1)$ such that $y = \lambda v^1 + (1-\lambda)v^2$. There exist $b^1, b^2 \in \mathcal{B}$, $\mu_1, \mu_2 \in [0,1]$ such that $y = \lambda \mu_1 b^1 + (1-\lambda)\mu_2 b^2$. We obtain $1 = \eta^T y = \lambda\mu_1 + (1-\lambda)\mu_2$. It follows that $\mu_1 = \mu_2 = 1$. Since y is a vertex of $\mathcal{B}$, we get $b^1 = b^2 = y$, whence $y = v^1 = v^2$. This means $y \in \operatorname{vert}([0,1]\mathcal{B})$.

Let $y^1, y^2 \in \mathcal{P}^\eta$ and $\lambda \in (0,1)$ such that $y = \lambda y^1 + (1-\lambda)y^2$ with $y^1 \neq y^2$. As a consequence of Lemma 5.22 and taking into account that $\mathbb{R}^q_+ \subseteq \mathcal{P}^h$ and $\mathcal{P}^\eta = [0,1]\mathcal{B} + \mathbb{R}^q_+$, we obtain

$$\mathcal{P}^\eta \cap \left\{ y \in \mathbb{R}^q \middle|\, \eta^T y \leq 1 \right\} = [0,1]\mathcal{B},$$

$$\mathcal{P}^\eta \cap \left\{ y \in \mathbb{R}^q \middle|\, \eta^T y > 1 \right\} = \mathcal{B} + \mathbb{R}^q_+ \setminus \{0\}.$$

This implies

$$([0,1]\mathcal{B}) \cap (\mathcal{B} + \mathbb{R}^q_+ \setminus \{0\}) = \emptyset \quad \text{and} \quad ([0,1]\mathcal{B}) \cup (\mathcal{B} + \mathbb{R}^q_+ \setminus \{0\}) = \mathcal{P}^\eta.$$

We consider three cases:

First let $y^1, y^2 \in [0,1]\mathcal{B}$. We get $y = y^1 = y^2$, since y is a vertex of $[0,1]\mathcal{B}$. Hence y is a vertex of $\mathcal{P}^\eta$.

Secondly let $y^1, y^2 \in \mathcal{B} + \mathbb{R}^q \setminus \{0\}$. We get the contradiction $1 = \eta^T y = \lambda \eta^T y^1 + (1-\lambda)\eta^T y^2 > 1$, i.e., this case cannot occur.

In the last case, we can assume that $y^1 \in [0,1]\mathcal{B}$ and $y^2 \in \mathcal{B} + \mathbb{R}^q_+ \setminus \{0\}$. There exist $b^1 \in \mathcal{B}$ and $\mu \in [0,1]$ such that $y^1 = \mu b^1$. Since $\eta^T y^2 > 1$, we get $\lambda\mu + (1-\lambda) < \lambda\mu\, \eta^T b^1 + (1-\lambda)\eta^T y^2 = \eta^T y = 1$ and hence $\mu < 1$. Setting

$$\bar\lambda := \lambda\mu \quad \text{and} \quad \bar b := \frac{1-\lambda}{1-\bar\lambda} y^2,$$

we get $y = \bar\lambda b^1 + (1-\bar\lambda)\bar b$. Moreover, we have

$$1 = \eta^T y = \bar\lambda \eta^T b^1 + (1-\bar\lambda)\eta^T \bar b = \bar\lambda + (1-\bar\lambda)\eta^T \bar b.$$

This implies $\eta^T \bar b = 1$. Furthermore, we have $y^2 \in \mathcal{P}^\eta \subseteq \mathcal{P}^h$ and hence $\bar b \in \mathcal{P}^h$. Consequently, $\bar b \in \mathcal{B}$. Since y is a vertex of $\mathcal{B}$, we get $b^1 = y = \bar b$.

We have shown that

$$y = \bar b = \frac{1-\lambda}{1-\bar\lambda} y^2.$$

There exist $b^2 \in \mathcal{B} \subseteq \mathcal{P}^h$ and $c \in \mathbb{R}^q_+ \setminus \{0\} \subseteq \mathcal{P}^h$ such that $y^2 = b^2 + c$. Since y is an extreme direction of $\mathcal{P}^h$ we get $b^2 = c$. It follows that $y \in \mathbb{R}^q_+$. Since $\mathbb{R}^q_+ \subseteq \mathcal{P}^h$, y is also an extreme direction of $\mathbb{R}^q_+$, i.e., y is a multiple of a unit vector e^i in $\mathbb{R}^q$. Since $\eta > 0$, the condition $\eta^T y = 1$ is satisfied for $y = \frac{1}{\eta_i} e^i$ and the desired statement follows. □

The geometric dual problem of (P^η), as introduced in Section 4.5 for arbitrary linear problems (P), can easily verified to be

$$\text{maximize } D^{*\eta} : \mathbb{R}^m \times \mathbb{R}^q \times \mathbb{R} \to \mathbb{R}^q \text{ with respect to } \le_K \text{ over } T^\eta, \quad (\mathrm{D}^{*\eta})$$

where the feasible set T^η is defined by

$$T^\eta := \left\{ (u,w,z) \in \mathbb{R}^m \times \mathbb{R}^q \times \mathbb{R} \,\middle|\, \begin{array}{c} (u,w,z) \ge 0,\ w^T e = 1, \\ B^T u = P^T(w + z\eta) \end{array} \right\},$$

and the objective function is

$$D^{*\eta} : \mathbb{R}^m \times \mathbb{R}^q \times \mathbb{R} \to \mathbb{R}^q, \quad D^{*\eta}(u,w,z) := (w_1, ..., w_{q-1}, -z)^T .$$

An example of the bounded problem (P^η) and its dual problem $(\mathrm{D}^{*\eta})$ is given next.

Example 5.24. Consider the homogeneous problem (P^h) with the data

$$P = \begin{pmatrix} 1 & 2 \\ 0 & 2 \end{pmatrix}, \qquad B = \begin{pmatrix} 1 & 3 \\ -1 & 2 \end{pmatrix}.$$

The feasible set of (P^h) is determined by

$$S^h = \left\{ x \in \mathbb{R}^2 \mid Bx \ge 0 \right\} = \left\{ x \in \mathbb{R}^2 \mid x_1 + 3x_2 \ge 0,\ -x_1 + 2x_2 \ge 0 \right\}.$$

Its image can be expressed as

$$P[S^h] = \left\{ Px \in \mathbb{R}^2 \mid Bx \ge 0 \right\} = \left\{ y \in \mathbb{R}^2 \mid 2y_1 + y_2 \ge 0,\ -y_1 + 2y_2 \ge 0 \right\}.$$

Both sets are displayed in Figure 5.4.

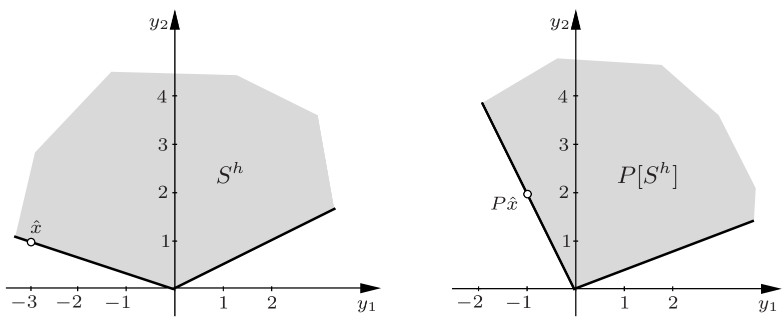

Fig. 5.4 The feasible set S^h and the image $P[S^h]$ of the homogeneous problem (P^h) in Example 5.24. Since $\hat{x} = (-3,1)^T$ is the only efficient direction (up to multiples), $(\{0\}, \{\hat{x}\})$ is a finitely generated solution to (P^h).

The upper image of (P^h) is the set

$$\mathcal{P}^h = P[S^h] + \mathbb{R}^q_+ = \left\{y \in \mathbb{R}^2 |\; 2y_1 + y_2 \geq 0,\; y_2 \geq 0\right\}.$$

The lower image of the geometric dual problem (D^{*h}) of (P^h) is easily obtained as

$$\mathcal{D}^{*h} = \mathrm{co}\left\{(0,0)^T, \left(\tfrac{2}{3}, 0\right)^T\right\} - K.$$

In order to formulate the corresponding bounded problem (P^η), we need a suitable vector η. The point $y^* = \left(\frac{1}{2}, -1\right)^T$ belongs to $\mathrm{int}\,\mathcal{D}^{*h}$ and $\eta = \left(\frac{1}{2}, \frac{1}{2}\right)^T \in \mathrm{int}\,(\mathcal{D}^{*h} + K)$ satisfies the condition $e^T\eta = 1$. The upper image $\mathcal{P}^\eta$ of the bounded problem (P^η) is obtained as

$$\begin{aligned}\mathcal{P}^\eta &= \left(\mathcal{P}^h \cap \left\{y \in \mathbb{R}^q |\; \eta^T y \leq 1\right\}\right) + \mathbb{R}^q_+ \\ &= \left\{y \in \mathbb{R}^q |\; y_1 \geq -2,\; y_2 \geq 0,\; 2y_1 + y_2 \geq 0\right\}.\end{aligned}$$

The lower image of the geometric dual problem $(\mathrm{D}^{*\eta})$ of (P^η) is the set

$$\mathcal{D}^{*\eta} = \mathrm{co}\left\{(0,0)^T, \left(\tfrac{2}{3}, 0\right)^T, (1,-2)^T\right\} - K.$$

The sets $\mathcal{P}^h$, $\mathcal{D}^{*h}$, $\mathcal{P}^\eta$ and $\mathcal{D}^{*\eta}$ are shown in Figure 5.5.

Let us point out some properties of the lower image $\mathcal{D}^{*\eta}$ of $(\mathrm{D}^{*\eta})$.

Theorem 5.25. *Let $\eta \in \mathrm{int}\,(\mathcal{D}^{*h} + K)$ such that $e^T\eta = 1$. The following statements are equivalent:*

(i) *y^* is a vertex of $\mathcal{D}^{*\eta}$ such that $y^*_q = 0$,*
(ii) *y^* is a vertex of $\mathcal{D}^{*h}$.*

Proof. (i) $\Rightarrow$ (ii). Let y^* be a vertex of $\mathcal{D}^{*\eta}$ with $y^*_q = 0$. There exists $(u,w,z) \in T^\eta$ such that $D^{*\eta}(u,w,z) = (w_1,\dots,w_{q-1},-z)^T \geq_K y^*$. We get $0 = y^*_q \leq -z$. But, we also have $z \geq 0$, whence $z = 0$. It follows $(u,w) \in T$ (feasibility for (D^{*h})) and $y^* \in \mathcal{D}^{*h}$.

Furthermore, it can easily be verified that $T \times \{0\} \subseteq T^\eta$ and hence $\mathcal{D}^{*h} \subseteq \mathcal{D}^{*\eta}$.

Let $y^{*1}, y^{*2} \in \mathcal{D}^{*h} \subseteq \mathcal{D}^{*\eta}$ and $\lambda \in (0,1)$ such that $y^* = \lambda y^{*1} + (1-\lambda)y^{*2}$. As y^* is vertex of $\mathcal{D}^{*\eta}$, we get $y^{*1} = y^{*2} = y^*$. Hence y^* is a vertex of $\mathcal{D}^{*h}$.

(ii) $\Rightarrow$ (i). Let y^* be a vertex of $\mathcal{D}^{*h}$. Taking into account (4.21), we get $y^*_q = 0$. Let $y^{*1}, y^{*2} \in \mathcal{D}^{*\eta}$ and $\lambda \in (0,1)$ such that $y^* = \lambda y^{*1} + (1-\lambda)y^{*2}$. There exist $(u^1,w^1,z^1), (u^2,w^2,z^2) \in T^\eta$ such that $y^{*1} \leq_K D^{*\eta}(u^1,w^1,z^1)$ and $y^{*2} \leq_K D^{*\eta}(u^2,w^2,z^2)$. We have $0 = \lambda y^{*1}_q + (1-\lambda)y^{*2}_q$, where $y^{*1}_q \leq -z^1 \leq 0$ and $y^{*2}_q \leq -z^2 \leq 0$. We conclude $y^{*1}_q = y^{*2}_q = z^1 = z^2 = 0$. Consequently, $(u^1,w^1), (u^2,w^2) \in T$ and thus $y^{*1}, y^{*2} \in \mathcal{D}^{*h}$. Since y^* is a vertex of $\mathcal{D}^{*h}$, we get $y^* = y^{*1} = y^{*2}$. Therefore, y^* is a vertex of $\mathcal{D}^{*\eta}$ and, as already shown, $y^*_q = 0$. $\square$

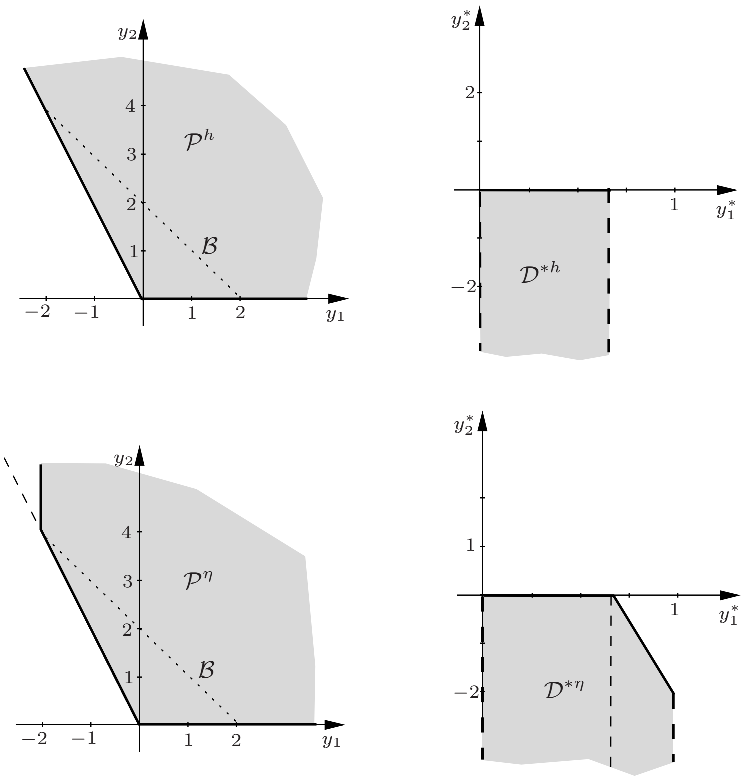

Fig. 5.5 The upper images of the homogeneous problem (P^h) and the corresponding bounded problem (P^η) from Example 5.24 are shown on the left. The lower images of the geometric dual problems are shown on the right.

We continue with a second example, which is illustrated in Figure 5.6.

Example 5.26. Consider the homogeneous problem (P^h) with the following data:

$$B = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ -1 & 1 & 0 & 0 & 0 \\ 0 & -1 & 1 & 0 & 0 \\ 0 & 0 & -1 & 1 & 0 \\ 0 & 0 & 0 & -1 & 1 \end{pmatrix}, \quad b = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \quad P = \begin{pmatrix} -4 & 0 & 0 & -4 & 4 \\ 4 & -4 & -4 & 4 & 0 \\ 0 & 0 & 4 & 0 & 0 \end{pmatrix}.$$

The vector $\eta = \left(\frac{1}{4}, \frac{1}{4}, \frac{1}{2}\right)^T$ belongs to $\operatorname{int}(\mathcal{D}^{*h} + K)$ and satisfies $e^T\eta = 1$. The corresponding bounded problem (P^η) is given by the data

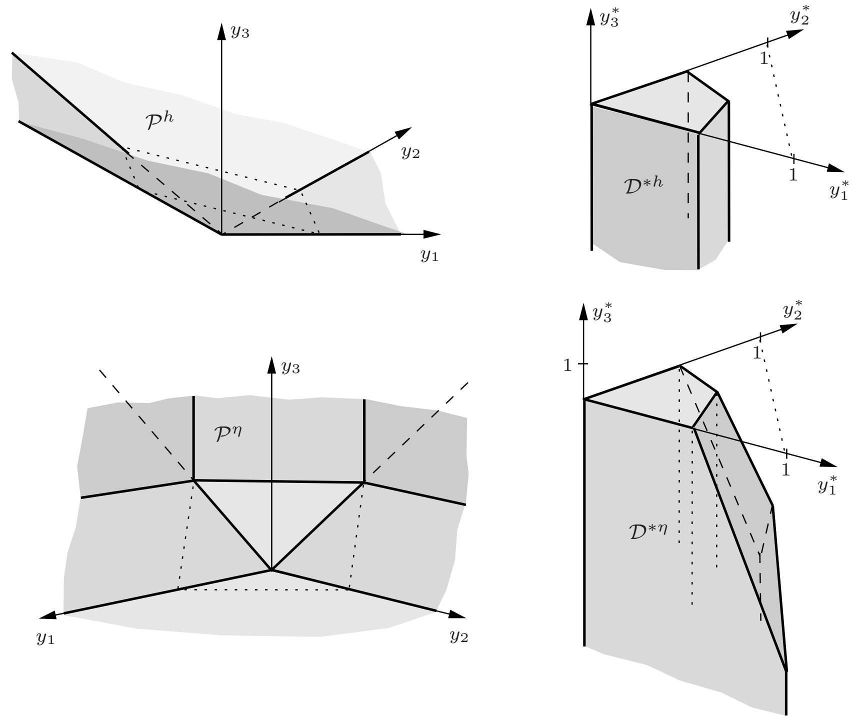

Fig. 5.6 Illustration of Example 5.26. The upper images of (P^h) and (P^η) are displayed on the left. The dotted line indicates the set $\mathcal{B}$. On the right, the lower images of (D^{*h}) and $(\mathrm{D}^{*\eta})$ are shown.

$$B = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ -1 & 1 & 0 & 0 & 0 \\ 0 & -1 & 1 & 0 & 0 \\ 0 & 0 & -1 & 1 & 0 \\ 0 & 0 & 0 & -1 & 1 \\ 0 & 1 & -1 & 0 & -1 \end{pmatrix}, \quad b = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ -1 \end{pmatrix}, \quad P = \begin{pmatrix} -4 & 0 & 0 & -4 & 4 \\ 4 & -4 & -4 & 4 & 0 \\ 0 & 0 & 4 & 0 & 0 \end{pmatrix}.$$

Problem (P^η) can be solved by Benson's algorithm (or by its dual variant), which yields (together with the geometric duality theorem) both a representation by vertices/directions and a representation by inequalities of both problems $\mathcal{P}^\eta$ and $\mathcal{D}^{*\eta}$:

(A) Vertices and extreme directions of $\mathcal{P}^\eta$:

$$y^0 = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix},\ y^1 = \begin{pmatrix} 0 \\ -4 \\ 4 \end{pmatrix},\ y^2 = \begin{pmatrix} -4 \\ 0 \\ 4 \end{pmatrix};$$

$$\bar{y}^1 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix},\ \bar{y}^2 = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix},\ \bar{y}^3 = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}.$$

(B) Inequality representation of $\mathcal{P}^\eta$:

$$\begin{aligned} y_3 &\geq 0 \\ y_1 + y_3 &\geq 0 \\ y_2 + y_3 &\geq 0 \\ y_1 + y_2 + y_3 &\geq 0 \\ y_1 + y_2 &\geq -4 \\ y_2 &\geq -4 \\ y_1 &\geq -4\,. \end{aligned}$$

(C) Vertices and extreme directions of $\mathcal{D}^{*\eta}$:

$$y^{*1} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix},\ y^{*2} = \frac{1}{2}\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix},\ y^{*3} = \frac{1}{2}\begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix},\ y^{*4} = \frac{1}{3}\begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix},$$

$$y^{*5} = \frac{1}{2}\begin{pmatrix} 1 \\ 1 \\ -4 \end{pmatrix},\ y^{*6} = \begin{pmatrix} 0 \\ 1 \\ -4 \end{pmatrix},\ y^{*7} = \begin{pmatrix} 1 \\ 0 \\ -4 \end{pmatrix};\ \bar{y}^{*1} = \begin{pmatrix} 0 \\ 0 \\ -1 \end{pmatrix}.$$

(D) Inequality representation of $\mathcal{D}^{*\eta}$:

$$\begin{aligned} - y_3^* &\geq 0 \\ -4y_1^* - 8y_2^* - y_3^* &\geq -4 \\ -8y_1^* - 4y_2^* - y_3^* &\geq -4 \\ y_1^* &\geq 0 \\ y_2^* &\geq 0 \\ -y_1^* - y_2^* &\geq -1\,. \end{aligned}$$

From (A) and Theorem 5.23 (iii) we get a superset of the extreme directions of $\mathcal{P}^h$ as:

$$\hat{y}^1 = \begin{pmatrix} 0 \\ -4 \\ 4 \end{pmatrix},\ \hat{y}^2 = \begin{pmatrix} -4 \\ 0 \\ 4 \end{pmatrix},\ \hat{y}^3 = \begin{pmatrix} 4 \\ 0 \\ 0 \end{pmatrix},\ \hat{y}^4 = \begin{pmatrix} 0 \\ 4 \\ 0 \end{pmatrix},\ \hat{y}^5 = \begin{pmatrix} 0 \\ 0 \\ 2 \end{pmatrix};$$

By geometric duality (Theorem 4.42 and Theorem 4.62), an inequality representation of $\mathcal{D}^{*h}$ is obtained as:

$$\begin{array}{rrrl} y_1^* & + 2y_2^* & & \leq 1 \\ 2y_1^* & + \ \ y_2^* & & \leq 1 \\ y_1^* & & & \geq 0 \\ & y_2^* & & \geq 0 \\ y_1^* & + \ \ y_2^* & & \leq 1 \\ & & y_3^* & \leq 0\,. \end{array}$$

Note that the first four inequalities correspond to the four extreme directions of $\mathcal{P}^h$ (Theorem 4.62). The fifth inequality is redundant and results from the direction $\hat{y}^5$, which is not extreme. The last inequality corresponds to the origin, which is the only vertex of $\mathcal{P}^h$ (Theorem 4.42).

From (C) and Theorem 5.25, we obtain the vertices of $\mathcal{D}^{*h}$:

$$\hat{y}^{*1} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}, \ \hat{y}^{*2} = \frac{1}{2}\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \ \hat{y}^{*3} = \frac{1}{2}\begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \ \hat{y}^{*4} = \frac{1}{3}\begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}.$$

By the geometric duality theorem, we obtain an inequality representation of $\mathcal{P}^h$:

$$\begin{array}{rrrl} & & y_3 & \geq 0 \\ y_1 & & + \ y_3 & \geq 0 \\ & y_2 & + \ y_3 & \geq 0 \\ y_1 & + \ y_2 & + \ y_3 & \geq 0\,. \end{array}$$

An overview of the connections between the problems (P^h), (D^{*h}), (P^η) and ($\mathrm{D}^{*\eta}$) in terms of the upper and lower images is given in Figure 5.7.

We close this section by relating finitely generated solutions of (P^η) and ($\mathrm{D}^{*\eta}$) to finitely generated solutions of the homogeneous problems (P^h) and (D^{*h}).

Theorem 5.27. *Let* $\eta \in \operatorname{int}(\mathcal{D}^{*h} + K)$ *such that* $e^T\eta = 1$. *If* $(\bar{S}^\eta, \emptyset)$ *is a finitely generated solution to* (P^η) *and*

$$\bar{S}^h := \left\{x \in \bar{S}^\eta \,|\; Px \neq 0\right\},$$

then $(\{0\}, \bar{S}^h)$ *is a finitely generated solution to* (P^h).

Proof. We show that the conditions (i) to (v) of Definition 4.4 are satisfied for (P^h). Note that we have $S = S^h = \{x \in \mathbb{R}^n |\; Bx \geq 0\}$.

Of course, $\bar{S} := \{0\}$ and $\bar{S}^h$ are finite subsets of S^h. Hence (i) and (ii) are satisfied.

Since $\operatorname{int} \mathcal{D}^{*h} \neq \emptyset$, $\mathcal{D}^{*h}$ has a K-maximal facet. By geometric duality, $\mathcal{P}^h$ has a vertex. As $\mathcal{P}^h$ is a cone, this vertex must be zero. Since every vertex of $\mathcal{P}^h$ is minimal (Corollary 4.67), we get $\{0\} = P[\{0\}] \subseteq \operatorname{Min} \mathcal{P}^h = \operatorname{Min} P[S^h]$, i.e., condition (iii) holds.

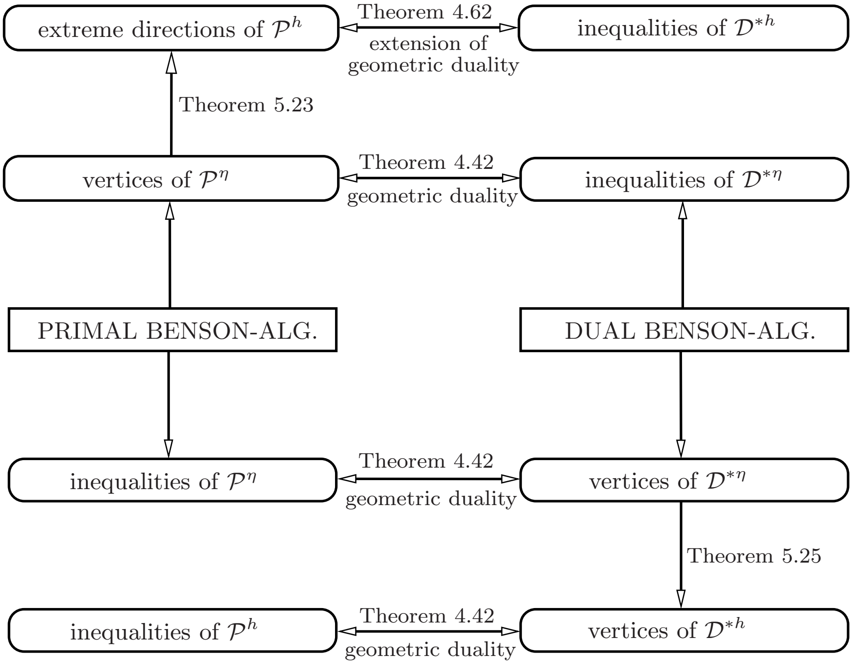

Fig. 5.7 Overview of the results in terms of upper and lower images. A finitely generated solution to (P^η) yields a representation of $\mathcal{P}^\eta$ by vectors and directions and, by geometric duality, a representation of $\mathcal{D}^{*\eta}$ by inequalities. A finitely generated solution to ($\mathrm{D}^{*\eta}$) yields a representation of $\mathcal{D}^{*\eta}$ by vectors and, by geometric duality, a representation of $\mathcal{P}^\eta$ inequalities. Finitely generated solution to both problems (P^η) and ($\mathrm{D}^{*\eta}$) can be obtained by Algorithm 1, but also by Algorithm 2. Information on the homogeneous problems (P^h) and (D^{*h}) can be derived.

For $x \in \bar{S}^\eta \subseteq S^\eta \subseteq S^h$, we have $Px \in \operatorname{Min} P[S^\eta]$. Assume that there is some $\tilde{x} \in S^h$ such that $P\tilde{x} \le Px$. Since $\eta \ge 0$, we get $\eta^T P\tilde{x} \le \eta^T Px \le 1$, i.e., $\tilde{x} \in S^\eta$. From $Px \in \operatorname{Min} P[S^\eta]$, we get $P\tilde{x} = Px$. Hence $Px \in \operatorname{Min} P[S^h]$, i.e., condition (iv) holds.

It remains to show that (v) of Definition 4.4 holds for (P^h), $\bar{S} = \{0\}$ and $\bar{S}^h$ as defined above, that is,

$$P[S^h] \subseteq \operatorname{cone} P[\bar{S}^h] + \mathbb{R}^q_+. \tag{5.3}$$

Condition (v) of Definition 4.4 holds for (P^η). This can be expressed as (recall that $\operatorname{cone} \emptyset = \{0\}$)

$$\mathcal{P}^\eta = \operatorname{co} P[\bar{S}^\eta] + \mathbb{R}^q_+, \tag{5.4}$$

where the equality is obtained using (i) of Definition 4.4. Moreover, we know that

$$\operatorname{vert}\mathcal{P}^\eta \subseteq P[\bar{S}^\eta]. \tag{5.5}$$

This follows from (5.4) and the fact that the vertices of $\mathcal{P}^\eta$ cannot be obtained as a nontrivial convex combination (by the definition of a vertex). Note further that $\mathcal{P}^h$ can be expressed by its extreme directions, where we can use the vertices of the bounded base $\mathcal{B}$ of $\mathcal{P}^h$, that is,

$$\mathcal{P}^h = \operatorname{cone}\operatorname{vert}\mathcal{B}.$$

Using the definition of $\bar{S}^h$, we obtain the following chain of inclusions

$$\begin{aligned} P[\bar{S}^h] \cup \left\{\tfrac{1}{\eta_1}e^1,\dots,\tfrac{1}{\eta_q}e^q\right\} \quad &= \quad P[\bar{S}^\eta]\setminus\{0\} \cup \left\{\tfrac{1}{\eta_1}e^1,\dots,\tfrac{1}{\eta_q}e^q\right\} \\ &\overset{(5.5)}{\supseteq} \quad (\operatorname{vert}\mathcal{P}^\eta)\setminus\{0\} \cup \left\{\tfrac{1}{\eta_1}e^1,\dots,\tfrac{1}{\eta_q}e^q\right\} \\ &\overset{\text{Th. 5.23 (iii)}}{\supseteq} \quad \operatorname{vert}\mathcal{B}. \end{aligned}$$

Together we get

$$P[S^h] \subseteq \mathcal{P}^h = \operatorname{cone}\operatorname{vert}\mathcal{B} \subseteq \operatorname{cone} P[\bar{S}^h] + \mathbb{R}^q_+,$$

i.e., (5.3) holds. □

The next result is the dual counterpart of the last one.

Theorem 5.28. *Let $\eta \in \operatorname{int}(\mathcal{D}^{*h} + K)$ such that $e^T\eta = 1$. If $\bar{T}^\eta$ is a finitely generated solution to* (D$^{*\eta}$)*, then*

$$\bar{T} := \left\{(u,w) \in \mathbb{R}^m \times \mathbb{R}^q \,\middle|\, (u,w,0) \in \bar{T}^\eta\right\}$$

is a finitely generated solution to (D^{*h}).

Proof. Of course, since $\bar{T}^\eta$ is a finite set, $\bar{T}$ is also a finite set. It is immediate that $(u,w,0) \in T^\eta$ implies $(u,w) \in T$. Hence, (i) of Definition 4.56 is satisfied.

Since $D^{*h}(u,w) = (w_1,\dots,w_{q-1},0)^T$, we have $D^{*h}(u,w) \in \operatorname{Max}_K D^{*h}[T]$ for every $(u,w) \in T$, hence (ii) of 4.56 holds for (D^{*h}).

It remains to show condition (iii), that is,

$$D^{*h}[T] \subseteq \operatorname{co} D^{*h}[\bar{T}] - K. \tag{5.6}$$

Condition (iii) of Definition 4.56 holds for (D$^{*\eta}$). This can be expressed (using (i)) as

$$\mathcal{D}^{*\eta} = \operatorname{co} D^{*\eta}[\bar{T}^\eta] - K. \tag{5.7}$$

Moreover, we know that

$$\operatorname{vert}\mathcal{D}^{*\eta} \subseteq D^{*\eta}[\bar{T}^\eta]. \tag{5.8}$$

This follows from (5.7) and the fact that the vertices of $\mathcal{D}^{*\eta}$ cannot be obtained as a nontrivial convex combination (definition of a vertex). Note further that $\mathcal{D}^{*h}$ can be expressed by its vertices as

$$\mathcal{D}^{*h} = \operatorname{co}\operatorname{vert}\mathcal{D}^{*h} - K.$$

Using the definition of $\bar{T}$, we obtain the following chain of inclusions

$$\begin{aligned} D^{*h}[\bar{T}] &= \left\{ y^* \in D^{*\eta}[\bar{T}^\eta] \,|\; y_q^* = 0 \right\} \\ &\overset{(5.8)}{\supseteq} \left\{ y^* \in \operatorname{vert}\mathcal{D}^{*\eta} \,|\; y_q^* = 0 \right\} \overset{\text{Th. 5.25}}{=} \operatorname{vert}\mathcal{D}^{*h}. \end{aligned}$$

Together, we get

$$D^{*h}[T] \subseteq \mathcal{D}^{*h} = \operatorname{co}\operatorname{vert}\mathcal{D}^{*h} - K \subseteq D^{*h}[\bar{T}] - K,$$

i.e., (5.6) holds. □

5.5 Computing an interior point of the lower image

Algorithms 1 and 2 as well as several statements about the connections between (P^h) and (P^η), and between their duals are based on the assumption that there exists an interior point of $\mathcal{D}^*$. Recall that

$$\operatorname{int}\mathcal{D}^{*h} \neq \emptyset \iff \operatorname{int}\mathcal{D}^* \neq \emptyset \iff \operatorname{int}(\mathcal{D}^* + K) \neq \emptyset$$

and

$$\mathcal{D}^* + K = \mathcal{D}^{*h} + K.$$

The following algorithm computes some $\eta \in \operatorname{int}(\mathcal{D}^* + K)$ or states that $\operatorname{int}\mathcal{D}^*$ is empty. In the first case, an interior point of $\mathcal{D}^*$ can be obtained as shown in Proposition 5.13. In the second case, we deduce from the geometric duality theorem that $\mathcal{P}$ has no vertex. From (Rockafellar, 1972, Corollary 18.5.3) we conclude that $\mathcal{P}$ contains a line. It is possible to consider the projection into the orthogonal space of the lineality space of $\mathcal{P}$ in order to obtain the desired properties, see e.g. (Rockafellar, 1972) for the technical details.

The following algorithm computes an interior point of a compact convex set $A \subseteq \mathbb{R}^k$. It is based on the well-known Gram-Schmidt orthogonalization method. The set A has to be given in such a way that the scalar optimization problem

$$\min c^T x \quad \text{subject to } \; x \in A \tag{Q(c)}$$

can be solved for every vector $c \in \mathbb{R}^k$.

Algorithm 3.

Input:
  A ... a convex, compact subset of $\mathbb{R}^k$;
Output:
  IntEmpty=false: $w \in \operatorname{int} A$;
  IntEmpty=true: $\operatorname{int} A = \emptyset$;
Variables:
  $v^0, \dots, v^k \in \mathbb{R}^k$, $c^0, \dots, c^k \in \mathbb{R}^k$, $z \in \mathbb{R}^k$, $w \in \mathbb{R}^k$;
  IntEmpty $\in$ {true, false};

01: begin
02:   IntEmpty $\leftarrow$ false;
03:   $z \leftarrow$ solve(Q(0));
04:   $v^0 \leftarrow 0$;
05:   for $i = 1$ to k do
06:   begin
07:     $c^{i-1} \leftarrow v^{i-1} - \sum_{j=1}^{i-2} \frac{\langle v^{i-1}, c^j \rangle}{\langle c^j, c^j \rangle} c^j$;
08:     $c^i \leftarrow$ solve $\{\langle c^1, x \rangle = 0, \dots, \langle c^{i-1}, x \rangle = 0,\ x \neq 0\}$;
09:     $v^i \leftarrow$ solve(Q(c^i)) $- z$;
10:     if $\langle c^i, v^i \rangle = 0$ then
11:       $v^i \leftarrow$ solve(Q($-c^i$)) $- z$;
12:     if $\langle c^i, v^i \rangle = 0$ then
13:     begin
14:       IntEmpty $\leftarrow$ true;
15:       stop;
16:     end;
17:   end;
18:   $w \leftarrow z + \frac{1}{k+1} \sum_{j=1}^k v^j$;
19: end.

Theorem 5.29. *Let A be convex and compact. Then Algorithm 3 either yields an element $w \in \operatorname{int} A$ or states that $\operatorname{int} A = \emptyset$.*

Proof. In line 03 we compute some $z \in A$. Note first that we perform in line 07 the Gram-Schmidt method which constructs an orthogonal system $\{c^1, \dots c^{k-1}\}$ from a system of linearly independent vectors $\{v^1, \dots v^{k-1}\}$ so that in each iteration $i \in \{1, \dots, k\}$ we have

$$\operatorname{span} \{c^1, \dots c^{i-1}\} = \operatorname{span} \{v^1, \dots v^{i-1}\}.$$

The condition $\langle c^i, v^i \rangle \neq 0$, which holds at the end of iteration i, ensures that the vectors $\{v^1, \dots, v^i\} \subseteq \mathbb{R}^k$ are linearly independent for each $i \in \{1, \dots, k\}$. This can be shown by induction. For $i = 1$, $\langle c^1, v^1 \rangle \neq 0$

implies $v^1 \neq 0$. Let $\{v^1, \dots, v^{i-1}\} \subseteq \mathbb{R}^k$ be a system of linearly independent vectors and assume that $v^i \in \operatorname{span}\{v^1, \dots, v^{i-1}\}$. We have $\operatorname{span}\{v^1, \dots, v^{i-1}\} = \operatorname{span}\{c^1, \dots, c^{i-1}\}$ and thus $v^i \in \operatorname{span}\{c^1, \dots, c^{i-1}\}$. Since $c^i \perp \operatorname{span}\{c^1, \dots, c^{i-1}\}$, which follows from line 08, we get $\langle c^i, v^i \rangle = 0$, a contradiction. Hence, $\{v^1, \dots, v^i\} \subseteq \mathbb{R}^k$ is a system of linearly independent vectors.

The vectors $z, z + v^1, \dots, z + v^k$, being solutions of (Q($c$)) for some $c \in \mathbb{R}^k$, belong to A. The same is true for their convex hull. The vector w as constructed in line 18 satisfies

$$w \in \operatorname{int} \operatorname{co} \{z, z + v^1, \dots, z + v^k\}.$$

It follows that $w \in \operatorname{int} A$.

If the algorithm stops in line 15 with IntEmpty = true, we have the situation

$$\min \{\langle c^i, x \rangle \mid x \in A\} = \langle c^i, z \rangle = \max \{\langle c^i, x \rangle \mid x \in A\}.$$

Since $c^i \neq 0$ (see line 08),

$$H = \{x \mid \langle c^i, x \rangle = \langle c^i, z \rangle\}$$

is a hyperplane. Since $A \subseteq H$, we have $\operatorname{int} A = \emptyset$. $\square$

In order to obtain an interior point of the set $\mathcal{D}^* + K \subseteq \mathbb{R}^q$, we consider the convex and compact set

$$A = \{(w_1, \dots, w_{q-1})^T \mid (u, w) \in T\} \subseteq \mathbb{R}^{q-1},$$

where T is the feasible set of (D*). We can solve Problem (Q(c)) by computing a solution $(\bar{u}, \bar{w})$ to the linear program

$$\text{minimize } c^T (w_1, \dots, w_{q-1})^T \text{ subject to } (u, w) \in T,$$

where $(\bar{w}_1, \dots, \bar{w}_{q-1})^T$ provides a solution to (Q(c)).

If $(w_1, \dots, w_{q-1}) \in \operatorname{int} A$, then $(w_1, \dots, w_{q-1}, \gamma) \in \operatorname{int}(\mathcal{D}^* + K)$ for arbitrary $\gamma \in \mathbb{R}$. Moreover, $\operatorname{int} A = \emptyset$ implies $\operatorname{int}(\mathcal{D}^* + K) = \emptyset$.

5.6 Degeneracy

Outer approximations of $\mathcal{P}$ and $\mathcal{D}^*$ are determined by primal and dual representations as introduced in Sections 5.1 and 5.2. We cannot always ensure that all the vectors and directions representing an outer approximation $\mathcal{T}$ are vertices (extreme points) and extreme directions, respectively; that is, redundant points and directions may occur. Likewise, redundant inequalities can occur in a dual representation of $\mathcal{T}$. To be more precise, we give the following definition.

Definition 5.30. Let $(\mathcal{T}^p, \hat{\mathcal{T}}^p)$ be a primal representation of a polyhedron $\mathcal{T} \subseteq \mathbb{R}^q$ with $\mathcal{T}_\infty \supseteq \mathbb{R}^q_+$. If every vector $y \in \mathcal{T}^p$ is a vertex of $\mathcal{T}$ and every direction $\hat{y} \in \hat{\mathcal{T}}^p$ is an extreme direction of $\mathcal{T}$, $(\mathcal{T}^p, \hat{\mathcal{T}}^p)$ is called *nondegenerate*. Otherwise, $(\mathcal{T}^p, \hat{\mathcal{T}}^p)$ is called *degenerate*.

Likewise, let $\mathcal{T}^p$ be a primal representation of a polyhedron $\mathcal{T} \subseteq \Delta$ with $\mathcal{T}_\infty = -K$. If every vector $y^* \in \mathcal{T}^p$ is a vertex of $\mathcal{T}$, $\mathcal{T}^p$ is called *nondegenerate*. Otherwise, $\mathcal{T}^p$ is called *degenerate*.

Definition 5.31. Let $\mathcal{T}^d$ be a dual representation of a polyhedron $\mathcal{T} \subseteq \mathbb{R}^q$ with $\mathcal{T}_\infty \supseteq \mathbb{R}^q_+$. If every vector $y^* \in \mathcal{T}^d$ belongs to an inequality which is not redundant, $\mathcal{T}^d$ is called *nondegenerate*. Otherwise, if less inequalities would be sufficient to represent $\mathcal{T}$, $\mathcal{T}^d$ is called *degenerate*. The same definition applies for a dual representation $(\mathcal{T}^d, \hat{\mathcal{T}}^d)$ of a polyhedron $\mathcal{T} \subseteq \Delta$ with $\mathcal{T}_\infty = -K$.

The occurrence of degeneracy in Algorithm 1 is demonstrated in the following example.

Example 5.32. Consider Problem (P) with the data

$$P = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad B = \begin{pmatrix} 2 & 1 \\ 1 & 2 \\ 3 & 3 \\ 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad b = \begin{pmatrix} 2 \\ 2 \\ 4 \\ 0 \\ 0 \end{pmatrix}.$$

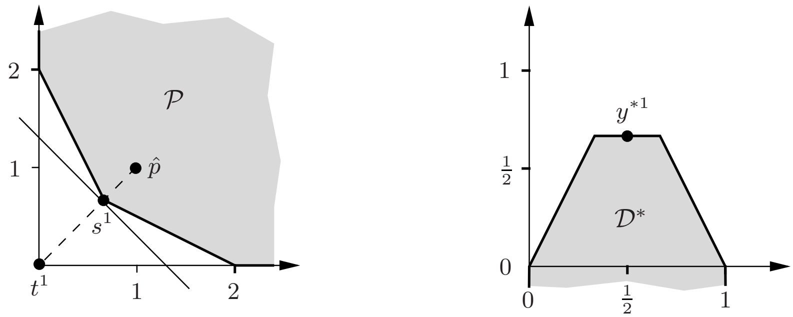

Fig. 5.8 Occurrence of degeneracy in Algorithm 1 (Example 5.32)

We apply Algorithm 1 for the choice $\hat{p} = (1,1)^T$. In the first iteration we obtain $t^1 = (0,0)^T$ and $\mathcal{T}^1 = \mathbb{R}^q_+$. We get $s^1 = (\frac{2}{3}, \frac{2}{3})^T$ and have to solve $(\mathrm{D}_2(s^1))$. This problem has three optimal extreme point solutions $(u^1, w^1)^T$, namely $(0, \frac{1}{3}, 0, 0, 0, \frac{1}{3}, \frac{2}{3})^T$, $(0, 0, \frac{1}{6}, 0, 0, \frac{1}{2}, \frac{1}{2})^T$ and $(\frac{1}{3}, 0, 0, 0, 0, \frac{2}{3}, \frac{1}{3})^T$. In case the algorithm selects the second one, we get the redundant inequality $3y_1 + 3y_2 \geq 4$. The corresponding hyperplane supports $\mathcal{P}$ not in a facet, but in the vertex s^1, see Figure 5.8. Also, for the choice $(u^1, w^1)^T = (0, 0, \frac{1}{6}, 0, 0, \frac{1}{2}, \frac{1}{2})^T$, the point $y^{*1} := D^*(u^1, w^1)$ is not a vertex of $\mathcal{D}^*$. This means, Algorithm 1 yields a degenerate dual representation of $\mathcal{P}$ and a degenerate primal representation of $\mathcal{D}^*$.

The following example illustrates the occurrence of degenerate representations of $\mathcal{P}$ and $\mathcal{D}^*$ in Algorithm 2.

Example 5.33. Consider the bounded problem (P) with the data

$$P = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}, \quad B = \begin{pmatrix} 7 & 21 & 9 \\ 0 & 0 & -1 \\ -7 & -42 & 3 \\ 1 & 7 & 0 \end{pmatrix}, \quad b = \begin{pmatrix} 30 \\ -1 \\ -39 \\ 6 \end{pmatrix}.$$

We apply Algorithm 2 for the choice $\hat{d} = (\frac{1}{2}, 0)^T$. In the initialization phase we solve $P_1(w(\hat{d}))$ and obtain the unique optimal solution $x^0 = (0, 1, 1)^T$. The initial outer approximation is therefore

$$\mathcal{T}^0 = \{y^* \in \mathbb{R}^q |\ 0 \leq y_1^* \leq 1,\ y_2^* \leq 1 - y_1^*\}.$$

There is exactly one vertex of $\mathcal{T}^1$ that does not belong to $\mathcal{D}^*$, namely $t^{*1} = (0, 1)^T$. In the first iteration we obtain $s^{*1} = (\frac{1}{8}, \frac{3}{4})^T$. Note that s^{*1} is not in the relative interior of a facet of $\mathcal{D}^*$ because it is a vertex of $\mathcal{D}^*$. We solve $(\mathrm{P}_1(w(s^{*1})))$ and obtain three optimal solutions x^1 that are extreme points of the feasible set S of $(\mathrm{P}_1(w(s^{*1})))$, namely, $(\frac{3}{4}, \frac{3}{4}, 1)^T$, $(3, \frac{3}{7}, 0)^T$ and $(6, 0, 1)^T$. In case the algorithm selects the second one, we get the redundant inequality $-\frac{18}{7} y_1^* + y_2^* \leq \frac{3}{7}$. The corresponding hyperplane supports $\mathcal{D}^*$ not in a facet but in the vertex s^{*1} (see Figure 5.9). For the choice $x^1 = (3, \frac{3}{7}, 0)^T$, the point $y^1 := Px^1$ is not a vertex of $\mathcal{P}$ (see Figure 5.9). This means, Algorithm 2 yields a degenerate dual representation of $\mathcal{D}^*$ and a degenerate primal representation of $\mathcal{P}$.

5.7 Notes on the literature

Many algorithms for linear vector optimization problems are based on extensions of the simplex method, see for example the contributions by Evans and Steuer (1973); Yu and Zeleny (1976); Steuer (1985, 1989); Armand and Malivert (1991); Armand (1993) and also the brief survey in (Ehrgott and Wiecek, 2005a). A classification of various methods to determine the whole set of efficient solutions can be found in (Pourkarimi *et al.*, 2009).

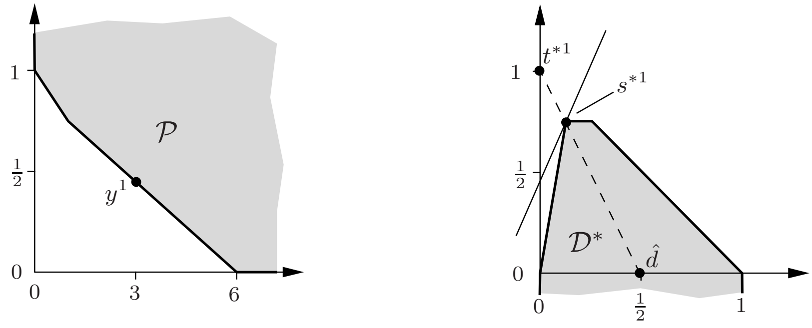

Fig. 5.9 Occurrence of degeneracy in Algorithm 2 (Example 5.33)

Benson (1998a,b) proposed a method which evaluates (weakly) minimal points in the objective space. Algorithm 1 is based on this method. In particular, Proposition 5.4, Theorem 5.6 and many ideas are due to Benson (1998b). In Benson's original variant of Algorithm 1, the feasible set S is supposed to be bounded. This assumption could be weakened by Ehrgott *et al.* (2007), where $\mathcal{P}$ is supposed to be $\mathbb{R}^q_+$-bounded below. Moreover, the mentioned paper contains a simplification of the algorithm. Using the upper image $\mathcal{P}$, the minimal vertices of $P[S]$ have been computed directly and the final step (Benson, 1998b, Theorem 3.2) to check whether a vertex is minimal or not could be omitted. The dual variant of Benson's algorithm first appeared in (Ehrgott *et al.*, 2007). Algorithm 2 is an extension of these results in the sense that the boundedness assumption has been omitted.

The considerations and examples on degeneracy in Section 5.6 are due to Ehrgott *et al.* (2007).

A treatment of approximate solutions can be found in (Shao and Ehrgott, 2008a,b). Moreover, an extended algorithm for nonlinear problems has been developed by Ehrgott *et al.* (2010). For applications and numerical investigations of the algorithms the reader is referred to (Ehrgott *et al.*, 2007; Shao and Ehrgott, 2008a,b) and (Werfel, 2009).

References

Adán, M. and Novo, V. (2005). Duality and saddle-points for convex-like vector optimization problems on real linear spaces. *Top*, **13**(2), 343–357.

Ait Mansour, M., Malivert, C., and Théra, M. (2007). Semicontinuity of vector-valued mappings. *Optimization*, **56**(1-2), 241–252.

Aliprantis, C. D. and Border, K. C. (1994). *Infinite dimensional analysis. A hitchhiker's guide.* Studies in Economic Theory. 4. Berlin: Springer-Verlag.

Anderson, B. and Annulis, J. (1973). A non-Archimedean vector lattice. *Am. Math. Mon.*, **80**(9), 1070–1071.

Armand, P. (1993). Finding all maximal efficient faces in multiobjective linear programming. *Math. Program., Ser. A*, **61**(3), 357–375.

Armand, P. and Malivert, C. (1991). Determination of the efficient set in multiobjective linear programming. *J. Optimization Theory Appl.*, **70**(3), 467–489.

Arrow, K., Barankin, E., and Blackwell, D. (1953). Admissible points of convex sets. Contrib. Theory of Games, II, Ann. Math. Stud. No. 28, 87-91 (1953).

Beer, G. (1993). *Topologies on closed and closed convex sets.* Mathematics and its Applications (Dordrecht). 268. Dordrecht: Kluwer Academic Publishers. xi.

Benson, H. (1998a). Further analysis of an outcome set-based algorithm for multiple-objective linear programming. *Journal of Optimization Theory and Applications*, **97**(1), 1–10.

Benson, H. (1998b). An outer approximation algorithm for generating all efficient extreme points in the outcome set of a multiple objective linear programming problem. *Journal of Global Optimization*, **13**, 1–24.

Benson, H. and Sun, E. (2000). Outcome space partition of the weight set in multiobjective linear programming. *Journal of Optimization Theory and Applications*, **105**(1), 17–36.

Birkhoff, G. (1979). *Lattice theory. Corr. repr. of the 1967 3rd ed.* American Mathematical Society Colloquium Publications, 25. Providence, R.I.: American Mathematical Society (AMS).

A. Löhne, *Vector Optimization with Infimum and Supremum*, Vector Optimization,
DOI 10.1007/978-3-642-18351-5, © Springer-Verlag Berlin Heidelberg 2011

Borwein, J. (1977). Multivalued convexity and optimization: A unified approach to inequality and equality constraints. *Math. Program.*, **13**, 183–199.

Borwein, J. (1980). The geometry of Pareto efficiency over cones. *Math. Operationsforsch. Stat., Ser. Optimization*, **11**, 235–248.

Borwein, J. M. and Lewis, A. S. (2000). *Convex analysis and nonlinear optimization. Theory and examples.* CMS Books in Mathematics/Ouvrages de Mathématiques de la SMC. 3. New York, NY: Springer.

Boţ, R. I., Grad, S.-M., and Wanka, G. (2009). *Duality in vector optimization.* Vector Optimization. Dordrecht: Springer.

Bourbaki, N. (1989). *Elements of mathematics. General topology. Chapters 1-4. Transl. from the French. 2nd printing.* Berlin etc.: Springer-Verlag.

Brumelle, S. (1981). Duality for multiple objective convex programs. *Math. Oper. Res.*, **6**, 159–172.

Chen, C. and Li, S. (2009). Different conjugate dual problems in vector optimization and their relations. *J. Optim. Theory Appl.*, **140**(3), 443–461.

Chen, G.-Y., Huang, X., and Yang, X. (2005). *Vector optimization. Set-valued and variational analysis.* Lecture Notes in Economics and Mathematical Systems 541. Berlin: Springer.

Chen, P.-C., Hansen, P., and Jaumard, B. (1991). On-line and off-line vertex enumeration by adjacency lists. *Oper. Res. Lett.*, **10**(7), 403–409.

Corley, H. (1987). Existence and Lagrangian duality for maximization of set-valued functions. *J. Optimization Theory Appl.*, **54**, 489–501.

Dolecki, S. and Malivert, C. (1993). General duality in vector optimization. *Optimization*, **27**(1-2), 97–119.

Ehrgott, M. (2000). *Multicriteria optimization.* Lecture Notes in Economics and Mathematical Systems. 491. Berlin: Springer.

Ehrgott, M. (2005). *Multicriteria Optimization.* Springer-Verlag, Berlin, 2nd edition.

Ehrgott, M. and Wiecek, M. (2005a). Multiobjective programming. In J. Figueira, S. Greco, and M. Ehrgott, editors, *Multicriteria Decision Analysis: State of the Art Surveys*, pages 667–722. Springer Science + Business Media, New York.

Ehrgott, M. and Wiecek, M. M. (2005b). Saddle points and Pareto points in multiple objective programming. *J. Glob. Optim.*, **32**(1), 11–33.

Ehrgott, M., Löhne, A., and Shao, L. (2007). A dual variant of Benson's outer approximation algorithm. Report 12, Departement of Mathematics, Martin-Luther-Universität Halle-Wittenberg.

Ehrgott, M., Shao, L., and Schöbel, A. (2010). An approximation algorithm for convex multi-objective programming problems. *manuscript*.

Eichfelder, G. (2008). *Adaptive scalarization methods in multiobjective optimization.* Vector Optimization. Berlin: Springer.

Ekeland, I. and Temam, R. (1976). *Convex analysis and variational problems. Translated by Minerva Translations, Ltd., London.* Studies in Mathematics and its Applications. Vol. 1. Amsterdam - Oxford: North-Holland Publishing Company; New York: American Elsevier Publishing Company.

Evans, J. and Steuer, R. (1973). A revised simplex method for linear multiple objective programs. *Math. Program.*, **5**, 54–72.

Ferro, F. (1999). A new ABB theorem in Banach spaces. *Optimization*, **46**(4), 353–362.

Fu, W. (1996). On the density of proper efficient points. *Proc. Am. Math. Soc.*, **124**(4), 1213–1217.

Gale, D., Kuhn, H. W., and Tucker, A. W. (1951). Linear programming and the theory of games. Activity Analysis of Production and Allocation, Chap. XIX, 317-329 (1951).

Gerritse, G. (1997). Lattice-valued semicontinuous functions. Vervaat, W. (ed.) et al., Probability and lattices. Amsterdam: Centrum voor Wiskunde en Informatica (CWI). CWI Tracts. 110, 93-125 (1997).

Gerstewitz (Tammer), C. (1983). Nichtkonvexe Dualität in der Vektoroptimierung.

Gerth (Tammer), C. and Weidner, P. (1990). Nonconvex separation theorems and some applications in vector optimization. *J. Optimization Theory Appl.*, **67**(2), 297–320.

Göpfert, A. (1973). *Mathematische Optimierung in allgemeinen Vektorräumen.* Mathematisch-Naturwissenschaftliche Bibliothek. Band 58. Leipzig: BSB B. G. Teubner Verlagsgesellschaft.

Göpfert, A. and Nehse, R. (1990). *Vektoroptimierung. Theorie, Verfahren und Anwendungen. (Vector optimization. Theory, methods and applications).* Mathematisch-Naturwissenschaftliche Bibliothek, 74. Leipzig: BSB B.G. Teubner.

Göpfert, A., Tammer, C., Riahi, H., and Zălinescu, C. (2003). *Variational methods in partially ordered spaces.* CMS Books in Mathematics/Ouvrages de Mathématiques de la SMC 17. New York, NY: Springer.

Göpfert, A., Riedrich, T., and Tammer, C. (2009). *Applied functional analysis. Motivations and methods for mathematicians and economists. (Angewandte Funktionalanalysis. Motivationen und Methoden für Mathematiker und Wirtschaftswissenschaftler.).* Studium. Studienbücher Wirtschaftsmathematik. Wiesbaden: Vieweg+Teubner.

Götz, A. and Jahn, J. (1999). The Lagrange multiplier rule in set-valued optimization. *SIAM J. Optim.*, **10**(2), 331–344.

Gros, C. (1978). Generalization of Fenchel's duality theorem for convex vector optimization. *Eur. J. Oper. Res.*, **2**, 368–376.

Grünbaum, B. (2003). *Convex Polytopes*, volume 221 of *Graduate Texts in Mathematics*. Springer, New York, 2nd edition. Prepared by V. Kaibel, V. Klee, and G.M. Ziegler.

Ha, T. X. D. (2005). Lagrange multipliers for set-valued optimization problems associated with coderivatives. *J. Math. Anal. Appl.*, **311**(2), 647–663.

Hamel, A. (2005). *Variational principles on metric and uniform spaces.* Habilitation thesis, Martin-Luther-Universität Halle-Wittenberg.

Hamel, A. (2009a). A duality theory for set-valued functions I: Fenchel conjugation theory. *J. Set-valued and Variational Analysis*, **17**, 153–182.

Hamel, A. (2009b). A Fenchel-Rockafellar duality theorem for set-valued optimization. *manuscript*.

Hamel, A. and Heyde, F. (2010). Duality for set-valued measures of risk. *SIAM Journal on Financial Mathematics*, **1**, 66–95.

Hamel, A. and Löhne, A. (2006). Minimal element theorems and Ekeland's principle with set relations. *J. Nonlinear Convex Anal.*, **7**(1), 19–37.

Hamel, A. and Rudloff, B. (2008). Continuity and finite-valuedness of set-valued risk measures. Tammer, Christiane (ed.) et al., Festschrift in celebration of Prof. Dr. Wilfried Grecksch's 60th birthday. Aachen: Shaker Verlag. Berichte aus der Mathematik, 49-64 (2008).

Hamel, A., Heyde, F., Löhne, A., Tammer, C., and Winkler, K. (2004). Closing the duality gap in linear vector optimization. *J. Convex Anal.*, **11**(1), 163–178.

Hernández, E. and Rodríguez-Marín, L. (2007). Existence theorems for set optimization problems.

Hernández, E., Rodríguez-Marín, L., and Sama, M. (2009). Some equivalent problems in set optimization. *Oper. Res. Lett.*, **37**(1), 61–64.

Heyde, F. and Löhne, A. (2008). Geometric duality in multiple objective linear programming. *SIAM J. Optim.*, **19**(2), 836–845.

Heyde, F. and Löhne, A. (2010). Solution concepts for vector optimization. A fresh look at an old story. *Optimization*. published online.

Heyde, F., Löhne, A., and Tammer, C. (2009a). Set-valued duality theory for multiple objective linear programs and application to mathematical finance. *Math. Methods Oper. Res.*, **69**(1), 159–179.

Heyde, F., Löhne, A., and Tammer, C. (2009b). The attainment of the solution of the dual program in vertices for vectorial linear programs. Barichard, Vincent (ed.) et al., Multiobjective programming and goal programming. Theoretical results and practical applications. Springer. Lecture Notes in Economics and Mathematical Systems 618, 13-24 (2009).

Isac, G. (1987). Supernormal cones and fixed point theory. *Rocky Mt. J. Math.*, **17**, 219–226.

Isermann, H. (1974a). *Lineare Vektoroptimierung.* Ph.D. thesis, Universität Regensburg.

Isermann, H. (1974b). Proper efficiency and the linear vector maximum problem. *Oper. Res.*, **22**, 189–191.

Isermann, H. (1978a). Duality in multiple objective linear programming. Multiple Criteria Probl. Solving, Proc. Conf. Buffalo 1977, Lect. Notes Econ. Math. Syst. 155, 274-285 (1978).

Isermann, H. (1978b). On some relations between a dual pair of multiple objective linear programs. *Z. Oper. Res., Ser. A*, **22**, 33–41.

Jahn, J. (1983). Duality in vector optimization. *Math. Program.*, **25**, 343–353.

Jahn, J. (1986). *Mathematical vector optimization in partially ordered linear spaces.* Methoden und Verfahren der Mathematischen Physik, Bd. 31. Frankfurt am Main-Bern-New York: Verlag Peter Lang.

Jahn, J. (1988). A generalization of a theorem of Arrow, Barankin, and Blackwell. *SIAM J. Control Optimization*, **26**(5), 999–1005.

Jahn, J. (2003). Grundlagen der Mengenoptimierung. Habenicht, Walter (ed.) et al., Multi-criteria and fuzzy systems in theory and practice. Wiesbaden: DUV, Deutscher Universitäts-Verlag. 37-71.

Jahn, J. (2004). *Vector optimization. Theory, applications, and extensions.* Berlin: Springer.

Kawasaki, H. (1981). Conjugate relations and weak subdifferentials of relations. *Math. Oper. Res.*, **6**, 593–607.

Kawasaki, H. (1982). A duality theorem in multiobjective nonlinear programming. *Math. Oper. Res.*, **7**, 95–110.

Kelley, J., Namioka, I., Donoghue, W., Lucas, K., Pettis, B., Poulsen, T., Price, G., Robertson, W., Scott, W., and Smith, K. (1963). *Linear topological spaces.* (The University Series in Higher Mathematics) Princeton, N.J. -Toronto- New York-London: D. van Nostrand Company.

Kelley, J. L. (1955). *General topology.* (The University Series in Higher Mathematics.) New York: D. van Nostrand Co.

Koopmans, T. C. (1951). Analysis of production as an efficient combination of activities. Activity Analysis of Production and Allocation, Chap. III, 33-97 (1951).

Kornbluth, J. (1974). Duality, indifference and sensitivity analysis in multiple objective linear programming. *Oper. Res. Quart.*, **25**, 599–614.

Köthe, G. (1966). *Topologische lineare Räume. I. 2., verbesserte Aufl.* Berlin-Heidelberg-New York: Springer-Verlag.

Kuhn, H. and Tucker, A. (1951). Nonlinear programming. Proc. Berkeley Sympos. math. Statist. Probability, California July 31 - August 12, 1950, 481-492 (1951).

Kuroiwa, D. (1998a). Lagrange duality of set-valued optimization with natural criteria. *RIMS Kokyuroku*, **1068**, 164–170.

Kuroiwa, D. (1998b). On natural criteria in set-valued optimization. *RIMS Kokyuroku*, **1048**, 86–92.

Li, S., Chen, C., and Wu, S. (2009). Conjugate dual problems in constraint set-valued optimization and applications. *Europ. J. Oper. Res.*, **196**, 21–32.

Li, Z. and Chen, G. (1997). Lagrangian multipliers, saddle points, and duality in vector optimization of set-valued maps. *J. Math. Anal. Appl.*, **215**(2), 297–316.

Li, Z. and Wang, S. (1994). Lagrange multipliers and saddle points in multiobjective programming. *J. Optimization Theory Appl.*, **83**(1), 63–81.

Liu, Y. and Luo, M. (1991). Lattice-valued mappings, completely distributive law and induced spaces. *Fuzzy Sets Syst.*, **42**(1), 43–56.

Löhne, A. (2005a). *Optimization with set relations.* Ph.D. thesis, Martin-Luther-Universität Halle-Wittenberg.

Löhne, A. (2005b). Optimization with set relations: conjugate duality. *Optimization*, **54**(3), 265–282.

Löhne, A. (2006). On convex functions with values in conlinear spaces. *J. Nonlinear Convex Anal.*, **7**(1), 115–122.

Löhne, A. (2008). On semicontinuity of convex-valued multifunctions and Cesari's property (Q). *J. Convex Anal.*, **15**(4), 803–818.

Löhne, A. (2010). *Vector optimization with infimum and supremum.* Habilitation thesis, Martin-Luther-Universität Halle-Wittenberg.

Löhne, A. and Tammer, C. (2007). A new approach to duality in vector optimization. *Optimization*, **56**(1-2), 221–239.

Löhne, A. and Zălinescu, C. (2006). On convergence of closed convex sets. *J. Math. Anal. Appl.*, **319**(2), 617–634.

Luc, D. T. (1988). *Theory of vector optimization.* Lecture Notes in Economics and Mathematical Sciences, 319. Berlin etc.: Springer-Verlag.

Martein, L. (1990). An approach to Lagrangian duality in vector optimization. Generalized convexity and fractional programming with economic applications, Proc. Workshop, Pisa/Italy 1988, Lect. Notes Econ. Math. Syst. 345, 242-251 (1990).

Martínez-Legaz, J.-E. and Singer, I. (1994). $*$-dualities. *Optimization*, **30**(4), 295–315.

Nieuwenhuis, J. (1980). Supremal points and generalized duality. *Math. Operationsforsch. Stat., Ser. Optimization*, **11**, 41–59.

Oettli, W. (1982). Optimality conditions for programming problems involving multivalued mappings. Modern applied mathematics, optimization and operations research, Lect. Summer Sch., Bonn 1979, 195-226 (1982).

Pallaschke, D. and Rolewicz, S. (1997). *Foundations of mathematical optimization: convex analysis without linearity.* Mathematics and its Applications (Dordrecht). 388. Dordrecht: Kluwer Academic Publishers.

Pascoletti, A. and Serafini, P. (1984). Scalarizing vector optimization problems. *J. Optimization Theory Appl.*, **42**, 499–524.

Penot, J. and Théra, M. (1982). Semi-continuous mappings in general topology. *Arch. Math.*, **38**, 158–166.

Peressini, A. (1967). *Ordered topological vector spaces.* New York-Evanston-London: Harper and Row Publishers.

Pourkarimi, L., Yaghoobi, M., and Mashinchi, M. (2009). Determining maximal efficient faces in multiobjective linear programming problem. *J. Math. Anal. Appl.*, **354**(1), 234–248.

Rockafellar, R. (1972). *Convex Analysis*. Princeton University Press, Princeton.

Rockafellar, R. (1974). *Conjugate duality and optimization.* CBMS-NSF Regional Conference Series in Applied Mathematics. 16. Philadelphia, Pa.: SIAM, Society for Industrial and Applied Mathematics. VI.

Rockafellar, R. and Wets, R. J.-B. (1998). *Variational analysis.* Grundlehren der Mathematischen Wissenschaften. 317. Berlin: Springer.

Rödder, W. (1977). A generalized saddlepoint theory. Its application to duality theory for linear vector optimum problems. *Eur. J. Oper. Res.*, **1**, 55–59.

Sawaragi, Y., Nakayama, H., and Tanino, T. (1985). *Theory of multiobjective optimization.* Mathematics in Science and Engineering, Vol. 176. Orlando etc.: Academic Press, Inc. (Harcourt Brace Jovanovich, Publishers).

Schaefer, H. H. (1980). *Topological vector spaces. 4th corr. printing.* Graduate Texts in Mathematics, 3. New York-Heidelberg-Berlin: Springer-Verlag.

Schrage, C. (2009). *Set-valued Convex Analysis*. Ph.D. thesis, Martin-Luther-Universität Halle-Wittenberg.

Shao, L. and Ehrgott, M. (2008a). Approximately solving multiobjective linear programmes in objective space and an application in radiotherapy treatment planning. *Math. Methods Oper. Res.*, **68**(2), 257–276.

Shao, L. and Ehrgott, M. (2008b). Approximating the nondominated set of an MOLP by approximately solving its dual problem. *Math. Methods Oper. Res.*, **68**(3), 469–492.

Song, W. (1997). Conjugate duality in set-valued vector optimization. *J. Math. Anal. Appl.*, **216**(1), 265–283.

Song, W. (1998). A generalization of Fenchel duality in set-valued vector optimization. *Math. Methods Oper. Res.*, **48**(2), 259–272.

Sonntag, Y. and Zălinescu, C. (1992). Scalar convergence of convex sets. *J. Math. Anal. Appl.*, **164**(1), 219–241.

Sonntag, Y. and Zălinescu, C. (2000). Comparison of existence results for efficient points. *J. Optimization Theory Appl.*, **105**(1), 161–188.

Stadler, W. (1979). A survey of multicriteria optimization or the vector maximum problem. I: 1776-1960. *J. Optimization Theory Appl.*, **29**, 1–52.

Steuer, R. (1985). *Multiple Criteria Optimization: Theory, Computation and Application*. John Wiley & Sons, New York, NY.

Steuer, R. (1989). Adbase multiple objective linear programming package. Technical report, University of Georgia, Athens, Georgia.

Tammer, C. (1991). Lagrange duality in vector optimization. (Lagrange-Dualität in der Vektoroptimierung.). *Wiss. Z. Tech. Hochsch. Ilmenau*, **37**(3), 71–88.

Tan, K., Yu, J., and Yuan, X. (1996). Existence theorems for saddle points of vector-valued maps. *J. Optimization Theory Appl.*, **89**(3), 731–747.

Tanaka, T. (1990). A characterization of generalized saddle points for vector-valued functions via scalarization. *Nihonkai Math. J.*, **1**(2), 209–227.

Tanaka, T. (1994). Generalized quasiconvexities, cone saddle points, and minimax theorem for vector-valued functions. *J. Optimization Theory Appl.*, **81**(2), 355–377.

Tanino, T. (1988). On supremum of a set in a multidimensional space. *J. Math. Anal. Appl.*, **130**(2), 386–397.

Tanino, T. (1992). Conjugate duality in vector optimization. *J. Math. Anal. Appl.*, **167**(1), 84–97.

Tanino, T. and Sawaragi, Y. (1980). Conjugate maps and duality in multiobjective optimization. *J. Optimization Theory Appl.*, **31**, 473–499.

Webster, R. (1994). *Convexity*. Oxford Science Publications, Oxford, Oxford University Press.

Weidner, P. (1990). An approach to different scalarizations in vector optimization. *Wiss. Z. Tech. Hochsch. Ilmenau*, **36**(3), 103–110.

Werfel, M. (2009). *Multikriterielle Portfolio-Optimierung mit dem Conditional Value at Risk.* Diploma thesis, Martin-Luther-Universität Halle-Wittenberg.

Yu, P. and Zeleny, M. (1976). Linear multiparametric programming by multicriteria simplex method. *Manage. Sci.*, **23**, 159–170.

Zowe, J. (1975). A duality theorem for a convex programming problem in order complete vector lattices. *J. Math. Anal. Appl.*, **50**, 273–287.

Zălinescu, C. (1983). Duality for vectorial nonconvex optimization by convexification and applications.

Zălinescu, C. (2002). *Convex analysis in general vector spaces.* Singapore: World Scientific. xx.

Index