2

Preliminaries on convex analysis and vector optimization

In this chapter we introduce some basic notions and results in convex analysis and vector optimization in order to make the book as self-contained as possible. The reader is supposed to have basic notions of functional analysis.

2.1 Convex sets

This section is dedicated mainly to the presentation of convex sets and their properties. With some exceptions the results we present in this section are given without proofs, as these can be found in the books and monographs on this topic mentioned in the bibliographical notes at the end of the chapter. All around this book we denote by $\mathbb{R}^n$ the *n-dimensional real vector space*, while by $\mathbb{R}^n_+ = \big\{x = (x_1, \ldots, x_n)^T \in \mathbb{R}^n : x_i \geq 0 \text{ for all } i = 1, \ldots, n\big\}$ we denote its *nonnegative orthant*. By $\mathbb{N} = \{1, 2, ...\}$ we denote the *set of natural numbers*, while $\emptyset$ is the *empty set*. All the vectors are considered as column vectors. An upper index ^{T} transposes a column vector to a row one and vice versa. By $\mathbb{R}^{m\times n}$ we denote the space of the *$m \times n$ matrices with real entries*. When we have a matrix $A \in \mathbb{R}^{m\times n}$, by A_i, $i = 1, \ldots, m$, we denote its rows and, naturally, by A^T its *transpose*. By $e^i \in \mathbb{R}^n$ we denote the *i-th unit vector of* $\mathbb{R}^n$, while by $e := \sum_{i=1}^n e^i \in \mathbb{R}^n$ we understand the vector having all entries equal to 1. If a function f takes everywhere the value $a \in \overline{\mathbb{R}}$ we write $f \equiv a$.

2.1.1 Algebraic properties of convex sets

Let X be a real nontrivial vector space. A *linear subspace* of X is a nonempty subset of it which is invariant with respect to the addition and the scalar multiplication on X. Note that an intersection of linear subspaces is itself a linear subspace. The *algebraic dual* space of X is defined as the set of all linear functionals on X and it is denoted by $X^\#$. Given any linear functional $x^\# \in X^\#$, we denote its value at $x \in X$ by $\langle x^\#, x\rangle$.

R.I. Boţ et al., *Duality in Vector Optimization*, Vector Optimization,
DOI 10.1007/978-3-642-02886-1_2,

For $x^\# \in X^\# \backslash \{0\}$ and $\lambda \in \mathbb{R}$ the set $\mathcal{H} := \{x \in X : \langle x^\#, x\rangle = \lambda\}$ is called *hyperplane*. The sets $\{x \in X : \langle x^\#, x\rangle \leq \lambda\}$ and $\{x \in X : \langle x^\#, x\rangle \geq \lambda\}$ are the *closed halfspaces* determined by the hyperplane $\mathcal{H}$, while $\{x \in X : \langle x^\#, x\rangle < \lambda\}$ and $\{x \in X : \langle x^\#, x\rangle > \lambda\}$ are the *open halfspaces* determined by $\mathcal{H}$. In order to simplify the presentation, the origins of all spaces will be denoted by 0, since the space where this notation is used always arises from the context. By Δ_{X^m} we denote the set $\{(x,\ldots,x) \in X^m : x \in X\}$, which is a linear subspace of $X^m := X \times ... \times X = \{(x_1,\ldots,x_m) : x_i \in X, i = 1,\ldots,m\}$.

If U and V are two subsets of X, their *Minkowski sum* is defined as $U + V := \{u + v : u \in U,\ v \in V\}$. For $U \subseteq X$ we define also $x + U = U + x := U + \{x\}$ when $x \in X$, $\lambda U := \{\lambda u : u \in U\}$ when $\lambda \in \mathbb{R}$ and $\Lambda U := \cup_{\lambda \in \Lambda} \lambda U$ when $\Lambda \subseteq \mathbb{R}$. According to these definitions one has that $U + \emptyset = \emptyset + U = \emptyset$ and $\lambda\emptyset = \emptyset$ whenever $U \subseteq X$ and $\lambda \in \mathbb{R}$. Moreover, if $U \subseteq V \subseteq X$ and $U \neq V$ we write $U \subsetneqq V$.

Some important classes of subsets of a real vector space X follow. Let be $U \subseteq X$. If $[-1,1]U \subseteq U$, then U is said to be a *balanced* set. When $U = -U$ we say that U is *symmetric*, while U is called *absorbing* if for all $x \in X$ there is some $\lambda > 0$ such that one has $x \in \lambda U$.

Affine and convex sets. Before introducing the notions of affine and convex sets, some necessary prerequisites follow. Taking some $x_i \in X$ and $\lambda_i \in \mathbb{R}$, $i = 1,\ldots,n$, the sum $\sum_{i=1}^n \lambda_i x_i$ is said to be a *linear combination* of the vectors $\{x_i : i = 1,\ldots,n\}$. The vectors $x_i \in X$, $i = 1,\ldots,n$, are called *linearly independent* if from $\sum_{i=1}^n \lambda_i x_i = 0$ follows $\lambda_i = 0$ for all $i = 1,\ldots,n$. The *linear hull* of a set $U \subseteq X$,

$$\text{lin}(U) := \left\{ \sum_{i=1}^n \lambda_i x_i :\ n \in \mathbb{N}, x_i \in U, \lambda_i \in \mathbb{R},\ i = 1,\ldots,n \right\},$$

is the intersection of all linear subspaces containing U, being the smallest linear subspace having U as a subset.

The set $U \subseteq X$ is called *affine* if $\lambda x + (1-\lambda)y \in U$ whenever $\lambda \in \mathbb{R}$. The intersection of arbitrarily many affine sets is affine, too. The smallest affine set containing U or, equivalently, the intersection of all affine sets having U as a subset is the *affine hull* of U,

$$\text{aff}(U) := \left\{ \sum_{i=1}^n \lambda_i x_i :\ n \in \mathbb{N}, x_i \in U, \lambda_i \in \mathbb{R},\ i = 1,\ldots,n, \sum_{i=1}^n \lambda_i = 1 \right\}.$$

A set $U \subseteq X$ is called *convex* if

$$\{\lambda x + (1-\lambda)y : \lambda \in [0,1]\} \subseteq U \text{ for all } x, y \in U.$$

Obviously, $\emptyset$ and the whole space X are convex sets, as well as the hyperplanes, linear subspaces, affine sets and any set containing a single element. An example of a convex set in $\mathbb{R}^n$ is the *standard* $(n-1)$*-simplex* which is the set $\Delta_n := \{x = (x_1,\ldots,x_n)^T \in \mathbb{R}^n_+ : \sum_{i=1}^n x_i = 1\}$. Given $x_i \in X$, $i = 1,\ldots,n$, and $(\lambda_1, \ldots, \lambda_n)^T \in \Delta_n$, the sum $\sum_{i=1}^n \lambda_i x_i$ is said to be a *convex combination* of the elements x_i, $i = 1, \ldots, n$. The intersection of arbitrarily many convex sets is convex, while in general the union of convex sets is not convex. Note also that when U and V are convex subsets of X, for all $\alpha, \beta \in \mathbb{R}$ the set $\alpha U + \beta V$ is convex, too.

If X_i, $i = 1, \ldots, m$, are nontrivial real vector spaces, then $U_i \subseteq X_i$, $i = 1, \ldots, m$, are convex sets if and only if $\prod_{i=1}^m U_i$ is a convex set in $\prod_{i=1}^m X_i$. When X and Y are nontrivial real vector spaces and $U \subseteq X \times Y$, the *projection of* U *on* X is the set $\mathrm{Pr}_X(U) := \{x \in X : \exists y \in Y \text{ such that } (x, y) \in U\}$. If U is convex then $\mathrm{Pr}_X(U)$ is convex, too.

When U is a subset of the real vector space X, the intersection of all convex sets containing U is the *convex hull* of U,

$$\mathrm{co}(U) := \left\{ \sum_{i=1}^{n} \lambda_i x_i : \ n \in \mathbb{N}, x_i \in U, \ i = 1, \ldots, n, (\lambda_1, \ldots, \lambda_n)^T \in \Delta_n \right\},$$

which is the smallest convex set with U as a subset. If U and V are subsets of X, for all $\alpha, \beta \in \mathbb{R}$ one gets $\mathrm{co}(\alpha U + \beta V) = \alpha\, \mathrm{co}(U) + \beta\, \mathrm{co}(V)$.

A special case of convex sets are the *polyhedral sets* which are finite intersections of closed halfspaces. If U and V are polyhedral sets, then for all $\lambda, \mu \in \mathbb{R}$ the set $\lambda U + \mu V$ is polyhedral, too.

Consider another nontrivial real vector space Y and let $T : X \to Y$ be a given mapping. The *image* of a set $U \subseteq X$ through T is the set $T(U) := \{T(u) : u \in U\}$, while the *counter image* of a set $W \subseteq Y$ through T is $T^{-1}(W) := \{x \in X : T(x) \in W\}$. The mapping A is called *linear* if $A(x+y) = Ax + Ay$ and $A(\lambda x) = \lambda Ax$ for all $x, y \in X$ and all $\lambda \in \mathbb{R}$ or, equivalently, if

$$A(\alpha x + \beta y) = \alpha Ax + \beta Ay \ \forall x, y \in X \ \forall \alpha, \beta \in \mathbb{R}.$$

If $A : X \to Y$ is a linear mapping and $U \subseteq X$ is a linear subspace, then $A(U)$ is a linear subspace, too. On the other hand, if $W \subseteq Y$ is a linear subspace, then $A^{-1}(W)$ is a linear subspace, too. A special linear mapping is the *identity function* on X, $\mathrm{id}_X : X \to X$ defined by $\mathrm{id}_X(x) = x$ for all $x \in X$.

The mapping $T : X \to Y$ is said to be *affine* if

$$T(\lambda x + (1 - \lambda)y) = \lambda T(x) + (1 - \lambda)T(y) \ \forall x, y \in X \ \forall \lambda \in \mathbb{R}.$$

If $T : X \to Y$ is an affine mapping and the set $U \subseteq X$ is affine (or convex), then $T(U)$ is affine (or convex), too. Moreover, if $W \subseteq Y$ is affine (or convex), then $T^{-1}(W)$ is affine (or convex), too.

Cones. A nonempty set $K \subseteq X$ which satisfies the condition $\lambda K \subseteq K$ for all $\lambda \geq 0$ is said to be a *cone*. Throughout this book we assume, as follows by the definition, that the considered cones always contain the origin. The intersection of a family of cones is a cone, too.

A *convex cone* is a cone which is a convex set. One can prove that a cone K is convex if and only if $K + K \subseteq K$. If K is a convex cone, then its *linearity* *space* $l(K) = K \cap (-K)$ is a linear subspace. A cone K is said to be *pointed* if $l(K) = \{0\}$. The cones $K = \{0\}$ and $K = X$ are called *trivial cones*. Typical examples of nontrivial cones which occur in optimization are, when $X = \mathbb{R}^n$, the nonnegative orthant $\mathbb{R}^n_+$ and the *lexicographic cone*

$$\mathbb{R}^n_{lex} := \{0\} \cup \{x \in \mathbb{R}^n : x_1 > 0\} \cup \\ \big\{x \in \mathbb{R}^n : \exists k \in \{2, ..., n\} \text{ such that } x_i = 0 \ \forall i \in \{1, \ldots, k-1\} \text{ and } x_k > 0\big\},$$

while for $X = \mathbb{R}^{n \times n}$ the *cone of symmetric positive semidefinite matrices* $\mathcal{S}^n_+ := \{A \in \mathbb{R}^{n \times n} : A = A^T, \langle x, Ax \rangle \geq 0 \ \forall x \in \mathbb{R}^n\}$. Note that in $\mathbb{R}$ one can find only four cones: $\{0\}$, $\mathbb{R}_+$, $-\mathbb{R}_+$ and $\mathbb{R}$.

The *conical hull* of a set $U \subseteq X$, denoted by $\operatorname{cone}(U)$, is the intersection of all the cones which contain U, being the smallest cone in X that contains U. One can show that $\operatorname{cone}(U) = \cup_{\lambda \geq 0} \lambda U$. When U is convex, then $\operatorname{lin}(U - x) = \operatorname{cone}(U - x)$ and, consequently, $\operatorname{aff}(U) = x + \operatorname{cone}(U - U)$, whenever $x \in U$.

The *convex conical hull* of a set $U \subseteq X$,

$$\operatorname{coneco}(U) := \left\{ \sum_{i=1}^{n} \lambda_i x_i : \ n \in \mathbb{N}, x_i \in U, \lambda_i \geq 0, \ i = 1, \ldots, n \right\},$$

is the intersection of all the convex cones that contain U, being the smallest convex cone having U as a subset. One has $\operatorname{coneco}(U) = \operatorname{cone}(\operatorname{co}(U)) = \operatorname{co}(\operatorname{cone}(U))$. Due to the *Minkowski-Weyl* theorem, a set $U \subseteq \mathbb{R}^n$ is polyhedral if and only if there are two finite sets $V, W \subseteq \mathbb{R}^n$ such that $U = \operatorname{co}(V) + \operatorname{coneco}(W)$.

If K is a nontrivial convex cone, then $U \subseteq K$ is called a *base of the cone* K if each $x \in K \setminus \{0\}$ has an unique representation of the form $x = \lambda u$ for some $\lambda > 0$ and $u \in U$. Each nontrivial convex cone with a base in a nontrivial real vector space is pointed.

If $K \subseteq X$ is a given cone, its *algebraic dual cone* is $K^\# := \{x^\# \in X^\# : \langle x^\#, x \rangle \geq 0 \text{ for all } x \in K\}$. The set $K^\#$ is a convex cone. If C and K are cones in X, one has $(C + K)^\# = C^\# \cap K^\# = (C \cup K)^\#$ and $C^\# + K^\# \subseteq (C \cap K)^\#$. If the two cones satisfy $C \subseteq K$, then $C^\# \supseteq K^\#$.

Given a set $U \subseteq X$ and $x \in U$ we consider the *normal cone to* U *at* x,

$$N(U, x) = \{x^\# \in X^\# : \langle x^\#, y - x \rangle \leq 0 \ \forall y \in U\},$$

which is a convex cone.

Partial orderings. Very important, not only in convex analysis, is to consider certain orderings on the spaces one works with. Let the nonempty set $R \subseteq X \times X$ be a so-called *binary relation* on X. The elements $x, y \in X$ are said in this case to be *in relation* R if $(x, y) \in R$ and we write also xRy. A *binary relation* R is said to be a *partial ordering* on the vector space X if it satisfies the following axioms

(i) *reflexivity*: for all $x \in X$ it holds xRx;

(ii) *transitivity*: for all $x, y, z \in X$ from xRy and yRz follows xRz;
(iii) *compatibility with the linear structure*:
- for all $x, y, z, w \in X$ from xRy and zRw follows $(x+z)R(y+w)$;
- for all $x, y \in X$ and $\lambda \in \mathbb{R}_+$ from xRy follows $(\lambda x)R(\lambda y)$.

In such a situation it is common to use the symbol "$\leq$" and to write $x \leq y$ for xRy. The partial ordering "$\leq$" is called *antisymmetric* if for $x, y \in X$ fulfilling $x \leq y$ and $y \leq x$ there is $x = y$. A real vector space equipped with a partial ordering is called a *partially ordered vector space*.

If there is a partial ordering "$\leq$" on X, then the set $\{x \in X : 0 \leq x\}$ is a convex cone. If the partial ordering "$\leq$" is moreover antisymmetric, this cone is also pointed. Vice versa, having a convex cone $K \subseteq X$, it induces on X a partial ordering relation "$\leqq_K$" defined as follows

$$\leqq_K := \big\{(x, y) \in X \times X : y - x \in K\big\}.$$

If K is pointed, then "$\leqq_K$" is antisymmetric. To write $x \leqq_K y$, also the notation $y \geqq_K x$ is used, while $x \not\leqq_K y$ means $y - x \notin K$. We denote also $x \leq_K y$ if $x \leqq_K y$ and $x \neq y$, while $x \not\leq_K y$ is used when $x \leq_K y$ is not fulfilled. A convex cone which induces a partial ordering on X is called *ordering cone*. For the natural partial ordering on $\mathbb{R}^n$, which is introduced by $\mathbb{R}^n_+$, we use "$\leqq$" instead of "$\leqq_{\mathbb{R}^n_+}$" and also "$\leq$" for "$\leq_{\mathbb{R}^n_+}$".

By $\overline{\mathbb{R}}$ we denote the *extended real space* which consists of $\mathbb{R} \cup \{\pm\infty\}$. The operations on $\overline{\mathbb{R}}$ are the usual ones on $\mathbb{R}$ to which we add the following natural ones: $\lambda + (+\infty) = (+\infty) + \lambda := +\infty\ \forall \lambda \in (-\infty, +\infty]$, $\lambda + (-\infty) = (-\infty) + \lambda := -\infty\ \forall \lambda \in [-\infty, +\infty)$, $\lambda \cdot (+\infty) := +\infty\ \forall \lambda \in (0, +\infty]$, $\lambda \cdot (+\infty) := -\infty\ \forall \lambda \in [-\infty, 0)$, $\lambda \cdot (-\infty) := -\infty\ \forall \lambda \in (0, +\infty]$ and $\lambda \cdot (-\infty) := +\infty\ \forall \lambda \in [-\infty, 0)$. We also assume by convention that

$$(+\infty) + (-\infty) = (-\infty) + (+\infty) := +\infty, 0(+\infty) := +\infty \text{ and } 0(-\infty) := 0.$$

In analogy to the extended real space we attach to X a *greatest* and a *smallest* element with respect to "$\leqq_K$", denoted by $+\infty_K$ and $-\infty_K$, respectively, which do not belong to X and let $\overline{X} := X \cup \{\pm\infty_K\}$. Then for $x \in \overline{X}$ it holds $-\infty_K \leqq_K x \leqq_K +\infty_K$. Similarly, we assume that $-\infty_K \leq_K x \leq_K +\infty_K$ for any $x \in X$. On $\overline{X}$ we consider the following operations, in analogy to the ones stated above for the extended real space: $x + (+\infty_K) = (+\infty_K) + x := +\infty_K$ $\forall x \in X \cup \{+\infty_K\}$, $x + (-\infty_K) = (-\infty_K) + x := -\infty_K\ \forall x \in X \cup \{-\infty_K\}$, $\lambda \cdot (+\infty_K) := +\infty_K\ \forall \lambda \in (0, +\infty]$, $\lambda \cdot (+\infty_K) := -\infty_K\ \forall \lambda \in [-\infty, 0)$, $\lambda \cdot (-\infty_K) := -\infty_K\ \forall \lambda \in (0, +\infty]$ and $\lambda \cdot (-\infty_K) := +\infty_K\ \forall \lambda \in [-\infty, 0)$. We consider also the following conventions

$$\begin{gathered}(+\infty_K) + (-\infty_K) = (-\infty_K) + (+\infty_K) := +\infty_K,\\ 0(+\infty_K) := +\infty_K \text{ and } 0(-\infty_K) := 0.\end{gathered} \tag{2.1}$$

Moreover, if $x^\# \in K^\#$ we let $\langle x^\#, +\infty_K\rangle := +\infty$.

Algebraic interiority notions. Even without assuming a topological structure on X, different algebraic interiority notions can be considered for its subsets, as follows. The *algebraic interior*, also called *core*, of a set $U \subseteq X$ is

$$\operatorname{core}(U) := \big\{x \in X : \text{ for every } y \in X \exists \delta > 0 \text{ such that } x + \lambda y \in U \forall \lambda \in [0, \delta]\big\}.$$

It is clear that $\operatorname{core}(U) \subseteq U$. The algebraic interior with respect to the affine hull of U is called the *intrinsic core* of U, being the set

$$\operatorname{icr}(U) := \big\{x \in X : \text{ for every } y \in \operatorname{aff}(U) \exists \delta > 0 \text{ such that } x + \lambda y \in U \forall \lambda \in [0, \delta]\big\}.$$

There is $\operatorname{core}(U) \subseteq \operatorname{icr}(U)$. If $x \in U$ and U is convex, then $x \in \operatorname{core}(U)$ if and only if $\operatorname{cone}(U - x) = X$ and, on the other hand, $x \in \operatorname{icr}(U)$ if and only if $\operatorname{cone}(U - x)$ is a linear subspace, or, equivalently, $\operatorname{cone}(U - x) = \operatorname{cone}(U - U)$.

Taking two subsets U and V of X we have $U + \operatorname{core}(V) \subseteq \operatorname{core}(U + V)$, with equality if $V = \operatorname{core}(V)$. The equality holds also in case U and V are convex and $\operatorname{core}(V) \neq \emptyset$, as proved in [176]. Note also that a set $U \subseteq X$ is absorbing if and only if $0 \in \operatorname{core}(U)$. If K is a cone in X with $\operatorname{core}(K) \neq \emptyset$, then $K - K = X$ and, consequently, $K^{\#}$ is pointed. When K is a convex cone then $\operatorname{core}(K) \cup \{0\}$ is a convex cone, too, and $\operatorname{core}(K) = \operatorname{core}(K) + K$. If K is a convex cone with nonempty algebraic interior, then one has $\operatorname{core}(K) = \big\{x \in X : \langle x^{\#}, x\rangle > 0 \ \forall x^{\#} \in K^{\#} \backslash \{0\}\big\}$.

2.1.2 Topological properties of convex sets

Further we consider X being a real *topological vector space*, i.e. a real vector space endowed with a topology $\mathcal{T}$ which renders continuous the following functions

$$(x, y) \mapsto x + y, \ \ x, y \in X \ \text{ and } (\lambda, x) \mapsto \lambda x, \ \ x \in X, \lambda \in \mathbb{R}.$$

Throughout the book, if we speak about (topological) vector spaces, we always mean real nontrivial (topological) vector spaces, this means not equal to $\{0\}$. Moreover we agree to omit further the word "real" in such contexts. A topological space for which any two different elements have disjoint neighborhoods is said to be *Hausdorff*. A topological vector space X is said to be *metrizable* if it can be endowed with a metric which is compatible with its topology. Every metrizable vector space is Hausdorff.

For a set $U \subseteq X$ we denote by $\operatorname{int}(U)$ the *interior* of U and by $\operatorname{cl}(U)$ its *closure*. Then $\operatorname{bd}(U) = \operatorname{cl}(U) \backslash \operatorname{int}(U)$ is called the *boundary* of U.

If Y is a topological vector space and $T : X \to Y$ is a linear mapping, then there is $T(\operatorname{cl}(U)) \subseteq \operatorname{cl}(T(U))$ for every $U \subseteq X$. If U is a convex subset of X, $x \in \operatorname{int}(U)$ and $y \in \operatorname{cl}(U)$, then $\{\lambda x + (1 - \lambda)y : \lambda \in (0, 1]\} \subseteq \operatorname{int}(U)$. For $U \subseteq X$ there is $\operatorname{int}(U) \subseteq \operatorname{core}(U)$. If U is convex and one of the following conditions is fulfilled: $\operatorname{int}(U) \neq \emptyset$; X is a Banach space and U is closed; X is finite dimensional, then $\operatorname{int}(U) = \operatorname{core}(U)$. If U is convex and $\operatorname{int}(U) \neq \emptyset$, then it holds $\operatorname{int}(U) = \operatorname{int}(\operatorname{cl}(U))$ and $\operatorname{cl}(\operatorname{int}(U)) = \operatorname{cl}(U)$. The interior and the closure of a convex set in a topological vector space are convex, too. If U is a subset of X, then the intersection of all closed convex sets containing U is the *closed convex hull* of U, denoted by $\overline{\operatorname{co}}(U)$, and it is the smallest closed convex set containing U. It is also the closure of the convex hull of U.

When $K \subseteq X$ is a convex cone with $\operatorname{core}(K) \neq \emptyset$ we denote $\widehat{K} := \operatorname{core}(K) \cup \{0\}$ and, for $x, y \in X$ which satisfy $y - x \in \operatorname{core}(K)$ we write $x <_K y$. When $\operatorname{int}(K) \neq \emptyset$, $x <_K y$ means $y - x \in \operatorname{int}(K)$. Concerning the elements $+\infty_K$ and $-\infty_K$ introduced in the previous subsection we assume that for all $x \in X$ one has $-\infty_K <_K x <_K +\infty_K$.

In $\mathbb{R}^n$ we work with the *Euclidean topology* induced by the *Euclidean norm*. The *open ball* centered in $x \in \mathbb{R}^n$ and with radius $\varepsilon > 0$ is denoted by $B(x, \varepsilon)$, while the *closed ball* centered in $x \in \mathbb{R}^n$ and with radius $\varepsilon > 0$ is denoted by $\overline{B}(x, \varepsilon)$.

Dual spaces. The set of all linear continuous mappings defined on X and taking values in the topological vector space Y is denoted by $\mathcal{L}(X, Y)$.

The topological vector space $\mathcal{L}(X, \mathbb{R})$ is said to be the *topological dual space* of X, being denoted by X^*. Further, we refer with "dual" to topological duals, not to algebraical ones, unless otherwise specified. Analogously to vector spaces, by $\langle x^*, x\rangle$ we denote the value taken at $x \in X$ by the linear continuous functional $x^* \in X^*$. The hyperplane $\mathcal{H} := \{x \in X : \langle x^\#, x\rangle = \lambda\}$ with $x^\# \in X^\#$ and $\lambda \in \mathbb{R}$ is closed if and only if $x^\#$ is continuous.

For a mapping $A \in \mathcal{L}(X, Y)$ we consider its *adjoint mapping* $A^* \in \mathcal{L}(Y^*, X^*)$ defined by $\langle A^*y^*, x\rangle := \langle y^*, Ax\rangle$ for all $x \in X$ and $y^* \in Y^*$. When $X = \mathbb{R}^n$ and $Y = \mathbb{R}^m$, A can be identified with an $m \times n$ matrix and A^* coincides with A^T.

For every $x^* \in X^*$ let the seminorm $p_{x^*} : X \to \mathbb{R}$ defined by $p_{x^*}(x) := |\langle x^*, x\rangle|$. The coarsest topology on X which makes all the seminorms p_{x^*}, for $x^* \in X^*$, continuous is called the *weak topology* on X induced by X^*, being denoted $w(X, X^*)$. Every *weakly closed* set in X, i.e. closed in the weak topology, is closed also in the original topology on X, while the reverse assertion does not always hold.

Considering for all $x \in X$ the seminorms $p_x : X^* \to \mathbb{R}$, $p_x(x^*) = |\langle x^*, x\rangle|$, one defines analogously a topology on X^*, called the *weak*[*] *topology*, denoted $w(X^*, X)$. When one works with X^* endowed with the topology $w(X^*, X)$, the *bidual space* X^{**} of X, defined as the topological dual of X^*, can be identified with X.

Locally convex spaces. By a *local base* $\mathcal{B}$ of the topological vector space X endowed with the topology $\mathcal{T}$ we understand a collection of neighborhoods of zero from $\mathcal{T}$ such that every neighborhood of zero contains an element of $\mathcal{B}$. Then a set belongs to $\mathcal{T}$ if and only if it can be written as a union of translates of members of $\mathcal{B}$. A topological vector space is called *locally convex* if it has a local base whose members are convex sets. In a Hausdorff locally convex space the weakly closed convex sets are identical with the closed convex sets.

A locally convex space is called *Fréchet* if it is complete and metrizable by a metric which is invariant to translations.

If X is a Hausdorff locally convex space, to a nonempty subset U of it one can introduce the *Bouligand tangent cone* at $x \in \text{cl}(U)$, which is

$$T(U,x) := \Big\{ y \in X : \exists (x_l)_{l \geq 1} \in U \text{ and } (\lambda_l)_{l \geq 1} > 0 \text{ such that } \lim_{l \to +\infty} x_l = x \text{ and } \lim_{l \to +\infty} \lambda_l (x_l - x) = y \Big\}.$$

Whenever $U \neq \emptyset$ and $x \in \text{cl}(U)$, $T(U,x)$ is a cone and $T(U,x) \subseteq \text{cl}(\text{cone}(U - x))$. For a convex set $U \subseteq X$ there is $\text{cone}(U - x) \subseteq T(U,x)$, which yields in this case that $\text{cl}(T(U,x)) = \text{cl}(\text{cone}(U - x))$. If X is metrizable, then $T(U,x)$ is closed and, thus, if U is convex one has $T(U,x) = \text{cl}(\text{cone}(U - x))$ for all $x \in \text{cl}(U)$.

Topological dual cones. Analogously to the algebraic dual cone used when working in vector spaces, one can consider a dual cone in topological vector spaces, too. When K is a cone in X, its *topological dual cone*, further called simply *dual cone*, is

$$K^* := \big\{ x^* \in X^* : \langle x^*, x \rangle \geq 0 \text{ for all } x \in K \big\}.$$

The cone K^* is always convex and weak* closed. If K is a convex cone with nonempty interior, then there is $\text{int}(K) = \big\{ x \in X : \langle x^*, x \rangle > 0 \ \forall x^* \in K^* \backslash \{0\} \big\}$. If C and K are convex closed cones in X, then $(C \cap K)^* = \text{cl}_{w(X^*,X)}(C^* + K^*)$ and the closure can be removed, for instance, when $C \cap \text{int}(K) \neq \emptyset$.

The *bidual cone* of a cone $K \subseteq X$ is

$$K^{**} := \{ x \in X : \langle x^*, x \rangle \geq 0 \text{ for all } x^* \in K^* \}.$$

Note that $K^{**} = \overline{\text{co}}(K)$. When X^* is endowed with the weak* topology then K^{**} is nothing but the dual cone of K^*.

Topological interiority notions. Let, unless otherwise specified, X be a Hausdorff locally convex space and X^* its topological dual space endowed with the weak* topology. Besides the already introduced interiority notions, which are defined only by algebraical means, we deal in this book also with topological notions of generalized interiors for a set.

The *quasi relative interior* of $U \subseteq X$ is

$$\text{qri}(U) := \big\{ x \in U : \text{cl}(\text{cone}(U - x)) \text{ is a linear subspace} \big\}.$$

If U is convex, then $x \in \text{qri}(U)$ if and only if $x \in U$ and $N(U,x)$ is a linear subspace of X^* (cf. [21]). The *quasi interior* of a set $U \subseteq X$ is the set

$$\text{qi}(U) := \big\{ x \in U : \text{cl}(\text{cone}(U - x)) = X \big\}.$$

Note that $\text{qi}(U)$ is a subset of $\text{qri}(U)$. When U is convex one has $x \in \text{qi}(U)$ if and only if $x \in U$ and $N(U,x) = \{0\}$ (cf. [26, 27]) and also that if $\text{qi}(U) \neq \emptyset$ then $\text{qi}(U) = \text{qri}(U)$. The next result provides a characterization for the quasi interior of the dual cone of a convex closed cone.

Proposition 2.1.1. *If $K \subseteq X$ is a convex closed cone, then*

$$\operatorname{qi}(K^*) = \big\{x^* \in K^* : \langle x^*, x\rangle > 0 \ \textit{ for all } x \in K\backslash\{0\}\big\}. \tag{2.2}$$

Proof. Assume first that there is some $x^* \in \operatorname{qi}(K^*)$ not belonging to set in the right-hand side of (2.2). Then there is some $x \in K\backslash\{0\}$ such that $\langle x^*, x\rangle = 0$. As $\langle y^*, x\rangle \geq 0$ for all $y^* \in K^*$, we obtain $\langle y^* - x^*, -x\rangle \leq 0$ for all $y^* \in K^*$, i.e. $-x \in N(K^*, x^*) = \{0\}$. As this cannot take place because $x \neq 0$, it follows that $\operatorname{qi}(K^*) \subseteq \{x^* \in K^* : \langle x^*, x\rangle > 0 \, \forall x \in K\backslash\{0\}\}$. Assume now the existence of some $x^* \in K^* \backslash \operatorname{qi}(K^*)$ which fulfills $\langle x^*, x\rangle > 0$ whenever $x \in K\backslash\{0\}$. Then there is some $y \in X\backslash\{0\}$ such that $\langle y^* - x^*, y\rangle \leq 0$ for all $y^* \in K^*$. This yields $\langle y^*, y\rangle \leq \langle x^*, y\rangle$ for all $y^* \in K^*$. Taking into consideration that K^* is a cone, this implies $\langle y^*, y\rangle \leq 0$ whenever $y^* \in K^*$, i.e. $y \in -K^{**}$. As K is convex and closed we get $y \in -K\backslash\{0\}$, thus $\langle x^*, y\rangle < 0$, which if false. Consequently, (2.2) holds. □

Whenever $K \subseteq X$ is a convex cone, even if not necessarily closed, the above proposition motivates the use of the name *quasi interior of the dual cone* of K for the set

$$K^{*0} := \big\{x^* \in K^* : \langle x^*, x\rangle > 0 \ \text{ for all } x \in K\backslash\{0\}\big\}.$$

In case X is a separable normed space and K is a pointed convex closed cone, the Krein-Rutman theorem guarantees the nonemptiness of K^{*0} (see [104, Theorem 3.38]). Considering $X = l^2$ and $K = l^2_+$, it can be noted that $(l^2_+)^{*0}$ is nonempty, different to $\operatorname{int}((l^2_+)^*) = \operatorname{int}(l^2_+)$ which is an empty set. If $K^{*0} \neq \emptyset$ then K is pointed. If K is closed and $\operatorname{int}_{w(X^*,X)}(K^*) \neq \emptyset$, then $\operatorname{int}_{w(X^*,X)}(K^*) = K^{*0}$.

The *strong quasi relative interior* of a set $U \subseteq X$ is

$$\operatorname{sqri}(U) := \big\{x \in U : \operatorname{cone}(U - x) \text{ is a closed linear subspace}\big\}.$$

It is known that $\operatorname{core}(U) \subseteq \operatorname{sqri}(U) \subseteq \operatorname{icr}(U)$. If U is convex, then $u \in \operatorname{sqri}(U)$ if and only if $u \in \operatorname{icr}(U)$ and $\operatorname{aff}(U - u)$ is a closed linear subspace. Assuming additionally that $X = \mathbb{R}^n$ and $U \subseteq \mathbb{R}^n$, there is $\operatorname{qi}(U) = \operatorname{int}(U)$ and $\operatorname{icr}(U) = \operatorname{sqri}(U) = \operatorname{qri}(U) = \operatorname{ri}(U)$, where

$$\operatorname{ri}(U) := \big\{x \in \operatorname{aff}(U) : \exists \varepsilon > 0 \text{ such that } B(x, \varepsilon) \cap \operatorname{aff}(U) \subseteq U\big\}$$

is the *relative interior* of the set U.

Separation theorems. Separation statements are very important in convex analysis and optimization, being crucial in the proofs of some of the basic results. In the following we present the ones which we need later in this book. We begin with a classical result in topological vector spaces followed by its version for real vector spaces and a consequence.

Theorem 2.1.2. *(Eidelheit) Let U and V be nonempty convex subsets of the topological vector space X with $\operatorname{int}(U) \neq \emptyset$. Then $\operatorname{int}(U) \cap V = \emptyset$ if and only if there are some $x^* \in X^* \backslash \{0\}$ and $\lambda \in \mathbb{R}$ such that*

$$\sup_{x \in U} \langle x^*, x \rangle \leq \lambda \leq \inf_{x \in V} \langle x^*, x \rangle$$

and $\langle x^*, x \rangle < \lambda$ for all $x \in \operatorname{int}(U)$.

Theorem 2.1.3. *Let U and V be nonempty convex subsets of a vector space X with $\operatorname{core}(U) \neq \emptyset$. Then $\operatorname{core}(U) \cap V = \emptyset$ if and only if there are some $x^\# \in X^\# \backslash \{0\}$ and $\lambda \in \mathbb{R}$ such that*

$$\sup_{x \in U} \langle x^\#, x \rangle \leq \lambda \leq \inf_{x \in V} \langle x^\#, x \rangle$$

and $\langle x^\#, x \rangle < \lambda$ for all $x \in \operatorname{core}(U)$.

Corollary 2.1.4. *Let U and V be nonempty convex subsets of the topological vector space X such that $\operatorname{int}(U - V) \neq \emptyset$. Then $0 \notin \operatorname{int}(U - V)$ if and only if there exists an $x^* \in X^* \backslash \{0\}$ such that*

$$\sup_{x \in U} \langle x^*, x \rangle \leq \inf_{x \in V} \langle x^*, x \rangle.$$

When working in locally convex spaces one has the following separation result.

Theorem 2.1.5. *(Tuckey) Let U and V be nonempty convex subsets of the locally convex space X, one compact and the other closed. Then $U \cap V = \emptyset$ if and only if there exists an $x^* \in X^* \backslash \{0\}$ such that*

$$\sup_{x \in U} \langle x^*, x \rangle < \inf_{x \in V} \langle x^*, x \rangle.$$

Corollary 2.1.6. *Let U and V be nonempty convex subsets of the locally convex space X. Then $0 \notin \operatorname{cl}(U - V)$ if and only if there exists an $x^* \in X^* \backslash \{0\}$ such that*

$$\sup_{x \in U} \langle x^*, x \rangle < \inf_{x \in V} \langle x^*, x \rangle.$$

In finite dimensional spaces, i.e. when $X = \mathbb{R}^n$, we also have the following separation statement involving relative interiors.

Theorem 2.1.7. *Let U and V be nonempty convex sets in $\mathbb{R}^n$. Then $\operatorname{ri}(U) \cap \operatorname{ri}(V) = \emptyset$ if and only if there exists an $x^* \in X^* \backslash \{0\}$ such that*

$$\sup_{x \in U} \langle x^*, x \rangle \leq \inf_{x \in V} \langle x^*, x \rangle$$

and

$$\inf_{x \in U} \langle x^*, x \rangle < \sup_{x \in V} \langle x^*, x \rangle.$$

2.2 Convex functions

In this section X and Y are considered, unless otherwise specified, Hausdorff locally convex spaces and X^* and Y^* their topological dual spaces, respectively. We list some well-known basic results on convex functions, but, as in the previous section, without proofs concerning the most of them. They can be found in different textbooks and monographs devoted to convex analysis, functional analysis, optimization theory, etc. (cf. [67, 90, 104, 157, 207]).

2.2.1 Algebraic properties of convex functions

We begin with some basic definitions and results.

Definition 2.2.1. *A function* $f : X \to \overline{\mathbb{R}}$ *is called convex if for all* $x, y \in X$ *and all* $\lambda \in [0,1]$ *one has*

$$f(\lambda x + (1-\lambda)y) \leq \lambda f(x) + (1-\lambda) f(y). \tag{2.3}$$

A function $f : X \to \overline{\mathbb{R}}$ *is said to be concave if* $(-f)$ *is convex.*

Remark 2.2.1. Given a convex set $U \subseteq X$ we say that a function $f : U \to \mathbb{R}$ is *convex on* U if (2.3) holds for all $x, y \in U$ and every $\lambda \in [0,1]$. The function f is said to be *concave on* U if $(-f)$ is convex on U. The *extension* of the function f to the whole space is the function

$$\tilde{f} : X \to \overline{\mathbb{R}},\ \tilde{f}(x) := \begin{cases} f(x), & \text{if } x \in U, \\ +\infty, & \text{otherwise.} \end{cases}$$

It is a simple verification to prove that $\tilde{f}$ is convex if and only if U is a convex set and f is convex on U. Thus the theory built for functions defined on the whole space X and having values in $\overline{\mathbb{R}}$ can be employed for real-valued functions defined on subsets of X, too.

In case $X = \mathbb{R}$ the following convexity criterion can be useful.

Remark 2.2.2. Consider $(a,b) \subseteq \mathbb{R}$ and the twice differentiable function $f : (a,b) \to \mathbb{R}$. Then f is convex (concave) on (a,b) if and only if $f''(x) \geq (\leq) 0$ for all $x \in (a,b)$.

Definition 2.2.2. *A function* $f : X \to \overline{\mathbb{R}}$ *is called strictly convex if for all* $x, y \in X$ *with* $x \neq y$ *and all* $\lambda \in (0,1)$ *one has* (2.3) *fulfilled as a strict inequality. A function* $f : X \to \overline{\mathbb{R}}$ *is called strictly concave if* $(-f)$ *is strictly convex.*

Example 2.2.1. (*a*) The *indicator function*

$$\delta_U : X \to \overline{\mathbb{R}},\ \delta_U(x) := \begin{cases} 0, & \text{if } x \in U, \\ +\infty, & \text{otherwise,} \end{cases}$$

of a set $U \subseteq X$ is convex if and only if U is convex.

(b) Let A be a $n \times n$ positive semidefinite matrix with real entries. Then the function $f : \mathbb{R}^n \to \mathbb{R}$, $f(x) = x^T Ax$, is convex. If A is positive definite, then f is strictly convex.

(c) If $\|\cdot\|$ denotes a norm on a vector space X, then $x \mapsto \|x\|$ is a convex function.

One can easily prove that a function $f : X \to \overline{\mathbb{R}}$ is convex if and only if for any $n \in \mathbb{N}$, $x_i \in X$ and $\lambda_i \in \mathbb{R}_+$, $i = 1, \dots, n$, such that $\sum_{i=1}^n \lambda_i = 1$, *Jensen's inequality* is satisfied, namely

$$f\left(\sum_{i=1}^n \lambda_i x_i\right) \le \sum_{i=1}^n \lambda_i f(x_i).$$

For a function $f : X \to \overline{\mathbb{R}}$ we consider the *(effective) domain* $\operatorname{dom} f := \{x \in X : f(x) < +\infty\}$ and the *epigraph* $\operatorname{epi} f := \{(x, r) \in X \times \mathbb{R} : f(x) \le r\}$. The *strict epigraph* of f is $\operatorname{epi}_s f := \{(x, r) \in X \times \mathbb{R} : f(x) < r\}$. A function $f : X \to \overline{\mathbb{R}}$ is called *proper* if $f(x) > -\infty$ for all $x \in X$ and $\operatorname{dom} f \neq \emptyset$. Otherwise f is said to be *improper*.

A characterization of the convexity of a function through the convexity of its epigraph is given in the next result.

Proposition 2.2.1. *Let the function $f : X \to \overline{\mathbb{R}}$. The following assertions are equivalent:*

(i) f is convex;
(ii) $\operatorname{epi} f$ is convex;
(iii) $\operatorname{epi}_s f$ is convex.

Remark 2.2.3. For $f : X \to \overline{\mathbb{R}}$ we have $\operatorname{Pr}_X(\operatorname{epi} f) = \operatorname{dom} f$. Thus, if f is convex, then its domain is a convex set.

If $f : X \to \overline{\mathbb{R}}$ and $\lambda \in \mathbb{R}$, we call $\{x \in X : f(x) \le \lambda\}$ the *level set* of f at λ and $\{x \in X : f(x) < \lambda\}$ is said to be the *strict level set* of f at λ. If f is convex, then the level sets and the strict level sets of f at λ are convex, for all $\lambda \in \mathbb{R}$. The opposite assertion is not true in general.

Definition 2.2.3. *A function $f : X \to \overline{\mathbb{R}}$ is called*

(a) subadditive if for all $x, y \in X$ one has $f(x + y) \le f(x) + f(y)$;
(b) positively homogenous if $f(0) = 0$ and for all $x \in X$ and all $\lambda > 0$ one has $f(\lambda x) = \lambda f(x)$;
(c) sublinear if it is subadditive and positively homogenous.

Example 2.2.2. Given a nonempty set $U \subseteq X$, its *support function* $\sigma_U : X^* \to \overline{\mathbb{R}}$ defined by $\sigma_U(x^*) := \sup\{\langle x^*, x\rangle : x \in U\}$ is sublinear.

Notice that a convex function $f : X \to \overline{\mathbb{R}}$ is sublinear if and only if it is also positively homogenous.

Let be given the convex functions $f, g : X \to \overline{\mathbb{R}}$. Then $f + g$ and λf, $\lambda \geq 0$, are convex. One should notice that, due to the way the operations on the extended real space are defined, there is $0f = \delta_{\operatorname{dom} f}$.

Proposition 2.2.2. *Given a family of functions $f_i : X \to \overline{\mathbb{R}}$, $i \in I$, where I is an arbitrary index set, one has* $\operatorname{epi}(\sup_{i \in I} f_i) = \cap_{i \in I} \operatorname{epi} f_i$. *Consequently, the pointwise supremum $f : X \to \overline{\mathbb{R}}$ of a family of convex functions $f_i : X \to \overline{\mathbb{R}}$, $i \in I$, defined by $f(x) = \sup_{i \in I} f_i(x)$ is a convex function, too.*

Consider the Hausdorff locally convex spaces X_i, $i = 1, \ldots, m$, and take $X = \prod_{i=1}^m X_i$. Given the convex functions $f_i : X_i \to \overline{\mathbb{R}}$, $i = 1, \ldots, m$, the function $f : X \to \overline{\mathbb{R}}$ defined by $f(x^1, \ldots, x^m) = \sum_{i=1}^m f_i(x^i)$ is convex, too. Obviously, $\operatorname{dom} f = \prod_{i=1}^m \operatorname{dom} f_i$.

Consider $U \subseteq X \times \mathbb{R}$ a given set. To U we associate the so-called *lower bound function* $\phi_U : X \to \overline{\mathbb{R}}$ defined as

$$\phi_U(x) := \inf\{t \in \mathbb{R} : (x, t) \in U\}.$$

For an arbitrary function $f : X \to \overline{\mathbb{R}}$ it holds $f(x) = \phi_{\operatorname{epi} f}(x)$ for all $x \in X$. If U is a convex set, then ϕ_U is a convex function. By means of the lower bound function we introduce in the following the *convex hull* of a function.

Definition 2.2.4. *Consider a function $f : X \to \overline{\mathbb{R}}$. The function* $\operatorname{co} f : X \to \overline{\mathbb{R}}$, *defined by*

$$\operatorname{co} f(x) := \phi_{\operatorname{co}(\operatorname{epi} f)}(x) = \inf\{t \in \mathbb{R} : (x, t) \in \operatorname{co}(\operatorname{epi} f)\},$$

is called the convex hull of f.

It is clear from the construction that the convex hull of a function $f : X \to \overline{\mathbb{R}}$ is convex and it is the greatest convex function less than or equal to f. Consequently,

$$\operatorname{co} f = \sup\{g : X \to \overline{\mathbb{R}} : g(x) \leq f(x) \text{ for all } x \in X \text{ and } g \text{ is convex}\}.$$

Thus, f is convex if and only if $f = \operatorname{co} f$. Regarding the convex hull of f we have also the following result.

Proposition 2.2.3. *Let the function $f : X \to \overline{\mathbb{R}}$ be given. Then the convex hull of its domain coincides with the domain of its convex hull, namely* $\operatorname{co}(\operatorname{dom} f) = \operatorname{dom}(\operatorname{co} f)$. *Moreover, there is* $\operatorname{epi}_s(\operatorname{co} f) \subseteq \operatorname{co}(\operatorname{epi} f) = \operatorname{epi}(\operatorname{co} f)$.

Next we consider some notions which extend the classical monotonicity to functions defined on partially ordered spaces.

Definition 2.2.5. *Let be the vector space V partially ordered by the convex cone K, a nonempty set $W \subseteq V$ and $g : V \to \overline{\mathbb{R}}$ a given function.*

(a) *If $g(x) \leq g(y)$ for all $x, y \in W$ such that $x \leqq_K y$, the function g is called K-increasing on W.*
(b) *If $g(x) < g(y)$ for all $x, y \in W$ such that $x \leq_K y$, the function g is called strongly K-increasing on W.*
(c) *If g is K-increasing on W, $\operatorname{core}(K) \neq \emptyset$ and for all $x, y \in W$ fulfilling $x <_K y$ follows $g(x) < g(y)$, the function g is called strictly K-increasing on W.*
(d) *When $W = V$ we call these classes of functions K-increasing, strongly K-increasing and strictly K-increasing, respectively.*

Remark 2.2.4. When $X = \mathbb{R}$, the $\mathbb{R}_+$-increasing functions are actually the *increasing functions*, while the strongly and the strictly $\mathbb{R}_+$-increasing functions are actually the *strictly increasing functions*.

Remark 2.2.5. For a cone $K \subseteq V$ with $\operatorname{core}(K) \neq \emptyset$, we defined $\widehat{K} := \operatorname{core}(K) \cup \{0\}$. Then the definition of the strictly K-increasing functions on a set $W \subseteq V$ coincide with the strongly $\widehat{K}$-increasing functions on W. When $\operatorname{int}(K) \neq \emptyset$ one has $\operatorname{int}(K) = \operatorname{core}(K)$, thus the core of the cone K can be replaced in Definition 2.2.5 by the interior of K.

Example 2.2.3. Consider a vector space V and a linear functional $v^\# \in V^\#$. If $v^\# \in K^\#$, then the definition of the algebraic dual cone secures that for all $v_1, v_2 \in V$ such that $v_1 \leqq_K v_2$ we have $\langle v^\#, v_2 - v_1 \rangle \geq 0$. Therefore $\langle v^\#, v_1 \rangle \leq \langle v^\#, v_2 \rangle$ and this means that the elements of $K^\#$ are actually K-increasing linear functions on the vector space V.

If $v^\# \in K^{\#0} := \{x^\# \in K^\# : \langle x^\#, x \rangle > 0 \text{ for all } x \in K \setminus \{0\}\}$, which can be seen as the analogous of K^{*0} in vector spaces, then for all $v_1, v_2 \in V$ such that $v_1 \leq_K v_2$ it holds $\langle v^\#, v_2 - v_1 \rangle > 0$. According to the previous definition this means that the elements of $K^{\#0}$ are strongly K-increasing linear functions on V.

On the other hand, if $\operatorname{core}(K) \neq \emptyset$, then, according to the representation $\operatorname{core}(K) = \{v \in V : \langle v^\#, v \rangle > 0 \ \forall v^\# \in K^\# \setminus \{0\}\}$, every $v^\# \in K^\# \setminus \{0\}$ is strictly K-increasing on V.

There are notions given for functions with extended real values that can be formulated also for functions mapping from X into vector spaces. Let V be Hausdorff locally convex space partially ordered by the convex cone K and $\overline{V} = V \cup \{\pm\infty_K\}$.

The *domain* of a vector function $h : X \to \overline{V}$ is the set $\operatorname{dom} h := \{x \in X : h(x) \neq +\infty_K\}$. When $h(x) \neq -\infty_K$ for all $x \in X$ and $\operatorname{dom} h \neq \emptyset$ we call h *proper*. The *K-epigraph* of a vector function $h : X \to \overline{V}$ is the set $\operatorname{epi}_K h := \{(x, v) \in X \times V : h(x) \leqq_K v\}$.

Definition 2.2.6. *A vector function $h : X \to \overline{V}$ is said to be K-convex if $\operatorname{epi}_K h$ is a convex set.*

One can easily prove that a function $h : X \to V \cup \{+\infty_K\}$ is K-convex if and only if

$$h(\lambda x + (1-\lambda)y) \leqq_K \lambda h(x) + (1-\lambda)h(y) \ \forall x, y \in X \ \forall \lambda \in [0,1].$$

For a convex set $U \subseteq X$ we say that the function $h : U \to V$ is *K-convex on U* if $h(\lambda x+(1-\lambda)y) \leqq_K \lambda h(x)+(1-\lambda)h(y)$ for all $x, y \in U$ and all $\lambda \in [0,1]$. Considering the function

$$\tilde{h} : X \to \overline{V},\ \tilde{h}(x) := \begin{cases} h(x), & \text{if } x \in U, \\ +\infty_K, & \text{otherwise,} \end{cases}$$

note that $\tilde{h}$ is K-convex if and only if U is convex and h is K-convex on U.

Having a set $U \subseteq X$, its *vector indicator function* is

$$\delta_U^V : X \to \overline{V},\ \delta_U^V(x) := \begin{cases} 0, & \text{if } x \in U, \\ +\infty_K, & \text{otherwise.} \end{cases}$$

Then δ_U^V is K-convex if and only if U is convex.

Remark 2.2.6. Let $h : X \to \overline{V}$ be a given vector function. For $v^* \in K^*$ we shall use the notation $(v^*h) : X \to \overline{\mathbb{R}}$ for the function defined by $(v^*h)(x) := \langle v^*, h(x)\rangle$ and one can easily notice that $\operatorname{dom}(v^*h) = \operatorname{dom} h$.

The proof of the following result is straightforward.

Theorem 2.2.4. *Let be the convex and K-increasing function $f : V \cup \{+\infty_K\} \to \overline{\mathbb{R}}$ defined with the convention $f(+\infty_K) = +\infty$ and consider the proper K-convex function $h : X \to \overline{V}$. Then the function $f \circ h : X \to \overline{\mathbb{R}}$ is convex.*

Corollary 2.2.5. *Let be the convex function $f : V \to \overline{\mathbb{R}}$ and the affine mapping $T : X \to V$. Then the function $f \circ T : X \to \overline{\mathbb{R}}$ is convex.*

Proof. For $K = \{0\}$, the mapping T is K-convex and the result follows by Theorem 2.2.4. □

Another important function attached to a given function $\Phi : X \times Y \to \overline{\mathbb{R}}$ is the so-called *infimal value function* to it, defined as follows

$$h : Y \to \overline{\mathbb{R}},\ h(y) := \inf\{\Phi(x, y) : x \in X\}.$$

Theorem 2.2.6. *Given a convex function $\Phi : X \times Y \to \overline{\mathbb{R}}$, its infimal value function is convex, too.*

Proof. One can prove that $\operatorname{epi}_s h = \operatorname{Pr}_{Y \times \mathbb{R}}(\operatorname{epi}_s \Phi)$. As the projection preserves the convexity and $\operatorname{epi}_s \Phi$ is convex, it follows that $\operatorname{epi}_s h$ is convex, too. By Proposition 2.2.1, h is convex. □

Remark 2.2.7. It can also be proven that

$$\mathrm{Pr}_{Y\times\mathbb{R}}(\mathrm{epi}\,\Phi) \subseteq \mathrm{epi}\,h \subseteq \mathrm{cl}(\mathrm{Pr}_{Y\times\mathbb{R}}(\mathrm{epi}\,\Phi)).$$

As a special case of Theorem 2.2.6 we obtain the following result.

Theorem 2.2.7. *Let be the convex function $f : X \to \overline{\mathbb{R}}$ and $T \in \mathcal{L}(X, V)$. Then the infimal function of f through T,*

$$Tf : V \to \overline{\mathbb{R}},\ (Tf)(y) := \inf\{f(x) : Tx = y\}$$

is convex, too.

For an arbitrary function $f : X \to \overline{\mathbb{R}}$ and $T \in \mathcal{L}(X, V)$ it holds $\mathrm{dom}(Tf) = T(\mathrm{dom}\,f)$.

The following notion can also be introduced as a particular instance of the infimal function of a given function through a suitable linear continuous mapping, as can be seen below. Though, we introduce it directly because of its importance in convex analysis and optimization.

Definition 2.2.7. *The infimal convolution of the functions $f_i : X \to \overline{\mathbb{R}}$, $i = 1, \ldots, m$, is the function*

$$f_1\square\ldots\square f_m : X \to \overline{\mathbb{R}}, (f_1\square\ldots\square f_m)(x) := \inf\left\{\sum_{i=1}^m f_i(x^i) : x^i \in X, \sum_{i=1}^m x^i = x\right\}.$$

When for $x \in X$ the infimum within is attained we say that the infimal convolution is exact at x. When the infimal convolution is exact everywhere we call it simply exact.

For $f_i : X \to \overline{\mathbb{R}}$, $i = 1, \ldots, m$, given functions, $f : X^m \to \overline{\mathbb{R}}$ defined by $f(x^1, \ldots, x^m) = \sum_{i=1}^m f_i(x^i)$ and $A \in \mathcal{L}(X^m, X)$, $A(x^1, \ldots, x^m) = \sum_{i=1}^m x^i$ it holds $Af = f_1\square\ldots\square f_m$. Thus $\mathrm{dom}(f_1\square\ldots\square f_m) = \sum_{i=1}^m \mathrm{dom}\,f_i$. By Theorem 2.2.7 it follows that if $f_i : X \to \overline{\mathbb{R}}$, $i = 1, \ldots, m$, are convex, as stated in the following theorem, their infimal convolution is convex, too.

Theorem 2.2.8. *Given the convex functions $f_i : X \to \overline{\mathbb{R}}$, $i = 1, \ldots, m$, then their infimal convolution $f_1\square\ldots\square f_m : X \to \overline{\mathbb{R}}$ is convex, too.*

The notion we introduce next is a generalization of the K-convexity (see Definition 2.2.6).

Definition 2.2.8. *A vector function $h : X \to V \cup \{+\infty_K\}$ is called K-convexlike if for all $x, y \in X$ and all $\lambda \in [0, 1]$ there is some $z \in X$ such that $h(z) \leqq_K \lambda h(x) + (1 - \lambda)h(y)$.*

It is easy to see that $h : X \to V \cup \{+\infty_K\}$ is K-convexlike if and only if $h(\operatorname{dom} h) + K$ is a convex set.

For $U \subseteq X$ a given nonempty set we call $h : U \to V$ *K-convexlike on U* if for all $x, y \in U$ and all $\lambda \in [0,1]$ there is some $z \in U$ such that $h(z) \leqq_K \lambda h(x) + (1-\lambda)h(y)$. Note that h is K-convexlike on U if and only if $h(U) + K$ is a convex set.

Remark 2.2.8. Every K-convex function $h : X \to V \cup \{+\infty_K\}$ is K-convexlike, but not all K-convexlike functions are K-convex. Consider, for instance, $\mathbb{R}^2$ partially ordered by the cone $\mathbb{R}^2_+$. Take the function $h : \mathbb{R} \to \mathbb{R}^2 \cup \{+\infty_{\mathbb{R}^2_+}\}$ defined by $h(x) = (x, \sin x)$ if $x \in [-\pi, \pi]$ and $h(x) = +\infty_{\mathbb{R}^2_+}$ otherwise. It can be proven that h is $\mathbb{R}^2_+$-convexlike, but not $\mathbb{R}^2_+$-convex.

2.2.2 Topological properties of convex functions

In this section we deal with topological notions for functions, which alongside the convexity endow them with special properties.

Definition 2.2.9. *A function $f : X \to \overline{\mathbb{R}}$ is called lower semicontinuous at $\bar{x} \in X$ if $\liminf_{x \to \bar{x}} f(x) \geq f(\bar{x})$. A function f is said to be upper semicontinuous at $\bar{x}$ if $(-f)$ is lower semicontinuous at $\bar{x}$. When a function f is lower (upper) semicontinuous at all $x \in X$ we call it lower (upper) semicontinuous.*

Obviously, $f : X \to \overline{\mathbb{R}}$ is continuous at $\bar{x} \in X$ if and only if f is lower and upper semicontinuous at $\bar{x} \in X$.

In the following we give some equivalent characterizations of the lower semicontinuity of a function.

Theorem 2.2.9. *Let be the function $f : X \to \overline{\mathbb{R}}$. The following statements are equivalent:*

(i) f is lower semicontinuous;
(ii) $\operatorname{epi} f$ *is closed;*
(iii) the level set $\{x \in X : f(x) \leq \lambda\}$ is closed for all $\lambda \in \mathbb{R}$.

Example 2.2.4. Given a set $U \subseteq X$, its indicator function δ_U is lower semicontinuous if and only if U is closed, while the support function σ_U is always weak* lower semicontinuous.

Proposition 2.2.10. *The pointwise supremum of a family of lower semicontinuous functions $f_i : X \to \overline{\mathbb{R}}$, $i \in I$, where I is an arbitrary index set, $f : X \to \overline{\mathbb{R}}$ defined by $f(x) = \sup_{i \in I} f_i(x)$ is lower semicontinuous, too.*

Proposition 2.2.11. *If $f, g : X \to \overline{\mathbb{R}}$ are lower semicontinuous at $x \in X$ and $\lambda \in (0, +\infty)$, then $f + g$ and λf are lower semicontinuous at x, too.*

Via the lower bound function one can introduce the *lower semicontinuous hull* of a function as follows.

Definition 2.2.10. *Consider a function $f : X \to \overline{\mathbb{R}}$. The function $\bar{f} : X \to \overline{\mathbb{R}}$, defined by*

$$\bar{f}(x) := \phi_{\mathrm{cl}(\mathrm{epi}\, f)}(x) = \inf\{t : (x,t) \in \mathrm{cl}(\mathrm{epi}\, f)\}$$

is called the lower semicontinuous hull of f.

Example 2.2.5. If $f : \mathbb{R} \to \overline{\mathbb{R}}$, $f = \delta_{(-\infty,0)}$, then obviously $\bar{f} = \delta_{(-\infty,0]}$.

Theorem 2.2.12. *Let be the function $f : X \to \overline{\mathbb{R}}$. Then the following statements are true*

(a) $\mathrm{epi}\, \bar{f} = \mathrm{cl}(\mathrm{epi}\, f)$*;*
(b) $\mathrm{dom}\, f \subseteq \mathrm{dom}\, \bar{f} \subseteq \mathrm{cl}(\mathrm{dom}\, f)$*;*
(c) $\bar{f}(x) = \liminf_{y \to x} f(y)$ *for all* $x \in X$*.*

Remark 2.2.9. For a given function $f : X \to \overline{\mathbb{R}}$ $\bar{f}$ is the greatest lower semicontinuous function less than or equal to f. Consequently,

$$\bar{f} = \sup\{g : X \to \overline{\mathbb{R}} : g(x) \le f(x)\ \forall x \in X \text{ and } g \text{ is lower semicontinuous}\}.$$

In the following we deal with functions that are both convex and lower semicontinuous and show some of the properties this class of functions is endowed with.

Theorem 2.2.13. *Let be $f : X \to \overline{\mathbb{R}}$ a convex function . Then f is lower semicontinuous if and only if it is weakly lower semicontinuous.*

Proposition 2.2.14. *If $f : X \to \overline{\mathbb{R}}$ is convex and lower semicontinuous, but not proper, then f cannot take finite values, i.e. f is everywhere equal to $+\infty$ or f takes the value $-\infty$ everywhere on its domain.*

Proposition 2.2.14 has as consequence the fact that if $f : X \to \overline{\mathbb{R}}$ is convex and lower semicontinuous and finite somewhere, then $f(x) > -\infty$ for all $x \in X$. By Proposition 2.2.1 and Theorem 2.2.12 follows that if $f : X \to \overline{\mathbb{R}}$ is convex then $\bar{f}$ is convex, too. Further, by Proposition 2.2.14 one has that if there is some $\bar{x} \in X$ such that $f(\bar{x}) = -\infty$, then $\bar{f}(x) = -\infty$ for all $x \in \mathrm{dom}\, \bar{f} \supseteq \mathrm{dom}\, f$.

We come now to a fundamental result linking a convex and lower semicontinuous function with the set of its affine minorants. A function $g : X \to \mathbb{R}$ is said to be *affine* if there are some $x^* \in X^*$ and $c \in \mathbb{R}$ such that $g(x) = \langle x^*, x\rangle + c$ for all $x \in X$. If $f : X \to \overline{\mathbb{R}}$ is a given function, then any affine function $g : X \to \mathbb{R}$ which fulfills $g(x) \le f(x)$ for all $x \in X$ is said to be an *affine minorant* of f.

Theorem 2.2.15. *Let be the given function $f : X \to \overline{\mathbb{R}}$. Then f is convex and lower semicontinuous and takes nowhere the value $-\infty$ if and only if its set of affine minorants is nonempty and f is the pointwise supremum of this set.*

Proof. The sufficiency is obvious, as a pointwise supremum of a family of affine functions is convex, by Proposition 2.2.2, and lower semicontinuous, via Proposition 2.2.10, noting that a function having an affine minorant cannot take the value $-\infty$.

To verify the necessity we first prove that the set

$$M := \big\{(x^*, \alpha) \in X^* \times \mathbb{R} : \langle x^*, x\rangle + \alpha \leq f(x)\ \forall x \in X\big\}$$

is nonempty. If $f \equiv +\infty$ then for all $x^* \in X^*$ and $\alpha \in \mathbb{R}$ we have $\langle x^*, x\rangle + \alpha \leq f(x)$ for all $x \in X$, i.e. $(x^*, \alpha) \in M$.

Otherwise, there must be at least an element $y \in X$ such that $f(y) \in \mathbb{R}$. Then $\operatorname{epi} f \neq \emptyset$ and $(y, f(y) - 1) \notin \operatorname{epi} f$. As the hypotheses guarantee that $\operatorname{epi} f$ is convex and closed, applying Theorem 2.1.5 follows the existence of some $x^* \in X^*$ and $\alpha \in \mathbb{R}$, $(x^*, \alpha) \neq (0, 0)$, such that

$$\langle x^*, y\rangle + \alpha(f(y) - 1) < \langle x^*, x\rangle + \alpha r\ \forall (x, r) \in \operatorname{epi} f.$$

As $(y, f(y)) \in \operatorname{epi} f$, it follows $\alpha > 0$ and $(1/\alpha)\langle x^*, y - x\rangle + f(y) - 1 < r$ for all $(x, r) \in \operatorname{epi} f$. Taking into consideration that whenever $x \in \operatorname{dom} f$ there is $(x, f(x)) \in \operatorname{epi} f$, the last inequality yields $(1/\alpha)\langle x^*, y - x\rangle + f(y) - 1 < f(x)$ for all $x \in \operatorname{dom} f$, and one can easily note that this inequality is valid actually for all $x \in X$. Consequently, the function $x \mapsto \langle(-1/\alpha)x^*, x\rangle + (1/\alpha)\langle x^*, y\rangle + f(y) - 1$ is an affine minorant of f, thus $M \neq \emptyset$ in this case, too.

For all $x \in X$ one has

$$f(x) \geq \sup\big\{\langle x^*, x\rangle + \alpha : (x^*, \alpha) \in X^* \times \mathbb{R},\ \langle x^*, z\rangle + \alpha \leq f(z)\ \forall z \in X\big\}$$

and next we prove that this inequality is always fulfilled as equality. Assume that there are some $\bar{x} \in X$ and $\bar{r} \in \mathbb{R}$ such that

$$f(\bar{x}) > \bar{r} > \sup\big\{\langle x^*, \bar{x}\rangle + \alpha : (x^*, \alpha) \in X^* \times \mathbb{R},\ \langle x^*, z\rangle + \alpha \leq f(z)\ \forall z \in X\big\}. \tag{2.4}$$

Then $(\bar{x}, \bar{r}) \notin \operatorname{epi} f$. Applying again Theorem 2.1.5 we obtain some $\bar{x}^* \in X^*$ and $\bar{\alpha} \in \mathbb{R}$, $(\bar{x}^*, \bar{\alpha}) \neq (0, 0)$, and an $\varepsilon > 0$ such that

$$\langle \bar{x}^*, x\rangle + \bar{\alpha} r > \langle \bar{x}^*, \bar{x}\rangle + \bar{\alpha}\bar{r} + \varepsilon\ \forall (x, r) \in \operatorname{epi} f. \tag{2.5}$$

For $(z, s) \in \operatorname{epi} f$ we get $(z, s + t) \in \operatorname{epi} f$ for all $t \geq 0$, thus $\bar{\alpha} \geq 0$. Assume that $f(\bar{x}) \in \mathbb{R}$. Then we obtain $\bar{\alpha}(f(\bar{x}) - \bar{r}) > \varepsilon$, which yields $\bar{\alpha} > 0$. Thus for all $x \in \operatorname{dom} f$ one has $f(x) > (1/\bar{\alpha})\langle \bar{x}^*, \bar{x} - x\rangle + \bar{r} + (1/\bar{\alpha})\varepsilon > (1/\bar{\alpha})\langle \bar{x}^*, \bar{x} - x\rangle + \bar{r}$. As the function $x \mapsto \langle(-1/\bar{\alpha})\bar{x}^*, x\rangle + \langle(1/\bar{\alpha})\bar{x}^*, \bar{x}\rangle + \bar{r}$ is an affine minorant of f taking at $x = \bar{x}$ the value $\bar{r}$, we obtain a contradiction to (2.4). Consequently, $f(\bar{x}) = +\infty$. Assuming $\bar{\alpha} > 0$ we reach again a contradiction, thus $\bar{\alpha} = 0$. Consider then the function $x \mapsto -\langle \bar{x}^*, x - \bar{x}\rangle + \varepsilon$. By (2.5) one gets $-\langle \bar{x}^*, x - \bar{x}\rangle + \varepsilon \leq 0$ for all $x \in \operatorname{dom} f$. As $M \neq \emptyset$, there are some $y^* \in X^*$ and $\beta \in \mathbb{R}$ such that $\langle y^*, z\rangle + \beta \leq f(z)$ whenever $z \in X$. Denote $\gamma := (\bar{r} - \langle y^*, \bar{x}\rangle - \beta)/\varepsilon$. It is clear that $\gamma > 0$ and that the function

$x \mapsto \langle y^* - \gamma\bar{x}^*, x\rangle + \langle \gamma\bar{x}^*, \bar{x}\rangle + \beta + \gamma\varepsilon$ is affine. For all $x \in \operatorname{dom} f$ there is $\langle y^* - \gamma\bar{x}^*, x\rangle + \langle \gamma\bar{x}^*, \bar{x}\rangle + \beta + \gamma\varepsilon = \langle y^*, x\rangle + \beta + \gamma(\langle -\bar{x}^*, x - \bar{x}\rangle + \varepsilon) \leq \langle y^*, x\rangle + \beta \leq f(x)$, thus $x \mapsto \langle y^* - \gamma\bar{x}^*, x\rangle + \langle \gamma\bar{x}^*, \bar{x}\rangle + \beta + \gamma\varepsilon$ is an affine minorant of f and for $x = \bar{x}$ one gets $\langle y^* - \gamma\bar{x}^*, \bar{x}\rangle + \langle \gamma\bar{x}^*, \bar{x}\rangle + \beta + \gamma\varepsilon = \bar{r}$, which contradicts (2.4). Thus f is the pointwise supremum of the set of its affine minorants. □

The lower bound function can be also used to introduce the *lower semicontinuous convex hull* of a function.

Definition 2.2.11. *Consider a function $f : X \to \overline{\mathbb{R}}$. The function $\overline{\operatorname{co}} f : X \to \overline{\mathbb{R}}$, defined by*

$$\overline{\operatorname{co}} f(x) := \phi_{\overline{\operatorname{co}}(\operatorname{epi} f)}(x) = \inf\{t : (x,t) \in \overline{\operatorname{co}}(\operatorname{epi} f)\}$$

is called the lower semicontinuous convex hull of f.

Some properties this notion is endowed with follow.

Theorem 2.2.16. *Let $f : X \to \overline{\mathbb{R}}$ be a given function. Then the following statements are true*

(a) $\operatorname{epi}(\overline{\operatorname{co}} f) = \overline{\operatorname{co}}(\operatorname{epi} f)$*;*
(b) $\operatorname{dom}(\operatorname{co} f) = \operatorname{co}(\operatorname{dom} f) \subseteq \operatorname{dom}(\overline{\operatorname{co}} f) \subseteq \operatorname{cl}(\operatorname{dom}(\operatorname{co} f)) = \overline{\operatorname{co}}(\operatorname{dom} f)$*.*

Remark 2.2.10. It is clear from the construction that the lower semicontinuous convex hull of a function $f : X \to \overline{\mathbb{R}}$ is convex and lower semicontinuous and it is the greatest convex lower semicontinuous function less than or equal to f. Consequently,

$$\overline{\operatorname{co}} f = \sup\{g : X \to \overline{\mathbb{R}} : g(x) \leq f(x) \text{ for all } x \in X \text{ and } g \text{ is convex and lower semicontinuous}\}.$$

Now we turn our attention to continuity properties of convex functions.

Theorem 2.2.17. *Let be $f : X \to \overline{\mathbb{R}}$ a convex function. The following statements are equivalent:*

(i) there is a nonempty open subset of X on which f is bounded from above by a finite constant and is not everywhere equal to $-\infty$;
(ii) f is proper and continuous on the interior of its effective domain, which is nonempty.

As a consequence of Theorem 2.2.17 it follows that a convex function $f : X \to \overline{\mathbb{R}}$ is continuous on $\operatorname{int}(\operatorname{dom} f)$ if and only if $\operatorname{int}(\operatorname{epi} f) \neq \emptyset$. If we take $X = \mathbb{R}^n$, every proper and convex function $f : \mathbb{R}^n \to \overline{\mathbb{R}}$ is continuous on $\operatorname{ri}(\operatorname{dom} f)$. Consequently, every convex function $f : \mathbb{R}^n \to \mathbb{R}$ is continuous.

Besides associating new functions to a given function, the lower bound function can be used to define the notion of a *gauge* of a given set. For $U \subseteq X$, consider the set $\operatorname{cone}(U \times \{1\}) = \{(\lambda x, \lambda) : \lambda \in \mathbb{R}_+, x \in U\}$. If U is convex and closed, then $\operatorname{cone}(U \times \{1\})$ is convex and closed, too.

Definition 2.2.12. *Given a convex absorbing subset U of a vector space X, the gauge (or Minkowski function) associated to it is the function $\gamma_U : X \to \mathbb{R}$ defined by*

$$\gamma_U(x) := \phi_{\mathrm{cone}(U \times \{1\})}(x) = \inf\{\lambda \geq 0 : x \in \lambda U\}.$$

In this situation, U is called the unit ball of the gauge γ_U.

Proposition 2.2.18. *(a) If X is a vector space and $U \subseteq X$ is absorbing and convex, then γ_U is sublinear and* $\mathrm{core}(U) = \{x \in X : \gamma_U(x) < 1\}$. *If U is moreover symmetric, then γ_U is a seminorm.*

(b) If X is a topological vector space and $U \subseteq X$ is a convex neighborhood of 0, *then γ_U is continuous,* $\mathrm{int}(U) = \{x \in X : \gamma_U(x) < 1\}$ *and* $\mathrm{cl}(U) = \{x \in X : \gamma_U(x) \leq 1\}$.

There are several extensions of the notion of lower semicontinuity for vector functions based on the properties of the lower semicontinuous functions. We recall here three of them, which are mostly used in convex optimization. Like before, V is a Hausdorff locally convex space partially ordered by the convex cone K.

Definition 2.2.13. *A function $h : X \to V \cup \{+\infty_K\}$ is called*

(a) K-lower semicontinuous at $x \in X$ if for any neighborhood W of zero in V and for any $b \in V$ satisfying $b \leqq_K h(x)$, there exists a neighborhood U of x in X such that $h(U) \subseteq b + W + K \cup \{+\infty_K\}$;

(b) star K-lower semicontinuous at $x \in X$ if $(k^* h)$ is lower semicontinuous at x for all $k^* \in K^*$.

Remark 2.2.11. The K-lower semicontinuity of a function $h : X \to V \cup \{+\infty_K\}$ was introduced by Penot ant Théra in [150], being later refined in [50]. For all $x \in \mathrm{dom}\, h$ the definition of the K-lower semicontinuity of h at x amounts to asking for any neighborhood W of zero in V the existence of a neighborhood U of x in X such that $h(U) \subseteq h(x) + W + K \cup \{+\infty_K\}$. The notion of star K-lower semicontinuity was first considered in [106].

Definition 2.2.14. *A function $h : X \to V \cup \{+\infty_K\}$ is called*

(a) K-lower semicontinuous if it is K-lower semicontinuous at every $x \in X$;

(b) star K-lower semicontinuous if it is star K-lower semicontinuous at every $x \in X$;

(c) K-epi closed if $\mathrm{epi}_K h$ *is closed.*

Proposition 2.2.19. *Let be the function $h : X \to V \cup \{+\infty_K\}$.*

(a) If h is K-lower semicontinuous at $x \in X$, then it is also star K-lower semicontinuous at x.

(b) If h is star K-lower semicontinuous, then it is also K-epi closed.

The following example shows that there are K-epi closed functions which are not star K-lower semicontinuous.

Example 2.2.6. Consider the function

$$h : \mathbb{R} \to \mathbb{R}^2 \cup \{+\infty_{\mathbb{R}^2_+}\},\ h(x) = \begin{cases} (\frac{1}{x}, x), & \text{if } x > 0, \\ +\infty_{\mathbb{R}^2_+}, & \text{otherwise.} \end{cases}$$

It can be verified that h is $\mathbb{R}^2_+$-convex and $\mathbb{R}^2_+$-epi-closed, but not star $\mathbb{R}^2_+$-lower semicontinuous. For instance, for $k^* = (0,1)^T \in (\mathbb{R}^2_+)^* = \mathbb{R}^2_+$ one has

$$((0,1)^T h)(x) = \begin{cases} x, & \text{if } x > 0, \\ +\infty, & \text{otherwise,} \end{cases}$$

which is not lower semicontinuous.

Remark 2.2.12. When $V = \mathbb{R}$ and $K = \mathbb{R}_+$, the notions of K-lower semicontinuity, star K-lower semicontinuity and K-epi closedness collapse into the classical notion of lower semicontinuity.

2.3 Conjugate functions and subdifferentiability

Throughout this entire section we consider X to be a Hausdorff locally convex space with its topological dual space X^* endowed with the weak* topology.

2.3.1 Conjugate functions

Let $f : X \to \overline{\mathbb{R}}$ be a given function. In the following we deal with the notion of conjugate function of f, a basic one in the theory of convex analysis and very important for establishing a general duality theory for convex optimization problems (see chapter 3 for more on this topic).

Definition 2.3.1. *The function*

$$f^* : X^* \to \overline{\mathbb{R}},\ f^*(x^*) := \sup_{x \in X}\{\langle x^*, x\rangle - f(x)\}$$

is said to be the (Fenchel) conjugate function of f.

Note that for all $x^* \in X^*$ it holds $f^*(x^*) = \sup_{x \in \operatorname{dom} f}\{\langle x^*, x\rangle - f(x)\}$.

Some of the investigations we make in this book will employ the *conjugate function of f with respect to the nonempty set $S \subseteq X$*, defined by

$$f^*_S : X^* \to \overline{\mathbb{R}}, f^*_S(x^*) := (f + \delta_S)^*(x^*) = \sup_{x \in S}\{\langle x^*, x\rangle - f(x)\}.$$

Lemma 2.3.1. *(a) If the function f is proper, then $f^*(x^*) > -\infty$ for all $x^* \in X^*$.*

(b) The function f^* *is proper if and only if* $\operatorname{dom} f \neq \emptyset$ *and* f *has an affine minorant.*

Proof. (a) If the function f is proper, then by definition there exists some $\bar{x} \in X$ such that $f(\bar{x}) \in \mathbb{R}$. Then for all $x^* \in X^*$ it holds $f^*(x^*) = \sup_{x \in X}\{\langle x^*, x\rangle - f(x)\} \geq \langle x^*, \bar{x}\rangle - f(\bar{x}) > -\infty$.

(b) Suppose first that the function f^* is proper. By definition there exists $\bar{x}^* \in X^*$ such that that $f^*(\bar{x}^*) \in \mathbb{R}$. Since $f^*(\bar{x}^*) = \sup_{x \in X}\{\langle \bar{x}^*, x\rangle - f(x)\} \geq \langle \bar{x}^*, x\rangle - f(x)$ for all $x \in X$, $x \mapsto \langle \bar{x}^*, x\rangle - f^*(\bar{x}^*)$ is an affine minorant of the function f. Assuming that $\operatorname{dom} f = \emptyset$ or, equivalently, $f \equiv +\infty$, one would have that $f^* \equiv -\infty$, which would contradict the assumption that f^* is proper. Thus $\operatorname{dom} f$ must be a nonempty set.

Assume now that $\operatorname{dom} f \neq \emptyset$ and that there exist $\bar{x}^* \in X^*$ and $c \in \mathbb{R}$ such that $f(x) \geq \langle \bar{x}^*, x\rangle + c$ for all $x \in X$. Obviously, the function f is proper and by (a) we get $f^* > -\infty$ on X^*. Moreover, $f^*(\bar{x}^*) = \sup_{x \in X}\{\langle \bar{x}^*, x\rangle - f(x)\} \leq -c$, whence f^* is proper. □

It is straightforward to verify that in case f is not proper one either has $f^* \equiv -\infty$ (if $f \equiv +\infty$) or $f^* \equiv +\infty$ (if there exists an $x \in X$ with $f(x) = -\infty$). In case $\operatorname{dom} f \neq \emptyset$ and f has an affine minorant $x \mapsto \langle x^*, x\rangle + c$, with $x^* \in X^*$ and $c \in \mathbb{R}$ it holds, as we have seen, $-c \geq f^*(x^*)$. Under these circumstances, $-f^*(x^*)$ represents the largest value $c \in \mathbb{R}$ for which $x \mapsto \langle x^*, x\rangle + c$ is an affine minorant of f.

Remark 2.3.1. It is a direct consequence of Definition 2.3.1 that f^* is the pointwise supremum of the family of affine functions $g_x : X^* \to \overline{\mathbb{R}}$, $g_x(x^*) = \langle x^*, x\rangle - f(x)$, $x \in \operatorname{dom} f$. Therefore f^* turns out to be a convex and lower semicontinuous function.

Next we collect some elementary properties of conjugate functions.

Proposition 2.3.2. *Let* f, g, $f_i : X \to \overline{\mathbb{R}}$, $i \in I$, *be given functions, where* I *is an arbitrary index set. Then the following statements hold*

(a) $f(x) + f^*(x^*) \geq \langle x^*, x\rangle \ \forall x \in X \ \forall x^* \in X^*$ *(Young-Fenchel inequality);*
(b) $\inf_{x \in X} f(x) = -f^*(0)$*;*
(c) $f \leq g$ *on* X *implies* $f^* \geq g^*$ *on* X^**;*
(d) $(\sup_{i \in I} f_i)^* \leq \inf_{i \in I} f_i^*$ *and* $(\inf_{i \in I} f_i)^* = \sup_{i \in I} f_i^*$*;*
(e) $(\lambda f)^*(x^*) = \lambda f^*((1/\lambda)x^*) \ \forall x^* \in X^* \ \forall \lambda > 0$*;*
(f) $(f + \beta)^* = f^* - \beta \ \forall \beta \in \mathbb{R}$*;*
(g) for $f_{x_0}(x) = f(x - x_0)$, *when* $x_0 \in X$, *there is* $(f_{x_0})^*(x^*) = f^*(x^*) + \langle x^*, x_0\rangle \ \forall x^* \in X^*$*;*
(h) for $f_{x_0^*}(x) = f(x) + \langle x_0^*, x\rangle$, *when* $x_0^* \in X^*$, *there is* $(f_{x_0^*})^*(x^*) = f^*(x^* - x_0^*) \ \forall x^* \in X^*$*;*
(i) for Y *a Hausdorff locally convex space and* $A : Y \to X$ *a linear continuous invertible mapping there is* $(f \circ A)^* = f^* \circ (A^{-1})^*$*;*
(j) $(f + g)^*(x^* + y^*) \leq f^*(x^*) + g^*(y^*) \ \forall x^*, y^* \in X^*$*;*

(k) $(\lambda f + (1-\lambda)g)^*(x^*) \leq \lambda f^*(x^*) + (1-\lambda)g^*(x^*)\ \forall x^* \in X^*\ \forall \lambda \in (0,1)$;

(l) for $f : X_1 \times \ldots \times X_m \to \overline{\mathbb{R}}$, $f(x_1, \ldots, x_m) = \sum_{i=1}^m f_i(x_i)$, *where* X_i *is a Hausdorff locally convex space and* $f_i : X_i \to \overline{\mathbb{R}}$, $i = 1, \ldots, m$, *there is* $f^*(x_1^*, \ldots, x_m^*) = \sum_{i=1}^m f_i^*(x_i^*)\ \forall (x_1^*, \ldots, x_m^*) \in X_1^* \times \ldots \times X_m^*$.

Proof. The verification of the above assertions is an obvious consequence of Definition 2.3.1. Therefore we confine ourselves only to point out the proof of the statements (d) and *(i)*.

(d) For all $j \in I$ and $x^* \in X^*$ we have $(\sup_{i\in I} f_i)^*(x^*) = \sup_{x\in X}\{\langle x^*, x\rangle - \sup_{i\in I} f_i(x)\} \leq \sup_{x\in X}\{\langle x^*, x\rangle - f_j(x)\} = f_j^*(x^*)$. Taking the infimum over $j \in I$ at the right-hand side of this inequality yields the wanted result.

In the second part of the statement, for all $x^* \in X^*$ we have $(\inf_{i\in I} f_i)^*(x^*) = \sup_{x\in X}\{\langle x^*, x\rangle - \inf_{i\in I} f_i(x)\} = \sup_{i\in I}\sup_{x\in X}\{\langle x^*, x\rangle - f_i(x)\} = \sup_{i\in I} f_i^*(x^*)$.

(i) Let $x^* \in X^*$ be arbitrarily taken. It holds $(f \circ A)^*(x^*) = \sup_{x\in X}\{\langle x^*, x\rangle - f(Ax)\} = \sup_{y\in Y}\{\langle x^*, A^{-1}y\rangle - f(y)\} = \sup_{y\in Y}\{\langle (A^{-1})^*x^*, y\rangle - f(y)\} = f^*((A^{-1})^*x^*) = (f^* \circ (A^{-1})^*)(x^*)$ and the desired relation is proved. □

Remark 2.3.2. The convention $(+\infty)+(-\infty) = (-\infty)+(+\infty) = +\infty$ ensures that the Young-Fenchel inequality applies also to improper functions.

For a function defined on the dual space X^* one can introduce its conjugate function analogously. More precisely, if we consider $g : X^* \to \overline{\mathbb{R}}$, then

$$g^* : X \to \overline{\mathbb{R}}, \quad g^*(x) = \sup_{x^*\in X^*}\{\langle x^*, x\rangle - g(x^*)\}$$

is the conjugate function of g. In particular, to the function $f : X \to \overline{\mathbb{R}}$ we can attach the so-called *biconjugate* function of f, which is defined as the conjugate function of the conjugate f^*, i.e.

$$f^{**} : X \to \overline{\mathbb{R}},\ f^{**}(x) := (f^*)^*(x) = \sup_{x^*\in X^*}\{\langle x^*, x\rangle - f^*(x^*)\}.$$

The next lemma is a direct consequence of the Young-Fenchel inequality.

Lemma 2.3.3. *For all* $x \in X$ *it holds* $f^{**}(x) \leq f(x)$.

Next the conjugates of some convex functions needed later are provided.

Example 2.3.1. Let $X = \mathbb{R}$ and $f : \mathbb{R} \to \overline{\mathbb{R}}$.

(a) If $f(x) = (1/2)x^2$, $x \in \mathbb{R}$, then $f^*(x^*) = (1/2)(x^*)^2$ for $x^* \in \mathbb{R}$.

(b) If $f(x) = \mathrm{e}^x$, $x \in \mathbb{R}$, then

$$f^*(x^*) = \begin{cases} x^*(\ln x^* - 1), & \text{if } x^* > 0, \\ 0, & \text{if } x^* = 0, \\ +\infty, & \text{if } x^* < 0. \end{cases}$$

(c) If

$$f(x) = \begin{cases} x(\ln x - 1), & \text{if } x > 0, \\ 0, & \text{if } x = 0, \\ +\infty, & \text{if } x < 0, \end{cases}$$

then $f^*(x^*) = e^{x^*}$ for $x^* \in \mathbb{R}$.

Example 2.3.2. Let be $f : X \to \mathbb{R}$, $f(x) = \langle y^*, x\rangle + c$, with $y^* \in X^*$ and $c \in \mathbb{R}$. Then

$$f^*(x^*) = \delta_{\{y^*\}}(x^*) - c = \begin{cases} -c, & \text{if } x^* = y^*, \\ +\infty, & \text{otherwise.} \end{cases}$$

Example 2.3.3. Let be $U \subseteq X$. Then for all $x^* \in X^*$ there is

$$\delta_U^*(x^*) = \sup_{x \in X}\{\langle x^*, x\rangle - \delta_U(x)\} = \sup_{x \in U}\langle x^*, x\rangle = \sigma_U(x^*).$$

Example 2.3.4. (a) Given a convex absorbing subset U of X, the conjugate of its gauge $\gamma_U : X \to \mathbb{R}$ at some $x^* \in X^*$ is

$$\begin{aligned} (\gamma_U)^*(x^*) &= \sup_{x \in X}\left\{\langle x^*, x\rangle - \inf\{\lambda \geq 0 : x \in \lambda U\}\right\} \\ &= \sup_{x \in X}\left\{\langle x^*, x\rangle + \sup_{\substack{\lambda \geq 0, \\ x \in \lambda U}}\{-\lambda\}\right\} = \sup_{\lambda \geq 0}\left\{ -\lambda + \sup_{y \in U}\langle x^*, \lambda y\rangle\right\} \\ &= \sup_{\lambda \geq 0}\left\{\lambda\left(\sup_{y \in U}\langle x^*, y\rangle - 1\right)\right\} = \begin{cases} 0, & \text{if } \sigma_U(x^*) \leq 1, \\ +\infty, & \text{otherwise.} \end{cases} \end{aligned}$$

(b) Let $(X, \|\cdot\|)$ be a normed vector space and $(X^*, \|\cdot\|_*)$ its topological dual space. The conjugate of the norm function can be deduced from the one of the gauge corresponding to the set $U = \{x \in X : \|x\| \leq 1\}$, since $\gamma_U = \|\cdot\|$ and $\sigma_U = \|\cdot\|_*$. Therefore for $x^* \in X^*$ one has

$$(\|\cdot\|)^*(x^*) = \begin{cases} 0, & \text{if } \|x^*\|_* \leq 1, \\ +\infty, & \text{otherwise.} \end{cases}$$

Example 2.3.5. Let X be a normed space and $f : X \to \mathbb{R}$, $f(x) = (1/p)\|x\|^p$, $1 < p < \infty$. Then $f^*(x^*) = (1/q)\|x^*\|_*^q$ for $x^* \in X^*$, where $(1/p) + (1/q) = 1$.

The results we prove next are necessary for deriving further statements, in particular concerning duality.

Proposition 2.3.4. *The following relations are always fulfilled*

(a) $f^* = (\bar{f})^* = (\overline{\text{co}} f)^*$ *on* X^**;*
(b) $f^{**} \leq \overline{\text{co}} f \leq \bar{f} \leq f$ *on* X*.*

Proof. (a) Taking a careful look at the way the functions $\overline{\text{co}} f$ and $\bar{f}$ are defined, it is not hard to see that on X the inequalities $\overline{\text{co}} f \leq \bar{f} \leq f$ are always fulfilled (see also Theorem 2.2.12(a)). Applying Proposition 2.3.2(c)

we get $(\overline{\text{co}} f)^* \geq \bar{f}^* \geq f^*$ on X^*. In order to get the desired conclusion, we prove that $(\overline{\text{co}} f)^* \leq f^*$. Let $x^* \in X^*$ be arbitrarily taken. We treat further three cases.

If $f^*(x^*) = +\infty$ we get $(\overline{\text{co}} f)^*(x^*) = (\bar{f})^*(x^*) = f^*(x^*) = +\infty$.

If $f^*(x^*) = -\infty$, then $f^{**} \equiv +\infty$ and, by Lemma 2.3.3, one has $f \equiv +\infty$. This implies further $\bar{f} = \overline{\text{co}} f = f \equiv +\infty$ and, consequently, $(\overline{\text{co}} f)^*(x^*) = (\bar{f})^*(x^*) = f^*(x^*) = -\infty$.

It remains to consider the case $f^*(x^*) \in \mathbb{R}$. Consider the function $g : X \to \mathbb{R}$, $g(x) = \langle x^*, x \rangle - f^*(x^*)$. According to the Young-Fenchel inequality for all $x \in X$ we have $g(x) \leq f(x)$, which is equivalent to $\text{epi}\, g \supseteq \text{epi}\, f$. Since g is an affine function, $\text{epi}\, g$ is a convex and closed set and it holds $\text{epi}\, g \supseteq \overline{\text{co}}(\text{epi}\, f)$. By Theorem 2.2.16$(a)$ we get $\text{epi}\, g \supseteq \text{epi}(\overline{\text{co}} f)$, and from here we deduce that $g(x) = \langle x^*, x \rangle - f^*(x^*) \leq \overline{\text{co}} f(x)$ for all $x \in X$. This implies $(\overline{\text{co}} f)^*(x^*) = \sup_{x \in X}\{\langle x^*, x \rangle - \overline{\text{co}} f(x)\} \leq f^*(x^*)$ and in this way the statement (a) has been verified.

(b) We only have to justify $f^{**} \leq \overline{\text{co}} f$ on X. As the equality $f^* = (\overline{\text{co}} f)^*$ is secured by (a), it holds $f^{**} = (\overline{\text{co}} f)^{**} \leq \overline{\text{co}} f$ on X, where the last inequality follows by Lemma 2.3.3. □

Because of the inequality $f^{**} \leq f$ on X, arises in a natural way the question when does the coincidence of f and f^{**} occur. The next statement gives an answer.

Theorem 2.3.5. *If $f : X \to \overline{\mathbb{R}}$ is proper, convex and lower semicontinuous, then f^* is proper and $f = f^{**}$.*

Proof. We prove first that f^* is proper. As f is proper, $\text{dom}\, f \neq \emptyset$ and, by Theorem 2.2.15, f has an affine minorant. Thus Lemma 2.3.1(b) guarantees the properness of f^*. We prove next that $f = f^{**}$. For all $x \in X$ we have

$$f^{**}(x) = \sup_{x^* \in X^*}\{\langle x^*, x \rangle - f^*(x^*)\} = \sup_{\substack{x^* \in X^*, c \in \mathbb{R}, \\ f^*(x^*) \leq -c}}\{\langle x^*, x \rangle + c\} =$$

$$\sup\{\langle x^*, x \rangle + c : x^* \in X^*, c \in \mathbb{R}, \langle x^*, z \rangle + c \leq f(z)\ \forall z \in X\}$$

and this is equal, again by Theorem 2.2.15, to $f(x)$. □

The well-known *Fenchel-Moreau theorem* follows as a direct consequence of Theorem 2.3.5. Because of its fame and historical importance we cite it here as a separate statement.

Theorem 2.3.6. *(Fenchel-Moreau) Let $f : X \to \overline{\mathbb{R}}$ be a proper function. Then $f = f^{**}$ if and only if f is convex and lower semicontinuous.*

Corollary 2.3.7. *Let $f : X \to \overline{\mathbb{R}}$. If $\overline{\text{co}} f > -\infty$, then $f^{**} = \overline{\text{co}} f$.*

Proof. If $\overline{\text{co}} f$ is proper, then the conclusion follows by Proposition 2.3.4(a) and Theorem 2.3.5. If $\overline{\text{co}} f \equiv +\infty$, then $\overline{\text{co}} f = f^{**} \equiv +\infty$ and the result holds also in this case. □

Remark 2.3.3. It is an immediate conclusion of Theorem 2.3.5 that for a convex function $f : X \to \overline{\mathbb{R}}$ its conjugate function f^* is proper if and only if $\bar{f}$ is proper.

Remark 2.3.4. Until now we have attached to a function $f : X \to \overline{\mathbb{R}}$ the conjugate function f^* and the biconjugate function f^{**}. It is natural to ask if it makes sense to consider the conjugate of the latter, namely $f^{***} : X^* \to \overline{\mathbb{R}}$ defined by $f^{***} = (f^{**})^*$. Since we always have $f^* = f^{***}$, this is not the case. In order to prove this we treat two cases.

Let us assume first that the function f^* is proper. Since f^* is also convex and lower semicontinuous, Theorem 2.3.5 secures the equality $f^{***} = (f^*)^{**} = f^*$. Assume now that the function f^* is not proper. If $f^* \equiv +\infty$ then $f^{**} \equiv -\infty$ and this implies $f^{***} \equiv +\infty$. If there is an x^* such that $f^*(x^*) = -\infty$, then $f^{**} \equiv +\infty$, which yields $f^{***} \equiv -\infty$. Moreover, by Lemma 2.3.3 it is obvious that $f \equiv +\infty$ and so $f^* \equiv -\infty$.

In convex analysis it is very natural and often also very useful to reformulate results employing functions in the language of their epigraphs. This applies also to conjugacy properties and the corresponding operations.

Let Y be another Hausdorff locally convex space whose topological dual space Y^* is endowed with the weak* topology. For $f : X \to \overline{\mathbb{R}}$ a given function and $A \in \mathcal{L}(X, Y)$ we calculate in the following the conjugate of the infimal function of f through A and derive from it the formula for the conjugate of the infimal convolution of a finite family of functions $f_i : X \to \overline{\mathbb{R}}$, $i = 1, \ldots, m$.

Proposition 2.3.8. *(a) Let $f : X \to \overline{\mathbb{R}}$ be a given function and $A \in \mathcal{L}(X, Y)$. Then it holds $(Af)^* = f^* \circ A^*$.*
(b) Let $f_i : X \to \overline{\mathbb{R}}$, $i = 1, \ldots, m$, be given functions. Then $(f_1 \Box \ldots \Box f_m)^* = \sum_{i=1}^m f_i^*$.

Proof. (a) By definition there holds for any $y^* \in Y^*$

$$(Af)^*(y^*) = \sup_{y \in Y}\{\langle y^*, y\rangle - (Af)(y)\} = \sup_{y \in Y}\{\langle y^*, y\rangle - \inf_{x \in X, Ax = y} f(x)\} =$$

$$\sup_{x \in X}\{\langle y^*, Ax\rangle - f(x)\} = \sup_{x \in X}\{\langle A^* y^*, x\rangle - f(x)\} = (f^* \circ A^*)(y^*).$$

(b) Taking $f : X^m \to \overline{\mathbb{R}}$, $f(x^1, \ldots, x^m) = \sum_{i=1}^m f_i(x^i)$ and $A \in \mathcal{L}(X^m, X)$, $A(x^1, \ldots, x^m) = \sum_{i=1}^m x^i$, we have seen that $Af = f_1 \Box \ldots \Box f_m$. Applying the result from (a) we get $(f_1 \Box \ldots \Box f_m)^* = f^* \circ A^*$. The conclusion follows by using Proposition 2.3.2(l) and the fact that $A^* x^* = (x^*, \ldots, x^*)$ for all $x^* \in X^*$. $\Box$

Of course, it is also of interest, to give a formula for the conjugate of the sum of a finite number of functions. A first calculation shows that for $x^{i*} \in X^*$, $i = 1, \ldots, m$, there is

$$\left(\sum_{i=1}^m f_i\right)^*\left(\sum_{i=1}^m x^{i*}\right) = \sup_{x\in X}\left\{\sum_{i=1}^m \langle x^{i*}, x\rangle - \sum_{i=1}^m f_i(x)\right\}$$

$$\leq \sum_{i=1}^m \sup_{x\in X}\{\langle x^{i*}, x\rangle - f_i(x)\} = \sum_{i=1}^m f_i^*(x^{i*}).$$

Consequently, for $x^* \in X^*$,

$$\left(\sum_{i=1}^m f_i\right)^*(x^*) \leq \inf\left\{\sum_{i=1}^m f_i^*(x^{i*}) : \sum_{i=1}^m x^{i*} = x^*\right\} = (f_1^*\Box\ldots\Box f_m^*)(x^*). \tag{2.6}$$

In a natural way the question of the coincidence of both sides of (2.6) arises. We can give first an equivalent characterization of this situation (see [38]).

Proposition 2.3.9. *Let $f_i : X \to \overline{\mathbb{R}}$, $i = 1,\ldots,m$, be proper functions such that $\cap_{i=1}^m \operatorname{dom} f_i \neq \emptyset$. Then the following statements are equivalent:*

(i) $\operatorname{epi}\left(\sum_{i=1}^m f_i\right)^* = \sum_{i=1}^m \operatorname{epi} f_i^*$;
(ii) $\left(\sum_{i=1}^m f_i\right)^* = f_1^*\Box\ldots\Box f_m^*$ *and the infimal convolution is exact.*

Proof. $(i) \Rightarrow (ii)$ Let $x^* \in X^*$ be arbitrarily taken. Then $\left(\sum_{i=1}^m f_i\right)^*(x^*) > -\infty$. If $\left(\sum_{i=1}^m f_i\right)^*(x^*) = +\infty$ then (ii) is automatically fulfilled, thus we consider further that $\left(\sum_{i=1}^m f_i\right)^*(x^*) < +\infty$, i.e. $\left(x^*, \left(\sum_{i=1}^m f_i\right)^*(x^*)\right) \in \operatorname{epi}\left(\sum_{i=1}^m f_i\right)^*$. By (i) there exist $(x^{i*}, r_i) \in \operatorname{epi} f_i^*$, $i = 1,\ldots,m$, such that $x^* = \sum_{i=1}^m x^{i*}$ and $\left(\sum_{i=1}^m f_i\right)^*(x^*) = \sum_{i=1}^m r_i$. This implies $f_i^*(x^{i*}) \leq r_i$, $i = 1,\ldots,m$, followed by $\sum_{i=1}^m f_i^*(x^{i*}) \leq \left(\sum_{i=1}^m f_i\right)^*(x^*)$. Consequently, $(f_1^*\Box\ldots\Box f_m^*)(x^*) \leq \left(\sum_{i=1}^m f_i\right)^*(x^*)$, which, combined with (2.6), yields (ii).

$(ii) \Rightarrow (i)$ Let the pairs $(x^{i*}, r_i) \in \operatorname{epi} f_i^*$, $i = 1,\ldots,m$, be given. Then $\left(\sum_{i=1}^m f_i\right)^*\left(\sum_{i=1}^m x^{i*}\right) \leq \sum_{i=1}^m f_i^*(x^{i*}) \leq \sum_{i=1}^m r_i$, i.e. $\left(\sum_{i=1}^m x^{i*}, \sum_{i=1}^m r_i\right) \in \operatorname{epi}\left(\sum_{i=1}^m f_i\right)^*$. Therefore $\operatorname{epi}\left(\sum_{i=1}^m f_i\right)^* \supseteq \sum_{i=1}^m \operatorname{epi} f_i^*$ and this inclusion is always valid. Taking now some arbitrary pair $(x^*, r) \in \operatorname{epi}\left(\sum_{i=1}^m f_i\right)^*$, we get $\left(\sum_{i=1}^m f_i\right)^*(x^*) \leq r$. By (ii) there exist some $x^{i*} \in X^*$, $i = 1,\ldots,m$, such that $\sum_{i=1}^m x^{i*} = x^*$ and $\sum_{i=1}^m f_i^*(x^{i*}) \leq r$. This yields the existence of some $r_i \in \mathbb{R}$, $i = 1,\ldots,m$, with $\sum_{i=1}^m r_i = r$, such that $f_i^*(x^{i*}) \leq r_i$ for all $i = 1,\ldots,m$. Then $(x^*, r^*) = \left(\sum_{i=1}^m x^{i*}, \sum_{i=1}^m r_i\right) \in \sum_{i=1}^m \operatorname{epi} f_i^*$ and the proof is complete. □

In order to give a sufficient condition for the equality in (2.6) note first that for any proper functions $f_i : X \to \overline{\mathbb{R}}$, $i = 1,\ldots,m$, fulfilling $\cap_{i=1}^m \operatorname{dom} f_i \neq \emptyset$, there is

$$\operatorname{cl}\left(\sum_{i=1}^m \operatorname{epi} f_i\right) \supseteq \operatorname{epi}(f_1\Box\ldots\Box f_m) \supseteq \sum_{i=1}^m \operatorname{epi} f_i,$$

which has as consequence that

$$\mathrm{cl}(\mathrm{epi}(f_1\Box\ldots\Box f_m)) = \mathrm{epi}\,\overline{f_1\Box\ldots\Box f_m} = \mathrm{cl}\left(\sum_{i=1}^m \mathrm{epi}\, f_i\right). \tag{2.7}$$

The following two results characterize the epigraph of the conjugate of the sum of finitely many functions.

Theorem 2.3.10. *Let be $f_i : X \to \overline{\mathbb{R}}$, $i = 1,\ldots,m$, proper, convex and lower semicontinuous functions fulfilling $\cap_{i=1}^m \mathrm{dom}\, f_i \neq \emptyset$. Then one has*

$$\left(\sum_{i=1}^m f_i\right)^* = \overline{f_1^*\Box\ldots\Box f_m^*},$$

and, consequently,

$$\mathrm{epi}\left(\sum_{i=1}^m f_i\right)^* = \mathrm{epi}\,\overline{f_1^*\Box\ldots\Box f_m^*} = \mathrm{cl}\left(\sum_{i=1}^m \mathrm{epi}\, f_i^*\right).$$

Proof. By Theorem 2.2.12(b) we get

$$\sum_{i=1}^m \mathrm{dom}\, f_i^* = \mathrm{dom}(f_1^*\Box\ldots\Box f_m^*) \subseteq \mathrm{dom}\,\overline{f_1^*\Box\ldots\Box f_m^*}.$$

Since Theorem 2.3.5 ensures that f_i^*, $i = 1,\ldots,m$, are proper functions, it holds $\sum_{i=1}^m \mathrm{dom}\, f_i^* \neq \emptyset$, consequently, $\mathrm{dom}\,\overline{f_1^*\Box\ldots\Box f_m^*} \neq \emptyset$. Assuming that there is some $x^* \in X^*$ such that $\overline{f_1^*\Box\ldots\Box f_m^*}(x^*) = -\infty$, we get $\sum_{i=1}^m f_i^{**} = \sum_{i=1}^m f_i \equiv +\infty$, which contradicts the hypothesis $\cap_{i=1}^m \mathrm{dom}\, f_i \neq \emptyset$. Therefore $\overline{f_1^*\Box\ldots\Box f_m^*}$ is a proper function. We can apply now Theorem 2.3.5, which yields $\overline{f_1^*\Box\ldots\Box f_m^*} = \left(\overline{f_1^*\Box\ldots\Box f_m^*}\right)^{**}$. Using Proposition 2.3.4(a) and Proposition 2.3.8(b), it follows $\overline{f_1^*\Box\ldots\Box f_m^*} = \left(f_1^*\Box\ldots\Box f_m^*\right)^{**} = \left(\sum_{i=1}^m f_i^{**}\right)^*$. Then the first formula follows via Theorem 2.3.5 and, together with (2.7), it yields the second one, too. □

Remark 2.3.5. For two proper, convex and lower semicontinuous functions $f, g : X \to \overline{\mathbb{R}}$ fulfilling $\mathrm{dom}\, f \cap \mathrm{dom}\, g \neq \emptyset$, Theorem 2.3.10 yields the classical *Moreau-Rockafellar formula*, namely

$$(f+g)^* = \overline{f^*\Box g^*}.$$

Turning to epigraphs, we get

$$\mathrm{epi}(f+g)^* = \mathrm{epi}\,\overline{f^*\Box g^*} = \mathrm{cl}(\mathrm{epi}\, f^* + \mathrm{epi}\, g^*).$$

Remark 2.3.6. As seen in Theorem 2.3.10 and Proposition 2.3.9, a sufficient condition to have equality in (2.6) for the proper functions $f_i : X \to \overline{\mathbb{R}}$, $i = 1,\ldots,m$, when $\cap_{i=1}^m \mathrm{dom}\, f_i \neq \emptyset$ and all these functions are convex and lower semicontinuous, is $\sum_{i=1}^m \mathrm{epi}\, f_i^*$ closed. For other sufficient conditions that guarantee the equality in (2.6) we refer to section 3.5.

We close this subsection by characterizing the conjugate of a K-increasing function, which will be useful in chapter 3 when dealing with composed convex optimization problems. Let V be a Hausdorff locally convex space, partially ordered by a convex cone $K \subseteq V$.

Proposition 2.3.11. *If $g : V \to \overline{\mathbb{R}}$ is a K-increasing function with* $\operatorname{dom} g \neq \emptyset$*, then $g^*(v^*) = +\infty$ for all $v^* \notin K^*$, i.e.* $\operatorname{dom} g^* \subseteq K^*$.

Proof. If $K = \{0\}$ the conclusion follows automatically. Assume that $K \neq \{0\}$ and take an arbitrary $v^* \notin K^*$. By definition there exists $\bar{v} \in K$ such that $\langle v^*, \bar{v} \rangle < 0$. Since for some arbitrary $\tilde{v} \in \operatorname{dom} g$ and for all $\alpha > 0$ we have $g(\tilde{v} - \alpha \bar{v}) \leq g(\tilde{v})$, it is straightforward to see that

$$g^*(v^*) = \sup_{v \in V} \{\langle v^*, v \rangle - g(v)\} \geq \sup_{\alpha > 0} \{\langle v^*, \tilde{v} - \alpha \bar{v} \rangle - g(\tilde{v} - \alpha \bar{v})\}$$

$$\geq \sup_{\alpha > 0} \{\langle v^*, \tilde{v} - \alpha \bar{v} \rangle - g(\tilde{v})\} = \langle v^*, \tilde{v} \rangle - g(\tilde{v}) + \sup_{\alpha > 0} \{-\alpha \langle v^*, \bar{v} \rangle\} = +\infty,$$

and the proof is complete. □

2.3.2 Subdifferentiability

In nondifferentiable convex optimization the classical *(Gâteaux) differentiability* may be replaced by the so-called *subdifferentiability*. To have a differentiability notion is extremely beneficial in analysis and optimization not only from the theoretical, but also from the numerical point of view. It allows, for instance, to formulate functional equations to describe mathematical objects and models for practical problems or to give optimality conditions in different fields of mathematical programming, variational calculus, for optimal control problems etc.

In this book we consider in the most situations scalar and multiobjective programming problems which involve convex sets and convex functions, without making use of the classical differentiability which is included as a special case of the general setting.

Definition 2.3.2. *Let $f : X \to \overline{\mathbb{R}}$ be a given function and take an arbitrary $x \in X$ such that $f(x) \in \mathbb{R}$. The set*

$$\partial f(x) := \{x^* \in X^* : f(y) - f(x) \geq \langle x^*, y - x \rangle \ \forall y \in X\}$$

is said to be the (convex) subdifferential of f at x. Its elements are called subgradients of f at x. We say that the function f is subdifferentiable at x if $\partial f(x) \neq \emptyset$.

If $f(x) \notin \mathbb{R}$ we consider by convention $\partial f(x) := \emptyset$.

Example 2.3.6. For $U \subseteq X$ and $f = \delta_U : X \to \overline{\mathbb{R}}$, one can easily show that for all $x \in U$ there is $\partial \delta_U(x) = N(U, x)$.

If $f : X \to \overline{\mathbb{R}}$ is subdifferentiable at x with $f(x) \in \mathbb{R}$ and $x^* \in \partial f(x)$, then the function $h : X \to \overline{\mathbb{R}}$, $h(y) = \langle x^*, y\rangle + f(x) - \langle x^*, x\rangle$ is an affine minorant of f. Moreover, this affine minorant coincides at x with f. In the following statement we give a characterization of the elements $x^* \in \partial f(x)$ according to the fact that for x^* and x the Fenchel-Young inequality is fulfilled as equality.

Theorem 2.3.12. *Let the function $f : X \to \overline{\mathbb{R}}$ be given and $x \in X$. Then $x^* \in \partial f(x)$ if and only if $f(x) + f^*(x^*) = \langle x^*, x\rangle$.*

Proof. Let $x^* \in \partial f(x)$. Then $f(x) \in \mathbb{R}$ and $f^*(x^*) = \sup\{\langle x^*, y\rangle - f(y) : y \in X\} \le \langle x^*, x\rangle - f(x)$. Since the opposite inequality is always true, $f(x) + f^*(x^*) = \langle x^*, x\rangle$ follows.

Vice versa, let $x \in X$ and $x^* \in X^*$ be such that $f(x) + f^*(x^*) = \langle x^*, x\rangle$. Then $f(x) \in \mathbb{R}$ and $\langle x^*, x\rangle - f(x) = f^*(x^*) = \sup_{y\in X}\{\langle x^*, y\rangle - f(y)\} \ge \langle x^*, y\rangle - f(y)$ for all $y \in X$, and hence $x^* \in \partial f(x)$. □

For $f : X \to \overline{\mathbb{R}}$ a given function one has that $x \in X$ with $f(x) \in \mathbb{R}$ is a solution of the optimization problem $\inf_{x\in X} f(x)$ if and only if $0 \in \partial f(x)$. In general one can express necessary and sufficient optimality conditions for optimization problems by means of subdifferentials as we shall see in section 3.3.

Proposition 2.3.13. *(a) For a given function $f : X \to \overline{\mathbb{R}}$, one has $\partial(f + \langle x^*, \cdot\rangle)(x) = \partial f(x) + x^*$ for all $x^* \in X^*$ and all $x \in X$.*
(b) For $f : X_1 \times \ldots \times X_m \to \overline{\mathbb{R}}$, $f(x^1, \ldots, x^m) = \sum_{i=1}^m f_i(x^i)$, where X_i is a Hausdorff locally convex space and $f_i : X_i \to \overline{\mathbb{R}}$, $i = 1, \ldots, m$, there is $\partial f(x^1, \ldots, x^m) = \prod_{i=1}^m \partial f_i(x^i)$ for all $(x^1, \ldots, x^m) \in X^1 \times \ldots \times X^m$.

Theorem 2.3.14. *Let $f : X \to \overline{\mathbb{R}}$ and $x \in X$. The subdifferential $\partial f(x)$ is a (possibly empty) convex and closed set in X^*.*

Proof. If $f(x) = \pm\infty$ there is nothing to prove. Let be $f(x) \in \mathbb{R}$. By the Young-Fenchel inequality and Theorem 2.3.12 it follows that $x^* \in \partial f(x)$ if and only if $f(x) + f^*(x^*) \le \langle x^*, x\rangle$. Therefore one can rewrite the subdifferential of the function f at x as the level set of the convex and lower semicontinuous function $x^* \mapsto -\langle x^*, x\rangle + f^*(x^*)$ at $-f(x)$, i.e. $\partial f(x) = \{x^* \in X^* : -\langle x^*, x\rangle + f^*(x^*) \le -f(x)\}$. This guarantees the convexity and, via Theorem 2.2.9, the closedness of $\partial f(x)$. □

The aim of the next theorem is to present some connections between the subdifferentials of the functions f, $\bar{f}$ and $\overline{\text{co}}f$.

Theorem 2.3.15. *Let be $f : X \to \overline{\mathbb{R}}$ and $x \in X$ be such that $\partial f(x) \ne \emptyset$. Then it holds*

(a) $\overline{\text{co}}f(x) = \bar{f}(x) = f(x)$ and the functions f, $\bar{f}$ and $\overline{\text{co}}f$ are proper and f is lower semicontinuous at x;
(b) $\partial(\overline{\text{co}}f)(x) = \partial\bar{f}(x) = \partial f(x)$;

(c) $f^{**} = \overline{\text{co}} f$.

Proof. (a) Let $x^* \in \partial f(x)$ be arbitrarily taken and consider the function $h : X \to \mathbb{R}$, $h(y) = \langle x^*, y\rangle + f(x) - \langle x^*, x\rangle$, which is an affine minorant of f. Note that $f(x) \in \mathbb{R}$. Since h is also convex and lower semicontinuous it holds $h \leq \overline{\text{co}} f \leq \bar{f} \leq f$. Taking into consideration that $f(x) = h(x)$ we deduce that $f(x) = h(x) \leq \overline{\text{co}} f(x) \leq \bar{f}(x) \leq f(x)$ and the desired equalities follow. This also implies that the function f is lower semicontinuous at x and the properness of f, $\bar{f}$ and $\overline{\text{co}} f$ follows easily.

(b) If $x^* \in \partial f(x)$ then, by definition, $f(y) \geq f(x) + \langle x^*, y - x\rangle$ for all $y \in X$. As $y \mapsto \langle x^*, y - x\rangle + f(x)$ is a convex and lower semicontinuous function which is everywhere less than or equal to f, using (a) we get $\overline{\text{co}} f(y) \geq \overline{\text{co}} f(x) + \langle x^*, y - x\rangle$ for all $y \in X$. Thus $x^* \in \partial(\overline{\text{co}} f)(x)$ and the inclusion $\partial f(x) \subseteq \partial(\overline{\text{co}} f)(x)$ follows. Assume now that $x^* \in \partial(\overline{\text{co}} f)(x)$. Because of (a), for all $y \in X$ we have $f(y) - f(x) \geq \overline{\text{co}} f(y) - \overline{\text{co}} f(x) \geq \langle x^*, y - x\rangle$, i.e. $x^* \in \partial f(x)$. Therefore $\partial(\overline{\text{co}} f)(x) \subseteq \partial f(x)$ and we actually have $\partial f(x) = \partial(\overline{\text{co}} f)(x)$. Following the same idea one can also prove that $\partial f(x) = \partial \bar{f}(x)$.

(c) The assertion follows from (a) and Corollary 2.3.7. □

Theorem 2.3.16. *Let be $f : X \to \overline{\mathbb{R}}$ and $x \in X$.*

(a) If $\partial f(x) \neq \emptyset$, then $f(x) = f^{**}(x)$.
(b) If $f(x) = f^{**}(x)$, then $\partial f(x) = \partial f^{**}(x)$.

Proof. (a) The statement follows directly from Theorem 2.3.15(a), (c).

(b) If $f(x) = f^{**}(x) = \pm\infty$, then by convention we have $\partial f(x) = \partial f^{**}(x) = \emptyset$. Otherwise, Theorem 2.3.12 and Remark 2.3.4 allow to conclude that $x^* \in \partial f(x) \Leftrightarrow f^*(x^*) = -f(x) + \langle x^*, x\rangle \Leftrightarrow f^{***}(x^*) = -f^{**}(x) + \langle x^*, x\rangle \Leftrightarrow x^* \in \partial f^{**}(x)$. □

Our next aim is to point out that the calculation rules which are available for the classical differential can be applied in general only partially to the subdifferential. Using Definition 2.3.2 it is easy to prove that for a given function $f : X \to \overline{\mathbb{R}}$ and $x \in X$ it holds

$$\partial(\lambda f)(x) = \lambda \partial f(x) \text{ for all } \lambda > 0.$$

Coming now to the sum, for some given arbitrary proper functions $f_i : X \to \overline{\mathbb{R}}$, $i = 1, \ldots, m$, one can only prove in general that for $x \in X$ it holds

$$\sum_{i=1}^{m} \partial f_i(x) \subseteq \partial\left(\sum_{i=1}^{m} f_i\right)(x). \tag{2.8}$$

We refer the reader to section 3.5 for sufficient conditions which guarantee, when the functions f_i, $i = 1, \ldots, m$, are convex, equality in (2.8).

The next result displays some connections between the subdifferential of a given function f and the one of its conjugate.

Theorem 2.3.17. *Let be $f : X \to \overline{\mathbb{R}}$ and $x \in X$.*

(a) If $x^* \in \partial f(x)$, then $x \in \partial f^*(x^*)$.
(b) If $f(x) = f^{**}(x)$, then $x^* \in \partial f(x)$ if and only if $x \in \partial f^*(x^*)$.
(c) If f is proper, convex and lower semicontinuous, then $x^* \in \partial f(x)$ if and only if $x \in \partial f^*(x^*)$.

Proof. (a) Since $x^* \in \partial f(x)$, according to Theorem 2.3.12 we have $f(x) + f^*(x^*) = \langle x^*, x\rangle$. But $f^{**}(x) \leq f(x)$, by Lemma 2.3.3, and thus $f^{**}(x) + f^*(x^*) \leq \langle x^*, x\rangle$. As the reverse inequality is always fulfilled, using once more Theorem 2.3.12, we get $x \in \partial f^*(x^*)$.

(b) Because of (a) only the sufficiency must be proven. For any $x \in \partial f^*(x^*)$, again by Theorem 2.3.12, it holds $\langle x^*, x\rangle = f^*(x^*) + f^{**}(x) = f^*(x^*) + f(x)$ and therefore $x^* \in \partial f(x)$.

(c) Theorem 2.3.5 yields $f = f^{**}$ and the equivalence follows from (b). □

A classical assertion on the existence of a subgradient is given in the following statement (cf. [67]).

Theorem 2.3.18. *Let the convex function $f : X \to \overline{\mathbb{R}}$ be finite and continuous at some point $x \in X$. Then $\partial f(x) \neq \emptyset$, i.e. f is subdifferentiable at x.*

Theorem 2.3.18 follows easily as a consequence of the Fenchel duality statement Theorem 3.2.6, which we give in the next chapter. For further results concerning subdifferential calculus we refer to section 3.5 and the book [207].

We conclude this subsection by resuming the relations between the subdifferentiability and the rather classical notion of Gâteaux differentiability accompanied by some further properties of the Gâteaux differential.

Definition 2.3.3. *Let $f : X \to \overline{\mathbb{R}}$ be a proper function and $x \in \operatorname{dom} f$. If the limit*

$$\lim_{t \downarrow 0} \frac{f(x + ty) - f(x)}{t}$$

exists we call it the directional derivative of f at x in the direction $y \in X$ and we denote it by $f'(x; y)$. If there exists an $x^* \in X^*$ such that $f'(x; y) = \langle x^*, y\rangle$ for all $y \in X$, then f is said to be Gâteaux differentiable at x, x^* is called the Gâteaux differential of f at x and it is denoted by $\nabla f(x)$, i.e. $f'(x; y) = \langle \nabla f(x), y\rangle$ for all $y \in X$.

We need to note that if f is proper and convex and $x \in \operatorname{dom} f$ then $f'(x; y)$ exists for all $y \in X$ (cf. [207, Theorem 2.1.12]). If, additionally, f is continuous at $x \in \operatorname{dom} f$ then for all $y \in X$ there is $f'(x; y) = \max\{\langle x^*, y\rangle : x^* \in \partial f(x)\}$. For convex functions the Gâteaux differentiability and the uniqueness of the subgradient are closely related, as stated below.

Proposition 2.3.19. *Let $f : X \to \overline{\mathbb{R}}$ be a proper and convex function and $x \in \operatorname{dom} f$.*

(a) *If* $x \in \operatorname{core}(\operatorname{dom} f)$ *and* f *is Gâteaux differentiable at* x*, then* f *is subdifferentiable at* x *and* $\partial f(x) = \{\nabla f(x)\}$.
(b) *If* f *is continuous at* x *and its subdifferential* $\partial f(x)$ *is a singleton, then* f *is Gâteaux differentiable at* x *and* $\partial f(x) = \{\nabla f(x)\}$.

In the next results the convexity of a Gâteaux differentiable function is characterized.

Proposition 2.3.20. *Let* $U \subseteq X$ *be a nonempty, open and convex set and* $f : U \to \mathbb{R}$ *a Gâteaux differentiable function on* U*. Then the function* f *is convex on* U *if and only if* $f(y) \geq f(x) + \langle \nabla f(x), y - x\rangle$ *for all* $x, y \in U$*. This is further equivalent to* $\langle \nabla f(y) - \nabla f(x), y - x\rangle \geq 0$ *for all* $x, y \in U$.

2.4 Minimal and maximal elements of sets

It is characteristic for *vector optimization problems* that, different to *scalar programming problems*, more than one conflicting objectives have to be taken into consideration. Thus, the objective values may be considered as vectors in a finite or even infinite dimensional vector space. As an example let us mention the *Markowitz portfolio optimization problem* which aims at finding an optimal portfolio of several risky securities, e.g. stocks and shares of companies. There are two reasonable objectives, the *expected return* which has to be maximized, and the *risk* (measured via the variance of the expected return or any other risk measure) that has to be minimized. These both objectives are conflicting because in general the risk grows if the expected return is increasing. This conflict must be reflected by a corresponding partial ordering relation in the two dimensional objective space. Such an ordering relation allows to compare different vector objective values in, at least, a partial sense. Based on the considered partial ordering one can define distinct types of solutions in connection to a vector optimization problem. Partial orderings in the sense considered in this book are defined by convex cones as we have already done in section 2.1. The basic terminology for such solutions is that of *efficiency*, i.e. we consider different types of so-called efficient solutions. For the first time efficient solutions have been considered by Edgeworth in [60] and Pareto in [148].

In this section we present different notions of minimality (maximality) for sets in vector spaces. In the next section these notions will be employed when introducing different efficiency solution concepts for vector optimization problems.

2.4.1 Minimality

Unless otherwise mentioned, in the following we consider V to be a vector space partially ordered by a convex cone $K \subseteq V$.

First of all let us define the usual notion of minimality for a nonempty set $M \subseteq V$ with respect to the partial ordering "$\leqq_K$" induced by K. Initially, we confine ourselves to the case where the ordering cone K is pointed, i.e. $l(K) = \{0\}$, in which case "$\leqq_K$" is antisymmetric, since this is the situation mostly encountered within this book and also in the majority of practical applications of vector optimization.

Definition 2.4.1. *An element $\bar{v} \in M$ is said to be a minimal element of M (regarding the partial ordering induced by K) if there is no $v \in M$ satisfying $v \leq_K \bar{v}$. The set of all minimal elements of M is denoted by $\mathrm{Min}(M, K)$ and it is called the minimal set of M (regarding the partial ordering defined by K).*

Remark 2.4.1. There are several obviously equivalent formulations for an element $\bar{v} \in M$ to be a minimal element of M. We list some of them in the following:

(i) there is no $v \in M$ such that $\bar{v} - v \in K \backslash \{0\}$;
(ii) from $v \leqq_K \bar{v}$, $v \in M$, follows $v = \bar{v}$;
(iii) from $v \leqq_K \bar{v}$, $v \in M$, follows $v \geqq_K \bar{v}$;
(iv) $(\bar{v} - K) \cap M = \{\bar{v}\}$;
(v) $(M - \bar{v}) \cap (-K) = \{0\}$;
(vi) for all $v \in M$ there is $v \nleq_K \bar{v}$.

We would like to underline that for the equivalence $(ii) \Leftrightarrow (iii)$ the pointedness of K is indispensable.

Example 2.4.1. An important case which occurs in practice is when $V = \mathbb{R}^k$ and $K = \mathbb{R}^k_+$. Let $M \subseteq \mathbb{R}^k$. Then $\bar{v} = (\bar{v}_1, \ldots, \bar{v}_k)^T \in M$ is a minimal element of M if there is no $v = (v_1, \ldots, v_k)^T \in M$ such that $v \neq \bar{v}$ and $v_i \leq \bar{v}_i$ for all $i = 1, \ldots, k$, i.e. there is no $v = (v_1, \ldots, v_k)^T \in M$ fulfilling $v_i \leq \bar{v}_i$ for all $i = 1, \ldots, k$ and $v_j < \bar{v}_j$ for at least one $j \in \{1, \ldots, k\}$.

In an analogous way one can define the notion of *maximal element* of a set M.

Definition 2.4.2. *An element $\bar{v} \in M$ is said to be a maximal element of M (regarding the partial ordering induced by K) if there is no $v \in M$ satisfying $v \geq_K \bar{v}$. The set of all maximal elements of M is denoted by $\mathrm{Max}(M, K)$ and it is called the maximal set of M (regarding the partial ordering defined by K).*

Remark 2.4.2. As in Remark 2.4.1 one can give the following equivalent formulations for the maximality of an element $\bar{v} \in M$ in M:

(i) there is no $v \in M$ such that $v - \bar{v} \in K \backslash \{0\}$;
(ii) from $v \geqq_K \bar{v}$, $v \in M$, follows $v = \bar{v}$;
(iii) from $v \geqq_K \bar{v}$, $v \in M$, follows $v \leqq_K \bar{v}$;
(iv) $(\bar{v} + K) \cap M = \{\bar{v}\}$;

(v) $(M - \bar{v}) \cap K = \{0\}$;
(vi) for all $v \in M$ there is $v \ngeqq_K \bar{v}$.

Remark 2.4.3. The problem of finding the maximal elements of the set M regarding the cone K may be reformulated as the problem of finding the minimal elements of the set $(-M)$ regarding K or, equivalently, as the problem of finding the minimal elements of the set M regarding the partial ordering induced by the cone $(-K)$. It holds $\operatorname{Max}(M, K) = \operatorname{Min}(M, -K) = -\operatorname{Min}(-M, K)$.

Although we mostly confine ourselves within this book to the most important framework of partial orderings induced by pointed convex cones, for the sake of completeness we present also the definition of minimality regarding a partial ordering induced by a convex but not pointed cone K. As noted in subsection 2.1.1, in this situation $l(K) = K \cap (-K)$ is a linear subspace of V not equal to $\{0\}$. If this situation occurs, then Definition 2.4.1 is not always suitable for defining minimal elements, and this because it may happen to have a $v \in M$, $v \neq \bar{v}$, such that $v \leqq_K \bar{v} \leqq_K v$. More precisely, if $\bar{v} - v \in l(K) \subseteq K$ then $v - \bar{v} \in l(K) \subseteq K$, too, and now it is clear that Definition 2.4.1 cannot be used if the ordering cone K is not pointed. This situation can be avoided if instead of Definition 2.4.1 one uses the following definition due to Borwein [19] (see also Remark 2.4.1(*iii*)).

Definition 2.4.3. *Let $K \subseteq V$ be an arbitrary ordering cone. An element $\bar{v} \in M$ is said to be a minimal element of M (regarding the partial ordering induced by K), if from $v \leqq_K \bar{v}$, $v \in M$, follows $v \geqq_K \bar{v}$.*

From this definition immediately follows that if $\bar{v} \in M$ is a minimal element of M then any $\tilde{v} \in M$ such that $\tilde{v} \leqq_K \bar{v}$ is also a minimal element of M. To see this take an arbitrary $v \in M$ such that $v \leqq_K \tilde{v}$. Since $\bar{v} \in \operatorname{Min}(M, K)$, it holds $v \geqq_K \bar{v} \geqq_K \tilde{v}$ and so $\tilde{v} \in \operatorname{Min}(M, K)$.

We observe further that $\bar{v}$ being minimal means that for all $v \in M$ fulfilling $v \leqq_K \bar{v}$ it is binding to have $\bar{v} - v \in l(K)$. If K is a pointed cone then $l(K) = \{0\}$ and in this case we have $\bar{v} = v$. Therefore Definition 2.4.3 applies to pointed cones K, too, while Definition 2.4.1 can be seen as a particular case of it.

Next we give some equivalent formulations to the notion of minimality in case the cone K is not assumed to be pointed. More precisely, $\bar{v} \in M$ is a minimal element of M if and only if one of the following conditions is fulfilled:

(i) there is no $v \in M$ such that $\bar{v} - v \in K \setminus l(K)$;
(ii) $(\bar{v} - K) \cap M \subseteq \bar{v} + K$;
(iii) $(-K) \cap (M - \bar{v}) \subseteq K$.

The *maximal elements* of the set M (in case the cone K is not assumed pointed) can be defined following the same idea as in Definition 2.4.3. Analogously to Remark 2.4.3 one has $\operatorname{Max}(M, K) = \operatorname{Min}(M, -K) = -\operatorname{Min}(-M, K)$.

The next result describes the relation between the minimal elements of the sets M and $M + K$.

Lemma 2.4.1. *(a) It holds* $\operatorname{Min}(M, K) \subseteq \operatorname{Min}(M + K, K)$*.*
(b) If K *is pointed, then* $\operatorname{Min}(M, K) = \operatorname{Min}(M + K, K)$*.*

Proof. (a) Take an arbitrary $\bar{v} \in \operatorname{Min}(M, K)$. By definition $\bar{v} \in M \subseteq M + K$. Let us prove now that for $v \in M + K$ such that $v \leqq_K \bar{v}$ the relation $\bar{v} \leqq_K v$ holds, too. Since $v \in M + K$ we have $v = \tilde{v} + k$ for some $\tilde{v} \in M$ and $k \in K$. Obviously $\tilde{v} = v - k \leqq_K \bar{v} - k \leqq_K \bar{v}$. But the minimality of $\bar{v}$ secures $\bar{v} \leqq_K \tilde{v}$ and, since $\tilde{v} = v - k \leqq_K v$, the desired conclusion follows.

(b) Assuming now K pointed, let $\bar{v} \in \operatorname{Min}(M+K, K)$. We have $\bar{v} \in M+K$ and we show that actually $\bar{v} \in M$. Assuming the contrary implies $\bar{v} = \tilde{v} + k$ with $\tilde{v} \in M$ and $k \in K \setminus \{0\}$. This yields $\tilde{v} \leq_K \bar{v}$ and as $\tilde{v} \in M+K$, one would get a contradiction to the minimality of $\bar{v}$ in $M + K$. Following a similar reasoning one can prove that in fact $\bar{v} \in \operatorname{Min}(M, K)$. Now (a) yields the desired conclusion. □

The next minimality notion we introduce is the so-called *strong minimality.* We work in the same setting, with V a vector space partially ordered by the (not necessarily pointed) convex cone K and M a nonempty subset of V.

Definition 2.4.4. *An element* $\bar{v} \in M$ *is said to be a strongly minimal element of* M *(regarding the partial ordering induced by* K*) if* $\bar{v} \leqq_K v$ *for all* $v \in M$*, i.e.* $M \subseteq \bar{v} + K$*.*

For vector optimization this definition is of secondary importance because in the most practical cases strongly minimal elements do not exist. If we consider the classical situation when $V = \mathbb{R}^k$ and $K = \mathbb{R}^k_+$, then the strong minimality of $\bar{v} \in M \subseteq \mathbb{R}^k$ means $\bar{v}_i \leq v_i$, $i = 1, \ldots, k$, for all $v = (v_1, \ldots, v_k)^T \in M$. Thus, in case of a multiobjective optimization problem, this must imply that all the k components of its objective function attain their minima at the same point, i.e. the objectives are not conflicting as it is typical for vector optimization.

Obviously, every strongly minimal element is minimal. A *strongly maximal element* $\bar{v} \in M$ is defined in analogous manner, namely one must have $\bar{v} \geqq_K v$ for all $v \in M$.

2.4.2 Weak minimality

Although from the practical point of view not so important as the minimal elements, the so-called *weakly minimal elements* of a given set are of theoretical interest, one of the arguments sustaining this assertion being that they allow a complete characterization by linear scalarization in the convex case, which is not always possible with minimal elements. We consider in this subsection V to be a vector space partially ordered by the convex cone $K \subseteq V$ fulfilling $\operatorname{core}(K) \neq \emptyset$ and $M \subseteq V$ being a nonempty set.

Definition 2.4.5. *An element $\bar{v} \in M$ is said to be a weakly minimal element of M (regarding the partial ordering induced by K) if $(\bar{v} - \operatorname{core}(K)) \cap M = \emptyset$. The set of all weakly minimal elements of M is denoted by* $\operatorname{WMin}(M, K)$ *and it is called the weakly minimal set of M (regarding the partial ordering induced by K).*

The relation $(\bar{v} - \operatorname{core}(K)) \cap M = \emptyset$ in Definition 2.4.5 is obviously equivalent to $(M - \bar{v}) \cap (-\operatorname{core}(K)) = \emptyset$. From here follows that $\operatorname{WMin}(M, V) = \emptyset$. Whenever the cone K is nontrivial one may also notice that if we consider as ordering cone $\widehat{K} = \operatorname{core}(K) \cup \{0\}$, then $\bar{v} \in \operatorname{WMin}(M, K)$ if and only if $(\bar{v} - \widehat{K}) \cap M = \{\bar{v}\}$, or, equivalently, $\bar{v} \in \operatorname{Min}(M, \widehat{K})$ (see Remark 2.4.1(iv)). Of course, if $K = \widehat{K}$, then the minimal and weakly minimal elements of M regarding the partial ordering induced by K coincide. This, however, is not the case in general.

If $K \neq V$, any minimal element of M is also weakly minimal, since $(\bar{v} - K) \cap M \subseteq \bar{v} + K$ implies $(\bar{v} - \operatorname{core}(K)) \cap M = \emptyset$. Indeed, notice that $(\bar{v} - \operatorname{core}(K)) \cap (\bar{v} + K) = \emptyset$, as in this situation $-\operatorname{core}(K) \cap K = \emptyset$. This result is summarized in the following statement.

Proposition 2.4.2. *If $K \neq V$, then* $\operatorname{Min}(M, K) \subseteq \operatorname{WMin}(M, K)$.

Next we provide a result, similar to Lemma 2.4.1, for the weakly minimal elements of a set $M \subseteq V$, the proof of which being relinquished to the reader (see, for instance, [104, Lemma 4.13]).

Lemma 2.4.3. *It holds*

(a) $\operatorname{WMin}(M, K) \subseteq \operatorname{WMin}(M + K, K)$;
(b) $\operatorname{WMin}(M + K, K) \cap M \subseteq \operatorname{WMin}(M, K)$.

Remark 2.4.4. When V is taken to be a topological vector space, in the above assertions the algebraic interior $\operatorname{core}(K)$ can be replaced with the topological interior $\operatorname{int}(K)$ when the latter is nonempty.

Weakly maximal elements may be defined analogously, namely an element $\bar{v} \in M$ is called a weakly maximal element of M (regarding the partial ordering induced by K) if $(\bar{v} + \operatorname{core}(K)) \cap M = \emptyset$. The set of all weakly maximal elements of M is denoted by $\operatorname{WMax}(M, K)$ and it is called the *weakly maximal set* of M. Also here it holds $\operatorname{WMax}(M, K) = \operatorname{WMin}(M, -K) = -\operatorname{WMin}(-M, K)$.

Consequently, one can formulate obvious variants of Lemma 2.4.1, Proposition 2.4.2 and Lemma 2.4.3 for maximal and weakly maximal elements of the set M, respectively.

2.4.3 Proper minimality

There is another very important notion of minimality subsumed under the category of *properly minimal elements*. Properly minimal elements turn out to be minimal elements with additional properties. There are a lot of different kinds of properly minimal elements. In the following we present an overview of their distinct definitions and establish some relations between them. The majority of these notions have been introduced in connection to some vector optimization problems under the name *proper efficiency*. We introduce them in this subsection as proper minimality notions for a given set by extending to this general situation the corresponding notions for vector optimization problems. To this end one has only to take in this situation M to be the image set of the feasible set through the objective function. Let us mention also that here we work only with properly minimal elements, for considering properly maximal elements one needs only replace the cone K by $-K$.

The first notion we present concerns a nonempty set $M \subseteq \mathbb{R}^k$ when the space $\mathbb{R}^k$ is partially ordered by the cone $K = \mathbb{R}^k_+$, being inspired by Geoffrion's paper [71].

Definition 2.4.6. *An element $\bar{v} = (\bar{v}_1, \ldots, \bar{v}_k)^T \in M$ is said to be a properly minimal element of M in the sense of Geoffrion if $\bar{v} \in \operatorname{Min}(M, \mathbb{R}^k_+)$ and if there exists a real number $N > 0$ such that for every $i \in \{1, \ldots, k\}$ and every $v = (v_1, \ldots, v_k)^T \in M$ satisfying $v_i < \bar{v}_i$ there exists at least one $j \in \{1, \ldots, k\}$ such that $\bar{v}_j < v_j$ and*

$$\frac{\bar{v}_i - v_i}{v_j - \bar{v}_j} \leq N.$$

The set of all properly minimal elements of M in the sense of Geoffrion is denoted by $\operatorname{PMin}_{Ge}(M, \mathbb{R}^k_+)$.

The definition above can be interpreted as follows: a decrease in one component relative to $\bar{v}$ entails an increase in at least another component such that the ratio of the absolute values of those differences is bounded. In multiobjective optimization this means that the trade-offs among the different components of the vector objective function are bounded. In economics unbounded trade-offs are mostly undesirable. However, not only with respect to the practical applications but also from the theoretical point of view, properly minimal (or maximal) elements have nice and beneficial properties as we will see in the next subsection.

One can establish an analogous lemma to Lemma 2.4.1 and Lemma 2.4.3 by replacing M with $M + \mathbb{R}^k_+$.

Lemma 2.4.4. *There is* $\operatorname{PMin}_{Ge}(M + \mathbb{R}^k_+, \mathbb{R}^k_+) = \operatorname{PMin}_{Ge}(M, \mathbb{R}^k_+)$.

Proof. Take an arbitrary $\bar{v} \in \operatorname{PMin}_{Ge}(M + \mathbb{R}^k_+, \mathbb{R}^k_+)$. By definition $\bar{v} \in \operatorname{Min}(M + \mathbb{R}^k_+, \mathbb{R}^k_+)$ and, by Lemma 2.4.1, it holds $\bar{v} \in \operatorname{Min}(M, \mathbb{R}^k_+)$. This implies, in particular, that $\bar{v} \in M$. Because $M \subseteq M + \mathbb{R}^k_+$, Definition 2.4.6 applies to any $v \in M$, i.e. $\bar{v} \in \operatorname{PMin}_{Ge}(M, \mathbb{R}^k_+)$.

Now let $\bar{v} \in \operatorname{PMin}_{Ge}(M, \mathbb{R}^k_+)$ and $N > 0$ be the positive constant provided by Definition 2.4.6. Clearly $\bar{v} \in M + \mathbb{R}^k_+$ and by Lemma 2.4.1, we get $\bar{v} \in \operatorname{Min}(M, \mathbb{R}^k_+) = \operatorname{Min}(M + \mathbb{R}^k_+, \mathbb{R}^k_+)$. Take an arbitrary $v = \tilde{v} + h \in M + \mathbb{R}^k_+$

where $\tilde{v} \in M$ and $h \in \mathbb{R}^k_+$ and $i \in \{1, \dots, k\}$ such that $v_i = \tilde{v}_i + h_i < \bar{v}_i$. Obviously $\tilde{v}_i < \bar{v}_i$ and, according to Definition 2.4.6, there exists at least one $j \in \{1, \dots, k\}$ such that $\bar{v}_j < \tilde{v}_j$ and

$$\frac{\bar{v}_i - \tilde{v}_i}{\tilde{v}_j - \bar{v}_j} \leq N.$$

But this implies $\bar{v}_j < \tilde{v}_j + h_j = v_j$ and

$$\frac{\bar{v}_i - v_i}{v_j - \bar{v}_j} = \frac{\bar{v}_i - \tilde{v}_i - h_i}{\tilde{v}_j + h_j - \bar{v}_j} \leq \frac{\bar{v}_i - \tilde{v}_i}{\tilde{v}_j - \bar{v}_j} \leq N.$$

Consequently, we get $\bar{v} \in \mathrm{PMin}_{Ge}(M + \mathbb{R}^k_+, \mathbb{R}^k_+)$. □

Already ten years earlier Hurwicz [93] has introduced a notion of proper efficiency for vector optimization problems which has been generalized in [83] to sets in partially ordered topological vector spaces.

Further we assume that V is a topological vector space partially ordered by the pointed convex cone K and $M \subseteq V$ is an arbitrary nonempty set.

Definition 2.4.7. *An element* $\bar{v} \in M$ *is said to be a properly minimal element of* M *in the sense of Hurwicz if* $\mathrm{cl}(\mathrm{coneco}((M - \bar{v}) \cup K)) \cap (-K) = \{0\}$. *The set of all properly minimal elements of* M *in the sense of Hurwicz is denoted by* $\mathrm{PMin}_{Hu}(M, K)$.

This definition seems to be in a certain manner natural if it is compared with the definition of minimality of $\bar{v} \in M$ in the equivalent formulation given in Remark 2.4.1(v) which states that $(M - \bar{v}) \cap (-K) = \{0\}$. Because $M - \bar{v} \subseteq \mathrm{cl}(\mathrm{coneco}((M - \bar{v}) \cup K))$ it is clear that $\mathrm{PMin}_{Hu}(M, K) \subseteq \mathrm{Min}(M, K)$.

Lemma 2.4.5. *There is* $\mathrm{PMin}_{Hu}(M + K, K) = \mathrm{PMin}_{Hu}(M, K)$.

Proof. Noting that $\mathrm{coneco}((M - \bar{v}) \cup K) = \mathrm{coneco}((M + K - \bar{v}) \cup K)$, we obtain the conclusion. □

Geoffrion's definition of proper efficiency is very illustrative concerning economical and geometrical aspects. But its drawback is the restriction to the ordering cone $K = \mathbb{R}^k_+$. To overcome this disadvantage Borwein proposed in [17] a notion of proper efficiency for vector maximization problems given in Hausdorff locally convex spaces partially ordered by a pointed convex closed cone which generalizes Geoffrion's definition. We employ Borwein's definition to sets and use the notion proper minimality instead of proper efficiency in accordance with our general context.

For the remaining part of this subsection we take V to be a Hausdorff locally convex space partially ordered by the pointed convex cone K and $M \subseteq V$ an arbitrary nonempty set.

Definition 2.4.8. *An element $\bar{v} \in M$ is said to be a properly minimal element of M in the sense of Borwein if* $\text{cl}(T(M+K,\bar{v})) \cap (-K) = \{0\}$*. The set of all properly minimal elements of M in the sense of Borwein is denoted by* $\text{PMin}_{Bo}(M,K)$*.*

Remark 2.4.5. The proper minimality in the sense of Borwein can be equivalently written as $0 \in \text{Min}(\text{cl}(T(M+K,\bar{v})),K)$. Observing that $T(M+K,\bar{v}) = T(M+K-\bar{v},0)$ one can see the affinity of this kind of proper minimality to the notion of minimality, via Remark 2.4.1(v) and Lemma 2.4.1. The element $\bar{v} \in M$ is minimal if and only if $(M+K-\bar{v}) \cap (-K) = \{\bar{v}\}$. Thus Borwein's definition of proper minimality is nothing else than additionally demanding $\text{cl}(T(M+K-\bar{v},0)) \cap (-K) = \{0\}$. Moreover, if V is metrizable, then the tangent cone is closed and in this situation one may omit the closure operation within Definition 2.4.8.

Remark 2.4.6. (a) Let us mention here that in the original definition in [17] the ordering cone was not explicitly assumed to be pointed. But this has to be assumed, otherwise $\text{PMin}_{Bo}(M,K)$ is the empty set. Indeed, if K is not pointed, then let $v \in K \cap (-K)$, $v \neq 0$, be arbitrarily chosen and $\bar{v} \in \text{PMin}_{Bo}(M,K)$. Setting $v_l = \bar{v} + (1/l)v \in M+K$ for $l \geq 1$, we get $\lim_{l\to+\infty} v_l = \bar{v}$ and $\lim_{l\to\infty} l(v_l - \bar{v}) = v$. But this means nothing else than $v \in T(M+K,\bar{v})$. Therefore $\text{cl}(T(M+K,\bar{v})) \cap (-K) \neq \{0\}$ and this means that in this situation $\text{PMin}_{Bo}(M,K) = \emptyset$.

(b) A second observation in this context is that in the original definition for a vector maximization problem a properly efficient solution is, additionally, assumed to be an efficient solution. This would mean to require in our definition that $\bar{v} \in \text{Min}(M,K)$. But this hypothesis is superfluous and turns out to be a consequence of the condition $\text{cl}(T(M+K,\bar{v})) \cap (-K) = \{0\}$. The conclusion is obvious if $K = \{0\}$. Assume that $K \neq \{0\}$ and that for a $\bar{v} \in \text{PMin}_{Bo}(M,K)$ it holds $\bar{v} \notin \text{Min}(M,K)$. Then there exists $v \in M$ such that $\bar{v} - v \in K \backslash \{0\}$. We show that $v - \bar{v} \in T(M+K,\bar{v})$. Setting $v_l = \bar{v} + (1/l)(v - \bar{v})$, $l \geq 1$, we can easily see that $v_l = v + ((l-1)/l)(\bar{v} - v) \in M+K$ for $l \geq 1$. Even more, as $\lim_{l\to+\infty} v_l = \bar{v}$ and $\lim_{l\to+\infty} l(v_l - \bar{v}) = v - \bar{v}$, it follows $v - \bar{v} \in T(M+K,\bar{v})$. Finally, since $0 \neq v - \bar{v} \in T(M+K,\bar{v}) \cap (-K) \subseteq \text{cl}(T(M+K,\bar{v})) \cap (-K)$, the equality $\text{cl}(T(M+K,\bar{v})) \cap (-K) = \{0\}$ fails, and this contradicts the assumption $\bar{v} \in \text{PMin}_{Bo}(M,K)$.

Since by the convexity of K it holds $(M+K)+K = M+K$, the following result follows easily via Lemma 2.4.1 and Remark 2.4.6(b).

Lemma 2.4.6. *There is* $\text{PMin}_{Bo}(M+K,K) = \text{PMin}_{Bo}(M,K)$*.*

Regarding the proper minimality in the sense of Geoffrion and in the sense of Borwein in case $V = \mathbb{R}^k$ and $K = \mathbb{R}^k_+$, we have that for a nonempty set $M \subseteq \mathbb{R}^k$ it holds $\text{PMin}_{Ge}(M,\mathbb{R}^k_+) \subseteq \text{PMin}_{Bo}(M,\mathbb{R}^k_+)$, which can be proven similarly as in [17, Proposition 1]. The following result, giving a sufficient condition for the coincidence of both notions, can also be proven similarly as for a corresponding assertion in [17] regarding a vector maximization problem under convexity assumptions.

Proposition 2.4.7. *If $M \subseteq \mathbb{R}^k$ is nonempty and $M + \mathbb{R}^k_+$ is convex, then* $\text{PMin}_{Ge}(M, \mathbb{R}^k_+) = \text{PMin}_{Bo}(M, \mathbb{R}^k_+)$.

The next proper minimality notion we consider here originates from Benson's paper [15] and it was introduced in order to extend Geoffrion's proper minimality.

Definition 2.4.9. *An element $\bar{v} \in M$ is said to be a properly minimal element of M in the sense of Benson if* $\text{cl}(\text{cone}(M + K - \bar{v})) \cap (-K) = \{0\}$. *The set of all properly minimal elements of M in the sense of Benson is denoted by* $\text{PMin}_{Be}(M, K)$.

Remark 2.4.7. In [15] the notion introduced above was given as proper efficiency for vector maximum problems in finite dimensional spaces, with the efficiency of the elements in discussion additionally assumed. This means in our situation to supplementary impose the condition $\bar{v} \in \text{Min}(M, K)$. But this is superfluous since $M - \bar{v} \subseteq \text{cl}(\text{cone}(M + K - \bar{v}))$ implies $\bar{v} \in \text{Min}(M, K)$ if $\text{cl}(\text{cone}(M + K - \bar{v})) \cap (-K) = \{0\}$, i.e. $(M - \bar{v}) \cap (-K) = \{0\}$, too.

The next result is a consequence of the convexity of K along with Lemma 2.4.1 and Remark 2.4.7.

Lemma 2.4.8. *There is* $\text{PMin}_{Be}(M + K, K) = \text{PMin}_{Be}(M, K)$.

As mentioned above, for $V = \mathbb{R}^k$ and $K = \mathbb{R}^k_+$, when $M \subseteq \mathbb{R}^k$ is an arbitrary nonempty set, it holds (cf. [15]) $\text{PMin}_{Ge}(M, \mathbb{R}^k_+) = \text{PMin}_{Be}(M, \mathbb{R}^k_+)$. By taking into consideration the way Borwein's and Benson's proper minimalities are defined, one has that $\text{PMin}_{Be}(M, K) \subseteq \text{PMin}_{Bo}(M, K)$ is always fulfilled. Further, let us notice that for $\bar{v} \in M$ it holds $\text{cone}(M + K - \bar{v}) \subseteq \text{coneco}((M - \bar{v}) \cup K)$, the two sets being equal if $M + K$ is convex. Thus we have in general that $\text{PMin}_{Hu}(M, K) \subseteq \text{PMin}_{Be}(M, K)$, while when $M + K$ is convex it follows that $\text{PMin}_{Hu}(M, K) = \text{PMin}_{Be}(M, K) = \text{PMin}_{Bo}(M, K)$.

In the following we introduce another proper minimality concept due to Borwein (cf. [18]), which is similar to Definition 2.4.9.

Definition 2.4.10. *An element $\bar{v} \in M$ is said to be a properly minimal element of M in the global sense of Borwein if* $\text{cl}(\text{cone}(M - \bar{v})) \cap (-K) = \{0\}$. *The set of all properly minimal elements of M in the global sense of Borwein is denoted by* $\text{PMin}_{GBo}(M, K)$.

If for $\bar{v} \in M$ Definition 2.4.10 is satisfied, then also $(M - \bar{v}) \cap (-K) = \{0\}$ and, due to Remark 2.4.1(v), we have $\bar{v} \in \text{Min}(M, K)$. Furthermore, it is always true that $\text{PMin}_{Be}(M, K) = \text{PMin}_{GBo}(M + K, K)$.

The next result relates $\text{PMin}_{GBo}(M + K, K)$ to $\text{PMin}_{GBo}(M, K)$.

Proposition 2.4.9. *There is* $\mathrm{PMin}_{GBo}(M+K,K) \subseteq \mathrm{PMin}_{GBo}(M,K)$.

Proof. Take an arbitrary $\bar{v} \in \mathrm{PMin}_{GBo}(M+K,K)$. Thus $\bar{v} \in M$ and $\mathrm{cl}(\mathrm{cone}(M+K-\bar{v})) \cap (-K) = \{0\}$. Since $M-\bar{v} \subseteq M+K-\bar{v}$ it follows that $\mathrm{cl}(\mathrm{cone}(M-\bar{v})) \cap (-K) \subseteq \mathrm{cl}(\mathrm{cone}(M+K-\bar{v})) \cap (-K) = \{0\}$ and, consequently, $\mathrm{cl}(\mathrm{cone}(M-\bar{v})) \cap (-K) = \{0\}$, which completes the proof. □

Obviously, we have that $\mathrm{PMin}_{Be}(M,K) \subseteq \mathrm{PMin}_{GBo}(M,K)$. On the other hand, no relation of inclusion between $\mathrm{PMin}_{Bo}(M,K)$ and $\mathrm{PMin}_{GBo}(M,K)$ can be given in general. Obviously, when $M+K$ is convex, then $\mathrm{PMin}_{Bo}(M,K) \subseteq \mathrm{PMin}_{GBo}(M,K)$.

A formally different minimality approach is the one introduced by Henig [88] and Lampe [122] by employing a nontrivial convex cone K' containing in its interior the given ordering cone K.

Definition 2.4.11. *An element* $\bar{v} \in M$ *is said to be a properly minimal element of* M *in the sense of Henig and Lampe if there exists a nontrivial convex cone* $K' \subseteq X$ *with* $K\backslash\{0\} \subseteq \mathrm{int}(K')$ *such that* $(M-\bar{v}) \cap (-K') = \{0\}$. *The set of all properly minimal elements of* M *in the sense of Henig and Lampe is denoted by* $\mathrm{PMin}_{He-La}(M,K)$.

If the cone K' is assumed also pointed, instead of $(M-\bar{v}) \cap (-K') = \{0\}$ one can write $\bar{v} \in \mathrm{Min}(M,K')$. It is an immediate consequence of this definition that $\bar{v} \in \mathrm{PMin}_{He-La}(M,K)$ implies $\bar{v} \in \mathrm{Min}(M,K)$.

Lemma 2.4.10. *There is* $\mathrm{PMin}_{He-La}(M+K,K) = \mathrm{PMin}_{He-La}(M,K)$.

Proof. Let $\bar{v} \in \mathrm{PMin}_{He-La}(M+K,K)$ be arbitrarily taken. Then one can easily show that $\bar{v} \in M$. Moreover, there exists a nontrivial convex cone K' such that $K\backslash\{0\} \subseteq \mathrm{int}(K')$ and $(M+K-\bar{v}) \cap (-K') = \{0\}$. Thus $(M-\bar{v}) \cap (-K') = \{0\}$ and so $\bar{v} \in \mathrm{PMin}_{He-La}(M,K)$.

Vice versa, take $\bar{v} \in \mathrm{PMin}_{He-La}(M,K)$. Then $\bar{v} \in M \subseteq M+K$ and there exists a nontrivial convex cone K' such that $K\backslash\{0\} \subseteq \mathrm{int}(K')$ and $(M-\bar{v}) \cap (-K') = \{0\}$. We prove that $(M+K-\bar{v}) \cap (-K') = \{0\}$. If we assume the contrary there would exist $v \in M$, $k \in K$ and $k' \in K'\backslash\{0\}$ such that $v+k-\bar{v} = -k'$. Then $v-\bar{v} = -(k+k') \in -(K+(K'\backslash\{0\})) \subseteq -(K' + (K'\backslash\{0\})) \subseteq -K'$ and so $k+k'=0$. Consequently, $-k \in K' \cap (-\mathrm{int}(K'))$, which leads to a contradiction. Therefore $\bar{v} \in \mathrm{PMin}_{He-La}(M+K,K)$ and the proof is complete. □

The following statement reveals the relation between the proper minimality notions in the sense of Benson and in the sense of Henig and Lampe.

Proposition 2.4.11. *There is* $\mathrm{PMin}_{He-La}(M,K) \subseteq \mathrm{PMin}_{Be}(M,K)$.

Proof. If $K=\{0\}$ the inclusion follows automatically. Assume that $K \neq \{0\}$. Let $\bar{v} \in \mathrm{PMin}_{He-La}(M,K)$. Then $\bar{v} \in M$ and there exists a nontrivial convex cone K' such that $K\backslash\{0\} \subseteq \mathrm{int}(K')$ and $(M-\bar{v}) \cap (-K') = \{0\}$. We prove that $\mathrm{cl}(\mathrm{cone}(M+K-\bar{v}))\cap(-K)=\{0\}$. To this end we assume the contrary, namely that there exists a $k\in K\backslash\{0\}$ such that $-k\in\mathrm{cl}(\mathrm{cone}(M+K-\bar{v}))$. Thus $-k\in\mathrm{int}(-K')$ and consequently there exist $\tilde{v}\in M$, $\tilde{k}\in K$ and $\tilde{\lambda}\geq 0$ such that $\tilde{\lambda}(\tilde{v}+\tilde{k}-\bar{v})\in\mathrm{int}(-K')$. Obviously, $\tilde{\lambda}\neq 0$ and so $\tilde{v}-\bar{v}\in -\mathrm{int}(K')-K'\subseteq -\mathrm{int}(K')$. This yields that $\tilde{v}-\bar{v}\neq 0$, contradicting the fact that $(M-\bar{v})\cap(-K')=\{0\}$. □

The following proper minimality notion, based on linear scalarization, allows the treatment of minimal elements as solutions of scalar optimization problems.

Definition 2.4.12. *An element $\bar{v}\in M$ is said to be a properly minimal element of M in the sense of linear scalarization if there exists a $v^*\in K^{*0}$ such that $\langle v^*,\bar{v}\rangle\leq\langle v^*,v\rangle$ for all $v\in M$. The set of properly minimal elements of M in the sense of linear scalarization is denoted by $\mathrm{PMin}_{LSc}(M,K)$.*

The properly minimal elements of M in the sense of linear scalarization are also minimal, as the next result shows.

Proposition 2.4.12. *There is $\mathrm{PMin}_{LSc}(M,K)\subseteq\mathrm{Min}(M,K)$.*

Proof. Take $\bar{v}\in\mathrm{PMin}_{LSc}(M,K)$ with the corresponding $v^*\in K^{*0}$. If $\bar{v}\notin\mathrm{Min}(M,K)$, then there exists $v\in M$ satisfying $v\leq_K\bar{v}$. As $v^*\in K^{*0}$ and $\bar{v}-v\in K\backslash\{0\}$, there is $\langle v^*,\bar{v}-v\rangle>0$, contradicting Definition 2.4.12. □

Simple examples illustrating that the opposite inclusion is in general not fulfilled can be found in [80]. Without any additional assumption on M, the properly minimal elements in the sense of linear scalarization of M and $M+K$ coincide. The simple proof of this assertion is left to the reader.

Lemma 2.4.13. *There is $\mathrm{PMin}_{LSc}(M+K,K)=\mathrm{PMin}_{LSc}(M,K)$.*

The connection between the properly minimal elements in the sense of linear scalarization and the properly minimal elements in the sense of Hurwicz and Henig and Lampe, respectively, is outlined in the following statements.

Proposition 2.4.14. *There is $\mathrm{PMin}_{LSc}(M,K)\subseteq\mathrm{PMin}_{Hu}(M,K)$.*

Proof. If $K=\{0\}$ there is nothing to be proven. Assume that $K\neq\{0\}$ and take $\bar{v}\in\mathrm{PMin}_{LSc}(M,K)$. Then $\bar{v}\in M$ and there exists $v^*\in K^{*0}$ such that $\langle v^*,v\rangle\geq 0$ for all $v\in M-\bar{v}$. This means that for all $v\in(M-\bar{v})\cup K$ it holds $\langle v^*,v\rangle\geq 0$ and, consequently, $\langle v^*,v\rangle\geq 0$ for all $v\in\mathrm{cl}(\mathrm{coneco}(M-\bar{v})\cup K)$.

Assuming that there exists a $k\in K\setminus\{0\}$ such that $-k\in\mathrm{cl}(\mathrm{coneco}(M-\bar{v})\cup K)$, we get $\langle v^*,k\rangle\leq 0$. On the other hand, since $v^*\in K^{*0}$ there is $\langle v^*,k\rangle>0$. □

Proposition 2.4.15. *There is $\mathrm{PMin}_{LSc}(M,K)\subseteq\mathrm{PMin}_{He-La}(M,K)$.*

Proof. Take an abitrary $\bar{v} \in \mathrm{PMin}_{LSc}(M, K)$. Then $\bar{v} \in M$ and there exists $v^* \in K^{*0}$ such that $\langle v^*, v \rangle \geq 0$ for all $v \in M - \bar{v}$. Let be $K' := \{v \in X : \langle v^*, v \rangle > 0\} \cup \{0\}$. Obviously, K' is a nontrivial convex cone and, since $v^* \in K^{*0}$, $K \backslash \{0\} \subseteq \mathrm{int}(K') = \{v \in X : \langle v^*, v \rangle > 0\}$.

We prove that $(M - \bar{v}) \cap (-K') = \{0\}$. Assuming the contrary yields that there exists $v \in M$ such that $\bar{v} - v \in K' \backslash \{0\}$, or, equivalently, $\langle v^*, \bar{v} - v \rangle > 0$. This contradicts the fact that $\langle v^*, \bar{v} \rangle \leq \langle v^*, v \rangle$ for all $v \in M$. □

Remark 2.4.8. In case the set $M + K$ is convex Lemma 2.4.13 allows to characterize the properly minimal elements $\bar{v} \in \mathrm{PMin}_{LSc}(M, K)$ as solutions of the scalar convex optimization problem

$$\min_{v \in M+K} \langle v^*, v \rangle$$

with an appropriate $v^* \in K^{*0}$. This again makes it possible to derive necessary and sufficient optimality conditions via scalar duality and also to construct vector dual problems in particular in the case when M is the image set of a feasible set through the objective function of a given vector optimization problem.

Summarizing the results proven above, we come to the following general scheme for the proper minimal sets introduced in this section. First we consider the general situation of an underlying Hausdorff locally convex space V partially ordered by the pointed convex cone K and let $M \subseteq V$ be a nonempty set.

Proposition 2.4.16. *There holds*

$$\mathrm{PMin}_{LSc}(M,K) \subseteq \begin{matrix} \mathrm{PMin}_{Hu}(M,K) \\ \mathrm{PMin}_{He-La}(M,K) \end{matrix} \subseteq \mathrm{PMin}_{Be}(M,K) \subseteq \begin{matrix} \mathrm{PMin}_{GBo}(M,K) \\ \mathrm{PMin}_{Bo}(M,K). \end{matrix}$$

If $M + K$ is convex, then

$$\mathrm{PMin}_{LSc}(M,K) \subseteq \mathrm{PMin}_{He-La}(M,K) \subseteq \mathrm{PMin}_{Hu}(M,K)$$

$$= \mathrm{PMin}_{Be}(M,K) = \mathrm{PMin}_{Bo}(M,K) \subseteq \mathrm{PMin}_{GBo}(M,K).$$

Under additional hypotheses some opposite inclusions hold, too. We begin with a statement that can be proven by considering some results from [83].

Proposition 2.4.17. *(a) If the ordering cone K is closed and it has a compact base, then* $\mathrm{PMin}_{LSc}(M,K) = \mathrm{PMin}_{Hu}(M,K)$.
(b) If V is normed, the ordering cone K is closed and it has a weakly compact base, then $\mathrm{PMin}_{He-La}(M,K) = \mathrm{PMin}_{Be}(M,K) = \mathrm{PMin}_{GBo}(M,K)$.

Whenever $V = \mathbb{R}^k$ and $K = \mathbb{R}^k_+$ more inclusions turn into equalities in the scheme considered in Proposition 2.4.16 one can be include the properly minimal elements in the sense of Geoffrion, too.

Proposition 2.4.18. *Let $V = \mathbb{R}^k$, $K = \mathbb{R}^k_+$ and $M \subseteq \mathbb{R}^k$ a nonempty set.*

(a) Then it holds

$$\mathrm{PMin}_{LSc}(M, \mathbb{R}^k_+) = \mathrm{PMin}_{Hu}(M, \mathbb{R}^k_+) \subseteq \mathrm{PMin}_{He-La}(M, \mathbb{R}^k_+) =$$

$$\mathrm{PMin}_{Be}(M, \mathbb{R}^k_+) = \mathrm{PMin}_{Ge}(M, \mathbb{R}^k_+) = \mathrm{PMin}_{GBo}(M, \mathbb{R}^k_+) \subseteq \mathrm{PMin}_{Bo}(M, \mathbb{R}^k_+).$$

(b) If, additionally, $M + \mathbb{R}^k_+$ is convex, then all the inclusions in (a) turn into equalities.

Remark 2.4.9. In section 4.4 we consider other minimality notions with respect to general increasing scalarization functions used only there, while in chapter 7 some minimality notions introduced in this section are extended for sets $M \subseteq \overline{V}$.

2.4.4 Linear scalarization

In this subsection we turn our attention to linear scalarization and its connections to the different minimality concepts introduced before. Scalarization in general allows us to associate a *scalar optimization problem* to a given *vector optimization problem*. This is very closely related to the monotonicity properties of the scalarizing function. In this context the dual cone and the quasi interior of the dual cone of the underlying ordering cone plays a crucial role. As noticed in Remark 2.4.8, characterizing minimal, weakly minimal or properly minimal elements of a given set by monotone scalarization, in particular linear scalarization, offers the possibility to investigate these notions by means of techniques which come from the scalar optimization.

Unless otherwise mentioned, in this subsection we consider V to be a vector space partially ordered by a convex cone $K \subseteq V$. If the pointedness of the cone K is required in some particular result, it will be explicitly mentioned. Moreover, the set $M \subseteq V$ is assumed to be nonempty.

Lemma 2.4.19. *Let $f : V \to \overline{\mathbb{R}}$ be a given function.*

(a) If f is K-increasing on M and there exists an uniquely determined element $\bar{v} \in M$ satisfying $f(\bar{v}) \leq f(v)$ for all $v \in M$, then $\bar{v} \in \mathrm{Min}(M, K)$.
(b) If f is strongly K-increasing on M and there exists $\bar{v} \in M$ satisfying $f(\bar{v}) \leq f(v)$ for all $v \in M$, then $\bar{v} \in \mathrm{Min}(M, K)$.

Proof. (a) Assuming $\bar{v} \notin \mathrm{Min}(M, K)$ yields the existence of $v \in M$ such that $v \leq_K \bar{v}$. Taking into consideration the fact that f is K-increasing we get $f(v) \leq f(\bar{v})$. Thus $f(v) = f(\bar{v})$ and this contradicts the uniqueness of $\bar{v}$ as solution of the problem $\min_{v \in M} f(v)$.

(b) Arguing as in part (a) in case $\bar{v} \notin \mathrm{Min}(M, K)$ one can find an element $v \in M$ such that $f(v) < f(\bar{v})$, but this contradicts the minimality of $f(\bar{v})$. □

For weakly minimal elements one has the following analogous characterization. Note that no difficulties arise if the ordering cone K is not pointed.

Lemma 2.4.20. *Suppose that* $\operatorname{core}(K) \neq \emptyset$ *and consider a function* $f : V \to \overline{\mathbb{R}}$ *which is strictly* K*-increasing on* M*. If there is an element* $\bar{v} \in M$ *fulfilling* $f(\bar{v}) \leq f(v)$ *for all* $v \in M$*, then* $\bar{v} \in \operatorname{WMin}(M, K)$*.*

Proof. If $\bar{v} \notin \operatorname{WMin}(M, K)$ then there exists $v \in (\bar{v} - \operatorname{core}(K)) \cap M$. This implies $f(v) < f(\bar{v})$, contradicting the assumption. □

The next scalarization result provides a necessary optimality condition for the minimal elements of M. One can notice the usefulness of the assumption of convexity for $M + K$, which allows giving such characterizations even if M is not a convex set. We refer to the previous subsections for the connections between the minimality properties of the sets M and $M + K$.

Theorem 2.4.21. *Assume that the ordering cone* K *is nontrivial and pointed,* $M + K$ *is convex and* $\operatorname{core}(M + K) \neq \emptyset$*. If* $\bar{v} \in \operatorname{Min}(M, K)$*, then there exists some* $v^{\#} \in K^{\#}\setminus\{0\}$ *such that* $\langle v^{\#}, \bar{v}\rangle \leq \langle v^{\#}, v\rangle$ *for all* $v \in M$*.*

Proof. If $\bar{v} \in \operatorname{Min}(M, K)$, then according to Lemma 2.4.1(*a*) we have $\bar{v} \in \operatorname{Min}(M + K, K)$, too, and this can be equivalently rewritten as $(M + K - \bar{v}) \cap (-K) = \{0\}$. Even more, as $M + K - \bar{v}$ and $(-K)$ are convex sets, $\operatorname{core}(M + K - \bar{v}) \neq \emptyset$ and $\operatorname{core}(M + K - \bar{v}) \cap (-K) = \emptyset$, Theorem 2.1.3 can be applied. Thus there exist $\bar{v}^{\#} \in V^{\#}\setminus\{0\}$ and $\lambda \in \mathbb{R}$ such that

$$\langle \bar{v}^{\#}, v + k_1 - \bar{v}\rangle \leq \lambda \leq \langle \bar{v}^{\#}, -k_2\rangle \ \forall v \in M \ \forall k_1, k_2 \in K. \tag{2.9}$$

If there exists $\bar{k} \in K\setminus\{0\}$ such that $\langle \bar{v}^{\#}, \bar{k}\rangle > 0$, then choosing $k_1 = \alpha\bar{k}$ for $\alpha > 0$, we obtain a contradiction to (2.9), as the left-hand side is unbounded for $\alpha \to +\infty$. Thus $\langle \bar{v}^{\#}, k\rangle \leq 0$ for all $k \in K\setminus\{0\}$ and this actually means that $\bar{v}^{\#} \in -K^{\#}$. Taking $k_1 = k_2 = 0$ and setting $v^{\#} := -\bar{v}^{\#} \in K^{\#}$ we get $\langle v^{\#}, \bar{v}\rangle \leq \langle v^{\#}, v\rangle$ for all $v \in M$, which completes the proof. □

It is clear that by means of the topological version of Eidelheit's separation theorem one can state an analogous scalarization result for the minimal elements of a subset M of a topological vector space.

Corollary 2.4.22. *Let* V *be a topological vector space partially ordered by a nontrivial pointed convex cone* K*. Moreover, assume that* $M + K$ *is convex and* $\operatorname{int}(M + K) \neq \emptyset$*. If* $\bar{v} \in \operatorname{Min}(M, K)$*, then there exists* $v^* \in K^*\setminus\{0\}$ *such that* $\langle v^*, \bar{v}\rangle \leq \langle v^*, v\rangle$ *for all* $v \in M$*.*

Now we present, again in the vector space setting, sufficient conditions for minimality which are immediate consequences of Lemma 2.4.19 and Example 2.2.3.

Theorem 2.4.23. *(a) If there exists* $v^{\#} \in K^{\#}$ *and* $\bar{v} \in M$ *such that* $\langle v^{\#}, \bar{v}\rangle < \langle v^{\#}, v\rangle$ *for all* $v \in M$, $v \neq \bar{v}$*, then* $\bar{v} \in \operatorname{Min}(M, K)$*.*

(b) If there exist $v^{\#} \in K^{\#0}$ *and* $\bar{v} \in M$ *such that* $\langle v^{\#}, \bar{v}\rangle \leq \langle v^{\#}, v\rangle$ *for all* $v \in M$*, then* $\bar{v} \in \operatorname{Min}(M, K)$*.*

Remark 2.4.10. (a) In Theorem 2.4.23(b) it is not necessary to impose the pointedness of K, because otherwise $K^{\#0} = \emptyset$.

(b) The necessary condition in Theorem 2.4.21 is not also sufficient because, as follows from Theorem 2.4.23(a), for this we need a strict inequality. Indeed, if $\langle v^{\#}, \bar{v}\rangle \leq \langle v^{\#}, v\rangle$ is for all $v \in M$ fulfilled, then $\bar{v}$ is weakly minimal to M, (see Theorem 2.4.25 below), but not necessarily minimal.

We would like to mention that in locally convex spaces partially ordered by a convex closed cone the strongly minimal elements can be as well equivalently characterized via linear scalarization by using linear continuous functionals from K^* (see [104, Theorem 5.6]). We omit giving this statement here, since strongly minimal elements are not interesting from the viewpoint of vector optimization and do not play any role in this book.

Next we turn our attention to necessary and sufficient optimality conditions characterizing via linear scalarization the weakly minimal elements of a nonempty subset of a vector space.

Theorem 2.4.24. *Let $K \subseteq V$ be such that* $\operatorname{core}(K) \neq \emptyset$ *and $M+K$ is convex. If $\bar{v} \in \operatorname{WMin}(M, K)$ then there exists $v^{\#} \in K^{\#}\setminus\{0\}$ such that $\langle v^{\#}, \bar{v}\rangle \leq \langle v^{\#}, v\rangle$ for all $v \in M$.*

Proof. The proof follows the lines of the proof of Theorem 2.4.21 using again the algebraic version of Eidelheit's separation theorem. □

Theorem 2.4.25. *Suppose that* $\operatorname{core}(K) \neq \emptyset$*. If there exist $v^{\#} \in K^{\#}\setminus\{0\}$ and $\bar{v} \in M$ such that for all $v \in M$ it holds $\langle v^{\#}, \bar{v}\rangle \leq \langle v^{\#}, v\rangle$, then $\bar{v} \in$* $\operatorname{WMin}(M, K)$.

Proof. The assertion is a straightforward conclusion of Lemma 2.4.20 and Example 2.2.3. □

Combining Theorem 2.4.24 and Theorem 2.4.25 one obtains an equivalent characterization via linear scalarization for weakly minimal elements.

Corollary 2.4.26. *Let $K \subseteq V$ be such that* $\operatorname{core}(K) \neq \emptyset$ *and $M+K$ is convex. Then $\bar{v} \in \operatorname{WMin}(M, K)$ if and only if there exists $v^{\#} \in K^{\#}\setminus\{0\}$ satisfying $\langle v^{\#}, \bar{v}\rangle \leq \langle v^{\#}, v\rangle$ for all $v \in M$.*

The following remark plays an important role when dealing with vector duality with respect to weakly minimal elements.

Remark 2.4.11. Assuming that V is a topological vector space partially ordered by the convex cone K with $\operatorname{int}(K) \neq \emptyset$ and $M \subseteq V$ is a nonempty set with $M + K$ convex, then by using the topological version of Eidelheit's separation theorem and the analog of Lemma 2.4.19 and Example 2.2.3 for topological vector spaces, one gets that $\bar{v} \in \operatorname{WMin}(M, K)$ if and only if there exists a $v^* \in K^*\setminus\{0\}$ such that $\langle v^*, \bar{v}\rangle \leq \langle v^*, v\rangle$ for all $v \in M$. Theorem 2.4.24 and Theorem 2.4.25 remain valid when formulated in a corresponding topological framework.

After characterizing minimal and weakly minimal elements of a set $M \subseteq V$ regarding the partial ordering induced by the convex cone $K \subseteq V$ via linear scalarization it is natural to ask whether it is possible to give analogous characterizations also for the properly minimal elements. First of all let us take a closer look at Definition 2.4.12, where we introduced $\text{PMin}_{LSc}(M, K)$, the set of properly minimal elements of M with respect to linear scalarization. This definition itself is already based on linear scalarization. If we look at Proposition 2.4.17(a) we see that under some additional hypotheses $\text{PMin}_{LSc}(M, K) = \text{PMin}_{Hu}(M, K)$, i.e. the properly minimal elements of M in the sense of Hurwicz may be characterized by linear scalarization using a functional $v^* \in K^{*0}$. Even more, as follows from Proposition 2.4.18(b), if $V = \mathbb{R}^k$, $K = \mathbb{R}^k_+$ and $M + K$ is a convex set, then all the properly minimal elements introduced in this section may be characterized in an equivalent manner by linear scalarization. But as far as properly minimal elements in the sense of Borwein are concerned, there exists a more general linear scalarization result, which can be proven like [104, Theorem 5.11 and Theorem 5.21].

Theorem 2.4.27. *Let V be a Hausdorff locally convex space partially ordered by the pointed convex closed cone K with $\text{int}_{w(V^*,V)}(K^*) \neq \emptyset$ and the nonempty set $M \subseteq V$ for which we assume that $M + K$ is convex. Then $\bar{v} \in \text{PMin}_{Bo}(M, K)$ if and only if there exists $v^* \in K^{*0}$ such that $\langle v^*, \bar{v} \rangle \leq \langle v^*, v \rangle$ for all $v \in M$.*

Remark 2.4.12. One should notice that for $V = \mathbb{R}^k$, $K = \mathbb{R}^k_+$ and $M \subseteq \mathbb{R}^k$ with $M + \mathbb{R}^k_+$ convex, when the hypotheses of Theorem 2.4.27 are fulfilled, there is

$$\text{PMin}_{LSc}(M, \mathbb{R}^k_+) = \text{PMin}_{Hu}(M, \mathbb{R}^k_+) = \text{PMin}_{He-La}(M, \mathbb{R}^k_+) =$$

$$\text{PMin}_{Be}(M, \mathbb{R}^k_+) = \text{PMin}_{Ge}(M, \mathbb{R}^k_+) = \text{PMin}_{GBo}(M, \mathbb{R}^k_+) = \text{PMin}_{Bo}(M, \mathbb{R}^k_+),$$

which is nothing but the assertion of Proposition 2.4.18(b).

2.5 Vector optimization problems

An optimization problem consisting in the minimization or maximization of several objective functions is a particular case of a *vector optimization problem*, for which one can find in the literature also the denotations *multiobjective* (or *multicriteria*) *optimization* (or *programming*) *problem* as well as *multiple objective optimization* (or *programming*) *problem*. The characteristic feature is the occurrence of several conflicting objectives, i.e. not all objectives under consideration attain their minimal or maximal values at the same element of the feasible set, which is a subset of the space where the objective functions are defined on, sometimes called *decision space* (or *input space*). In most real life decisions it is much more realistic to take into account not only one objective but different ones. For instance, if we look at an investment decision on the capital market it is reasonable to consider at least two objectives, namely the expected return which has to be maximized and the risk of an investment in a security or a portfolio of securities which should be minimized. In other situations one wants to minimize the cost and to maximize different features of quality of a product or a production process or to minimize the production time and to maximize the production capacity etc. It is obvious that such objectives often appear in a conflicting manner or as conflicting interests between different persons, groups of people or within a single decision-maker itself.

A widely used way of assessing the multiple objectives is on the base of partial ordering relations induced by convex cones. This allows to compare different vector objective values in the sense that an objective value is preferred if it is less than (if we consider a minimization problem) or greater than (if case of a maximization problem) another one with respect to the considered partial ordering induced by the underlying convex cone. The solutions are defined by those objective values that cannot be improved by another one in the sense of this preference notion. Thus, one immediately sees that the notions of minimal elements for sets introduced in section 2.4 turn out to be natural solution concepts in vector optimization. Although in many practical applications the number of considered objectives is finite, from a mathematical point of view the *objective space* (or *image space*), sometimes also called *outcome space*, may be an infinite dimensional space. So, for the sake of generality, we will define the vector optimization problem initially by considering general vector spaces for the decision and outcome spaces, the latter partially ordered by a convex cone.

Let X and V be vector spaces and assume that V is partially ordered by the convex cone $K \subseteq V$. For a given proper function $h : X \to \overline{V} = V \cup \{\pm\infty_K\}$ we investigate the *vector optimization problem* formally denoted by

$$(PVG) \quad \operatorname*{Min}_{x \in X} h(x).$$

It consists in determining the minimal, weakly minimal or properly minimal elements of the image set of X through h, also called *outcome set* (or *image set*), $h(\operatorname{dom} h) = \{v \in V : \exists x \in \operatorname{dom} h, v = h(x)\}$. In other words, we are interested in determining the sets $\operatorname{Min}(h(\operatorname{dom} h), K)$, $\operatorname{WMin}(h(\operatorname{dom} h), K)$ or $\operatorname{PMin}(h(\operatorname{dom} h), K)$, where PMin is a generic notation for all sets of properly minimal elements. On the other hand, we are also interested in finding the so-called *efficient*, *weakly efficient* or *properly efficient solutions* to (PVG).

Definition 2.5.1. *An element $\bar{x} \in X$ is said to be*

(a) an efficient solution to the vector optimization problem (PVG) if $\bar{x} \in \operatorname{dom} h$ and $h(\bar{x}) \in \operatorname{Min}(h(\operatorname{dom} h), K)$;
(b) a weakly efficient solution to the vector optimization problem (PVG) if $\bar{x} \in \operatorname{dom} h$ and $h(\bar{x}) \in \operatorname{WMin}(h(\operatorname{dom} h), K)$;

(c) a properly efficient solution to the vector optimization problem (PVG) *if* $\bar{x} \in \operatorname{dom} h$ *and* $h(\bar{x}) \in \operatorname{PMin}(h(\operatorname{dom} h), K)$.

The set containing all the efficient solutions to (PVG) is called the *efficiency set* of (PVG), the one containing all the weakly efficient solutions to (PVG) is said to be the *weak efficiency set* of (PVG), while the name used for the one containing all the properly efficient solutions to (PVG) is the *proper efficiency set* of (PVG).

It is worth mentioning that in many cases in practice a decision-maker is only interested to have a subset or even a single element of one of these efficiency sets. This is a direct consequence of the practical requirements in applications.

Frequently, one looks for efficient elements in a nonempty subset $\mathcal{A} \subseteq X$, where the objective function is $h : \mathcal{A} \to V$. This problem can be reformulated in the form of (PVG) by considering as objective function $\tilde{h} : X \to \overline{V}$,

$$\tilde{h}(x) = \begin{cases} h(x), & \text{if } x \in \mathcal{A}, \\ +\infty_K, & \text{otherwise.} \end{cases}$$

Although we have just defined the efficient solutions via the minimality notions for the image set, for the sake of convenience let us state them in an explicit manner.

Definition 2.5.2. *An element* $\bar{x} \in X$ *is said to be an efficient solution to the vector optimization problem* (PVG) *if* $\bar{x} \in \operatorname{dom} h$ *and for all* $x \in \operatorname{dom} h$ *from* $h(x) \leqq_K h(\bar{x})$ *follows* $h(\bar{x}) \leqq_K h(x)$. *The set of efficient solutions to* (PVG) *is denoted by* $\operatorname{Eff}(PVG)$.

As pointed out in the previous section, there are several equivalent formulations for $\bar{x} \in Eff(PVG)$, like, for example, $(h(\bar{x}) - K) \cap h(\operatorname{dom} h) \subseteq h(\bar{x}) + K$ and, in case K is pointed, $(h(\bar{x}) - K) \cap h(\operatorname{dom} h) = \{h(\bar{x})\}$.

Definition 2.5.3. *Suppose that* $\operatorname{core}(K) \neq \emptyset$. *An element* $\bar{x} \in X$ *is said to be a weakly efficient solution to the vector optimization problem* (PVG) *if* $\bar{x} \in \operatorname{dom} h$ *and* $(h(\bar{x}) - \operatorname{core}(K)) \cap h(\operatorname{dom} h) = \emptyset$. *The set of weakly efficient solutions to* (PVG) *is denoted by* $\operatorname{WEff}(PVG)$.

One can see that $\bar{x} \in \operatorname{WEff}(PVG)$ if and only if $\bar{x} \in \operatorname{dom} h$ and there is no $x \in \operatorname{dom} h$ satisfying $h(x) <_K h(\bar{x})$.

Taking into consideration Proposition 2.4.2, whenever $\operatorname{core}(K) \neq \emptyset$ and $K \neq V$, we have $\operatorname{Eff}(PVG) \subseteq \operatorname{WEff}(PVG)$. In section 2.4 we have pointed out the close connection between the different types of minimal elements to the sets M and $M + K$, when $M \subseteq X$ is a nonempty set. These results are important in the context of scalarization since we have seen that the property of $M + K$ to be convex is sufficient for the characterization of minimal elements of M by means of linear scalarization. We may transfer this to the vector optimization problem in an obvious manner.

From section 2.2 we know that if $h : X \to \overline{V}$ is a proper function, the assumption that $h(\operatorname{dom} h) + K$ is convex is equivalent to the property that the function h is K-convexlike. This allows to establish scalarization results for vector optimization problems with K-convexlike and indirectly with K-convex objective functions.

In an analogous manner one can deliver explicitly definitions for the different notions of properly efficient solutions. We restrict ourselves here only to the properly efficient solutions with respect to linear scalarization based on Definition 2.4.12 because this type of properly efficient solutions will be later involved in different duality statements for vector optimization problems.

Let us suppose that V is a Hausdorff locally convex space partially ordered by a pointed convex cone K. One can alternatively define the properly efficient solutions in the sense of linear scalarization in the following manner.

Definition 2.5.4. *An element $\bar{x} \in X$ is said to be a properly efficient solution to (PVG) in the sense of linear scalarization if there exists $v^* \in K^{*0}$ such that $(v^*h)(\bar{x}) \leq (v^*h)(x)$ for all $x \in X$. The set of properly efficient solutions in the sense of linear scalarization to (PVG) is denoted by* $\mathrm{PEff}_{LSc}(PVG)$.

In other words, $\bar{x} \in \mathrm{PEff}_{LSc}(PVC)$ if and only if $\bar{x}$ is an optimal solution to the scalar optimization problem

$$\min_{x \in X} (v^*h)(x).$$

The results in this section can be used for providing corresponding characterizations for the properly efficient solutions to (PVG) by means of linear scalarization.

Bibliographical notes

Convex analysis established itself as a distinct area of mathematics after the publishing of Rockafellar's seminal book [157] where one can find most of the known results concerning convex sets and functions in finite dimensional spaces. Besides it, for our exposition on convex sets and functions we used mainly the books of Hiriart-Urruty and Lemarechal [90] and Borwein and Lewis [22] when working in finite dimensional spaces and the ones of Ekeland and Temam [67] and Zălinescu [207] for infinite dimensional spaces. Most of the statements we give here without proofs are demonstrated in these books. We also refer to the books [104, 125] for notions and results concerning vector functions and cones and to [20, 27, 207] for some results involving generalized interiors.

Conjugate functions have been introduced in finite dimensional spaces by Fenchel in [68]. The further development of this theory in finite dimensional spaces can be pursued in Rockafellar's book [157] and in topological vector spaces in the books of Ekeland and Temam [67] and Zălinescu [207]. A good and detailed treatment of the finite dimensional theory can be found also in the books [91,92] of Hiriart-Urruty and Lemaréchal. Our presentation contains many results, partially with some modified proofs and some extensions, the reader can find in the mentioned books.

As far as subdifferentiability is concerned we refer to the mentioned books, too. Subdifferentiability is closely related to conjugacy as we have conveyed in this chapter and as an appropriate reference for this notion we quote here the book [159] of Rockafellar. Concerning partially ordered vector spaces we would like to refer to the initiating paper [110] of Kantorovitch, which is more than seventy years old. Books in this field are due to Nachbin [141], Peressini [151], Jameson [105] and others. Fuchssteiner and Lusky investigated convex cones in [69].

In vector optimization one of the essential and basic notions is that of minimal elements of a set regarding the partial ordering induced by a convex cone. The first definition of proper minimal elements has been proposed by Kuhn and Tucker [119] for finite dimensional vector optimization problems under differentiability assumptions, followed by the one due to Hurwicz [93]. Geoffrion introduced his widely used and economically inspired definition in [71], also for finite dimensional vector maximum problems. Later, more general notions of properly efficient elements, also for infinite dimensional vector optimization problems, came into discussion in papers written by Borwein [17], Benson [15], etc.

Very beneficial regarding the connections to scalar optimization is the characterization of different types of properly efficient and weakly efficient solutions by linear scalarization. For contributions on the relations between different kinds of properly minimal elements of a set we refer to [83,87] and the book [81], while linear scalarization results for minimal, properly and weakly minimal (or efficient) solutions can be found in Jahn's book [104].

Linear scalarization results for properly efficient elements in the sense of Borwein can be found in [17]. In Jahn's book [104] one can find also so-called norm scalarization results. An overview about different linear and other scalarization methods for finite dimensional vector optimization problems including sensitivity results and numerical methods is given in [66]. Vector optimization has attracted the attention and the research activities of many scientists. Let us mention here only some textbooks and monographs which allow to get an introduction as well as a deeper insight into the field (cf. [49, 64, 65, 80, 81, 104, 108, 125, 130, 163, 181, 209]).

Springer

http://www.springer.com/978-3-642-02885-4

Duality in Vector Optimization
Bot, R.I.; Grad, S.-M.; Wanka, G.
2009, XVI, 400 p., Hardcover
ISBN: 978-3-642-02885-4