Chapter 1

INTRODUCTION TO CONVEX AND QUASICONVEX ANALYSIS

Johannes B.G. Frenk
Department of Economics, Econometric Institute
Erasmus University
3000DR Rotterdam, The Netherlands
frenk@few.eur.nl

Gábor Kassay
Faculty of Mathematics and Computer Science
Babes Bolyai University
3400 Cluj, Romania
kassay@math.ubbcluj.ro

Abstract In the first chapter of this book the basic results within convex and quasiconvex analysis are presented. In Section 2 we consider in detail the algebraic and topological properties of convex sets within $\mathbb{R}^n$ together with their primal and dual representations. In Section 3 we apply the results for convex sets to convex and quasiconvex functions and show how these results can be used to give primal and dual representations of the functions considered in this field. As such, most of the results are well known with the exception of Subsection 3.4 dealing with dual representations of quasiconvex functions. In Section 3 we consider applications of convex analysis to noncooperative game and minimax theory, Lagrangian duality in optimization and the properties of positively homogeneous evenly quasiconvex functions. Among these result an elementary proof of the well-known Sion's minimax theorem concerning quasiconvex-quasiconcave bifunctions is presented, thereby avoiding the less elementary fixed point arguments. Most of the results are proved in detail and the authors have tried to make these proofs as transparent as possible. Remember that convex analysis deals with the study of convex cones and convex sets and these objects are generalizations of linear subspaces and affine sets, thereby extending the field of linear algebra. Although some of the proofs are technical, it is possible to give a clear geometrical interpretation of the main ideas of convex analysis. Finally in Section 5 we list a short and probably incomplete overview on the history of convex and quasiconvex analysis.

Keywords: Convex Analysis, Quasiconvex Analysis, Noncooperative games, Minimax, Optimization theory.

1. Introduction

In this chapter the fundamental questions studied within the field of convex and quasiconvex analysis are discussed. Although some of these questions can also be answered within infinite dimensional real topological vector spaces, our universe will be the finite dimensional real linear space $\mathbb{R}^n$ equipped with the well-known Euclidean norm $\|.\|$. Since convex and quasiconvex analysis can be seen as the study of certain sets, we consider in Section 2 the basic sets studied in this field and list with or without proof the most important algebraic and topological properties of those sets. In this section a proof based on elementary calculus of the important separation result for disjoint convex sets in $\mathbb{R}^n$ will be given. In Section 3 we introduce the so-called convex and quasiconvex functions and show that the study of these functions can be reduced to the study of the sets considered in Section 2. As such, the formulation of the separation result for disjoint convex sets is now given by the dual representation of a convex or quasiconvex function. In Section 4 we will discuss important applications of convex and quasiconvex analysis to optimization theory, game theory and the study of positively homogeneous evenly quasiconvex functions. Finally in Section 5 we consider some of the historical developments within the field of convex and quasiconvex analysis.

2. Sets studied within convex and quasiconvex analysis

In this section the basic sets studied within convex and quasiconvex analysis in $\mathbb{R}^n$ are discussed and their most important properties listed. Since in some cases these properties are well-known we often mention them without any proof. We introduce in Subsection 2.1 the definition of a linear subspace, an affine set, a cone and a convex set in $\mathbb{R}^n$ together with their so-called primal representation. Also the important concept of a hull operation applied to an arbitrary set is considered. In Subsection 2.2 the topological properties of the sets considered in Subsection 2.1 are listed and in Subsection 2.3 we prove the well-known separation result for disjoint convex sets. Finally in Subsection 2.4 this separation result is applied to derive the so-called dual representation of a closed convex set. In case proofs are included we have tried to make these proofs as transparent and simple as possible. Also in some cases these proofs can be easily adapted, if our universe is an infinite dimensional real topological vector space. Most of the material in this section together with the proofs can be found in Lancaster and Tismenetsky (cf. [47]) for the linear algebra part, while for the convex analysis part the reader is referred to Rockafellar (cf. [63]) and Hiriart-Urruty and Lemaréchal (cf. [34], [35]).

2.1 Algebraic properties of sets

As already observed our universe will always be the n-dimensional Euclidean space $\mathbb{R}^n$ and any element of $\mathbb{R}^n$ is denoted by the vector $\mathbf{x} = (x_1, ..., x_n)^\top, x_i \in \mathbb{R}$ or $\mathbf{y} = (y_1, ..., y_n)^\top, y_i \in \mathbb{R}$. The inner product $< ., . >: \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ is then given by

$$< \mathbf{x}, \mathbf{y} >:= \sum_{i=1}^{n} x_i y_i = \mathbf{x}^\top \mathbf{y},$$

while the Euclidean norm $\|.\|$ is defined by

$$\|\mathbf{x}\| := \sqrt[2]{< \mathbf{x}, \mathbf{x} >}.$$

To simplify the notation, we also introduce for the sets $A, B \subseteq \mathbb{R}^n$ and $\alpha, \beta \in \mathbb{R}$ the *Minkowsky sum* $\alpha A + \beta B$ given by

$$\alpha A + \beta B := \{\alpha\mathbf{x} + \beta\mathbf{y} : \mathbf{x} \in A, \mathbf{y} \in B\}.$$

The first sets to be introduced are the main topic of study within *linear algebra* (cf. [47]).

Definition 1.1 *A set $L \subseteq \mathbb{R}^n$ is called a linear subspace if L is nonempty and $\alpha L + \beta L \subseteq L$ for every α, $\beta \in \mathbb{R}$. Moreover, a set $M \subseteq \mathbb{R}^n$ is called affine if $\alpha M + (1-\alpha)M \subseteq M$ for every $\alpha \in \mathbb{R}$.*

The empty set $\varnothing$ and $\mathbb{R}^n$ are extreme examples of an affine set. Also it can be shown that the set M is affine and $\mathbf{0} \in M$ if and only if M is a linear subspace and for each nonempty affine set M there exists a unique linear subspace L_M satisfying

$$M = L_M + \mathbf{x} \tag{1.1}$$

for any given $\mathbf{x} \in M$ (cf. [63]).

Since $\mathbb{R}^n$ is a linear subspace, we can apply to any nonempty set $S \subseteq \mathbb{R}^n$ the so-called *linear hull operation* and construct the set

$$lin(S) := \cap\{L : S \subseteq L \text{ and } L \text{ a linear subspace}\}. \tag{1.2}$$

For any collection of linear subspaces $L_i, i \in I$ containing S it is obvious that the intersection $\cap_{i \in I} L_i$ is again a linear subspace containing S and this shows that the set $lin(S)$ is the smallest linear subspace containing S. The set $lin(S)$ is called the *linear hull generated by the set S* and if S has a finite number of elements the linear hull is called *finitely generated.* By a similar argument one can construct, using the so-called *affine hull operation,* the smallest affine set containing S. This set, denoted by $aff(S)$, is called the *affine hull generated by the set S* and is given by

$$aff(S) := \cap\{M : S \subseteq M \text{ and } M \text{ an affine set}\}. \tag{1.3}$$

If the set S has a finite number of elements, the affine hull is called *finitely generated.* Since any linear subspace is an affine set, it is clear that $aff(S) \subseteq lin(S)$. To give a so-called *primal representation* of these sets we introduce the next definition.

Definition 1.2 *A vector* $\mathbf{x}$ *is a linear combination of the vectors* $\mathbf{x}_1, ..., \mathbf{x}_k$ *if*

$$\mathbf{x} = \sum_{i=1}^{k} \alpha_i \mathbf{x}_i, \alpha_i \in \mathbb{R}, 1 \leq i \leq k.$$

A vector $\mathbf{x}$ *is an affine combination of the vectors* $\mathbf{x}_1, ..., \mathbf{x}_k$ *if*

$$\mathbf{x} = \sum_{i=1}^{k} \alpha_i \mathbf{x}_i, \alpha_i \in \mathbb{R}, 1 \leq i \leq k \text{ and } \sum_{i=1}^{k} \alpha_i = 1.$$

A linear combination of the nonempty set S is given by the set $\sum_{i=1}^{k} \alpha_i S$ *with* $\alpha_i \in \mathbb{R}, 1 \leq i \leq k$, *while an affine combination of the same set is given by the set* $\sum_{i=1}^{k} \alpha_i S$ *with* $\sum_{i=1}^{k} \alpha_i = 1$ *and* $\alpha_i \in \mathbb{R}, 1 \leq i \leq k$.

A trivial consequence of Definitions 1.1 and 1.2 is given by the next result which also holds in infinite dimensional linear spaces.

Lemma 1.1 *A nonempty set* $L \subseteq \mathbb{R}^n$ *is a linear subspace if and only if it contains all linear combinations of the set L. Moreover, a nonempty set* $M \subseteq \mathbb{R}^n$ *is an affine set if and only if it contains all affine combinations of the set M.*

The result in Lemma 1.1 yields a primal representation of a linear subspace and an affine set. In particular, we obtain from Lemma 1.1

that the set $lin(S)$ ($aff(S)$) with $S \subseteq \mathbb{R}^n$ nonempty equals all linear (affine) combinations of the set S. This means

$$lin(S) = \cup_{k=1}^{\infty}\{\sum_{i=1}^{k} \alpha_i S : \alpha_i \in \mathbb{R}\} \tag{1.4}$$

and

$$aff(S) = \cup_{k=1}^{\infty}\{\sum_{i=1}^{k} \alpha_i S : \alpha_i \in \mathbb{R} \text{ and } \sum_{i=1}^{k} \alpha_i = 1\}. \tag{1.5}$$

For any nonempty sets $S_1 \subseteq \mathbb{R}^n$ and $S_2 \subseteq \mathbb{R}^m$ one can now show using relation (1.4) that

$$lin(S_1 \times S_2) = lin(S_1) \times lin(S_2) \tag{1.6}$$

and using relation (1.5) that

$$aff(S_1 \times S_2) = aff(S_1) \times aff(S_2). \tag{1.7}$$

Also, for $A : \mathbb{R}^n \to \mathbb{R}^m$ a linear mapping, it is easy to verify that

$$A(lin(S)) = lin(A(S)) \tag{1.8}$$

and for $A : \mathbb{R}^n \to \mathbb{R}^m$an affine mapping, that

$$A(aff(S)) = aff(A(S)). \tag{1.9}$$

Recall a mapping $A : \mathbb{R}^n \to \mathbb{R}^m$ is called linear if

$$A(\alpha\mathbf{x} + \beta\mathbf{y}) = \alpha A(\mathbf{x}) + \beta A(\mathbf{y})$$

for every $\alpha, \beta \in \mathbb{R}$ and $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$ and it is called affine if

$$A(\alpha\mathbf{x} + (1-\alpha)\mathbf{y}) = \alpha A(\mathbf{x}) + (1-\alpha)A(\mathbf{y})$$

for every $\alpha \in \mathbb{R}$ and $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$. Moreover, in case we apply relation (1.7) to the affine mapping $A : \mathbb{R}^{2n} \to \mathbb{R}^n$, given by $A(\mathbf{x}, \mathbf{y}) = \alpha\mathbf{x} + \beta\mathbf{y}$, with $\alpha, \beta \in \mathbb{R}$ and use relation (1.9) the following rule for the affine hull of the sum of sets is easy to verify.

Lemma 1.2 *For any nonempty sets $S_1, S_2 \subseteq \mathbb{R}^n$ and $\alpha, \beta \in \mathbb{R}$ it follows that*

$$aff(\alpha S_1 + \beta S_2) = \alpha aff(S_1) + \beta aff(S_2).$$

Another application of relations (1.4) and (1.5) yields the next result.

Lemma 1.3 *For any nonempty set* $S \subseteq \mathbb{R}^n$ *and* $\mathbf{x}_0$ *belonging to* $aff(S)$ *it follows that* $aff(S) = \mathbf{x}_0 + lin(S - \mathbf{x}_0)$.

An improvement of Lemma 1.1 is given by the observation that any linear subspace (affine set) of $\mathbb{R}^n$ can be written as the linear or affine hull of a *finite* subset $S \subseteq \mathbb{R}^n$. To show this improvement one needs to introduce the next definition (cf. [47]).

Definition 1.3 *The vectors* $\mathbf{x}_1, ..., \mathbf{x}_k$ *are called linearly independent if*

$$\sum_{i=1}^{k} \alpha_i \mathbf{x}_i = \mathbf{0} \text{ and } \alpha_i \in \mathbb{R} \Rightarrow \alpha_i = 0,\ 1 \leq i \leq k.$$

Moreover, the vectors $\mathbf{x}_1, ..., \mathbf{x}_k$ *are called affinely independent if*

$$\sum_{i=1}^{k} \alpha_i \mathbf{x}_i = \mathbf{0} \text{ and } \sum_{i=1}^{k} \alpha_i = 0 \Rightarrow \alpha_i = 0,\ 1 \leq i \leq k.$$

For $k \geq 2$ an equivalent characterization of affinely independent vectors is given by the observation that the vectors $\mathbf{x}_1, ..., \mathbf{x}_k$ are affinely independent if and only if the vectors $\mathbf{x}_2 - \mathbf{x}_1, ..., \mathbf{x}_k - \mathbf{x}_1$ are linear independent (cf. [34]). To explain the name linearly and affinely independent we observe that the vectors $\mathbf{x}_1, ..., \mathbf{x}_k$ are linearly independent if and only if any vector $\mathbf{x}$ belonging to the linear hull $lin(\{\mathbf{x}_1, ..., \mathbf{x}_k\})$ can be written as a unique linear combination of the vectors $\mathbf{x}_1, ..., \mathbf{x}_k$. Moreover, the vectors $\mathbf{x}_1, ..., \mathbf{x}_k$ are affinely independent if and only if any vector $\mathbf{x}$ belonging to the affine hull $aff(\{\mathbf{x}_1, ..., \mathbf{x}_k\})$ can be written as a unique affine combination of the vectors $\mathbf{x}_1, ..., \mathbf{x}_k$. The improvement of Lemma 1.1 is given by the following result well-known within linear algebra (cf. [47]).

Lemma 1.4 *For any linear subspace* $L \subseteq \mathbb{R}^n$ *containing nonzero elements there exists a set of linearly independent vectors* $\mathbf{x}_1, ..., \mathbf{x}_k$, $k \leq n$ *satisfying* $lin(\{\mathbf{x}_1, ..., \mathbf{x}_k\}) = L$. *Also for any nonempty affine set* $M \subseteq \mathbb{R}^n$ *there exists a set of affinely independent vectors* $\mathbf{x}_0, ..., \mathbf{x}_k, k \leq n$ *satisfying* $aff(\{\mathbf{x}_0, ..., \mathbf{x}_k\}) = M$.

By Lemma 1.4 any linear subspace $L \subseteq \mathbb{R}^n$ containing nonzero elements can be represented as the linear hull of $k \leq n$ linearly independent vectors. If this holds, the dimension $\dim(L)$ of the linear subspace L is given by k. Since any $\mathbf{x}$ belonging to L can be written as a unique linear combination of linearly independent vectors this shows (cf. [47]) that $\dim(L)$ is well defined for L containing nonzero elements. If $L = \{\mathbf{0}\}$ the dimension $\dim(L)$ is by definition equal to 0. To extend this to affine sets we observe by relation (1.1) that any nonempty affine set M is parallel to its unique subspace L_M and the dimension $\dim(M)$ of a nonempty affine set M is now given by $\dim(L_M)$. By definition the dimension of the empty set $\varnothing$ equals -1. Finally, the dimension $\dim(S)$ of an arbitrary set $S \subseteq \mathbb{R}^n$ is given by $\dim(\operatorname{aff}(S))$. In the next definition we will introduce the sets which are the main objects of study within the field of convex and quasiconvex analysis.

Definition 1.4 *A set $C \subseteq \mathbb{R}^n$ is called convex if $\alpha C + (1-\alpha)C \subseteq C$ for every $0 < \alpha < 1$. Moreover, a set $K \subseteq \mathbb{R}^n$ is called a cone if $\alpha K \subseteq K$ for every $\alpha > 0$.*

The empty set $\varnothing$ is an extreme example of a convex set and a cone. An affine set is clearly a convex set but it is obvious that not every convex set is an affine set. This shows that convex analysis is an extension of linear algebra. Moreover, it is easy to show for every cone K that

$$K \text{ convex} \Leftrightarrow K + K \subseteq K. \tag{1.10}$$

Finally, for $A : \mathbb{R}^n \to \mathbb{R}^m$ an affine mapping and $C \subseteq \mathbb{R}^n$ a nonempty convex set it follows that the set $A(C)$ is convex, while for $A : \mathbb{R}^n \to \mathbb{R}^m$ a linear mapping and $K \subseteq \mathbb{R}^n$ a nonempty cone the set $A(K)$ is a cone.

To relate convex sets to convex cones we observe for $\mathbb{R}_+ := [0, \infty)$ and any nonempty set $S \subseteq \mathbb{R}^n$ that the set

$$\mathbb{R}_+(S \times \{1\}) := \{(\alpha \mathbf{x}, \alpha) : \alpha \geq 0, \mathbf{x} \in S\} \subseteq \mathbb{R}^{n+1}$$

is a cone. This implies by relation (1.10) that the set $\mathbb{R}_+(C \times \{1\})$ is a convex cone for any convex set $C \subseteq \mathbb{R}^n$. It is now clear for any nonempty set $S \subseteq \mathbb{R}^n$ that

$$\mathbb{R}_+(S \times \{1\}) \cap (\mathbb{R}^n \times \{1\}) = S \times \{1\} \tag{1.11}$$

and so any convex set C can be seen as an intersection of the convex cone $\mathbb{R}_+(C \times \{1\})$ and the affine set $\mathbb{R}^n \times \{1\}$. This shows that convex sets are closely related to convex cones and by relation (1.11) one can study convex sets by only studying affine sets and convex cones containing $\mathbf{0}$. We will not pursue this approach but only remark that the above relation is sometimes useful. Introducing an important subclass of convex sets, let $\mathbf{a}$ be a nonzero vector belonging to $\mathbb{R}^n$ and $b \in \mathbb{R}$ and

$$H^<(\mathbf{a}, b) := \{\mathbf{x} \in \mathbb{R}^n : \mathbf{a}^\intercal \mathbf{x} < b\}. \tag{1.12}$$

The set $H^<(\mathbf{a}, b)$ is called a halfspace and clearly this halfspace is a convex set. Moreover, the set $H^\leq(\mathbf{a}, b) := \{\mathbf{x} \in \mathbb{R}^n : \mathbf{a}^\intercal \mathbf{x} \leq b\}$ is also called a halfspace and this set is also a convex set. Another important subclass of convex sets useful within the study of quasiconvex functions is given by the following definition (cf. [23]).

Definition 1.5 *A set $C_e \subseteq \mathbb{R}^n$ is called evenly convex if $C_e = \mathbb{R}^n$ or C_e is the intersection of a collection of halfspaces $H^{<}(\mathbf{a}, b)$.*

Clearly the empty set $\varnothing$ is evenly convex and since any halfspace $H^{\leq}(\mathbf{a}, b)$ can be obtained by intersecting the halfspaces $H^{<}(\mathbf{a}, b+\frac{1}{n}), n \geq 1$ it also follows that any halfspace $H^{\leq}(\mathbf{a}, b)$ is evenly convex. In Subsection 2.3 it will be shown that any closed or open convex set is evenly convex. However, there exist convex sets which are not evenly convex.

Example 1.1 If $C := \{(x_1, x_2) : 0 \leq x_1 \leq 1, 0 \leq x_2 < 1\} \cup \{(1,1)\}$, then it follows that C is convex but not evenly convex.

Since $\mathbb{R}^n$ is a convex set, we can apply to any nonempty set $S \subseteq \mathbb{R}^n$ the so-called *convex hull operation* and construct the nonempty set

$$co(S) := \cap\{C : S \subseteq C \text{ and } C \text{ a convex set}\}. \tag{1.13}$$

For any collection of convex sets $C_i, i \in I$ containing S it is obvious that the intersection $\cap_{i \in I} C_i$ is again a convex set containing S and this shows that the set $\text{co}(S)$ is the smallest convex set containing S. The set $co(S)$ is called the *convex hull generated by the set S* and if S has a finite number of elements the convex hull is called *finitely generated.* Since $\mathbb{R}^n$ is by definition evenly convex one can construct by a similar argument using the so-called *evenly convex hull operation* the smallest evenly convex set containing the nonempty set S. This set, denoted by $ec(S)$, is called the *evenly convex hull generated by the set S* and is given by

$$ec(S) := \cap \{C_e : S \subseteq C_e \text{ and } C_e \text{ an evenly convex set}\}. \tag{1.14}$$

Since any evenly convex set is convex it follows that $co(S) \subseteq ec(S)$.

By the so-called *canonic hull operation* one can also construct the smallest convex cone containing the nonempty set S, and the smallest convex cone containing $S \cup \{\mathbf{0}\}$. The last set is given by

$$cone(S) := \cap\{K : S \cup \{\mathbf{0}\} \subseteq K \text{ and } K \text{ a convex cone}\}. \tag{1.15}$$

Unfortunately this set is called the *convex cone generated by S* (cf. [63]). Clearly the set $cone(S)$ is in general not equal to the smallest convex cone containing S unless the zero element belongs to S. To give an alternative characterization of the above sets we introduce the next definition.

Definition 1.6 *A vector* $\mathbf{x}$ *is a canonical combination of the vectors* $\mathbf{x}_1, ..., \mathbf{x}_k$ *if*

$$\mathbf{x} = \sum_{i=1}^{k} \alpha_i \mathbf{x}_i, \alpha_i \geq 0, 1 \leq i \leq k.$$

The vector $\mathbf{x}$ *is called a strict canonical combination of the same vectors if* $\alpha_i > 0, 1 \leq i \leq k$. *A vector* $\mathbf{x}$ *is a convex combination of the vectors* $\mathbf{x}_1, ..., \mathbf{x}_k$ *if*

$$\mathbf{x} = \sum_{i=1}^{k} \alpha_i \mathbf{x}_i \text{ and } \sum_{i=1}^{k} \alpha_i = 1, \alpha_i > 0.$$

A canonical combination of the nonempty set S *is given by the set* $\sum_{i=1}^{k} \alpha_i S$ *with* $\alpha_i \geq 0, 1 \leq i \leq k$, *while a strict canonical combination of the same set is given by* $\sum_{i=1}^{k} \alpha_i S$ *with* $\alpha_i > 0, 1 \leq i \leq k$. *Finally a convex combination of the set* S *is given by the set* $\sum_{i=1}^{k} \alpha_i S$ *with* $\sum_{i=1}^{k} \alpha_i = 1, \alpha_i > 0, 1 \leq i \leq k$.

A trivial consequence of Definitions 1.4 and 1.6 is given by the next result which also holds in infinite dimensional linear spaces.

Lemma 1.5 *A nonempty set* $K \subseteq \mathbb{R}^n$ *is a convex cone (convex cone containing* $\mathbf{0}$*) if and only if it contains all strict canonical (canonical) combinations of the set K. Moreover, a nonempty set* $C \subseteq \mathbb{R}^n$ *is a convex set if and only if it contains all convex combinations of the set C.*

The result in Lemma 1.5 yields a primal representation of a convex cone and a convex set. In particular, we obtain from Lemma 1.5 that the set $cone(S)$ $(co(S))$ with $S \subseteq \mathbb{R}^n$ nonempty equals all canonical (convex) combinations of the set S. This means

$$cone(S) = \cup_{k=1}^{\infty} \{\sum_{i=1}^{k} \alpha_i S : \alpha_i \geq 0\} \tag{1.16}$$

and

$$co(S) = \cup_{k=1}^{\infty} \{\sum_{i=1}^{k} \alpha_i S : \sum_{i=1}^{k} \alpha_i = 1, \alpha_i > 0\}. \tag{1.17}$$

We observe that the representations of $cone(S)$ and $co(S)$, listed in relations (1.16) and (1.17), are the "convex equivalences" of the representation of $lin(S)$ and $aff(S)$ given by relations (1.4) and (1.5). Moreover, to relate the above representations, it is easy to see that

$$cone(S) = \mathbb{R}_+(co(S)). \tag{1.18}$$

Since by relations (1.16) and (1.17) a convex cone containing $\mathbf{0}$ (convex set) can be seen as a generalization of a linear subspace (affine set) one might wonder whether a similar result as in Lemma 1.4 holds. Hence we wonder whether any convex cone containing $\mathbf{0}$ (convex set) can be seen as a canonical (convex) combination of a finite set S.

Example 1.2 Contrary to linear subspaces it is not true that any convex cone containing $\mathbf{0}$ is a canonical combination of a finite set. An example is given by the so-called L^2 or ice-cream cone $K = \{(\mathbf{x}, t) : \|\mathbf{x}\| \leq t\} \subseteq \mathbb{R}^{n+1}$.

Despite this negative result it is possible in finite dimensional linear spaces to improve for canonical hulls and convex hulls the representation given by relations (1.16) and (1.17). In the next result it is shown that any element belonging to cone(S) with S containing nonzero elements can be written as a canonical combination of at most n linearly independent vectors belonging to S. This is called *Caratheodory's theorem* for canonical hulls. Using this result and relation (1.11) a related result holds for convex hulls and in this case linearly independent is replaced by affinely independent and at most n is replaced by at most $n+1$. Clearly this result (cf. [63]) is the "convex equivalence" of Lemma 1.4.

Lemma 1.6 *If $S \subseteq \mathbb{R}^n$ is a set containing nonzero elements, then for any $\mathbf{x}$ belonging to $cone(S)$ there exists a set of linearly independent vectors $\mathbf{x}_1, ..., \mathbf{x}_k, k \leq n$ belonging to S such that $\mathbf{x}$ can be written as a canonical combination of these vectors. Moreover, for any $\mathbf{x} \in co(S)$ there exists a set of affinely independent vectors $\mathbf{x}_1, ..., \mathbf{x}_k, k \leq n+1$ belonging to S such that $\mathbf{x}$ can be written as a convex combination of these vectors.*

Proof. Clearly for $\mathbf{0} \in cone(S)$ the desired result holds and so $\mathbf{x} \in cone(S)$ should be nonzero. By relation (1.16) there exists some finite set $\{\mathbf{x}_1, ..., \mathbf{x}_k\}$ and $\alpha_i > 0, 1 \leq i \leq k$ satisfying $\mathbf{x} = \sum_{i=1}^k \alpha_i \mathbf{x}_i$. If the vectors $\mathbf{x}_i, 1 \leq i \leq k$ are linearly independent, then clearly $k \leq n$ and we are done. Otherwise, there exists a nonzero sequence $\beta_i, 1 \leq i \leq k$ satisfying $\mathbf{0} = \sum_{i=1}^k \beta_i \mathbf{x}_i$ and without loss of generality we may assume that the set $I := \{1 \leq i \leq k : \beta_i > 0\}$ is nonempty. If $\epsilon := \min\{\frac{\alpha_i}{\beta_i} : i \in I\} > 0$ and $i^* := \arg\min\{\frac{\alpha_i}{\beta_i} : i \in I\}$ we obtain that

$$\mathbf{x} = \sum_{i=1}^k (\alpha_i - \epsilon\beta_i)\mathbf{x}_i = \sum_{i=1, i \neq i^*}^k (\alpha_i - \epsilon\beta_i)\mathbf{x}_i$$

and so $\mathbf{x}$ can be written as a strict canonical combination of at most $k-1$ vectors. Applying now the same procedure again until we have identified a subset of $\{\mathbf{x}_1, ..., \mathbf{x}_k\}$ consisting of linearly independent vectors the first part follows. To show the result for convex hulls it follows for any $\mathbf{x} \in co(S)$ that $(\mathbf{x}, 1)$ belongs to $co(S) \times \{1\} \subseteq \mathbb{R}^+(co(S) \times \{1\}\} \subseteq \mathbb{R}^{n+1}$.

By relation (1.18) the set $\mathbb{R}^+(co(S) \times \{1\}\}$ is the convex cone generated by $S \times \{1\}$ and by the first part the vector $(\mathbf{x}, 1)$ can be written as a canonical combination of at most $n+1$ linearly independent vectors $(\mathbf{x}_i, 1) \in S \times \{1\}$. Hence one can find positive scalars α_i satisfying $\mathbf{x} = \sum_{i=1}^k \alpha_i \mathbf{x}_i, k \leq n+1$ and $\sum_{i=1}^k \alpha_i = 1$ and since the vectors $(\mathbf{x}_i, 1) \in S \times \{1\}$ are linearly independent if and only if the vectors $\mathbf{x}_i \in S$ are affinely independent the desired result follows. □

Although in the above lemma $k \leq n$ for cones and $k \leq n+1$ for convex hulls it is easy to see that n can be replaced by $\dim(S) \leq n$. This concludes our discussion on algebraic properties of linear subspaces, affine sets, convex sets, and convex cones. In the next subsection we investigate topological properties of these sets.

2.2 Topological properties of sets

In this subsection we focus on the topological properties of the different classes of sets used within linear algebra and convex analysis. To start with affine sets one can show the following result. This result can be easily verified using Lemma 1.4 (cf. [46]).

Lemma 1.7 *Any affine set $M \subseteq \mathbb{R}^n$ is closed.*

An important consequence of Lemma 1.7 is given by the following observation. For a given set $S \subseteq \mathbb{R}^n$ let $int(S)$ and $cl(S)$ denote the *interior,* respectively the *closure* of the set *S*. By Lemma 1.7 we obtain $cl(S) \subseteq aff(S) \subseteq lin(S)$ and this yields by the monotonicity of the hull operation that

$$aff(cl((S))) = aff(S) \text{ and } lin(cl(S)) = lin(S). \tag{1.19}$$

Opposed to affine sets it is not true that convex cones and convex sets are closed. However, as will be shown later, the algebraic property *convexity* and the topological property *closed* are necessary and sufficient to give a so-called *dual representation* of a set. Due to this important representation one needs beforehand easy sufficient conditions on a convex set to be closed. Recall that every affine set can be seen as the *affine hull* of a *finite set of affinely independent vectors* and this property implies that every affine set is closed. By this observation it seems reasonable to consider convex sets which are the *convex hull* of a smaller set and identify which property on the smaller set *S* one needs to guarantee that the convex set $co(S)$ is closed. Looking at the following counterexample it is not sufficient to impose that the set *S* is closed and this implies that we need a stronger property on *S*.

Example 1.3 If $S = \{\mathbf{0}\} \cup \{(x, 1) : x \geq 0\}$, then S is closed and its convex hull given by $co(S) = \{(x_1, x_2) : 0 < x_2 \leq 1, x_1 \geq 0\} \cup \{\mathbf{0}\}$ is clearly not closed.

In the above counterexample the closed set S is unbounded and this prevents $co(S)$ to be closed. Imposing now the additional property that the closed set S is bounded or equivalently compact, one can show that $co(S)$ is compact and hence closed. Using relation (1.18) this also yields a way to identify for which sets S the set $cone(S)$ is closed. So finiteness of the generator S for affine sets should be replaced by compactness of S for convex hulls. To prove the next result we first introduce the so-called *unit simplex*

$$\Delta_{n+1} := \{\alpha : \sum_{i=1}^{n+1} \alpha_i = 1 \text{ and } \alpha_i \geq 0\} \subseteq \mathbb{R}^{n+1}.$$

If the function $f : \Delta_{n+1} \times S^{n+1} \to \mathbb{R}^n$ with S^k denoting the k-fold Cartesian product of the set $S \subseteq \mathbb{R}^n$ is given by

$$f(\alpha, \mathbf{x}_1, ..., \mathbf{x}_{n+1}) = \sum_{i=1}^{n+1} \alpha_i \mathbf{x}_i,$$

then by Lemma 1.6 it follows that

$$co(S) = f(\Delta_{n+1} \times S^{n+1}). \tag{1.20}$$

Using relation (1.20) one can now show the following result (cf. [34]).

Lemma 1.8 *If the nonempty set $S \subseteq \mathbb{R}^n$ is compact, then the set $co(S)$ is compact. Moreover, if S is compact and $\mathbf{0}$ does not belong to $co(S)$, then the set $cone(S)$ is closed.*

Proof. It is well known, that the set $\Delta_{n+1} \times S^{n+1}$ is compact (cf. [64]) and this shows by relation (1.20) and f a continuous function that $co(S)$ is compact. To verify the second part we observe by relation (1.18) that $cone(S) = \mathbb{R}_+(co(S))$ and so we need to show that the set $\mathbb{R}_+(co(S))$ is closed. Consider now an arbitrary sequence $t_n\mathbf{x}_n, n \in \mathbb{N}$ belonging to $\mathbb{R}_+(co(S))$ satisfying $\lim_{n\uparrow\infty} t_n\mathbf{x}_n = \mathbf{y}$. This implies $\lim_{n\uparrow\infty} t_n\|\mathbf{x}_n\| = \|\mathbf{y}\|$ and since $\mathbf{0} \notin co(S)$ and $co(S)$ is compact there exists a subsequence $N_0 \subseteq \mathbb{N}$ with $\lim_{n\in N_0\uparrow\infty} \mathbf{x}_n = \mathbf{x}_\infty \in co(S)$ and $\mathbf{x}_\infty \neq \mathbf{0}$. Hence we obtain

$$0 \leq \lim_{n\in N_0\uparrow\infty} t_n = \lim_{n\in N_0\uparrow\infty} \frac{t_n \| \mathbf{x}_n \|}{\| \mathbf{x}_n\|} = \frac{\|\mathbf{y}\|}{\| \mathbf{x}_\infty \|} := t_\infty < \infty.$$

and so $\mathbf{y} = t_\infty \mathbf{x}_\infty \in \mathbb{R}_+(\text{co}(S))$, showing the desired result. □

The following example shows that the condition $\mathbf{0} \notin co(S)$ cannot be omitted in Lemma 1.8.

Example 1.4 If the condition $\mathbf{0} \notin S$ is omitted in Lemma 1.8, then the set $cone(S)$ might not be closed as shown by the following example. Let $S = \{(x_1, x_2) : (x_1 - 1)^2 + x_2^2 \leq 1\}$. Clearly S is compact and $\mathbf{0} \in S$. Moreover, by relation (1.18) it follows that $cone(S) = \{(x_1, x_2) : x_1 > 0\} \cup \{\mathbf{0}\}$ and this set is not closed.

An immediate consequence of Caratheodory's theorem (Lemma 1.6) and Lemma 1.8 is given by the next result for convex cones generated by some nonempty set S.

Lemma 1.9 *If the set $S \subseteq \mathbb{R}^n$ contains a finite number of elements, then the set $cone(S)$ is closed.*

Proof. For the finite set S we consider the finite set $V := \{I : I \subseteq S$ and the set I consists of linearly independent vectors$\}$. By Lemma 1.6 it follows that $cone(S) = \cup_{I \in V} cone(I)$. Since each I belonging to V is a finite set of linearly independent vectors the set I is compact and $\mathbf{0}$ does not belong to co(I). This shows by Lemma 1.8 that $cone(I)$ is closed for every I belonging to V and since V is a finite set the result follows. □

Next we introduce within a finite dimensional linear space the definition of a relative interior point, generalizing the notion of an interior point. A similar notion can also be defined within a so-called (infinite dimensional) locally convex topological vector space (cf. [58]).

Definition 1.7 *If $E := \{\mathbf{x} \in \mathbb{R}^n : \|\mathbf{x}\| < 1\}$, a vector $\mathbf{x} \in \mathbb{R}^n$ is called a relative interior point of the set $S \subseteq \mathbb{R}^n$ if $\mathbf{x}$ belongs to $aff(S)$ and there exists some $\epsilon > 0$ such that*

$$(\mathbf{x} + \epsilon E) \cap aff(S) \subseteq S.$$

The relative interior $ri(S)$ of any set S is given by $ri(S) := \{\mathbf{x} \in \mathbb{R}^n : \mathbf{x}$ is a relative interior point of $S\}$. The set $S \subseteq \mathbb{R}^n$ is called relatively open if S equals $ri(S)$ and it is called regular if $ri(S)$ is nonempty.

As shown by the next example it is quite natural to assume that $\mathbf{x}$ belongs to $aff(S)$. This assumption implies that $ri(S) \subseteq S$.

Example 1.5 Consider the set $S = \{0\} \times [-1, 1] \subseteq \mathbb{R}^2$ and let $\mathbf{x} = (1,0)$. Clearly the set $aff(S)$ is given by $\{0\} \times \mathbb{R}$ and for $\epsilon = 1$ it follows that $(\mathbf{x} + E) \cap aff(S) \subseteq S$. If one would delete in the definition of a relative interior point the condition that $\mathbf{x}$ must belong to $aff(S)$, then according to this, the vector $(1, 0)$ would be a relative interior point of the set S. However, the vector $(1, 0)$ is not an element of S and so this definition is not natural.

By Definition 1.7 it is clear for $S \subseteq \mathbb{R}^n$ full dimensional or equivalently $aff(S) = \mathbb{R}^n$ that relative interior means interior and hence relative refers to relative with respect to $aff(S)$. By the same definition, we also obtain that every affine set is relatively open. Moreover, since by Lemma 1.7 the set $aff(S)$ is closed it follows that $cl(S) \subseteq aff(S)$ and so it is useless to introduce closure relative to the affine hull of a given set S. Contrary to the different hull operations the relative interior operator is not a monotone operator. This means that $S_1 \subseteq S_2$ does not imply that $ri(S_1) \subseteq ri(S_2)$.

Example 1.6 If $C_1 = \{0\}$ and $C_2 = [0, 1]$, then both sets are convex and $ri(C_1) = \{0\}$ and $ri(C_2) = (0, 1)$. This shows $C_1 \subseteq C_2$ and $ri(C_1) \nsubseteq ri(C_2)$.

To guarantee that the relative interior operator is monotone we need to impose the additional condition that $aff(S_1) = aff(S_2)$. If this holds it is easy to check that

$$S_1 \subseteq S_2 \Rightarrow ri(S_1) \subseteq ri(S_2). \tag{1.21}$$

By the above observation it is important to know which different sets cannot be distinguished by the affine hull operator. The next lemma shows that this holds for the sets S, $cl(S)$, $co(S)$ and $cl(co(S))$. This result can be easily verified using $cl(co(S)) \subseteq aff(S)$

Lemma 1.10 *It follows for every nonempty set $S \subseteq \mathbb{R}^n$ that*

$$aff(S) = aff(cl(S)) = aff(co(S)) = aff(cl(co(S))).$$

By relation (1.21) and Lemma 1.10 we obtain $ri(S) \subseteq ri(cl(S)) \subseteq ri(cl(co(S)))$ and $ri(S) \subseteq ri(co(S))$ for arbitrary sets $S \subseteq \mathbb{R}^n$. Moreover, by relation (1.7) it is easy to verify that

$$ri(S_1 \times S_2) = ri(S_1) \times ri(S_2). \tag{1.22}$$

Since we also like to show $aff(ri(S)) = aff(S)$ an alternative definition of a relative interior point is given by the next lemma.

Lemma 1.11 *If the set $S \subseteq \mathbb{R}^n$ is regular, then the vector $\mathbf{x}$ is a relative interior point of the set S if and only if $\mathbf{x}$ belongs to $aff(S)$ and there exists some $\epsilon > 0$ such that $(\mathbf{x} + \epsilon E) \cap aff(S) \subseteq ri(S)$.*

Proof. We only need to verify the if implication. Let $\mathbf{x}$ be a relative interior point of the set S. This means $\mathbf{x} \in aff(S)$ and there exists some $\epsilon > 0$ such that $(\mathbf{x}+\epsilon E) \cap aff(S) \subseteq S$. Since $\mathbf{x} \in aff(S)$ we obtain that

$(\mathbf{x}+\delta E)\cap aff(S)$ is nonempty for every $\delta > 0$ and so we may consider any point $\mathbf{y}$ belonging to $(\mathbf{x}+\frac{\epsilon}{2}E)\cap aff(S)$. Clearly $\mathbf{y}\in aff(S)$ and $\mathbf{y}+\frac{\epsilon}{2}E \subseteq \mathbf{x}+\epsilon E$ and this shows $(\mathbf{y}+\frac{\epsilon}{2}E)\cap aff(S)\subseteq S$. Hence $\mathbf{y}$ belongs to $ri(S)$ and we have verified that $(\mathbf{x}+\frac{\epsilon}{2}E)\cap aff(S)\subseteq ri(S)$. □

The next result shows for regular sets $S \subseteq \mathbb{R}^n$ that the affine hull operation cannot distinguish the sets $ri(S)$ and S and so this lemma can be seen as an extension of Lemma 1.10.

Lemma 1.12 *If the set $S \subseteq \mathbb{R}^n$ is regular, then it follows that*

$$aff(ri(S)) = aff(S).$$

Proof. It is clear that $aff(ri(S)) \subseteq aff(S)$ and to show the converse inclusion it is sufficient to verify that $S\backslash ri(S) \subseteq aff(ri(S))$. Let $\mathbf{x}\in S\backslash ri(S)$. Since the set S is regular one can find some $\mathbf{y}\in ri(S)\subseteq S$ and so by Lemma 1.11 there exists some $\epsilon > 0$ satisfying

$$(\mathbf{y}+\epsilon E)\cap aff(S)\subseteq ri(S). \tag{1.23}$$

Clearly the set $[\mathbf{y},\mathbf{x}] := \{(1-\alpha)\mathbf{y}+\alpha\mathbf{x} : 0 \leq \alpha \leq 1\}$ belongs to $co(S) \subseteq aff(S)$ and this implies by relation (1.23) that $(\mathbf{y}+\epsilon E)\cap[\mathbf{y},\mathbf{x}]\subseteq ri(S)$. This means that the halfline starting in $\mathbf{y}$ and passing through $\mathbf{x}_1 \in (\mathbf{y}+\epsilon E)\cap[\mathbf{y},\mathbf{x}]\subseteq ri(S)$ is a subset of $aff(ri(S))$ and contains $\mathbf{x}$. Hence $\mathbf{x}$ belongs to $aff(ri(S))$ and so $S\backslash ri(S)\subseteq aff(ri(S))$. □

An immediate consequence of Lemmas 1.11 and 1.12 is given by the observation that for any regular set $S \subseteq \mathbb{R}^n$ it follows that $\mathbf{x}$ is a relative interior point of S if and only if $\mathbf{x}$ belongs to $aff(ri(S))$ and there exists some $\epsilon > 0$ satisfying $(\mathbf{x}+\epsilon E)\cap aff(ri(S))\subseteq ri(S)$. This implies for every regular set $S \subseteq \mathbb{R}^n$ that $ri(ri(S)) = ri(S)$, and since by definition $ri(\emptyset) = \emptyset$, we obtain for any set $S \subseteq \mathbb{R}^n$ that

$$ri(ri(S)) = ri(S). \tag{1.24}$$

Keeping in mind the close relationship between affine hulls and convex sets and the observation that nonempty affine sets are regular (in fact $ri(M) = M$!) we might wonder whether convex sets are regular. This is indeed the case as the following result shows (cf. [63]).

Lemma 1.13 *Every nonempty convex set $C \subseteq \mathbb{R}^n$ is regular.*

Although convexity is not a necessary condition for a set to be regular it follows by the definition of a regular set that at least around any relative interior point the set must be "locally" convex. A set, which clearly violates this condition, is the set $\mathbb{Q}$ of rational numbers and this set is therefore not regular. Besides convexity of the set C the proof of Lemma 1.13 uses also that C is a subset of a finite dimensional linear space. If the last condition does not hold and C is an infinite dimensional convex subset of a locally convex topological vector space, then the above result might not hold. We will now list some important properties of relative interiors. To start with this, we first verify the following technical result.

Lemma 1.14 *If $S_1, S_2 \subseteq \mathbb{R}^n$ are nonempty sets, then it follows for every $0 < \alpha < 1$ that*

$$(\alpha S_1 + (1-\alpha)S_2) \cap aff(S_1) \subseteq \alpha S_1 + (1-\alpha)(S_2 \cap aff(S_1)).$$

Proof. Consider for $0 < \alpha < 1$ the vector $\mathbf{y} = \alpha \mathbf{x}_1 + (1-\alpha)\mathbf{x}_2$ with $\mathbf{x}_i \in S_i$, $i = 1, 2$ and $\mathbf{y} \in aff(S_1)$. It is now necessary to verify that $\mathbf{x}_2$ belongs to $S_2 \cap aff(S_1)$. By the definition of $\mathbf{y}$ and $0 < \alpha < 1$ we obtain that

$$\mathbf{x}_2 = \frac{1}{1-\alpha}\mathbf{y} - \frac{\alpha}{1-\alpha}\mathbf{x}_1 \in \frac{1}{1-\alpha} aff(S_1) - \frac{\alpha}{1-\alpha} S_1,$$

and so it follows that $\mathbf{x}_2$ belongs to $aff(S_1)$. Hence the vector $\mathbf{x}_2$ belongs to $S_2 \cap aff(S_1)$ and this shows the desired result. □

Applying now Lemma 1.14, the next important result for convex sets can be shown. This result will play an prominent role in verifying the topological properties of convex sets.

Lemma 1.15 *If $C \subseteq \mathbb{R}^n$ is a nonempty convex set, then it follows for every $0 \leq \alpha < 1$ that*

$$\alpha cl(C) + (1-\alpha) ri(C) \subseteq ri(C).$$

Proof. To prove the above result it is sufficient to show that $\alpha cl(C) + (1-\alpha)\mathbf{x}_2 \subseteq ri(C)$ for any $\mathbf{x}_2 \in ri(C)$ and $0 < \alpha < 1$. Clearly this set is a subset of $aff(C)$ and since $\mathbf{x}_2$ belongs to $ri(C) \subseteq C$ there exists some $\epsilon > 0$ satisfying

$$(\mathbf{x}_2 + \frac{(1+\alpha)\epsilon}{1-\alpha} E) \cap aff(C) \subseteq C. \tag{1.25}$$

Moreover, since $cl(C) \subseteq \cap_{\epsilon > 0}(C + \epsilon E)$ it follows that $cl(C) \subseteq C + \epsilon E$, and this implies

$$\alpha cl(C) + (1-\alpha)\mathbf{x}_2 + \epsilon E \subseteq \alpha C + (1-\alpha)(\mathbf{x}_2 + \frac{(1+\alpha)\epsilon}{1-\alpha} E).$$

Applying now Lemma 1.14 and relation (1.25) we obtain by the convexity of the set C that

$$(\alpha cl(C) + (1-\alpha)\mathbf{x}_2 + \epsilon E) \cap aff(C) \subseteq \alpha C + (1-\alpha)C \subseteq C$$

and this shows the result. □

By Lemmas 1.13 and 1.15 it follows for any nonempty convex set C that the set $ri(C)$ is nonempty and convex. Also, since $cl(C) = \cap_{\epsilon>0}(C + \epsilon E)$ we obtain that $cl(C)$ is a convex set. An easy and important consequence of Lemma 1.15 is given by the observation that the relative interior operator cannot distinguish the convex sets C and $cl(C)$. A similar observation holds for the closure operator applied to the convex sets $ri(C)$ and C. The next result also plays an important role in the proof of the weak separation result to be discussed in Subsection 2.3.

Lemma 1.16 *If $C \subseteq \mathbb{R}^n$ is a nonempty convex set, then it follows that*

$$cl(ri(C)) = cl(C) \text{ and } ri(C) = ri(cl(C)).$$

Proof. To prove the first formula we only need to check that $cl(C) \subseteq cl(ri(C))$. To verify this we consider $\mathbf{x} \in cl(C)$ and select some $\mathbf{y}$ belonging to ri(C). By Lemma 1.15 the half-open line segment $[\mathbf{y}, \mathbf{x})$ belongs to $ri(C)$ and this implies that the vector $\mathbf{x}$ belongs to $cl(ri(C))$. Hence $cl(C) \subseteq cl(ri(C))$ and the first formula is verified. To prove the second formula, it follows immediately by relation (1.21) that $ri(C) \subseteq ri(cl(C))$. To verify $ri(cl(C)) \subseteq ri(C)$ consider an arbitrary $\mathbf{x}$ belonging to $ri(cl(C))$ and so one can find some $\epsilon > 0$ satisfying

$$(\mathbf{x} + \epsilon E) \cap aff(cl(C)) \subseteq cl(C). \tag{1.26}$$

Moreover, since $ri(C)$ is nonempty, construct for some $\mathbf{y} \in ri(C)$ the line $T := \{(1-t)\mathbf{x} + t\mathbf{y} : t \in \mathbb{R}\}$ through the points $\mathbf{x}$ and $\mathbf{y}$. Since $\mathbf{x} \in ri(cl(C))$ and $\mathbf{y} \in ri(C)$ it follows that $T \subseteq aff(cl(C))$ and so by relation (1.26) there exists some $\mu < 0$ satisfying $\mathbf{y}_1 := (1-\mu)\mathbf{x} + \mu\mathbf{y} \in cl(C)$. This shows

$$\mathbf{x} = \frac{1}{1-\mu}\mathbf{y}_1 - \frac{\mu}{1-\mu}\mathbf{y}, \tag{1.27}$$

and since $\mathbf{y}_1 \in cl(C)$ and $\mathbf{y} \in ri(C)$ this implies by Lemma 1.15 and relation (1.27) that $\mathbf{x} \in ri(C)$. Hence it follows that $ri(cl(C)) \subseteq ri(C)$, and this proves the second formula. □

In the above lemma one might wonder whether the convexity of the set C is necessary. In the following example we present a regular set

S with $ri(S)$ and $cl(S)$ convex and S not convex and this set does not satisfy the result of Lemma 1.16.

Example 1.7 Let $S = [0,1] \cup ((1,2] \cap \mathbb{Q})$. This set is clearly not convex and $ri(S) = (0,1)$ while $cl(S) = [0,2]$. Moreover, $ri(cl(S)) \neq ri(S)$ and $cl(ri(S)) \neq cl(S)$.

We will now give a primal representation of the relative interior of a convex set S (cf. [63]).

Lemma 1.17 *If $S \subseteq \mathbb{R}^n$ is a nonempty convex set, then it follows that*

$$ri(S) = \{\mathbf{x} \in \mathbb{R}^n : \forall_{\mathbf{y} \in cl(S)}\ \exists_{\mu < 0} \text{ such that } (1-\mu)\mathbf{x} + \mu\mathbf{y} \in S\}.$$

The above result is equivalent to the geometrically obvious fact that for S a convex set and any $\mathbf{x} \in ri(S)$ and $\mathbf{y} \in S$ the line segment $[\mathbf{y}, \mathbf{x}]$ can be extended beyond $\mathbf{x}$ without leaving S. Also, by relation (1.24) and Lemma 1.16 another primal representation of $ri(S)$ with S a convex set is given by

$$ri(S) = \{\mathbf{x} \in \mathbb{R}^n : \forall_{\mathbf{y} \in cl(S)}\ \exists_{\mu < 0} \text{ such that } (1-\mu)\mathbf{x} + \mu\mathbf{y} \in ri(S)\}.$$

Since affine mappings preserve convexity it is also of interest to know how the relative interior operator behaves under an affine mapping. Using Lemma 1.17 one can show the next result (cf. [63]).

Lemma 1.18 *If $A : \mathbb{R}^n \to \mathbb{R}^m$ is an affine mapping and $C \subseteq \mathbb{R}^n$ is a nonempty convex set, then it follows that $A(ri(C)) = ri(A(C))$. Moreover, if $C \subseteq \mathbb{R}^m$ is a nonempty convex set satisfying $A^{-1}(ri(C)) := \{\mathbf{x} \in \mathbb{R}^n : A(\mathbf{x}) \in ri(C)\}$ is nonempty, then $ri(A^{-1}(C)) = A^{-1}(ri(C))$.*

As shown by the following counterexample the condition $A^{-1}(ri(C))$ is nonempty cannot be omitted in the previous lemma.

Example 1.8 Let $A : \mathbb{R} \to \mathbb{R}$ given by $A(x) = 1$ for all $x \in \mathbb{R}$ and let $C := [0,1] \subset \mathbb{R}$. Then clearly $ri(C) = (0,1)$, $A^{-1}(ri(C)) = \emptyset$ and $ri(A^{-1}(C)) = \mathbb{R}$.

An immediate consequence of Lemma 1.18 is given by the observation that

$$ri(\alpha S_1 + \beta S_2) = \alpha ri(S_1) + \beta ri(S_2), \tag{1.28}$$

for any $\alpha, \beta \in \mathbb{R}$ and $S_i \subseteq \mathbb{R}^n, i = 1,2$ convex sets. To conclude our discussion on topological properties for sets we finally mention the following result (cf. [63]).

Lemma 1.19 *If the sets C_i, $i \in I$ are convex and $\cap_{i\in I} ri(C_i)$ is nonempty, then it follows that $cl(\cap_{i\in I} C_i) = \cap_{i\in I} cl(C_i)$. Moreover, if the set I is finite, we obtain $ri\,(\cap_{i\in I} C_i) = \cap_{i\in I} ri(C_i)$.*

As shown by the next counterexample it is necessary to assume in Lemma 1.19 that the intersection $\cap_{i\in I} ri(C_i)$ is nonempty.

Example 1.9 Let $C_1 = \{\mathbf{x} \in \mathbb{R}^2 : x_1 > 0, x_2 > 0\} \cup \{\mathbf{0}\}$ and $C_2 = \{\mathbf{x} \in \mathbb{R}^2 : x_2 = 0\}$. It is obvious that $ri(C_1) = \{\mathbf{x} : x_1 > 0, x_2 > 0\}$ and $ri(C_2) = C_2$, and so we obtain $ri(C_1) \cap ri(C_2) = \varnothing$ and $ri(C_1 \cap C_2) \neq ri(C_1) \cap ri(C_2)$. For the same example it is also easy to see that $cl(C_1 \cap C_2) \neq cl(C_1) \cap cl(C_2)$.

In the following counterexample we show that the second result listed in Lemma 1.19 does not hold if the set I is not finite.

Example 1.10 Let $I = (0, \infty)$ and $C_\alpha = [0, 1+\alpha], \alpha > 0$. For this example it follows $ri(\cap_{\alpha>0} C_\alpha) = ri([0,1]) = (0,1)$, and since $ri(C_\alpha) = (0, 1+\alpha)$ for each $\alpha > 0$, we obtain $\cap_{\alpha>0} ri(C_\alpha) = (0,1]$.

This last example concludes our discussion of topological properties of convex sets. In the next subsection we will discuss basic separation results for those sets.

2.3 Separation of convex sets

For a nonempty convex set $C \subseteq \mathbb{R}^n$ consider for any $\mathbf{y} \in \mathbb{R}^n$ the so-called *minimum norm problem* given by

$$v(\mathbf{y}) := \inf\{\| \mathbf{x} - \mathbf{y} \|^2 : \mathbf{x} \in C\}. \tag{$P(\mathrm{y})$}$$

If additionally C is closed, a standard application of the Weierstrass theorem (cf. [64]) shows that for every $\mathbf{y}$ the optimal objective value $v(\mathbf{y})$ in the above optimization problem is attained. To verify that the minimum norm problem has a unique solution, observe for any $\mathbf{z}_1, \mathbf{z}_2$ belonging to $\mathbb{R}^n$ that

$$\|\mathbf{z}_1 + \mathbf{z}_2\|^2 + \|\mathbf{z}_1 - \mathbf{z}_2\|^2 = 2\|\mathbf{z}_1\|^2 + 2\|\mathbf{z}_2\|^2. \tag{1.29}$$

For every $\mathbf{x}_1 \neq \mathbf{x}_2$ belonging to C it follows by relation (1.29) with $\mathbf{z}_i$ replaced by $\mathbf{x}_i - \mathbf{y}$ for $i = 1, 2$ that

$$\|\frac{1}{2}(\mathbf{x}_1 + \mathbf{x}_2) - \mathbf{y}\|^2 < \frac{1}{2}\|\mathbf{x}_1 - \mathbf{y}\|^2 + \frac{1}{2}\|\mathbf{x}_2 - \mathbf{y}\|^2,$$

and so for $\mathbf{x}_i, i = 1, 2$ different optimal solutions of the minimum norm problem $(P(\mathrm{y}))$ we obtain that $\|\frac{1}{2}(\mathbf{x}_1 + \mathbf{x}_2) - \mathbf{y}\|^2 < v(\mathbf{y})$. Since the set C

is convex and hence $\frac{1}{2}(\mathbf{x}_1 + \mathbf{x}_2)$ belongs to C, this yields a contradiction and the optimal solution is therefore unique. Denoting now this optimal solution by $p_C(\mathbf{y})$ one can show the following result (cf. [34]).

Lemma 1.20 *For any* $\mathbf{y} \in \mathbb{R}^n$ *and* $C \subseteq \mathbb{R}^n$ *a nonempty closed convex set it follows that*

$$\mathbf{z} = p_C(\mathbf{y}) \Leftrightarrow \mathbf{z} \in C \textit{ and } (\mathbf{z} - \mathbf{y})^{\intercal}(\mathbf{x} - \mathbf{z}) \geq 0 \textit{ for every } \mathbf{x} \in C.$$

Moreover, for every $\mathbf{x} \in C$ *the triangle inequality*

$$\|\mathbf{x} - p_C(\mathbf{y})\|^2 + \|p_C(\mathbf{y}) - \mathbf{y}\|^2 \leq \|\mathbf{x} - \mathbf{y}\|^2$$

holds.

Proof. To show the only if implication we observe that

$$0 \leq (\mathbf{z} - \mathbf{y})^{\intercal}(\mathbf{x} - \mathbf{z}) = -\|\mathbf{z} - \mathbf{y}\|^2 + (\mathbf{z} - \mathbf{y})^{\top}(\mathbf{x} - \mathbf{y})$$

and this shows by the Cauchy-Schwarz inequality (cf. [46])

$$0 \leq (\mathbf{z} - \mathbf{y})^{\intercal}(\mathbf{x} - \mathbf{z}) \leq -\|\mathbf{z} - \mathbf{y}\|^2 + \|\mathbf{z} - \mathbf{y}\|\|\mathbf{x} - \mathbf{y}\| \tag{1.30}$$

for every $\mathbf{x} \in C$. If $\mathbf{y} \in C$ we obtain, substituting $\mathbf{x} = \mathbf{y}$ in relation (1.30), that $0 \leq -\|\mathbf{z} - \mathbf{y}\|^2$ and by the nonnegativity of $\|.\|^2$ this yields $0 = \|\mathbf{z} - \mathbf{y}\|^2$. Also, using $\mathbf{y} \in C$, we obtain $\mathbf{y} = p_C(\mathbf{y})$ and so $\mathbf{z} = \mathbf{y} = p_C(\mathbf{y})$. Moreover, if $\mathbf{y} \notin C$, then $\|\mathbf{z} - \mathbf{y}\| > 0$ and this implies by relation (1.30) that $\|\mathbf{z} - \mathbf{y}\| \leq \|\mathbf{x} - \mathbf{y}\|$ for every $\mathbf{x} \in C$. Hence $\mathbf{z}$ is an optimal solution and by the uniqueness of this solution we obtain $\mathbf{z} = p_C(\mathbf{y})$. To verify the if implication, it follows for $\mathbf{z} = p_C(\mathbf{y})$ that $\mathbf{z} \in C$ and since C is convex this shows

$$\|\mathbf{z} - \mathbf{y}\|^2 \leq \|\alpha\mathbf{x} + (1 - \alpha)\mathbf{z} - \mathbf{y}\|^2 = \|\mathbf{z} - \mathbf{y} + \alpha(\mathbf{x} - \mathbf{z})\|^2 \tag{1.31}$$

for every $\mathbf{x} \in C$ and $0 < \alpha < 1$. Rewriting relation (1.31) we obtain for every $0 < \alpha < 1$ that $2(\mathbf{z} - \mathbf{y})^{\top}(\mathbf{x} - \mathbf{z}) + \alpha\|\mathbf{x} - \mathbf{z}\|^2 \geq 0$ and letting $\alpha \downarrow 0$ the desired inequality follows. To show the triangle inequality, we observe using $\|\mathbf{z}_1\|^2 - \|\mathbf{z}_2\|^2 =< \mathbf{z}_1 - \mathbf{z}_2, \mathbf{z}_1 + \mathbf{z}_2 >$ for every $\mathbf{z}_1, \mathbf{z}_2$ that

$$\|\mathbf{x} - p_C(\mathbf{y})\|^2 - \|\mathbf{x} - \mathbf{y}\|^2 =< \mathbf{y} - p_C(\mathbf{y}), 2\mathbf{x} - \mathbf{y} - p_C(\mathbf{y}) > .$$

The last term equals $-\|p_C(\mathbf{y}) - \mathbf{y})\|^2 + 2 < \mathbf{y} - p_C(\mathbf{y}), \mathbf{x} - p_C(\mathbf{y}) >$ and applying now the first part yields the desired inequality. □

Actually the above result is nothing else than the first order necessary and sufficient condition for a minimum of a convex function on a closed convex set. We will now prove one of the most fundamental results in convex analysis. This result has an obvious geometric interpretation and serves as a basic tool in deriving dual representations. In infinite dimensional locally convex topological vector spaces the next result is also known as the Hahn-Banach theorem (cf. [65]).

Theorem 1.1 *If* $C \subseteq \mathbb{R}^n$ *is a nonempty convex set and* $\mathbf{y}$ *does not belong to the set* $cl(C)$, *then there exists some nonzero vector* $\mathbf{y}_0 \in \mathbb{R}^n$ *and* $\epsilon > 0$ *with* $\mathbf{y}_0^\top \mathbf{x} \geq \mathbf{y}_0^\top \mathbf{y} + \epsilon$ *for every* $\mathbf{x}$ *belonging to* $cl(C)$. *In particular, the vector* $\mathbf{y}_0 := p_{cl(C)}(\mathbf{y}) - \mathbf{y}$ *satisfies this inequality.*

Proof. By Lemma 1.20 we obtain for every $\mathbf{x} \in cl(C)$ and the nonzero vector $\mathbf{y}_0 := p_{cl(C)}(\mathbf{y}) - \mathbf{y}$ that $\mathbf{y}_0^\top \mathbf{x} \geq \mathbf{y}_0{}^\top p_{cl(C)}(\mathbf{y})$. This shows

$$\mathbf{y}_0^\top \mathbf{x} \geq \|\mathbf{y}_0\|^2 + \mathbf{y}_0^\top \mathbf{y} \tag{1.32}$$

and since $\mathbf{y}_0 \neq \mathbf{0}$ the desired result follows. □

The nonzero vector $\mathbf{y}_0$ belonging to $cl(C) - \mathbf{y}$ is called the *normal vector* of the *separating hyperplane*

$$H^{=}(\mathbf{a}, a) := \{\mathbf{x} \in \mathbb{R}^n : \mathbf{a}^\top \mathbf{x} = a\},$$

$\mathbf{a} = \mathbf{y}_0$ and $a = \mathbf{y}_0{}^\top \mathbf{y} + \frac{\epsilon}{2}$, and this hyperplane strongly separates the closed convex set $cl(C)$ and $\mathbf{y}$. Since $\mathbf{y}_0 \neq \mathbf{0}$ we may take as a normal vector of the hyperplane the vector $\mathbf{y}_0\|\mathbf{y}_0\|^{-1}$ and this vector has norm 1 and belongs to $cone(cl(C) - \mathbf{y})$.

The strong separation result of Theorem 1.1 can be used to prove the following "weaker" separation result valid under a weaker condition on the point $\mathbf{y}$. Instead of $\mathbf{y}$ does not belong to $cl(C)$ we assume that $\mathbf{y}$ does not belong to $\text{ri}(C)$. By Theorem 1.1 it is clear that we may assume without loss of generality that $\mathbf{y}$ belongs to the *relative boundary* $rbd(C) := cl(C) \backslash ri(C)$ of the convex set $C \subseteq \mathbb{R}^n$.

Theorem 1.2 *If* $C \subseteq \mathbb{R}^n$ *is a nonempty convex set and* $\mathbf{y}$ *does not belong to* $ri(C)$, *then there exists some nonzero vector* $\mathbf{y}_0$ *belonging to the unique linear subspace* $L_{aff(C)}$ *satisfying* $\mathbf{y}_0^\top \mathbf{x} \geq \mathbf{y}_0^\top \mathbf{y}$ *for every* $\mathbf{x} \in C$. *Moreover, for the vector* $\mathbf{y}_0$ *there exists some* $\mathbf{x}_0 \in C$ *such that* $\mathbf{y}_0^\top \mathbf{x}_0 > \mathbf{y}_0^\top \mathbf{y}$.

Proof. Consider for every $n \in \mathbb{N}$ the set $(\mathbf{y} + n^{-1}E) \cap aff(cl(C))$. By Lemma 1.16 it follows that $\mathbf{y}$ does not belong to $ri(cl(C))$ and so there exists some vector $\mathbf{y}_n$ satisfying

$$\mathbf{y}_n \notin cl(C) \text{ and } \mathbf{y}_n \in (\mathbf{y} + n^{-1}E) \cap aff(cl(C)). \tag{1.33}$$

The set $cl(C)$ is a closed convex set and by relation (1.33) and Theorem 1.1 one can find some vector $\mathbf{y}_n^* \in \mathbb{R}^n$ satisfying

$$\| \mathbf{y}_n^* \| = 1,\ \mathbf{y}_n^* \in cone(cl(C) - \mathbf{y}_n) \subseteq L_{aff(C)} \text{ and } \mathbf{y}_n^{*\top}\mathbf{x} \geq \mathbf{y}_n^{*\top}\mathbf{y}_n \quad (1.34)$$

for every $\mathbf{x} \in cl(C)$. The sequence $\{\mathbf{y}_n^* : n \in \mathbb{N}\}$ belongs to a compact set and so there exists a convergent subsequence $\{\mathbf{y}_n^* : n \in N_0\}$ with

$$\lim_{n \in N_0 \uparrow \infty} \mathbf{y}_n^* = \mathbf{y}_0. \quad (1.35)$$

This implies by relations (1.33), (1.34) and (1.35) that

$$\mathbf{y}_0^\top \mathbf{x} = \lim_{n \in N_0 \uparrow \infty} \mathbf{y}_n^{*\top}\mathbf{x} \geq \lim_{n \in N_0 \uparrow \infty} \mathbf{y}_n^{*\top}\mathbf{y}_n = \mathbf{y}_0^\top \mathbf{y} \quad (1.36)$$

for every $\mathbf{x} \in cl(C)$ and

$$\mathbf{y}_0 \in L_{aff(C)} \text{ and } \| \mathbf{y}_0 \| = 1. \quad (1.37)$$

Suppose now that there does not exist some $\mathbf{x}_0 \in C$ satisfying $\mathbf{y}_0^\top \mathbf{x}_0 > \mathbf{y}_0^\top \mathbf{y}$. By relation (1.36) this implies that $\mathbf{y}_0^\top(\mathbf{x} - \mathbf{y}) = 0$ for every $\mathbf{x} \in C$ and since $\mathbf{y}$ belongs to $\text{cl}(C) \subseteq \text{aff}(C)$ we obtain by relation (1.4) and Lemma 1.3 that $\mathbf{y}_0^\top \mathbf{z} = 0$ for every $\mathbf{z}$ belonging to $L_{aff(C)}$. Since by relation (1.37) the vector $\mathbf{y}_0$ belongs to $L_{aff(C)}$ this implies $\| \mathbf{y}_0 \|^2 = 0$ and so we contradict $\| \mathbf{y}_0 \| = 1$. Hence it must follow that there exists some $\mathbf{x}_0 \in C$ satisfying $\mathbf{y}_0^\top \mathbf{x}_0 > \mathbf{y}_0^\top \mathbf{y}$ and this proves the desired result.

□

The separation of Theorem 1.2 is called a *proper separation* between the set C and the vector $\mathbf{y}$. One can also introduce proper separation between two convex sets.

Definition 1.8 *The convex sets $C_1, C_2 \subseteq \mathbb{R}^n$ are called properly separated if there exist some $\mathbf{y}_0 \in \mathbb{R}^n$ satisfying*

$$\inf_{x \in C_1} \mathbf{y}_0^\top \mathbf{x} \geq \sup_{\mathbf{x} \in C_2} \mathbf{y}_0^\top \mathbf{x} \text{ and } \mathbf{y}_0^\top \mathbf{x}_1 > \mathbf{y}_0^\top \mathbf{x}_2$$

for some $\mathbf{x}_1 \in C_1$ and $\mathbf{x}_2 \in C_2$.

An immediate consequence of Theorem 1.2 is given by the next result.

Theorem 1.3 *If the convex sets $C_1, C_2 \subseteq \mathbb{R}^n$ satisfy $ri(C_1) \cap ri(C_2) = \emptyset$, then the two sets can be properly separated.*

Proof. By relation (1.28) we obtain for $\alpha = 1$ and $\beta = -1$ that $ri(C_1 - C_2) = ri(C_1) - ri(C_2)$, and this shows $ri(C_1) \cap ri(C_2) = \emptyset$ if and only if

$\mathbf{0} \notin ri(C_1 - C_2)$. Applying now Theorem 1.2 with $\mathbf{y} = \mathbf{0}$ and the convex set given by $C_1 - C_2$, the result follows. □

The above separation results are the corner stones of convex and quasiconvex analysis. Observe in infinite dimensional locally convex topological vector spaces one can show similar separation results under stronger assumptions on the convex sets C_1 and C_2 (cf. [65],[17],[58]). An easy consequence of the separation results is given by the observation that closed convex sets and relatively open convex sets are evenly convex. These convex sets play an important role in duality theory for quasiconvex functions.

Lemma 1.21 *If the nonempty convex set $C \subseteq \mathbb{R}^n$ is closed or relatively open, then C is evenly convex.*

Proof. If $C = \mathbb{R}^n$, the result follows by definition and so we may suppose that the closed set C is a proper subset of $\mathbb{R}^n$. Hence there exists some $\mathbf{y} \notin C$ and this implies by Theorem 1.1 that there exists some $\mathbf{a} \in \mathbb{R}^n$ and $b \in \mathbb{R}$ satisfying $C \subseteq H^<(\mathbf{a}, b)$. This shows that the set $\mathcal{H}_C$ of all open halfspaces H satisfying $C \subseteq H$ is nonempty and by the definition of $\mathcal{H}_C$ it is clear that $C \subseteq \cap\{H : H \in \mathcal{H}_C\}$. Again by Theorem 1.1 one can show using contradiction that C equals $\cap\{H : H \in \mathcal{H}_C\}$ and this shows that every closed convex set is evenly convex. To verify the second result, we observe $\mathcal{H}_{cl(C)} \subseteq \mathcal{H}_C$ and since $\mathcal{H}_{cl(C)}$ is nonempty by the first part, it follows that $\mathcal{H}_C$ is nonempty and $C \subseteq \cap\{H : H \in \mathcal{H}_C\}$. To show that $C = \cap\{H : H \in \mathcal{H}_C\}$ we assume by contradiction that there exists some $\mathbf{y} \notin C$ with $\mathbf{y} \in H$ for every $H \in \mathcal{H}_C$. Due to $\mathbf{y} \notin C$ it follows by Theorem 1.2 that there exists some nonzero $\mathbf{y}_0 \in L_{aff(C)}$ satisfying

$$\mathbf{y}_0^\top \mathbf{x} \geq \mathbf{y}_0^\top \mathbf{y} \tag{1.38}$$

for every $\mathbf{x} \in C$. Since the convex set C is relatively open there exists for every $\mathbf{x} \in C$ some $\epsilon > 0$ satisfying $\mathbf{x} - \epsilon\mathbf{y}_0 \in C$ and so by relation (1.38) we obtain for every $\mathbf{x} \in C$ that $\mathbf{y}_0^\top \mathbf{x} = \mathbf{y}_0^\top(\mathbf{x} - \epsilon\mathbf{y}_0) + \epsilon\|\mathbf{y}_0\|^2 > \mathbf{y}_0^\top \mathbf{y}$. Hence the open halfspace $H^<(\mathbf{a}, b) := \{\mathbf{x} \in \mathbb{R}^n : \mathbf{a}^\top \mathbf{x} < b\}$ with $\mathbf{a} := -\mathbf{y}_0$ and $b := -\mathbf{y}_0^\top \mathbf{y}$ belongs to $\mathcal{H}_C$ and since $\mathbf{y} \notin H^<(\mathbf{a}, b)$ this contradicts $\mathbf{y} \in H$ for every H belonging to $\mathcal{H}_C$. □

This concludes our discussion of separation results of convex sets. In the next subsection we will use these separation results to derive dual representations for convex sets.

2.4 Dual representations of convex sets

In contrast to the primal representation of a linear subspace, affine set, convex cone and convex set discussed in Subsection 2.1 we can also give a so-called *dual representation* of these sets. From a geometrical point of view a primal representation is a representation from "within" the set, while a dual representation turns out to be a representation from "outside" the set. Such a characterization can be seen as an improvement of the hull operation given by relations (1.2), (1.3), (1.16) and (1.17). We start with linear subspaces or affine sets (cf. [47]).

Definition 1.9 *If $S \subseteq \mathbb{R}^n$ is some nonempty set, then the nonempty set $S^{\perp} \subseteq \mathbb{R}^n$ given by $S^{\perp} := \{\mathbf{x}^* \in \mathbb{R}^n : \mathbf{x}^\intercal \mathbf{x}^* = 0$ for every $\mathbf{x} \in S\}$ is called the orthogonal complement of the set S.*

It is easy to verify that the orthogonal complement $S^{\perp}$ of the set S is a linear subspace. Moreover, a basic result (cf. [47]) in linear algebra is given by the following.

Lemma 1.22 *For any linear subspace L it follows that $(L^{\perp})^{\perp} = L$.*

By Lemma 1.22 a so-called dual representation of any linear hull $lin(S)$ with S nonempty can be constructed. Using $S \subseteq lin(S)$ it follows by Lemma 1.22 that $(S^{\perp})^{\perp} \subseteq (lin(S)^{\perp})^{\perp} = lin(S)$. Since $lin(S)$ is the smallest linear subspace containing S and $(S^{\perp})^{\perp}$ is clearly a linear subspace containing S the previous inclusion implies

$$(S^{\perp})^{\perp} = lin(S). \tag{1.39}$$

The alternative representation of $lin(S)$ in relation (1.39) is called a dual representation. To construct a dual representation for an affine hull we observe by Lemma 1.3 and the dual representation of a linear hull that $aff(S) = \mathbf{x}_0 + ((S - \mathbf{x}_0)^{\perp})^{\perp}$ for $\mathbf{x}_0$ belonging to $aff(S)$. Since it is easy to verify that $(S - \mathbf{x}_0)^{\perp} = (S - \mathbf{x}_1)^{\perp}$ for every $\mathbf{x}_1 \in aff(S)$ we obtain for affine hulls the dual representation

$$aff(S) = \mathbf{x}_0 + ((S - \mathbf{x}_1)^{\perp})^{\perp} \tag{1.40}$$

for every $\mathbf{x}_0, \mathbf{x}_1 \in aff(S)$.

Next we discuss the dual representation of a closed convex set containing $\mathbf{0}$ and a closed convex cone. This dual representation will be verified by means of the strong separation result listed in Theorem 1.1. Recall first the definition of a support function.

Definition 1.10 *If* $S \subseteq \mathbb{R}^n$ *is* some *nonempty set, then the function* $\sigma_S : \mathbb{R}^n \to (-\infty, \infty]$ *given by* $\sigma_S(\mathbf{s}) := \sup\{\mathbf{s}^\intercal \mathbf{x} : \mathbf{x} \in S\}$ *is called the support function of the set S.*

An equivalent formulation of Theorem 1.1 involving the support function of the closed convex set C is given by the following result.

Theorem 1.4 *If* $C \subseteq \mathbb{R}^n$ *is a proper nonempty convex set, then it follows that* $\mathbf{x}_0 \in cl(C)$ *if and only if* $\mathbf{s}^\top \mathbf{x}_0 \leq \sigma_{cl(C)}(\mathbf{s})$ *for every* $\mathbf{s} \in \mathbb{R}^n$.

Proof. Clearly $\mathbf{x}_0 \in cl(C)$ implies that $\mathbf{s}^\intercal \mathbf{x}_0 \leq \sigma_{cl(C)}(\mathbf{s})$ for every $\mathbf{s}$ belonging to $\mathbb{R}^n$. To show the reverse implication let $\mathbf{s}^\intercal \mathbf{x}_0 \leq \sigma_{cl(C)}(\mathbf{s})$ for every $\mathbf{s} \in \mathbb{R}^n$ and suppose by contradiction that $\mathbf{x}_0 \notin cl(C)$. By Theorem 1.1 there exists some nonzero vector $\mathbf{y}_0 \in \mathbb{R}^n$ and $\epsilon > 0$ satisfying $-\mathbf{y}_0^\top \mathbf{x} \leq -\mathbf{y}_0^\top \mathbf{x}_0 - \epsilon$ for every $\mathbf{x}$ belonging to $cl(C)$. This implies $\sigma_{cl(C)}(-\mathbf{y}_0) \leq -\mathbf{y}_0^\top \mathbf{x}_0 - \epsilon < -\mathbf{y}_0^\top \mathbf{x}_0$, contradicting our initial assumption and so it must follow that $\mathbf{x}_0$ belongs to $cl(C)$. □

To generalize the dual representation of linear subspaces in Lemma 1.22 to the larger class of closed convex sets containing $\mathbf{0}$ we need to generalize the orthogonality relation given in Definition 1.9.

Definition 1.11 *If* $S \subseteq \mathbb{R}^n$ is a *nonempty set, then the set* S^0, *given by* $S^0 := \{\mathbf{x}^* \in \mathbb{R}^n : \mathbf{x}^\intercal \mathbf{x}^* \leq 1 \text{ for every } \mathbf{x} \in S\}$, *is called the polar of the set S. Moreover, the bipolar* S^{00} *of the set S is defined by* $S^{00} := (S^0)^0$.

The polar S^0 of a nonempty set $S \subseteq \mathbb{R}^n$ is a nonempty closed convex set and satisfies $S^0 = (cl(S))^0$. If the nonempty set $K \subseteq \mathbb{R}^n$ is a convex cone, then it is easy to show that $K^0 = \{\mathbf{x}^* \in \mathbb{R}^n : \mathbf{x}^\intercal \mathbf{x}^* \leq 0 \text{ for every } \mathbf{x} \in K\}$ and K^0 is a closed convex cone, while for L a linear subspace it follows that $L^0 = L^\perp$. Hence the *polar* operator applied to a linear subspace reduces to the *orthogonal* operator and can therefore be seen as a generalization of this operator. To prove a generalization of Lemma 1.22 it is convenient to introduce the so-called Minkowski functional (cf. [65]). Recall in the next definition that $\inf\{\emptyset\} := \infty$.

Definition 1.12 *The Minkowski functional or gauge of the nonempty set* $S \subseteq \mathbb{R}^n$ *is given by the function* $\gamma_S : \mathbb{R}^n \to [0, \infty]$ *defined by*

$$\gamma_S(\mathbf{s}) := \inf\{t > 0 : \mathbf{s} \in tS\}.$$

As shown by the next result the support function of any set S containing the zero vector $\mathbf{0}$ equals the gauge of the closed convex polar S^0.

Lemma 1.23 *If $S \subseteq \mathbb{R}^n$ is a nonempty set containing* $\mathbf{0}$, *then it follows that $\sigma_{cl(S)}(\mathbf{s}) = \gamma_{S^0}(\mathbf{s})$ for every $\mathbf{s} \in \mathbb{R}^n$.*

Proof. Since $\mathbf{0}$ belongs to $cl(S)$, it follows that the support function $\sigma_{cl(S)}$ of the set $cl(S)$ is nonnegative. Consider now the following two cases. If $\sigma_{cl(S)}(\mathbf{s}_0) = 0$, we obtain $t^{-1}\mathbf{s}_0^\top \mathbf{x} \leq 0$ for every $t > 0$ and $\mathbf{x} \in S$. This shows $t^{-1}\mathbf{s}_0 \in S^0$ for every $t > 0$ and so $\gamma_{S^0}(\mathbf{s}_0) = 0 = \sigma_{cl(S)}(\mathbf{s}_0)$. Moreover, if $\sigma_{cl(S)}(\mathbf{s}_0) > 0$, we obtain using $\sigma_{cl(S)} = \sigma_S$ that

$$0 < \sigma_S(\mathbf{s}_0) = \inf\{t > 0 : \mathbf{s}_0^\top \mathbf{x} \leq t, \mathbf{x} \in S\} = \inf\{t > 0 : \frac{\mathbf{s}_0}{t} \in S^0\}$$

and this shows the desired result. □

Finally we can prove the so-called *bipolar* theorem for closed convex sets containing $\mathbf{0}$, generalizing Lemma 1.22. This representation can be seen as a so-called dual representation of a closed convex set containing $\mathbf{0}$.

Theorem 1.5 *If $C \subseteq \mathbb{R}^n$ is a nonempty convex set with $\mathbf{0} \in cl(C)$, then it follows that $C^{00} = cl(C)$.*

Proof. It is obvious that $cl(C) \subseteq C^{00}$ and so we only need to verify the reverse inclusion. Since for any $\mathbf{s} \in \mathbb{R}^n$ satisfying $\gamma_{C^0}(\mathbf{s}) < \infty$ it follows that

$$(\gamma_{C^0}(\mathbf{s}) + \epsilon)^{-1}\mathbf{s} \in C^0$$

for every $\epsilon > 0$, we obtain for every $\mathbf{x}_0 \in C^{00}$ that $\mathbf{s}^\top \mathbf{x}_0 \leq \gamma_{C^0}(\mathbf{s}) + \epsilon$. This implies $\mathbf{s}^\top \mathbf{x}_0 \leq \gamma_{C^0}(\mathbf{s})$ and since this inequality trivially holds for $\gamma_{C^0}(\mathbf{s}) = \infty$ we obtain by Lemma 1.23 that $\mathbf{s}^\top \mathbf{x}_0 \leq \sigma_{cl(C)}(\mathbf{s})$ for every $\mathbf{s}$. Applying now Theorem 1.4 shows $\mathbf{x}_0 \in cl(C)$ and we have checked that $C^{00} \subseteq cl(C)$. □

By a similar approach as used after Lemma 1.22 it is easy to construct a dual representation of the convex set $co(S \cup \{\mathbf{0}\})$ with S a nonempty set. First we observe by the definition of the polar operator and using Theorem 1.5 that $S^{00} \subseteq (co(S \cup \{\mathbf{0}\})^{00} = cl(co(S \cup \{\mathbf{0}\}))$. Since S^{00} is a closed convex set containing $S \cup \{\mathbf{0}\}$ and $cl(co(S \cup \{\mathbf{0}\}))$ is the smallest closed convex set containing $S \cup \{\mathbf{0}\}$ we obtain by the previous inclusion the general formula

$$S^{00} = cl(co(S \cup \{\mathbf{0}\}\}). \tag{1.41}$$

The formula, listed in relation (1.41), is called the *bipolar theorem* for arbitrary sets $S \subseteq \mathbb{R}^n$. Replacing Theorem 1.4 by its equivalent version valid in locally convex topological vector spaces one can verify using a similar proof the bipolar theorem (cf. [10], [37]) in locally convex topological vector spaces.

An important special case of Theorem 1.5 is given by $K^{00} = cl(K)$ with K a convex cone. By means of similar proof techniques (cf. [67]) it is also possible to give a dual representation of the relative interior $ri(K)$ of a convex cone K. Without proof we now list the following result. For related results, valid in infinite dimensional topological vector spaces, the reader should consult [38].

Theorem 1.6 For *any nonempty convex cone* $K \subseteq \mathbb{R}^n$ *it follows that*

$$\mathbf{x} \in ri(K) \Leftrightarrow \mathbf{x} \in (K^{\perp})^{\perp} \textit{ and } \mathbf{x}^{*\top}\mathbf{x} < 0 \textit{ for } \mathbf{x}^* \in K^0 \cap (K^{\perp})^{\perp} \backslash \{0\}.$$

This concludes our section on sets. In the next section we will consider functions studied within convex and quasiconvex analysis.

3. Functions studied within convex and quasiconvex analysis

In this section we first introduce in Subsection 3.1 the different classes of functions studied within convex and quasiconvex analysis and derive their algebraic properties. These algebraic properties are an easy consequence of two important relations between functions and sets and the properties of sets derived in Subsection 2.1. Also from Subsection 2.1 we know how to apply hull operations to sets and using this it is also possible to construct so-called hull functions. These different hull functions are also introduced in Subsection 3.1 and their properties will be derived. In Subsection 3.2 topological properties of functions are introduced together with some of the "topological" hull functions. It will turn out that especially the class of lower semicontinuous functions is extremely important in this field. Finally in Subsections 3.3 and 3.4 dual characterizations of the considered functions will be derived. The key results in these sections are the Fenchel-Moreau theorem within convex analysis and its generalization to the so-called evenly quasiconvex and lower semicontinuous quasiconvex functions.

3.1 Algebraic properties of functions

In this subsection we relate functions to sets and use the algebraic properties of sets given in Subsection 2.1 to derive algebraic properties of functions. To start with this approach, let $f : \mathbb{R}^n \to [-\infty, \infty]$ be an extended real valued function and associate with f its so-called *epigraph*

$$epi(f) := \{(\mathbf{x}, r) \in \mathbb{R}^{n+1} : f(\mathbf{x}) \leq r\}. \tag{1.42}$$

A related set is the strict *epigraph*

$$\widetilde{epi}(f) := \{(\mathbf{x}, r) \in \mathbb{R}^{n+1} : f(\mathbf{x}) < r\}. \tag{1.43}$$

Within convex analysis it is now useful to represent a function f by the obvious relation (cf. [63])

$$f(\mathbf{x}) = \inf\{r : (\mathbf{x}, r) \in epi(f)\}. \tag{1.44}$$

By definition $\inf\{\emptyset\} = \infty$ and this only happens if the vector $\mathbf{x}$ does not belong to the so-called *effective domain*

$$dom(f) := \{\mathbf{x} \in \mathbb{R}^n : f(\mathbf{x}) < \infty\} \tag{1.45}$$

of the function f. By this observation it follows that $dom(f)$ is nonempty if and only if $epi(f)$ is nonempty and if this holds we obtain

$$dom(f) = A(epi(f)) \tag{1.46}$$

with A the projection of $\mathbb{R}^{n+1}$ onto $\mathbb{R}^n$ given by $A(\mathbf{x}, r) = \mathbf{x}$. As shown by the following definition, the representation of the function f given by relation (1.44) is useful in the study of convex functions.

Definition 1.13 *The function $f : \mathbb{R}^n \to [-\infty, \infty]$ is called convex if the set $epi(f)$ is convex. Moreover, the function $f : \mathbb{R}^n \to [-\infty, \infty]$ is called positively homogeneous if the set $epi(f)$ is a cone.*

An equivalent definition of a convex function is given by the next result, which is easy to verify.

Lemma 1.24 *A function $f : \mathbb{R}^n \to [-\infty, \infty]$ is convex if and only if the set $\widetilde{epi}(f)$ is convex.*

Using Lemma 1.24 we obtain that a function $f : \mathbb{R}^n \to [-\infty, \infty]$ is convex if and only if for every $0 < \alpha < 1$

$$f(\alpha\mathbf{x}_1 + (1-\alpha)\mathbf{x}_2) < \alpha r_1 + (1-\alpha) r_2 \tag{1.47}$$

whenever $f(\mathbf{x}_i) < r_i \in \mathbb{R}$. In case we know additionally that $f > -\infty$ we obtain by relation (1.44) that f is convex if and only if for every $0 < \alpha < 1$

$$f(\alpha\mathbf{x}_1 + (1-\alpha)\mathbf{x}_2) \leq \alpha f(\mathbf{x}_1) + (1-\alpha) f(\mathbf{x}_2) \tag{1.48}$$

and so we recover the more familiar definition of a convex function. An important special case satisfying relation (1.48) is given by $f > -\infty$ and

$dom(f)$ is nonempty. If this holds the function f is called *proper.* Also the next result is easy to verify.

Lemma 1.25 *The function* $f : \mathbb{R}^n \to [-\infty, \infty]$ *is positively homogeneous if and only if* $f(\alpha\mathbf{x}) = \alpha f(\mathbf{x})$ *for every* $\mathbf{x} \in \mathbb{R}^n$ *and* $\alpha > 0$.

To investigate under which operations on convex functions this property is preserved we observe for any collection of functions f_i, $i \in I$ that

$$epi(\sup_{i\in I} f_i) = \cap_{i\in I} epi(f_i). \tag{1.49}$$

Since the intersection of convex sets is again convex we obtain by relation (1.49) that the function $\sup_{i\in I} f_i$ is convex if f_i is convex for every $i \in I$. Moreover, by relation (1.48), it follows that any strict canonical combination of the convex functions $f_i > -\infty$, $i = 1, 2$ is again convex.

In case we use the representation of a function f, given by relation (1.44), and the various hull operations on a set defined in Subsection 2.1 it is easy to introduce the various so-called hull functions of f. The first hull function is given by the next definition (cf. [63]). In this volume the various hull functions, given in this subsection and the next, are also discussed by Crouzeix (cf. [11]).

Definition 1.14 *For any function* $f : \mathbb{R}^n \to [-\infty, \infty]$ *the function* $f_c : \mathbb{R}^n \to [-\infty, \infty]$, *given by* $f_c(\mathbf{x}) := \inf\{r : (\mathbf{x}, r) \in co(epi(f))\}$, *is called the convex hull function of the function* f.

The next result yields an interpretation of the convex hull function of a function f. Recall that the convex hull of the empty set is again the empty set.

Lemma 1.26 *For any function* $f : \mathbb{R}^n \to [-\infty, \infty]$ *the convex hull function* f_c *is the greatest convex function majorized by* f. *Moreover, it follows that* $epi(f_c)) \subseteq co(epi(f)) \subseteq epi(f_c)$ *and* $dom(f_c) = co(dom(f))$.

Proof. Without loss of generality we may assume that $epi(f)$ or equivalently $dom(f)$ is nonempty. Since $co(epi(f))$ is a convex set we obtain by Definition 1.14 for every $r_i > f_c(\mathbf{x}_i), i = 1, 2$ that $f_c(\alpha\mathbf{x}_1 + (1-\alpha)\mathbf{x}_2) < \alpha r_1 + (1-\alpha)r_2$ for every $0 < \alpha < 1$. This shows by relation (1.47) that the function f_c is convex. Moreover, if $h \leq f$ and h is convex, then $co(epi(f)) \subseteq co(epi(h)) = epi(h)$ and so f_c is the greatest convex function majorized by f. Using again Definition 1.14 it is also easy to verify that $epi(f_c) \subseteq co(epi(f)) \subseteq epi(f_c)$. To show the last part of this lemma, let $\mathbf{x} \in dom(f_c)$ and so $(\mathbf{x}, r) \in co(epi(f))$ for every $r > f_c(\mathbf{x})$. This implies by relation (1.46) that $\mathbf{x} \in A(co(epi(f)) = co(A(epi(f)) =$

$co(dom(f))$ with A the projection of $\mathbb{R}^{n+1}$ onto $\mathbb{R}^n$ and we have verified $dom(f_c) \subseteq co(dom(f))$. Also, for $\mathbf{x} \in co(dom(f))$, we obtain by relation (1.46) that $\mathbf{x} \in A(co(epi(f)))$ and so $\mathbf{x}$ belongs to $dom(f_c)$ showing the reverse inclusion. □

In general it follows that $epi(f_c) \neq co(epi(f))$. A direct consequence of Lemma 1.26 and the fact that $\sup_{i \in I} f_i$ is convex for f_i, $i \in I$ a collection of convex functions, is the often used representation of the function f_c given by

$$f_c(\mathbf{x}) = \sup\{h(\mathbf{x}) : h \leq f \text{ and } h \text{ is a convex function}\}. \tag{1.50}$$

Next to the epigraph of a function $f : \mathbb{R}^n \to [-\infty, \infty]$ one also considers the so-called *lower-level set* $L(f,r), r \in \mathbb{R}$ of the function f given by

$$L(f,r) := \{\mathbf{x} \in \mathbb{R}^n : f(\mathbf{x}) \leq r\}. \tag{1.51}$$

A related set is the *strict lower-level set* of the function f of level r represented by

$$\widetilde{L}(f,r) := \{\mathbf{x} \in \mathbb{R}^n : f(\mathbf{x}) < r\}. \tag{1.52}$$

Within quasiconvex analysis it is now useful to represent a function f by the obvious relation (cf. [15])

$$f(\mathbf{x}) = \inf\{r : \mathbf{x} \in L(f,r)\}. \tag{1.53}$$

As shown by the following definition, the representation of the function f, given by relation (1.53), is useful in the study of quasiconvex functions.

Definition 1.15 *The function $f : \mathbb{R}^n \to [-\infty, \infty]$ is called quasiconvex if for every $r \in \mathbb{R}$ the lower-level set $L(f,r)$ is convex. Moreover, the function f is called evenly quasiconvex if for every $r \in \mathbb{R}$ the lower level set $L(f,r)$ is evenly convex.*

To derive the relation between convex and quasiconvex functions we observe that $epi(f) \cap (\mathbb{R}^n \times \{r\}) = L(f,r) \times \{r\}$ for every $r \in \mathbb{R}$. This implies that a convex function is also a quasiconvex function. Since each monotonic (increasing or decreasing) function $f : \mathbb{R} \to \mathbb{R}$ is quasiconvex, but not necessarily convex, the converse is not true. For quasiconvex functions a similar result as in Lemma 1.24 can be easily verified.

Lemma 1.27 *A function $f : \mathbb{R}^n \to [-\infty, \infty]$ is quasiconvex if and only if the set $\widetilde{L}(f,r)$ is convex for every $r \in \mathbb{R}$.*

To recover a more familiar representation of a quasiconvex function it can be shown easily (cf. [2]) that a function $f : \mathbb{R}^n \to [-\infty, \infty]$ is quasiconvex if and only if $f(\alpha \mathbf{x}_1 + (1-\alpha)\mathbf{x}_2) \leq \max\{f(\mathbf{x}_1), f(\mathbf{x}_2)\}$, for every $0 < \alpha < 1$.

As for convex functions, one is interested under which operations on quasiconvex functions this property is preserved. Clearly for any collection offunctions $f_i, i \in I$ it follows that

$$L(\sup_{i\in I} f_i, r) = \cap_{i\in I} L(f_i, r) \tag{1.54}$$

and this shows that the function $\sup_{i\in I} f_i$ is quasiconvex if f_i is quasiconvex for every $i \in I$. Opposed to convex functions, it is not true that a strict canonical combination of quasiconvex functions is quasiconvex and this is shown by the following example.

Example 1.11 Let $f_i : \mathbb{R} \to \mathbb{R}, i = 1,2$ be given by $f_1(x) = x$ and

$$f_2(x) = x^2 \text{ for } |x| \leq 1 \text{ and } f_2(x) = 1 \text{ otherwise.}$$

These functions are quasiconvex, but it is easy to verify by means of a picture that the sum of the two functions is not quasiconvex.

Using relation (1.53), one can apply the different hull operations to the lower level set. The first hull function constructed in this way is listed in the next definition (cf. [15], [11]).

Definition 1.16 *For any function* $f : \mathbb{R}^n \to [-\infty, \infty]$ *the function* $f_q : \mathbb{R}^n \to [-\infty, \infty]$, *given by* $f_q(\mathbf{x}) := \inf\{r : \mathbf{x} \in co(L(f, r))\}$, *is called the quasiconvex hull function of the function* f.

The next result (cf. [15]) yields an interpretation of the quasiconvex hull function of a function f.

Lemma 1.28 *For any function* $f : \mathbb{R}^n \to [-\infty, \infty]$ *the quasiconvex hull function* f_q *is the greatest quasiconvex function majorized by* f. *Moreover, it follows that* $L(f_q, r) = \cap_{\beta > r} co(L(f, \beta))$ *for every* $r \in \mathbb{R}$.

Proof. Again we may assume without loss of generality that $dom(f)$ is nonempty. By Definition 1.16 it follows that $L(f_q, r) \subseteq \cap_{\beta>r} co(L(f, \beta))$. Since it is obvious that the reverse inclusion holds, we obtain $L(f_q, r) = \cap_{\beta>r} co(L(f, \beta))$. By this relation it is clear that the function f_q is quasiconvex and applying a similar argument as in Lemma 1.26 to lower level sets it can be shown that this function is the greatest quasiconvex function majorized by the function f. □

A direct consequence of Lemma 1.28 and the fact that $\sup_{i\in I} f_i$ is quasiconvex for f_i, $i \in I$ a collection of quasiconvex functions, is the often used representation of f_q given by

$$f_q(\mathbf{x}) = \sup\{h(\mathbf{x}) : h \leq f \text{ and } h \text{ is a quasiconvex function}\}. \quad (1.55)$$

To conclude this subsection, we consider a hull function based on evenly convex sets (cf. [55], [11]). It will turn out that this function plays an important role in duality theory for quasiconvex functions.

Definition 1.17 *For any function* $f : \mathbb{R}^n \to [-\infty, \infty]$ *the function* $f_{ec} : \mathbb{R}^n \to [-\infty, \infty]$, *given by* $f_{ec}(\mathbf{x}) := \inf\{r : \mathbf{x} \in ec(L(f, r))\}$, *is called the evenly quasiconvex hull function of the function* f.

As done for the quasiconvex hull function one can show by a similar proof the following result (cf. [55]).

Lemma 1.29 *For any function* $f : \mathbb{R}^n \to [-\infty, \infty]$ *the evenly quasiconvex hull function* f_{ec} *is the greatest evenly quasiconvex function majorized by* f. *Moreover, it follows that* $L(f_{ec}, r) = \cap_{\beta > r} ec(L(f, \beta))$ *for every* $r \in \mathbb{R}$.

A direct consequence of Lemma 1.29 and the fact that $\sup_{i \in I} f_i$ is evenly quasiconvex for f_i, $i \in I$ a collection of evenly quasiconvex functions, is the often used representation of f_{ec} given by

$$f_{ec}(\mathbf{x}) = \sup\{h(\mathbf{x}) : h \leq f,\ h \text{ evenly quasiconvex function}\}. \quad (1.56)$$

Since an evenly quasiconvex function is clearly a quasiconvex function it holds that $f_{ec} \leq f_q$. This concludes our discussion of algebraic properties of convex and quasiconvex functions. In the next subsection we will consider topological properties of functions.

3.2 Topological properties of functions

In this subsection we first introduce the class of lower semicontinuous functions. These functions play an important role within the theory of convex functions.

Definition 1.18 *If* $f : \mathbb{R}^n \to [-\infty, \infty]$ *is some function, then this function is called lower semicontinuous at* $\mathbf{x} \in \mathbb{R}^n$ *if* $\liminf_{\mathbf{y} \to \mathbf{x}} f(\mathbf{y}) = f(\mathbf{x})$ *with*

$$\liminf_{\mathbf{y} \to \mathbf{x}} f(\mathbf{y}) := \sup_{\epsilon > 0} \inf\{f(\mathbf{y}) : \mathbf{y} \in \mathbf{x} + \epsilon E\}. \quad (1.57)$$

Moreover, the function $f : \mathbb{R}^n \to [-\infty, \infty]$ *is called upper semicontinuous at* $\mathbf{x} \in \mathbb{R}^n$ *if the function* $-f$ *is lower semicontinuous at* $\mathbf{x}$ *and it is* *called continuous at* $\mathbf{x}$ *if it is both lower and upper semicontinuous at* $\mathbf{x}$. *The function* $f : \mathbb{R}^n \to [-\infty, \infty]$ *is called lower semicontinuous (upper semicontinuous) if* f *is lower semicontinuous (upper semicontinuous) at every* $\mathbf{x} \in \mathbb{R}^n$ *and it is called continuous if it is both upper and lower semicontinuous.*

We mostly abbreviate lower semicontinuous by l.s.c.. To relate the above definition of liminf to the liminf of a sequence we observe for every sequence $\mathbf{y}_k, k \in \mathbb{N}$ that $\liminf_{k\uparrow\infty} f(\mathbf{y}_k) := \lim_{n\uparrow\infty} \inf_{k\geq n} f(\mathbf{y}_k)$. Using this definition one can easily show the following result.

Lemma 1.30 *The function* $f : \mathbb{R}^n \to [-\infty, \infty]$ *is l.s.c. at* $\mathbf{x} \in \mathbb{R}^n$ *if and only if* $\liminf_{k\uparrow\infty} f(\mathbf{y}_k) \geq f(\mathbf{x})$ *for every sequence* $\mathbf{y}_k, k \in \mathbb{N}$ *satisfying* $\lim_{k\uparrow\infty} \mathbf{y}_k = \mathbf{x} \in \mathbb{R}^n$.

Using Lemma 1.30 the following important characterization of l.s.c. functions can be proved (cf. [63], [1]).

Theorem 1.7 *If* $f : \mathbb{R}^n \to [-\infty, \infty]$ *is an extended real valued function, then the following conditions are equivalent:*

1. *The function* f *is l.s.c..*
2. *The set* $epi(f)$ *is closed.*
3. *The set* $L(f, r)$ *is closed for every* $r \in \mathbb{R}$.

It is useful to know under which operations on l.s.c. functions this property is preserved. Since $epi(\sup_{i\in I} f_i) = \cap_{i\in I} epi(f_i)$ and the intersection of closed sets is again a closed set we obtain by Theorem 1.7 that the function $\sup_{i\in I} f_i$ is l.s.c. if each function $f_i, i \in I$ is l.s.c.. Also it follows for every finite set I that $epi(\min_{i\in I} f_i) = \cup_{i\in I} epi(f_i)$ and this shows by Theorem 1.7 and the fact that a finite union of closed sets is closed, that the function $\min_{i\in I} f_i$ is l.s.c. if each f_i, $i \in I$ is l.s.c.. Finally, for arbitrary functions $f_i : \mathbb{R}^n \to (-\infty, \infty], i = 1, 2$ we obtain that

$$L(f_1 + f_2, r)^c = \cup_{q\in\mathbb{Q}} (L(f_1, r - q)^c \cup L(f_2, q)^c)$$

with A^c denoting the complement of the set $A \subseteq \mathbb{R}^n$ and this implies using Theorem 1.7 that the function $\alpha f_1 + \beta f_2$ is l.s.c. for every $\alpha, \beta \geq 0$, if the functions $f_i > -\infty, i = 1, 2$ are l.s.c..

To verify the next theorem we introduce for any function $f : \mathbb{R}^n \to [-\infty, \infty]$ the (possibly empty) set of continuous real valued minorants $\mathcal{C}_f$ of f given by

$$\mathcal{C}_f := \{h : h \leq f \text{ and } h \text{ is a real valued continuous function}\}.$$

In the next result it is now shown that any l.s.c. function can be seen as a pointwise limit of an increasing sequence of real valued continuous functions.

Theorem 1.8 *For any function $f : \mathbb{R}^n \to (-\infty, \infty]$ the following conditions are equivalent:*

1 The function f is l.s.c..

2 There exists an increasing sequence of continuous functions $(h_m)_{m \in \mathbb{N}}$ satisfying $f(\mathbf{x}) = \lim_{m \uparrow \infty} h_m(\mathbf{x})$ for every $\mathbf{x} \in \mathbb{R}^n$.

3 $f(\mathbf{x}) = \sup\{h(\mathbf{x}) : h \in \mathcal{C}_f\}$ with $\mathcal{C}_f$ nonempty.

Proof. We only give a proof of 1 $\Rightarrow$ 2 since the other implications are obvious. We first show the desired result for a nonnegative uniformly bounded function f. Actually, if the function f is nonnegative and uniformly bounded, then the sequence $f_m : \mathbb{R}^n \to [0, \infty), m \in \mathbb{N}$ given by $f_m(\mathbf{x}) := \inf\{f(\mathbf{z}) + m\|\mathbf{x} - \mathbf{z}\| : \mathbf{z} \in \mathbb{R}^n\}$ is increasing, converges pointwise to $f(\mathbf{x})$ and each f_m is continuous (actually Lipschitz continuous with Lipschitz constant m!). To reduce the general case of a proper l.s.c. function f to this special case, replace the proper l.s.c. function f by the nonnegative uniformly bounded l.s.c. function $g = k \circ f$, where $k(x) = \frac{1}{2}\pi + \arctan(x)$ and apply the first part. Hence there exists an increasing sequence g_m of continuous functions converging pointwise to g. Use now that the function $k : (-\infty, \infty] \to (0, \pi]$ is one-to-one, strictly increasing and continuous with a continuous inverse $k^{\leftarrow}$and select the sequence $h_m := k^{\leftarrow} \circ g_m$. □

By Theorem 1.8 we obtain that the set of l.s.c. functions is the smallest set of functions, which are closed under taking the sup operation to any collection of functions belonging to this set and which contain the set of continuous real valued functions.

As in the previous subsection, we are going to introduce hull operations related to functions. In this case topological properties will be involved. First we consider the so-called l.s.c. hull function of a function f (cf. [63], [11]).

Definition 1.19 *For any function $f : \mathbb{R}^n \to [-\infty, \infty]$ the function $\overline{f} : \mathbb{R}^n \to [-\infty, \infty]$, given by $\overline{f}(\mathbf{x}) := \inf\{r : (\mathbf{x}, r) \in cl(epi(f))\}$, is called the l.s.c. hull function of the function f.*

In the next result an interpretation of the l.s.c. hull function of a function f is given.

Lemma 1.31 *For any function* $f : \mathbb{R}^n \to [-\infty, \infty]$ *the l.s.c. hull function* $\overline{f}$ *is the greatest l.s.c. function majorized by* f*. Moreover, its epigraph equals* $cl(epi(f))$ *and* $dom(f) \subseteq dom(\overline{f}) \subseteq cl(dom(f))$*. If additionally* $dom(f)$ *is a convex set, it follows that* $ri(dom(\overline{f})) = ri(dom(f))$*.*

Proof. By Definition 1.19 we obtain $(\mathbf{x}, r) \in epi(\overline{f}) \Leftrightarrow \forall_{\epsilon>0}\ (\mathbf{x}, r+\epsilon) \in cl(epi(f)) \Leftrightarrow (\mathbf{x}, r) \in cl(epi(f))$. This means that $epi(\overline{f})$ equals $cl(epi(f))$ and by Theorem 1.7 the function $\overline{f}$ is l.s.c.. Moreover, if $h \leq f$ and h is l.s.c., then by Theorem 1.7 we obtain $cl(epi(f)) \subseteq cl(epi(h)) = epi(h)$ and so it follows that $h \leq \overline{f}$. To verify the last part we may assume without loss of generality that $dom(f)$ is nonempty. Since $\overline{f} \leq f$ it follows that $dom(f) \subseteq dom(\overline{f})$ and by relation (1.46) we obtain $dom(\overline{f}) = A(cl(epi(f)) \subseteq cl(A(epi(f)) = cl(dom(f))$. Finally, if $dom(f)$ is a nonempty convex set it follows by Lemma 1.16 that $ri(dom(f)) = ri(cl(dom(f)))$ and since $dom(f) \subseteq dom(\overline{f}) \subseteq cl(dom(f))$ we obtain $ri(dom(\overline{f})) = ri(dom(f))$. □

A direct consequence of Lemma 1.31 and the fact that $\sup_{i \in I} f_i$ is l.s.c. for f_i, $i \in I$ a collection of l.s.c. functions, is the often used representation of $\overline{f}$ given by

$$\overline{f}(\mathbf{x}) = \sup\{h(\mathbf{x}) : h \leq f \text{ and } h \text{ is a l.s.c. function }\}. \tag{1.58}$$

For nondecreasing functions $f : \mathbb{R} \to [-\infty, \infty]$ it is possible to give a more detailed description of the l.s.c. hull function $\overline{f}$ of f. To show this result we first introduce the next definition.

Definition 1.20 *For any function* $f : \mathbb{R} \to [-\infty, \infty]$ *the function* $f^{\diamond} : \mathbb{R} \to [-\infty, \infty]$ *is given by*

$$f^{\diamond}(t) := \sup_{s<t} f(s).$$

The next result is needed in the proof of a dual representation of a l.s.c. quasiconvex function.

Lemma 1.32 *For any nondecreasing function* $f : \mathbb{R} \to [-\infty, \infty]$ *it follows that* $\overline{f}(t) = f^{\diamond}(t)$ *for every* $t \in \mathbb{R}$.

Proof. Since the function f is nondecreasing, it is easy to verify that $f^{\diamond}$ is nondecreasing and $f^{\diamond} \leq f$. We now verify that the function $f^{\diamond}$ is l.s.c. and so by Theorem 1.7 we need to check that the lower-level set $L(f^{\diamond}, r)$ is closed for every $r \in \mathbb{R}$. Assume now by contradiction that there exists some $r_0 \in \mathbb{R}$ such that the set $L(f^{\diamond}, r_0)$ is not closed. Hence there exists a sequence $\{t_n : n \in \mathbb{N}\} \subseteq L(f^{\diamond}, r_0)$ with $t_\infty := \lim_{n\uparrow\infty} t_n$ and t_∞ does not belong to $L(f^{\diamond}, r_0)$. Since $f^{\diamond}$ is nondecreasing and $f^{\diamond}(t_{\infty}) > r_0$ it follows that $t_n < t_\infty$ for every $n \in \mathbb{N}$ and by Definition 1.20 one can find some $s_0 < t_\infty$ satisfying $f(s_0) > r_0$. This implies that there exists some t_n satisfying $s_0 < t_n < t_\infty$ and so $f^\diamond(t_n) \geq f(s_0) > r_0$ contradicting t_n belongs to $L(f^\diamond, r_0)$. Therefore $f^\diamond$ is l.s.c. and using $f^\diamond \leq f$ it follows by relation (1.58) that $f^\diamond \leq \overline{f}$. Suppose now by contradiction that $f^\diamond(t_0) < \overline{f}(t_0)$ for some t_0. By relation (1.57) and $\overline{f}$ is l.s.c. this implies that there exists some $\epsilon > 0$ satisfying $\overline{f}(t) > f^\diamond(t_0)$ for every $t_0 - \epsilon \leq t \leq t_0 + \epsilon$ and so

$$f(t_0 - \epsilon) \geq \overline{f}(t_0 - \epsilon) > f^\diamond(t_0) \geq f(t_0 - \epsilon).$$

This yields a contradiction and the result is proved. □

The next result relates $\overline{f}$ to f and this result is nothing else than a "function value translation" of the original definition of the l.s.c. hull function $\overline{f}$ of f.

Lemma 1.33 *For any function* $f : \mathbb{R}^n \to [-\infty, \infty]$ *and* $\mathbf{x} \in \mathbb{R}^n$ *it follows that* $\overline{f}(\mathbf{x}) = \liminf_{\mathbf{y} \to \mathbf{x}} f(\mathbf{y})$.

Proof. Since $\overline{f} \leq f$ and $\overline{f}$ is a l.s.c. function we obtain that $\overline{f}(\mathbf{x}) = \liminf_{\mathbf{y} \to \mathbf{x}} \overline{f}(\mathbf{y}) \leq \liminf_{\mathbf{y} \to \mathbf{x}} f(\mathbf{y})$. Suppose now by contradiction that $\overline{f}(\mathbf{x}) < \liminf_{\mathbf{y} \to \mathbf{x}} f(\mathbf{y})$. If this holds, then clearly $\overline{f}(\mathbf{x}) < \infty$ and by the definition of liminf there exists some finite γ and $\epsilon > 0$ satisfying $f(\mathbf{x}+\mathbf{y}) > \gamma > \overline{f}(\mathbf{x})$ for every $\mathbf{y} \in \epsilon E$. This implies that the open set $(\mathbf{x}+\epsilon E) \times (-\infty, \gamma)$ containing the point $(\mathbf{x}, \overline{f}(\mathbf{x}))$ has an empty intersection with epi(f). However, by Lemma 1.31 it follows that $(\mathbf{x}, \overline{f}(\mathbf{x}))$ belongs to $cl(epi(f))$ and so every open set containing $(\mathbf{x}, \overline{f}(\mathbf{x}))$ must have a nonempty intersection with $epi(f)$. Hence we obtain a contradiction and so the result is proved. □

By Lemma 1.33 and Definition 1.18 it follows immediately that

$$f \text{ is l.s.c. at } \mathbf{x} \Leftrightarrow \overline{f}(\mathbf{x}) = f(\mathbf{x}). \tag{1.59}$$

Using Theorem 1.7 and Lemmas 1.31 and 1.33 one can show that the l.s.c. hull operation applied to functions preserves the convexity and quasiconvexity property.

Lemma 1.34 *If the function* $f : \mathbb{R}^n \to [-\infty, \infty)$ *is convex (quasiconvex), then also the l.s.c. hull function* $\overline{f}$ *of* f *is convex (quasiconvex).*

Proof. If the function f is convex, then $epi(f)$ is a convex set and hence also $cl(epi(f))$ is a convex set. Since by Lemma 1.31 the epigraph of $\overline{f}$ is given by $cl(epi(f))$ this shows that $\overline{f}$ is a convex function. To verify that $\overline{f}$ is quasiconvex for f quasiconvex we need to verify by Lemma 1.27 that the set $\widetilde{L}(\overline{f},r)$ is convex for every $r \in \mathbb{R}$. If the vectors $\mathbf{x}_i, i = 1,2$ belong to $\widetilde{L}(\overline{f},r)$ it follows by Lemma 1.33 that

$$\inf\{f(\mathbf{y}) : \mathbf{y} \in \mathbf{x}_i + \epsilon E\} \leq \overline{f}(\mathbf{x}_i) < r$$

for every $\epsilon > 0$ and $i = 1,2$. This implies for every $\epsilon > 0$ and $i = 1,2$ that there exists some vector $\mathbf{y}_{i,\epsilon} \in \mathbf{x}_i + \epsilon E$ satisfying

$$f(\mathbf{y}_{i,\epsilon}) \leq r_0 := \frac{1}{2}(\max\{\overline{f}(\mathbf{x}_1), \overline{f}(\mathbf{x}_2)\} + r) < r$$

Applying now the quasiconvexity of the function f we obtain for every $0 < \alpha < 1$ that

$$f(\alpha\mathbf{y}_{1,\epsilon} + (1-\alpha)\mathbf{y}_{2,\epsilon}) \leq \max\{f(\mathbf{y}_{1,\epsilon}), f(\mathbf{y}_{2,\epsilon})\} \leq r_0$$

and since the vector $\alpha\mathbf{y}_{1,\epsilon} + (1-\alpha)\mathbf{y}_{2,\epsilon}$ belongs to the set $\alpha\mathbf{x}_1 + (1-\alpha)\mathbf{x}_2 + \epsilon E$ this yields

$$\inf\{f(\mathbf{y}) : \mathbf{y} \in \alpha\mathbf{x}_i + (1-\alpha)\mathbf{x}_2 + \epsilon E\} \leq r_0 < r$$

for every $\epsilon > 0$. Using again Lemma 1.33 we obtain

$$\overline{f}(\alpha\mathbf{x}_1 + (1-\alpha)\mathbf{x}_2) = \liminf_{\mathbf{y}\to\alpha\mathbf{x}_1+(1-\alpha)\mathbf{x}_2} f(\mathbf{y}) \leq r_0 < r$$

and it follows that $\alpha\mathbf{x}_1 + (1-\alpha)\mathbf{x}_2$ belongs to $\widetilde{L}(\overline{f},r)$. □

To improve Lemma 1.33 for convex functions f we need to give a representation of the relative interior of the epigraph of a convex function. This representation is an immediate consequence of the following observation. If $f : \mathbb{R}^n \to [-\infty,\infty]$ is a convex function and $f(\mathbf{x})$ is finite for some $\mathbf{x}$, then clearly $\{\mathbf{x}\} \times (f(\mathbf{x}),\infty) = ri(\{\mathbf{x}\} \times [f(\mathbf{x}),\infty))$ and so

$$\{\mathbf{x}\} \times (f(\mathbf{x}),\infty) = ri((\{\mathbf{x}\} \times \mathbb{R}) \cap epi(f)). \tag{1.60}$$

A similar observation also holds for $f(\mathbf{x}) = -\infty$ and this shows that relation (1.60) is valid for every $\mathbf{x} \in dom(f)$. Also by relation (1.46) and Lemma 1.18 we obtain

$$ri(dom(f)) = ri(A(epi(f)) = A(ri(epi(f))) \tag{1.61}$$

with $A : \mathbb{R}^{n+1} \to \mathbb{R}^n$ the projection on $\mathbb{R}^n$ and so it follows by relation (1.61) that

$$\mathbf{x} \in ri(dom(f)) \Leftrightarrow (\{\mathbf{x}\} \times \mathbb{R}) \cap ri(epi(f)) \neq \emptyset. \tag{1.62}$$

Since the set $\{\mathbf{x}\} \times \mathbb{R}$ is affine and therefore relatively open we obtain by relation (1.62) that the conditions of Lemma 1.19 hold and hence by relation (1.60) we obtain

$$\{\mathbf{x}\} \times (f(\mathbf{x}), \infty) = (\{\mathbf{x}\} \times \mathbb{R}) \cap ri(epi(f)) \tag{1.63}$$

for every $\mathbf{x} \in ri(dom(f))$. Using this equality the next representation is easy to verify.

Lemma 1.35 *If the function* $f : \mathbb{R}^n \to [-\infty, \infty]$ *is convex and* $dom(f)$ *is nonempty, then the set* $ri(epi(f))$ *is nonempty and*

$$ri(epi(f)) = \{(\mathbf{x}, r) \in \mathbb{R}^{n+1} : f(\mathbf{x}) < r, \mathbf{x} \in ri(dom(f))\}.$$

Proof. If $\mathbf{x}$ belongs to $ri(dom(f))$ and $f(\mathbf{x}) < r$ it follows by relation (1.63) that $(\mathbf{x}, r) \in ri(epi(f))$. To show the reverse inclusion we proceed as follows. If $(\mathbf{x}, r)$ belongs to $ri(epi(f))$ then by relation (1.61) we obtain $\mathbf{x} \in ri(dom(f))$. Applying now relation (1.63) yields $f(\mathbf{x}) < r$. □

In case f is a convex function with $dom(f)$ nonempty, the result of Lemma 1.33 can be improved as follows.

Lemma 1.36 *If the function* $f : \mathbb{R}^n \to [-\infty, \infty]$ *is convex and* $dom(f)$ *is nonempty, then* $\overline{f}(\mathbf{x}) = \lim_{t \downarrow 0} f(\mathbf{x}+t(\mathbf{y}-\mathbf{x}))$ *for every* $\mathbf{y} \in ri(dom(f))$. *Moreover, if* $\mathbf{x} \in ri(dom(f))$, *then it follows that* $\overline{f}(\mathbf{x}) = f(\mathbf{x})$.

Proof. By Lemma 1.33 it is obvious that $\overline{f}(\mathbf{x}) \leq \liminf_{t \downarrow 0} f(\mathbf{x}+t(\mathbf{y}-\mathbf{x}))$. If $\overline{f}(\mathbf{x}) = \infty$ then the result holds by the previous inequality and so we assume $\overline{f}(\mathbf{x}) < \infty$. This implies that $(\mathbf{x}, r) \in epi(\overline{f}) = cl(epi(f))$ for every $r > \overline{f}(\mathbf{x})$ and since $\mathbf{y} \in ri(dom(f))$ it follows by Lemma 1.35 that $(\mathbf{y}, r_1) \in ri(epi(f))$ for every $r_1 > f(\mathbf{y})$. Applying now Lemma 1.15 we obtain for every $0 < t < 1$ that $((1-t)\mathbf{x}+t\mathbf{y}, (1-t)r+tr_1) \in epi(f)$ and this shows $f(\mathbf{x} + t(\mathbf{y} - \mathbf{x})) = f(t\mathbf{y} + (1-t)\mathbf{x}) \leq tr_1 + (1-t)r$. Hence it follows that $\limsup_{t \downarrow 0} f(\mathbf{x} + t(\mathbf{y} - \mathbf{x})) \leq r$ and since $r > \overline{f}(\mathbf{x})$ we obtain $\limsup_{t \downarrow 0} f(\mathbf{x}+t(\mathbf{y}-\mathbf{x})) \leq \overline{f}(\mathbf{x})$. This proves the first part and to verify the second part we first observe that the convex set $dom(f)$ is nonempty and so by Lemma 1.13 the set $ri(dom(f))$ is nonempty. By Lemma 1.31 and 1.35 and $\overline{f}$ is convex it now follows that

$$\{(\mathbf{x}, r) \in \mathbb{R}^{n+1} : \overline{f}(\mathbf{x}) < r, \mathbf{x} \in ri(dom(f))\} = ri(cl(epi(f)).$$

This implies using Lemma 1.16 and f is convex that

$$\{(\mathbf{x}, r) \in \mathbb{R}^{n+1} : \overline{f}(\mathbf{x}) < r, \mathbf{x} \in ri(dom(f))\} \subseteq epi(f), \tag{1.64}$$

and by contradiction we obtain $\overline{f}(\mathbf{x}) \geq f(\mathbf{x})$ for every $\mathbf{x} \in ri(dom(f))$. Since always $\overline{f}(\mathbf{x}) \leq f(\mathbf{x})$ the proof is completed. □

We now introduce the most important hull function used within the field of convex analysis (cf. [63], [11]).

Definition 1.21 *For any function $f : \mathbb{R}^n \to [-\infty, \infty]$ the function $f_{\overline{c}} : \mathbb{R}^n \to [-\infty, \infty]$, given by $f_{\overline{c}}(\mathbf{x}) := \inf\{r : (\mathbf{x}, r) \in cl(co(epi(f)))\}$, is called the l.s.c. convex hull function of the function f.*

Using now a similar approach as in Lemma 1.31 one can prove the following result.

Lemma 1.37 *For any function $f : \mathbb{R}^n \to [-\infty, \infty]$ the l.s.c. convex hull function $f_{\overline{c}}$ is the greatest l.s.c. convex function majorized by f. Moreover, it follows that $epi(f_{\overline{c}}) = cl(co(epi(f)))$, $dom(f_c) \subseteq dom(f_{\overline{c}})) \subseteq cl(dom(f_c))$ and $ri(dom(f_c)) = ri(dom(f_{\overline{c}}))$.*

A direct consequence of Lemma 1.37 and the fact that $\sup_{i \in I} f_i$ is a l.s.c. convex function for f_i, $i \in I$ a collection of l.s.c. convex functions, is the often used representation of $f_{\overline{c}}$ given by

$$f_{\overline{c}}(\mathbf{x}) = \sup\{h(\mathbf{x}) : h \leq f \text{ and } h \text{ is a l.s.c. convex function}\}. \tag{1.65}$$

To relate the various hull functions based on relation (1.44) we observe by Lemmas 1.26 and 1.34 that the function $\overline{f_c}$ is convex and l.s.c. Since $\overline{f_c} \leq \overline{f} \leq f$ this shows by Lemma 1.37 that $\overline{f_c} \leq f_{\overline{c}}$. Also by Lemmas 1.26 and 1.37 it holds that the l.s.c. function $f_{\overline{c}}$ is bounded from above by f_c. This implies by Lemma 1.31 that $f_{\overline{c}} \leq \overline{f_c}$ and combining both inequalities yields

$$f_{\overline{c}}(\mathbf{x}) = \overline{f_c}(\mathbf{x}) \tag{1.66}$$

for every $\mathbf{x} \in \mathbb{R}^n$. An immediate consequence of relation (1.66) is now given by the chain of inequalities

$$f_{\overline{c}}(\mathbf{x}) \leq f_c(\mathbf{x}) \leq f(\mathbf{x}) \text{ and } f_{\overline{c}}(\mathbf{x}) \leq \overline{f}(\mathbf{x}) \leq f(\mathbf{x}). \tag{1.67}$$

for every $\mathbf{x} \in \mathbb{R}^n$. We finally consider hull functions based on the lower level set (cf. [15],[11]).

Definition 1.22 *For any function $f : \mathbb{R}^n \to [-\infty, \infty]$ the function $f_{\overline{q}} : \mathbb{R}^n \to [-\infty, \infty]$, given by $f_{\overline{q}}(\mathbf{x}) := \inf\{r : \mathbf{x} \in cl(co(L(f, r)))\}$, is called the l.s.c. quasiconvex hull function of the function f.*

Using a similar approach as in Lemma 1.28 one can show the following result.

Lemma 1.38 *For any function $f : \mathbb{R}^n \to [-\infty, \infty]$ the l.s.c. quasiconvex hull function $f_{\overline{q}}$ is the greatest l.s.c. quasiconvex function majorized by f. Moreover, it follows that $L(f_{\overline{q}}, r) = \cap_{\beta > r} cl(co(L(f, \beta)))$ for every $r \in \mathbb{R}$.*

A direct consequence of Lemma 1.38 and the fact that $\sup_{i \in I} f_i$ is a l.s.c. quasiconvex function for $f_i, i \in I$ a collection of l.s.c. quasiconvex functions, is the often used representation of $f_{\overline{q}}$ given by

$$f_{\overline{q}}(\mathbf{x}) = \sup\{h(\mathbf{x}) : h \leq f \text{ and } h \text{ is a l.s.c. quasiconvex function}\}. \tag{1.68}$$

To relate the various hull functions based on relation (1.53) we first observe by Lemma 1.21 that every closed convex set is evenly convex and so it follows that

$$f_{\overline{q}}(\mathbf{x}) \leq f_{ec}(\mathbf{x}) \leq f_q(\mathbf{x}) \leq f(\mathbf{x}). \tag{1.69}$$

for every $\mathbf{x} \in \mathbb{R}^n$. Moreover, using $\overline{f_q} \leq \overline{f} \leq f$, relation (1.68) and Lemma 1.34 we obtain $\overline{f_q} \leq f_{\overline{q}}$ and since by relation (1.69) and Lemma 1.38 also $f_{\overline{q}} = \overline{f_{\overline{q}}} \leq \overline{f_{ec}} \leq \overline{f_q}$, this finally yields

$$f_{\overline{q}}(\mathbf{x}) = \overline{f_{ec}}(\mathbf{x}) = \overline{f_q}(\mathbf{x}) \tag{1.70}$$

for every $\mathbf{x} \in \mathbb{R}^n$. The above representations of the hull functions do not depend on the fact that the domain is finite dimensional and so we can also introduce the same hull functions in linear topological vector spaces (cf. [56]). In the next two subsections we consider the dual representations of some of the hull functions.

3.3 Dual representations of convex functions

In this subsection we will consider in detail properties of convex functions, which can be derived using the strong and weak separation results for nonempty convex sets. In particular, we will discuss a dual representation of a l.s.c. convex function f satisfying $f > -\infty$. As always in mathematics one likes to approximate complicated functions by simpler functions. For convex functions these simpler functions are given by the so-called affine minorants.

Definition 1.23 *For any function $f : \mathbb{R} \to [-\infty, \infty]$ the affine function $a : \mathbb{R}^n \to \mathbb{R}$, given by $a(\mathbf{x}) = \mathbf{a}^\intercal \mathbf{x} + \alpha$, with $\mathbf{a} \in \mathbb{R}^n$ and $\alpha \in \mathbb{R}$ is called an affine minorant of the function f if $f(\mathbf{x}) \geq a(\mathbf{x})$ for every $\mathbf{x}$ belonging to $\mathbb{R}^n$. Moreover, the possibly empty set of affine minorants of the function f is denoted by $\mathcal{A}_f$.*

Since any affine minorant a of a function f is continuous and convex it is easy to verify the following result.

Lemma 1.39 *For any function* $f : \mathbb{R}^n \to [-\infty, \infty]$ *it follows that* $\mathcal{A}_f = \mathcal{A}_{f_c} = \mathcal{A}_{f_{\bar{c}}} = \mathcal{A}_{\overline{f_c}}$.

Proof. We only give a proof of the above result for $\mathcal{A}_f$ nonempty. Since by relations (1.67) and (1.66) we know that $f_{\bar{c}} = \overline{f_c} \leq f_c \leq f$ it follows immediately that $\mathcal{A}_{f_{\bar{c}}} = \mathcal{A}_{\overline{f_c}} \subseteq \mathcal{A}_{f_c} \subseteq \mathcal{A}_f$. Moreover, if the function a belongs to $\mathcal{A}_f$, then clearly $a \leq f$ and a is continuous and convex. This implies by relation (1.65) that $a \leq f_{\bar{c}}$ and hence the affine function a belongs to $\mathcal{A}_{f_{\bar{c}}}$. □

Since an affine function is always finite valued the set $\mathcal{A}_f$ is empty if there exists some $\mathbf{x} \in \mathbb{R}^n$ satisfying $f(\mathbf{x}) = -\infty$ and so it is necessary to consider functions $f : \mathbb{R}^n \to (-\infty, \infty]$. In Theorem 1.9 necessary and sufficient conditions are given for $\mathcal{A}_f$ to be nonempty. To prove this result we first need to verify the next important lemma.

Lemma 1.40 *If* $f : \mathbb{R} \to [-\infty, \infty]$ *is an arbitrary function and* $f_{\bar{c}}(\mathbf{x}_0)$ *is finite for some* $\mathbf{x}_0$, *then the set* $\mathcal{A}_f$ *is nonempty.*

Proof. It follows that the vector $(\mathbf{x}_0, f_{\bar{c}}(\mathbf{x}_0) - 1)$ does not belong to the set $epi(f_{\bar{c}})$. By Lemma 1.37 the nonempty set $epi(f_{\bar{c}})$ is convex and closed and applying Theorem 1.1, there exists some nonzero vector $(\mathbf{y}_0, \beta)$ satisfying

$$\mathbf{y}_0^\top \mathbf{x} + \beta r > \mathbf{y}_0^\top \mathbf{x}_0 + \beta(f_{\bar{c}}(\mathbf{x}_0) - 1)$$

for every $(\mathbf{x}, r) \in epi(f_{\bar{c}})$. Since $(\mathbf{x}_0, f_{\bar{c}}(\mathbf{x}_0))$ belongs to $epi(f_{\bar{c}})$ this implies $\beta > 0$ and so for every $(\mathbf{x}, r) \in epi(f_{\bar{c}})$ the inequality

$$r > -\beta^{-1}\mathbf{y}_0^\top(\mathbf{x} - \mathbf{x}_0) + f_{\bar{c}}(\mathbf{x}_0) - 1 \tag{1.71}$$

holds. By relation (1.71) it follows by contradiction that $f_{\bar{c}}(\mathbf{x}) > -\infty$ for every $\mathbf{x} \in dom(f_{\bar{c}})$ and this yields using $dom(f) \subseteq dom(f_{\bar{c}})$ that $(\mathbf{x}, f_{\bar{c}}(\mathbf{x})) \in epi(f_{\bar{c}})$ for every $\mathbf{x} \in dom(f)$. Substituting this into relation (1.71) we obtain

$$f(\mathbf{x}) \geq f_{\bar{c}}(\mathbf{x}) > -\beta^{-1}\mathbf{y}_0^\top(\mathbf{x} - \mathbf{x}_0) + f_{\bar{c}}(\mathbf{x}_0) - 1$$

for every $\mathbf{x} \in dom(f)$. Since the previous inequality trivially holds for $\mathbf{x} \notin dom(f)$ the function $a(\mathbf{x}) := -\beta^{-1}\mathbf{y}_0^\top(\mathbf{x} - \mathbf{x}_0) + f_{\bar{c}}(\mathbf{x}_0) - 1$ is an affine minorant of f and the desired result is proved. □

Using Lemma 1.40 one can show the following theorem.

Theorem 1.9 *For any function* $f : \mathbb{R}^n \to [-\infty, \infty]$ *the following conditions are equivalent:*

1 The set $\mathcal{A}_f$ *is nonempty.*

2 $f_c > -\infty$.

3 $f_{\bar{c}} > -\infty$.

Proof. If the set $\mathcal{A}_f$ is nonempty then for any $a \in \mathcal{A}_f$ we obtain by relation (1.65) that $f_{\bar{c}}(\mathbf{x}) \geq a(\mathbf{x})$ for every $\mathbf{x} \in \mathbb{R}^n$, and this shows the implication 1 $\Rightarrow$ 3. Due to $f_{\bar{c}} \leq f_c$ the implication 3 $\Rightarrow$ 2 is obvious. To show the implication 2 $\Rightarrow$ 1 consider some f satisfying $f_c > -\infty$. In case $dom(f_c)$ is empty it follows that $f \equiv \infty$ and so trivially $\mathcal{A}_f$ is nonempty. Therefore assume that $dom(f_c)$ is nonempty. By Lemma 1.26 this is a nonempty convex set and so by Lemma 1.13 one can find some $\mathbf{x}_0 \in ri(dom(f_c))$. Since $f_c > -\infty$ is a convex function it follows by Lemma 1.36 that $-\infty < f_c(\mathbf{x}_0) = \overline{f_c}(\mathbf{x}_0) = f_{\bar{c}}(\mathbf{x}_0) < \infty$ and so we have found some $\mathbf{x}_0$ satisfying $f_{\bar{c}}(\mathbf{x}_0)$ is finite. Applying now Lemma 1.40 yields $\mathcal{A}_f$ is nonempty and the result is proved. □

As shown by the following example it is not true that $\mathcal{A}_f$ is nonempty for $f > -\infty$.

Example 1.12 For the concave function $f : \mathbb{R} \to \mathbb{R}$ given by $f(x) = -x^2$ it is easy to verify that $co(epi(f)) = \mathbb{R}^2$ and $f > -\infty$. Hence we obtain that $\mathcal{A}_{f_c}$ is empty and this yields by Lemma 1.39 that $\mathcal{A}_f$ is empty.

To prove an important representation for a subclass of convex functions we introduce the following definition.

Definition 1.24 *The function* $f : \mathbb{R}^n \to [-\infty, \infty]$ *belongs to the set* $\Gamma(\mathbb{R}^n)$ *iff is convex and l.s.c. and* $f > -\infty$.

It is now possible to prove the following representation for the set $\Gamma(\mathbb{R}^n)$. This result is known as Minkowski's theorem.

Theorem 1.10 *For any function* $f : \mathbb{R}^n \to [-\infty, \infty]$ *it follows that*

$$f \in \Gamma(\mathbb{R}^n) \Leftrightarrow f(\mathbf{x}) = \sup\{a(\mathbf{x}) : a \in \mathcal{A}_f\} \text{ and the set } \mathcal{A}_f \text{ is nonempty.}$$

Proof. If the function $f : \mathbb{R}^n \to [-\infty, \infty]$ has the representation $f(\mathbf{x}) = \sup\{a(\mathbf{x}) : a \in \mathcal{A}_f\}$ and the set $\mathcal{A}_f$ is nonempty, then clearly the function f is l.s.c., convex and $f > -\infty$ and so f belongs to $\Gamma(\mathbb{R}^n)$. To prove the reverse implication, we observe for $f \in \Gamma(\mathbb{R}^n)$ that $f_c = f > -\infty$ and this shows by Theorem 1.9 that the set $\mathcal{A}_f$ is nonempty and hence $f(\mathbf{x}) \geq \sup\{a(\mathbf{x}) : a \in \mathcal{A}_f\} > -\infty$. Suppose now by contradiction that $f(\mathbf{x}_0) > \sup\{a(\mathbf{x}_0) : a \in \mathcal{A}_f\}$ for some $\mathbf{x}_0 \in \mathbb{R}^n$. Hence one can find some $\gamma \in \mathbb{R}$ satisfying

$$f(\mathbf{x}_0) > \gamma > \sup\{a(\mathbf{x}_0) : a \in \mathcal{A}_f\}, \tag{1.72}$$

and so $(\mathbf{x}_0, \gamma) \notin epi(f)$. If $epi(f)$ is empty, then the affine function $a(\mathbf{x}) = \gamma$ is an affine minorant of f and this contradicts relation (1.72). Therefore we assume that $epi(f)$ is nonempty and since this set is closed and convex there exists by Theorem 1.1 a nonzero vector $(\mathbf{y}_0, \beta)$ and $\epsilon > 0$ satisfying

$$\mathbf{y}_0^\top \mathbf{x} + \beta r \geq \mathbf{y}_0^\top \mathbf{x}_0 + \beta\gamma + \epsilon \tag{1.73}$$

for every $(\mathbf{x}, r) \in epi(f)$. Since for $(\mathbf{x}, r) \in epi(f)$ and $h > 0$ the vector $(\mathbf{x}, r + h)$ belongs to $epi(f)$ it follows by relation (1.73) that $\beta \geq 0$. Consider now the two cases $f(\mathbf{x}_0) < \infty$ and $f(\mathbf{x}_0) = \infty$. If $f(\mathbf{x}_0) < \infty$ we obtain by relation (1.73) replacing $(\mathbf{x}, r)$ by $(\mathbf{x}_0, f(\mathbf{x}_0))$ that $\beta(f(\mathbf{x}_0) - \gamma) \geq \epsilon$ and this implies using relation (1.72) that $\beta > 0$. Hence by relation (1.73) it holds that

$$f(\mathbf{x}) \geq a(\mathbf{x}) := -\beta^{-1}\mathbf{y}_0^\top(\mathbf{x} - \mathbf{x}_0) + \gamma$$

for every $\mathbf{x}$ belonging to $dom(f)$ and we have found some $a \in \mathcal{A}_f$ satisfying $a(\mathbf{x}_0) = \gamma$ contradicting relation (1.72). If $f(\mathbf{x}_0) = \infty$ and $\beta > 0$ in relation (1.73), then by the same proof we obtain a contradiction and so we consider the last case $f(\mathbf{x}_0) = \infty$ and $\beta = 0$. Introduce now the affine function $a_0 : \mathbb{R}^n \to \mathbb{R}$, given by

$$a_0(\mathbf{x}) = -\mathbf{y}_0^\top(\mathbf{x} - \mathbf{x}_0) + \epsilon.$$

By relation (1.73) $a_0(\mathbf{x}) \leq 0$ for every $\mathbf{x} \in dom(f)$ and $a_0(\mathbf{x}_0) > 0$. Since $\mathcal{A}_f$ is nonempty, select some $a \in \mathcal{A}_f$ and by relation (1.72) it follows that $\lambda_0 := a_0(\mathbf{x}_0)^{-1}(\gamma - a(\mathbf{x}_0) > 0$. Introducing now the affine function $a_{\lambda_0} : \mathbb{R}^n \to \mathbb{R}$ given by

$$a_{\lambda_0}(\mathbf{x}) := a(\mathbf{x}) + \lambda_0 a_0(\mathbf{x})$$

we obtain $a_{\lambda_0}(\mathbf{x}_0) = \gamma$ and since $a_0(\mathbf{x}) \leq 0$ for every $\mathbf{x} \in dom(f)$ and $a \in \mathcal{A}_f$ we also obtain $a_{\lambda_0} \in \mathcal{A}_f$. Hence a_{λ_0} is an affine minorant of f satisfying $a_{\lambda_0}(\mathbf{x}_0) = \gamma$ and this contradicts relation (1.72) showing the desired result. □

An immediate consequence of Minkowski's theorem and Lemma 1.39 is listed in the next result.

Theorem 1.11 *If* $\underline{f} : \mathbb{R}^n \to [-\infty, \infty]$ *is a function satisfying* $f_{\overline{c}} > -\infty$, *then it follows that* $\overline{f_c}(\mathbf{x}) = f_{\overline{c}}(\mathbf{x}) = \sup\{a(\mathbf{x}) : a \in \mathcal{A}_f\}$ *and the set* $\mathcal{A}_f$ *is nonempty.*

Proof. By relation (1.66) and Theorem 1.10 we obtain that $\overline{f_c}(\mathbf{x}) = f_{\overline{c}}(\mathbf{x}) = \sup\{a(\mathbf{x}) : a \in \mathcal{A}_{f_{\overline{c}}}\}$ with the set $\mathcal{A}_{f_{\overline{c}}}$ is nonempty. Applying now Lemma 1.39 the desired result follows. □

In Theorem 1.11 we only guarantee that any function $f_{\overline{c}} > -\infty$ can be approximated from below by affine functions. However, it is sometimes useful to derive an approximation formula in terms of the original function f. This formula was first constructed in its general form by Fenchel (cf. [21]) and it has an easy geometrical interpretation (cf. [27]).

Definition 1.25 *For any function* $f : \mathbb{R}^n \to [-\infty, \infty]$ *the function* $f^* : \mathbb{R}^n \to [-\infty, \infty]$ *given by* $f^*(\mathbf{a}) := \sup\{\mathbf{a}^\top\mathbf{x} - f(\mathbf{x}) : \mathbf{x} \in \mathbb{R}^n\}$ *is called the conjugate function of the function* f. *The function* $f^{**} : \mathbb{R}^n \to [-\infty, \infty]$ *given by* $f^{**}(\mathbf{x}) := \sup\{\mathbf{a}^\top\mathbf{x} - f^*(\mathbf{a}) : \mathbf{a} \in \mathbb{R}^n\}$ *is called the biconjugate function of* f.

By the above definition it is immediately clear that the conjugate function f^* is convex and l.s.c.. Moreover, if the function $f : \mathbb{R}^n \to [-\infty, \infty]$ is proper and the set $\mathcal{A}_f$ of affine minorants is nonempty, then it is easy to verify that the function f^* is also proper. As shown by the next result the biconjugate function has a clear geometrical interpretation.

Lemma 1.41 *If* $f : \mathbb{R}^n \to [-\infty, \infty]$ *is an arbitrary function satisfying* $\mathcal{A}_f$ *is nonempty, then it follows that* $(\mathbf{a}, r) \in epi(f^*)$ *if and only if* $a \in \mathcal{A}_f$ *with* $a(\mathbf{x}) = \mathbf{a}^\top\mathbf{x} - r$. *Additionally, it holds that* $f^{**}(\mathbf{x}) = \sup\{a(\mathbf{x}) : a \in \mathcal{A}_f\}$ *for every* $\mathbf{x} \in \mathbb{R}^n$.

Proof. To verify the equivalence relation we observe for $a(\mathbf{x}) = \mathbf{a}^\top\mathbf{x} - r \leq f(\mathbf{x})$ for every $\mathbf{x} \in \mathbb{R}^n$ that $r \geq f^*(\mathbf{a}) = \sup\{\mathbf{a}^\top\mathbf{x} - f(\mathbf{x}) : \mathbf{x} \in \mathbb{R}^n\}$ or $(\mathbf{a}, r) \in epi(f^*)$. Moreover, if $(\mathbf{a}, r) \in epi(f^*)$ we obtain $r \geq f^*(\mathbf{a})$ and this implies for every $\mathbf{x} \in \mathbb{R}^n$ that $a(\mathbf{x}) = \mathbf{a}^\top\mathbf{x} - r \leq f(\mathbf{x})$. To prove the relation for the biconjugate function it follows by the definition of $epi(f^*)$ that $f^{**}(\mathbf{x}) = \sup\{\mathbf{a}^\top\mathbf{x} - r : (\mathbf{a}, r) \in epi(f^*)\}$. Since by the first part $(\mathbf{a}, r) \in epi(f^*)$ if and only if $a(\mathbf{x}) = \mathbf{a}^\top\mathbf{x} - r$ is an affine minorant of f, this shows that $f^{**}(\mathbf{x}) = \sup\{a(\mathbf{x}) : a \in \mathcal{A}_f\}$ for every $\mathbf{x} \in \mathbb{R}^n$ and hence the equality for the biconjugate function is verified. □

To prove one of the most important theorems in convex analysis we introduce the definition of the closure of the function f.

Definition 1.26 *If* $f : \mathbb{R}^n \to [-\infty, \infty]$ *is an arbitrary function, then the closure* $cl(f) : \mathbb{R}^n \to [-\infty, \infty]$ *of the function* f *is given by*

$$cl(f) := \begin{cases} \overline{f} & \text{if } \overline{f} > -\infty \\ -\infty & \text{otherwise} \end{cases} .$$

Clearly the function $cl(f)$ is l.s.c. and satisfies $cl(f) \leq \overline{f}$. Also it is easy to verify by Lemma 1.41, Theorem 1.9 and using $\mathcal{A}_f = \mathcal{A}_{\overline{f}}$ that

$$cl(f)^* = f^* \tag{1.74}$$

for any convex function f. The next result is known as the Fenchel-Moreau theorem and is one of the most important results in convex analysis.

Theorem 1.12 *For any function* $f : \mathbb{R}^n \to [-\infty, \infty]$ *it follows that* $f^{**}(\mathbf{x}) = cl(f_c)(\mathbf{x})$ *for every* $\mathbf{x} \in \mathbb{R}^n$.

Proof. If $f_{\overline{c}}(\mathbf{x}_0) = -\infty$ for some $\mathbf{x}_0 \in \mathbb{R}^n$ then $f^* \equiv \infty$. To show this, suppose by contradiction that $f^*(\mathbf{a}_0) < \infty$ for some $\mathbf{a}_0$. This implies the existence of some $r \in \mathbb{R}$ satisfying $r \geq \mathbf{a}_0^\top \mathbf{x} - f(\mathbf{x})$ for every $\mathbf{x} \in \mathbb{R}^n$ and so the function $a(\mathbf{x}) = \mathbf{a}_0^\top \mathbf{x} - r$ is an affine minorant of f. Hence by relation (1.65) we obtain that $f_{\overline{c}}(\mathbf{x}_0) > -\infty$ and this contradicts our initial assumption. Since $f^* \equiv \infty$ we obtain $f^{**} \equiv -\infty$ and by Definition 1.26 we obtain $f^{**} = cl(f_c)$. In case $f_{\overline{c}} > -\infty$ the result follows by Theorem 1.11 and Lemma 1.41. □

An important consequence of the Fenchel-Moreau theorem is given by the following result. Recall a function is *sublinear,* if it is positively homogeneous and convex.

Lemma 1.42 *Any l.s.c. sublinear function* $f : \mathbb{R}^n \to (-\infty, \infty]$ *has the representation*

$$f(\mathbf{x}) = \sup\{\mathbf{a}^\top \mathbf{x} : \mathbf{a} \in C\}$$

with $C = \{\mathbf{a} \in \mathbb{R}^n : f^*(\mathbf{a}) \leq 0\}$ *a nonempty closed convex set.*

Proof. By the Fenchel Moreau theorem it follows that

$$f(\mathbf{x}) = f^{**}(\mathbf{x}) = \sup_{\mathbf{a} \in \mathbb{R}^n} \{\mathbf{a}^\top \mathbf{x} - f^*(\mathbf{a})\}.$$

Since f is positively homogeneous we obtain by Lemma 1.25 that

$$\alpha f^*(\mathbf{a}) = \sup_{\mathbf{x} \in \mathbb{R}^n} \{\mathbf{a}^\top (\alpha \mathbf{x}) - f(\alpha \mathbf{x})\} = f^*(\mathbf{a})$$

for every $\alpha > 0$ and $\mathbf{a} \in \mathbb{R}^n$ and this shows that $f^*(\mathbf{a}) \in \{\infty, -\infty, 0\}$. If $f^*(\mathbf{a}) = \infty$ for every $\mathbf{a} \in \mathbb{R}^n$, then $f^{**}(\mathbf{x}) = -\infty$ for every $\mathbf{x}$ and this shows by the Fenchel Moreau theorem that $f(\mathbf{x}) = -\infty$ for every $\mathbf{x}$, contradicting $f > -\infty$. Therefore f^* is not identically ∞ and this yields that the set C is not empty. Again by the Fenchel Moreau theorem we obtain

$$f(\mathbf{x}) = f^{**}(\mathbf{x}) = \sup_{\mathbf{a}\in\mathbb{R}^n}\{\mathbf{a}^\top\mathbf{x} - f^*(\mathbf{a})\} = \sup_{\mathbf{a}\in C}\mathbf{a}^\top\mathbf{x}.$$

and since the function f^* is l.s.c. and convex the nonempty set C is closed and convex. □

Finally we introduce the so-called subgradient set of a function at a point.

Definition 1.27 *For any function $f : \mathbb{R}^n \to [-\infty, \infty]$ and $\mathbf{x}_0 \in \mathbb{R}^n$ the subset of $\mathbb{R}^n$ consisting of those vectors $\mathbf{a}_0$ satisfying $f(\mathbf{x}) \geq f(\mathbf{x}_0) + \mathbf{a}_0^\top(\mathbf{x} - \mathbf{x}_0)$ for every $\mathbf{x} \in \mathbb{R}^n$ is called the subgradient set of the function f at the point $\mathbf{x}_0$. This set is denoted by $\partial f(\mathbf{x}_0)$ and its elements are called subgradients.*

If $f(\mathbf{x}_0) = -\infty$, then clearly $\partial f(\mathbf{x}_0) = \mathbb{R}^n$ and so it is sufficient to consider those $\mathbf{x}_0 \in \mathbb{R}^n$ satisfying $f(\mathbf{x}_0) > -\infty$. Moreover, if $f(\mathbf{x}_0) > -\infty$ and dom(f) is empty, then again $\partial f(\mathbf{x}_0) = \mathbb{R}^n$ and hence we only need to consider $f(\mathbf{x}_0) > -\infty$ and dom(f) is not empty. If $\mathbf{x}_0 \notin dom(f)$ or $f(\mathbf{x}_0) = \infty$, then this implies, using dom(f) is nonempty, that $\partial f(\mathbf{x}_0) = \emptyset$ and so the only interesting case which remains is given by $f(\mathbf{x}_0)$ finite. It is now relatively easy to prove for $f(\mathbf{x}_0)$ finite that $\partial f(\mathbf{x}_0) \neq \emptyset$ is equivalent to another condition related to the conjugate function.

Lemma 1.43 *If $f : \mathbb{R}^n \to [-\infty, \infty]$ is an arbitrary function satisfying $f(\mathbf{x}_0)$ is finite for some $\mathbf{x}_0$, then it follows that $\mathbf{a}_0 \in \partial f(\mathbf{x}_0)$ if and only if $f(\mathbf{x}_0) + f^*(\mathbf{a}_0) = \mathbf{a}_0^\top\mathbf{x}_0$.*

Proof. If $\mathbf{a}_0 \in \partial f(\mathbf{x}_0)$ then by definition $f(\mathbf{x}) \geq f(\mathbf{x}_0) + \mathbf{a}_0^\top(\mathbf{x} - \mathbf{x}_0)$ for every $\mathbf{x}$ and this implies using $f(\mathbf{x}_0)$ is finite that $\mathbf{a}_0^\top\mathbf{x}_0 - f(\mathbf{x}_0) \geq \mathbf{a}_0^\top\mathbf{x} - f(\mathbf{x})$ for every $\mathbf{x}$. Hence we obtain that $\mathbf{a}_0^\top\mathbf{x}_0 - f(\mathbf{x}_0) = f^*(\mathbf{a}_0)$ and this shows the equality. To verify the reverse implication is trivial and so we omit its proof. □

Up to now we did not show any existence result for the subgradient set of f at $\mathbf{x}_0$ in case $f(\mathbf{x}_0)$ is finite. Such a result will be given by the next theorem.

Theorem 1.13 *If the function $f : \mathbb{R}^n \to [-\infty, \infty]$ is convex and $f(\mathbf{x}_0)$ is finite for some $\mathbf{x}_0 \in ri(dom(f))$, then the set $\partial f(\mathbf{x}_0)$ is nonempty.*

Proof. If $\mathbf{x}_0 \in ri(dom(f))$ and $f(\mathbf{x}_0)$ is finite we obtain by Lemma 1.35 that $(\mathbf{x}_0, f(\mathbf{x}_0)) \notin ri(epi(f))$. This implies by the convexity of the set $epi(f)$ and Theorem 1.2 that there exists some nonzero vector $(\mathbf{y}_0, \beta) \in L_{aff(epi(f))}$ satisfying

$$\mathbf{y}_0^\intercal \mathbf{x} + \beta r \geq \mathbf{y}_0^\intercal \mathbf{x}_0 + \beta f(\mathbf{x}_0) \tag{1.75}$$

for $(\mathbf{x}, r) \in epi(f)$. Moreover, using $(\mathbf{x}_0, f(\mathbf{x}_0) + h)$ belongs to epi(f) for every $h \geq 0$, we obtain $\beta \geq 0$ and to show that $\beta > 0$ assume by contradiction that $\beta = 0$. Hence it follows by relation (1.75) that

$$\mathbf{y}_0^\intercal \mathbf{x} \geq \mathbf{y}_0^\intercal \mathbf{x}_0 \tag{1.76}$$

for every $(\mathbf{x}, r) \in epi(f)$. Since $aff(epi(f)) = aff(dom(f)) \times \mathbb{R}$ and so

$$L_{aff(epi(f))} = L_{aff(dom(f))} \times \mathbb{R}$$

we know that $\mathbf{y}_0$ belongs to $L_{aff(dom(f))}$. This implies, using $\mathbf{x}_0$ belongs to $ri(dom(f))$, that there exists some $\epsilon > 0$ satisfying $\mathbf{x}_0 - \epsilon \mathbf{y}_0 \in dom(f)$ and applying now relation (1.76) with $\mathbf{x}$ replaced by $\mathbf{x}_0 - \epsilon \mathbf{y}_0$ yields $-\epsilon \|\mathbf{y}_0\|^2 \geq 0$. Hence it follows that $(\mathbf{y}_0, \beta) = \mathbf{0}$ and we obtain a contradiction. Therefore it must hold that $\beta > 0$ and dividing now the inequality in relation (1.75) by $\beta > 0$ and using that $f(\mathbf{x})$ is finite for every $\mathbf{x} \in dom(f)$ yields

$$f(\mathbf{x}) \geq f(\mathbf{x}_0) - \beta^{-1} \mathbf{y}_0^\intercal (\mathbf{x} - \mathbf{x}_0)$$

for every $\mathbf{x} \in dom(f)$. This shows that the vector $\mathbf{a}_0 = -\beta^{-1} \mathbf{y}_0$ is a subgradient of the function f at the point $\mathbf{x}_0$ and so $\partial f(\mathbf{x}_0)$ is a nonempty set. □

In case $\mathbf{x}_0$ does not belong to $ri(dom(f))$ for some convex function f it might happen that f does not have a subgradient at the point $\mathbf{x}_0$. This is shown by the following example.

Example 1.13 Consider the convex function $f : \mathbb{R} \to (-\infty, \infty]$ given by $f(x) = -\sqrt{x}$ for $x \geq 0$ and $f(x) = \infty$ otherwise. Clearly 0 belongs to the relative boundary of $dom(f)$ but $\partial f(0)$ is empty.

In case the function $f > -\infty$ is a sublinear function one can show the following improvement of Theorem 1.13 replacing the condition $\mathbf{0} \in ri(dom(f))$ by the condition $\mathbf{0} \in dom(f)$.

Theorem 1.14 *If the function $f : \mathbb{R}^n \to (-\infty, \infty]$ is sublinear and $\mathbf{0} \in dom(f)$, then the set $\partial f(\mathbf{0})$ is nonempty and $\partial f(\mathbf{0}) = \{\mathbf{a} \in \mathbb{R}^n : f^*(\mathbf{a}) \leq 0\}$.*

Proof. Since f is convex it follows that $f_c = \overline{f} > -\infty$ and this implies by Theorem 1.9 that $\mathcal{A}_f$ is nonempty and so $\overline{f}$ is a proper function. Since by Definition 1.13 and Lemma 1.31 the function $\overline{f}$ is also sublinear one may apply Lemma 1.42 and this shows $\overline{f}(\mathbf{x}) = \sup\{\mathbf{a}^\top \mathbf{x} : \mathbf{a} \in C\}$ with $C = \{\mathbf{a} \in \mathbb{R}^n : \overline{f}^*(\mathbf{a}) \leq 0\}$ a nonempty closed convex set. By relation (1.74) and $\overline{f} = cl(f)$ it follows that $\overline{f}^* = f^*$ and so $C = \{\mathbf{a} \in \mathbb{R}^n : f^*(\mathbf{a}) \leq 0\}$. We will now verify that $\partial f(\mathbf{0}) = C$. By the definition of f^* we obtain for $\mathbf{a} \in C$ that $f(\mathbf{x}) \geq \mathbf{a}^\top \mathbf{x}$ for every $\mathbf{x} \in \mathbb{R}^n$. Since $f(0)$ is finite and f positively homogeneous it follows that $f(0) = 0$ and so it follows that $\mathbf{a} \in \partial f(\mathbf{0})$. This shows $C \subseteq \partial f(\mathbf{0})$ and to verify the reverse inclusion we observe for every $\mathbf{a} \in \partial f(\mathbf{0})$ that $f(\mathbf{x}) \geq \mathbf{a}^\top \mathbf{x}$ for every $\mathbf{x}$. This implies $f^*(\mathbf{a}) \leq 0$ and so $\mathbf{a}$ belongs to C. Hence $C = \partial f(\mathbf{0})$ is nonempty and the proof is completed. □

In Theorem 1.14 we actually show for $f : \mathbb{R}^n \to (-\infty, \infty]$ sublinear and $\mathbf{0} \in dom(f)$ that

$$\overline{f}(\mathbf{x}) = \sup\{\mathbf{a}^\top \mathbf{x} : \mathbf{a} \in \partial f(\mathbf{0})\} \text{ and } \partial f(\mathbf{0}) \neq \emptyset. \tag{1.77}$$

A nice implication of Theorem 1.13 is the observation that convex functions have remarkable continuity properties. Before showing this result we need the following technical lemmas.

Lemma 1.44 *If the vectors* $\mathbf{z}_i \in \mathbb{R}^n, 1 \leq i \leq k \leq n$ *form an orthonormal system and the set* P *is the convex hull generated by the set* $S = \{\mathbf{z}_1, ..., \mathbf{z}_k, -\mathbf{z}_1, ..., -\mathbf{z}_k\}$, *then it follows that*

$$k^{-\frac{1}{2}} E \cap lin(\{\mathbf{z}_1, ..., \mathbf{z}_k\}) \subseteq P.$$

Proof. Since the vectors $\mathbf{z}_i, 1 \leq i \leq k$ form an orthonormal system we obtain for any vector $\alpha^\top = (\alpha_1, ..., \alpha_k) \in \mathbb{R}^k$ that

$$\|\sum_{i=1}^k \alpha_i \mathbf{z}_i\|^2 = \|\alpha\|^2. \tag{1.78}$$

Applying now the Cauchy-Schwartz inequality to the inner product of the vectors $|\alpha|^\top := (|\alpha_1|, ..., |\alpha_k|)$ and $\mathbf{e}^\top = (1, ..., 1)$ it follows that

$$\sum_{i=1}^k |\alpha_i| = < |\alpha|, \mathbf{e} > \leq \|\alpha\| k^{\frac{1}{2}}$$

and this implies by relation (1.78) that

$$\|\sum_{i=1}^k \alpha_i \mathbf{z}_i\| \geq k^{-\frac{1}{2}} \sum_{i=1}^k |\alpha_i|. \tag{1.79}$$

Consider now an arbitrary vector $\mathbf{y}$ belonging to $k^{-\frac{1}{2}}E \cap lin(\{\mathbf{z}_1,...,\mathbf{z}_k\})$. Since the vectors $\mathbf{z}_i, 1 \leq i \leq k$ are independent, there exists a unique vector $\alpha^\top = (\alpha_1,...,\alpha_k) \in \mathbb{R}^k$ such that

$$\mathbf{y} = \sum_{i \in I} \alpha_i \mathbf{z}_i + \sum_{i \notin I} -\alpha_i(-\mathbf{z}_i)$$

with $I := \{1 \leq i \leq k : \alpha_i > 0\}$. Applying now the inequality in relation (1.79) it follows that

$$k^{-\frac{1}{2}} \geq \|\mathbf{y}\| \geq k^{-\frac{1}{2}} \left(\sum_{i \in I} \alpha_i + \sum_{i \notin I} -\alpha_i\right)$$

and this shows that the vector $\mathbf{y}$ belongs to P. □

Another result which is needed in the proof of Theorem 1.15 is given by the following lemma.

Lemma 1.45 *If the function* $f : \mathbb{R}^n \to (-\infty, \infty]$ *is convex and for some* $\mathbf{x}_0$ *and* $\delta > 0$ *there exists some finite constants* m, M *satisfying* $m \leq f(\mathbf{x}) \leq M$ *for every* $\mathbf{x}$ *belonging to* $(\mathbf{x}_0 + 2\delta E) \cap dom(f)$, *then one can find some* $L > 0$ *satisfying*

$$|f(\mathbf{x}_1) - f(\mathbf{x}_2)| \leq L\|\mathbf{x}_1 - \mathbf{x}_2\|$$

for every $\mathbf{x}_1, \mathbf{x}_2$ *belonging to* $(\mathbf{x}_0 + \delta E) \cap ri(dom(f))$.

Proof. Let $\mathbf{x}_1$ and $\mathbf{x}_2$ be two different vectors belonging to $(\mathbf{x}_0 + \delta E) \cap ri(dom(f))$. This yields that the vector $\mathbf{x}_1 - \mathbf{x}_2$ belongs to $L_{aff(dom(f))}$ and since $\mathbf{x}_1$ is a relative interior point of the convex set $dom(f)$ one can find some $0 < \epsilon < \delta$ satisfying

$$\mathbf{x}_3 := \mathbf{x}_1 + \epsilon\|\mathbf{x}_1 - \mathbf{x}_2\|^{-1}(\mathbf{x}_1 - \mathbf{x}_2) \in dom(f). \tag{1.80}$$

Hence the vector $\mathbf{x}_3$ belongs to $(\mathbf{x}_0 + 2\delta E) \cap dom(f)$ and by relation (1.80) we obtain

$$\mathbf{x}_1 = \frac{\|\mathbf{x}_1 - \mathbf{x}_2\|}{\|\mathbf{x}_1 - \mathbf{x}_2\| + \epsilon}\mathbf{x}_3 + \frac{\epsilon}{\|\mathbf{x}_1 - \mathbf{x}_2\| + \epsilon}\mathbf{x}_2.$$

Using now relation (1.48) and the fact that the function f is bounded from above and below on $(\mathbf{x}_0 + 2\delta E) \cap dom(f)$ it follows for $L := \epsilon^{-1}(M - m)$ that

$$f(\mathbf{x}_1) - f(\mathbf{x}_2) \leq \frac{\|\mathbf{x}_1 - \mathbf{x}_2\|}{\|\mathbf{x}_1 - \mathbf{x}_2\| + \epsilon}(f(\mathbf{x}_3) - f(\mathbf{x}_2)) \leq L\|\mathbf{x}_1 - \mathbf{x}_2\|.$$

Reversing the roles of $\mathbf{x}_1$ and $\mathbf{x}_2$ yields a similar bound for $f(\mathbf{x}_2) - f(\mathbf{x}_1)$ and the desired inequality is verified. □

The above property of the function f is called *Lipschitz continuity* on the set $(\mathbf{x}_0 + \delta E) \cap ri(dom(f))$. Using Lemmas 1.45, 1.44 and Theorem 1.13 one can now show the next result, which is an improvement of Lemma 1.36.

Theorem 1.15 *If* $f : \mathbb{R}^n \to (-\infty, \infty]$ *is a convex function, then it follows that* f *is continuous on* $ri(dom(f))$ *and Lipschitz continuous on every compact subset of* $ri(dom(f))$.

Proof. If $\mathbf{x}_0 \in ri(dom(f))$ one can find some $\epsilon > 0$ satisfying

$$(\mathbf{x}_0 + 2\epsilon E) \cap aff(dom(f)) \subseteq dom(f). \tag{1.81}$$

To give a more detailed characterization of $aff(dom(f))$ we observe by Lemma 1.4, that there exists a set of $k \leq n$ linearly independent vectors $\mathbf{z}_1, ..., \mathbf{z}_k$ satisfying $L_{aff(dom(f))} = lin(\{\mathbf{z}_1, ..., \mathbf{z}_k\})$ and so

$$aff(dom(f)) = \mathbf{x}_0 + lin(\{\mathbf{z}_1, ..., \mathbf{z}_k\}). \tag{1.82}$$

Without loss of generality (Use the well-known Gram-Schmidt orthogonalization process (cf. [47])) we may assume that the set $\{\mathbf{z}_1, ..., \mathbf{z}_k\}$ is an orthonormal system. By relations (1.82) and (1.81) and $dom(f)$ a convex set it follows that the set $\mathbf{x}_0 + P$ with P the convex hull generated by the set $S = \{\epsilon\mathbf{z}_1, ..., \epsilon\mathbf{z}_k, -\epsilon\mathbf{z}_1, ..., -\epsilon\mathbf{z}_k\}$ belongs to $dom(f)$. Also by the convexity of the function f and relation (1.48) we obtain that

$$f(\mathbf{x}) \leq \max\{f(\mathbf{x}_0 + \epsilon\mathbf{z}_i), f(\mathbf{x}_0 - \epsilon\mathbf{z}_i), 1 \leq i \leq k\} < \infty$$

for every $\mathbf{x} \in \mathbf{x}_0 + P$. Since by Lemma 1.44 there exists some $\gamma > 0$ satisfying

$$(\mathbf{x}_0 + \gamma E) \cap aff(dom(f)) \subseteq P.$$

this shows that the function f is bounded from above on $(\mathbf{x}_0 + \gamma E) \cap dom(f)$. Using Theorem 1.13 we also obtain that the function f is bounded from below on $(\mathbf{x}_0 + \gamma E) \cap dom(f)$ and applying now Lemma 1.45 with 2δ replaced by γ yields the desired result. □

This concludes our discussion on dual representations and conjugation for convex functions. In the next subsection we consider the same topic for quasiconvex functions.

3.4 Dual representations of quasiconvex functions

In this section we study dual representations of evenly quasiconvex and l.s.c. quasiconvex functions. Most of the results of this section can be found in [56]. Unfortunately in [56] no geometrical interpretation of the results are given and for such an interpretation the reader should consult [27]. In [56] it is shown, that one can use the same approach as in convex analysis and this results in proving that certain subsets of quasiconvex functions can be approximated from below by so-called c-affine functions with $c : \mathbb{R} \to [-\infty, \infty]$ belonging to a given class $\mathcal{C}$ of extended real valued univariate functions. Recall that a function is called univariate if its domain is given by $\mathbb{R}$. As in convex analysis the used approximations and the generalized biconjugate functions have a clear geometrical interpretation (cf. [27]). To start with this approach we introduce in the next definition the class of c-affine functions. More general classes of so-called coupling functions a are discussed in this volume by Martínez-Legaz (cf. [49]).

Definition 1.28 *For a given univariate function* $c : \mathbb{R} \to [-\infty, \infty]$ *the function* $a : \mathbb{R}^n \to [-\infty, \infty]$ *is called a* c*-affine function, if there exist some* $\mathbf{a} \in \mathbb{R}^n$ *and* $r \in \mathbb{R}$ *such that* $a(\mathbf{x}) = c(\mathbf{a}^\top \mathbf{x}) + r$ *for every* $\mathbf{x} \in \mathbb{R}^n$*. If* $\mathcal{C}$ *denotes a subset of the set of extended real valued univariate functions the function* a *is called a* $\mathcal{C}$*-affine function, if for some* $c \in \mathcal{C}$ *the function* a *is a* c*-affine function. The function* a *is called a* $\mathcal{C}$*-affine minorant of the function* $f : \mathbb{R}^n \to [-\infty, \infty]$ *if* $a(\mathbf{x}) \leq f(\mathbf{x})$ *for every* $\mathbf{x} \in \mathbb{R}^n$ *and* a *is a* $\mathcal{C}$*-affine function. The set* $\mathcal{C}\mathcal{A}_f$ *denotes now the (possibly empty) set of* $\mathcal{C}$*-affine minorants of* f.

To specify the set $\mathcal{C}$ we first consider the set $\mathcal{C}_0$ of extended real valued nondecreasing univariate functions $c : \mathbb{R} \to [-\infty, \infty]$ and the proper subset $\mathcal{C}_1 \subseteq \mathcal{C}_0$ of extended real valued nondecreasing l.s.c. univariate functions. Since for any $c \in \mathcal{C}_i, i = 0, 1$ and $r \in \mathbb{R}$ also the function $c^* : \mathbb{R} \to [-\infty, \infty]$, given by $c^*(t) = c(t) + r$, belongs to $\mathcal{C}_i$, $i = 0, 1$, we observe for these classes of extended real valued univariate functions that the class of $\mathcal{C}_i$-affine functions, $i = 0, 1$ reduces to the set of functions $a : \mathbb{R}^n \to [-\infty, \infty]$, given by $a(\mathbf{x}) = c(\mathbf{a}^\top \mathbf{x})$ for some $\mathbf{a} \in \mathbb{R}^n$ and $c \in \mathcal{C}_i$. Clearly $\mathcal{C}_1\mathcal{A}_f \subseteq \mathcal{C}_0\mathcal{A}_f$ and since the function $a : \mathbb{R}^n \to [-\infty, \infty]$ with $a(\mathbf{x}) = -\infty$ for every $\mathbf{x} \in \mathbb{R}^n$ belongs to the set $\mathcal{C}_1$, we obtain that $\mathcal{C}_1\mathcal{A}_f$ is nonempty for every $f : \mathbb{R}^n \to [-\infty, \infty]$. This is a major difference with the set of affine minorants of a function f, since this set might be empty. Observe in Theorem 1.9 we showed that this set is nonempty if and only if $f_c > -\infty$. One can now show the following result for $\mathcal{C}$-affine functions with $\mathcal{C}$ either equal to $\mathcal{C}_1$ or $\mathcal{C}_0$.

Lemma 1.46 *If $a : \mathbb{R}^n \to [-\infty, \infty]$ is $\mathcal{C}_0$-affine, then the function a is evenly quasiconvex. Moreover, if a is $\mathcal{C}_1$-affine, then the function a is l.s.c. and quasiconvex.*

Proof. If a is a $\mathcal{C}_0$-affine function, then there exists some $c \in \mathcal{C}_0$ and $\mathbf{a} \in \mathbb{R}^n$ such that $L(a,r) = \{\mathbf{x} \in \mathbb{R}^n : \mathbf{a}^\top \mathbf{x} \in L(c,r)\}$ for every $r \in \mathbb{R}$ with $L(c,r)$ the lower level set of the function c. Since c is nondecreasing, this lower level set is either empty or an interval given by $(-\infty, \beta_r)$ or $(-\infty, \beta_r]$ with $\beta_r := \sup\{t \in \mathbb{R} : c(t) \leq r\}$. Hence the set $L(a,r)$ is either empty or an open or closed halfspace and this shows that $L(a,r)$ is evenly convex. Similarly for $c \in \mathcal{C}_1$ we obtain, using Theorem 1.7, that $L(c,r)$ is empty or $(-\infty, \beta_r]$ and hence $L(a,r)$ is empty or a closed halfspace. This shows that the function a is quasiconvex and by Theorem 1.7 it is also l.s.c.. □

By Lemma 1.28, 1.29, 1.38 and 1.46 and $f_{\overline{q}} = \overline{f_q} \leq f_{ec} \leq f_q \leq f$ (see relations (1.69) and 1.70).) one can show, applying a similar proof as in Lemma 1.39, that the following result holds.

Lemma 1.47 *For any function $f : \mathbb{R}^n \to [-\infty, \infty]$ it follows that $\mathcal{C}_0\mathcal{A}_f = \mathcal{C}_0\mathcal{A}_{f_{ec}}$ and $\mathcal{C}_1\mathcal{A}_f = \mathcal{C}_1\mathcal{A}_{f_q} = \mathcal{C}_1\mathcal{A}_{\overline{f_q}} = \mathcal{C}_1\mathcal{A}_{f_{\overline{q}}}$.*

Contrary to functions studied in convex analysis, we do not have to determine for which extended real valued functions the sets $\mathcal{C}_i\mathcal{A}_f, i = 1,2$ are nonempty and so we can start generalizing Minkowsky's theorem (see Theorem 1.10) to evenly quasiconvex and l.s.c. quasiconvex functions. In the proof of this generalization and in the remainder of this subsection an important role is played by the following functions.

Definition 1.29 *For any function $f : \mathbb{R}^n \to [-\infty, \infty]$ and $\mathbf{a} \in \mathbb{R}^n$, let $c_{\mathbf{a}} : \mathbb{R} \to [-\infty, \infty]$ denote the function $c_{\mathbf{a}}(t) := \inf\{f(\mathbf{y}) : \mathbf{a}^\top \mathbf{y} \geq t\}$.*

It is now possible to show the following result.

Theorem 1.16 *If $f : \mathbb{R}^n \to [-\infty, \infty]$ is an evenly quasiconvex function, then $f(\mathbf{x}) = \sup\{a(\mathbf{x}) : a \in \mathcal{C}_0\mathcal{A}_f\}$ for every $\mathbf{x} \in \mathbb{R}^n$. Moreover, if f is an l.s.c. quasiconvex function, then $f(\mathbf{x}) = \sup\{a(\mathbf{x}) : a \in \mathcal{C}_1\mathcal{A}_f\}$ for every $\mathbf{x} \in \mathbb{R}^n$.*

Proof. Since the set $\mathcal{C}_0\mathcal{A}_f$ is nonempty, we obtain by the definition of $\mathcal{C}_0\mathcal{A}_f$ that $f(\mathbf{x}) \geq \sup\{a(\mathbf{x}) : a \in \mathcal{C}_0\mathcal{A}_f\}$ for every $\mathbf{x} \in \mathbb{R}^n$. Suppose now by contradiction that $f(\mathbf{x}_0) > \sup\{a(\mathbf{x}_0) : a \in \mathcal{C}_0\mathcal{A}_f\}$ for some $\mathbf{x}_0$ and so there exists some $\gamma \in \mathbb{R}$ satisfying

$$f(\mathbf{x}_0) > \gamma > \sup\{a(\mathbf{x}_0) : a \in \mathcal{C}_0\mathcal{A}_f\}. \tag{1.83}$$

If the set $L(f,\gamma)$ is empty, it follows that $f(\mathbf{x}) > \gamma$ for every $\mathbf{x} \in \mathbb{R}^n$ and choosing $c(t) = \gamma$ for every $t \in \mathbb{R}$ and $a(\mathbf{x}) = c(\mathbf{a}^\top \mathbf{x})$ with $\mathbf{a} \in \mathbb{R}^n$ arbitrary, we obtain that $a \in \mathcal{C}_1\mathcal{A}_f \subseteq \mathcal{C}_0\mathcal{A}_f$ contradicting relation (1.83). Therefore the set $L(f,\gamma)$ is nonempty and since the function f is evenly quasiconvex one can find a collection of vectors $(\mathbf{a}_i, b_i)_{i\in I}$ satisfying

$$L(f,\gamma) = \cap_{i\in I} H^{<}(\mathbf{a}_i, b_i). \tag{1.84}$$

By relation (1.83) the vector $\mathbf{x}_0$ does not belong to $L(f,\gamma)$ and this shows by relation (1.84) that there exists some $i \in I$ with a nonzero $\mathbf{a}_i$ satisfying $\mathbf{a}_i^\top \mathbf{x}_0 \geq b_i$. This implies again by relation (1.84) that

$$\{\mathbf{y} \in \mathbb{R}^n : \mathbf{a}_i^\top \mathbf{y} \geq \mathbf{a}_i^\top \mathbf{x}_0\} \subseteq \{\mathbf{y} \in \mathbb{R}^n : f(\mathbf{y}) > \gamma\}. \tag{1.85}$$

Since the vector $\mathbf{a}_i$ is nonzero, the function $c_{\mathbf{a}_i}$, given in Definition 1.29, is nondecreasing and so the function $a(\mathbf{x}) := c_{\mathbf{a}_i}(\mathbf{a}_i^\top \mathbf{x})$ is $\mathcal{C}_0$-affine and by relation (1.85) it satisfies $a(\mathbf{x}_0) \geq \gamma$. Also for every $\mathbf{x} \in \mathbb{R}^n$ we obtain that $a(\mathbf{x}) \leq f(\mathbf{x})$ and so we have constructed a $\mathcal{C}_0$-affine minorant a of the function f satisfying $a(\mathbf{x}_0) \geq \gamma$. This contradicts relation (1.83) and hence we have shown that $f(\mathbf{x}) = \sup\{a(\mathbf{x}) : a \in \mathcal{C}_0\mathcal{A}_f\}$ for every $\mathbf{x} \in \mathbb{R}^n$. To verify the representation for f quasiconvex and l.s.c. we again assume by contradiction that there exists some $\gamma \in \mathbb{R}$ satisfying

$$f(\mathbf{x}_0) > \gamma > \sup\{a(\mathbf{x}_0) : a \in \mathcal{C}_1\mathcal{A}_f\} \tag{1.86}$$

for some $\mathbf{x}_0$. If the convex set $L(f,\gamma)$ is empty then as in the first part we obtain a contradiction. Therefore the closed convex set $L(f,\gamma)$ is nonempty and since by relation (1.86) it holds that $\mathbf{x}_0$ does not belong to $L(f,\gamma)$, there exist by Theorem 1.1 some nonzero vector $\mathbf{a}_0 \in \mathbb{R}^n$ and $\beta \in \mathbb{R}$ satisfying $\mathbf{a}_0^\top \mathbf{x} < \beta < \mathbf{a}_0^\top \mathbf{x}_0$ for every $\mathbf{x} \in L(f,\gamma)$. This implies for every $\mathbf{y}$ satisfying $\mathbf{a}_0^\top \mathbf{y} \geq \beta$ that $f(\mathbf{y}) > \gamma$ and so $c_{\mathbf{a}_0}(\beta) \geq \gamma$. Introducing now the function $a(\mathbf{x}) := c_{\mathbf{a}_0}^{\diamond}(\mathbf{a}_0^\top \mathbf{x})$ with $c_{\mathbf{a}_0}^{\diamond}(t)$ listed in Definition 1.20 this implies

$$a(\mathbf{x}_0) = c_{\mathbf{a}_0}^{\diamond}(\mathbf{a}_0^\top \mathbf{x}_0) = \sup_{s<\mathbf{a}_0^\top \mathbf{x}_0} c_{\mathbf{a}_0}(s) \geq c_{\mathbf{a}_0}(\beta) \geq \gamma.$$

By Lemma 1.32 the function $c_{\mathbf{a}_0}^{\diamond}$ is l.s.c. and $c_{\mathbf{a}_0}^{\diamond}(\mathbf{a}_0^\top \mathbf{x}) \leq c_{\mathbf{a}_0}(\mathbf{a}_0^\top \mathbf{x}) \leq f(\mathbf{x})$ for every $\mathbf{x}$. Hence we have constructed a $\mathcal{C}_1$-affine minorant a of the function f satisfying $a(\mathbf{x}_0) \geq \gamma$ and this contradicts relation (1.86).

Therefore $f(\mathbf{x}) = \sup\{a(\mathbf{x}) : a \in \mathcal{C}_1\mathcal{A}_f\}$ for every $\mathbf{x} \in \mathbb{R}^n$ and the proof is completed. □

By Theorem 1.16 it is clear that the set of $\mathcal{C}_1$-affine functions ($\mathcal{C}_0$-affine functions) play the same role for l.s.c. quasiconvex functions (evenly quasiconvex functions) as the affine functions do for l.s.c. convex functions. However, besides this observation, it is also interesting to investigate the question whether these sets of $\mathcal{C}$-affine minorants are the smallest possible class satisfying the above property. In this section we will also pay attention to this question. An immediate consequence of Theorem 1.16 and Lemma 1.47 is given by the next result.

Theorem 1.17 *For any function* $f : \mathbb{R}^n \to [-\infty, \infty]$ *it follows that* $f_{ec}(\mathbf{x}) = \sup\{a(\mathbf{x}) : a \in \mathcal{C}_0\mathcal{A}_f\}$ *and* $f_{\overline{q}}(\mathbf{x}) = \overline{f_q}(\mathbf{x}) = \sup\{a(\mathbf{x}) : a \in \mathcal{C}_1\mathcal{A}_f\}$ *for every* $\mathbf{x} \in \mathbb{R}^n$.

Proof. By Theorem 1.16 we obtain $f_{ec}(\mathbf{x}) = \sup\{a(\mathbf{x}) : a \in \mathcal{C}_0\mathcal{A}_{f_{ec}}\}$ and since by Lemma 1.47 it holds that $\mathcal{C}_0\mathcal{A}_f = \mathcal{C}_0\mathcal{A}_{f_{ec}}$ the first formula follows. The second formula can be verified similarly. □

Studying the proof of Theorem 1.16 for evenly quasiconvex functions one can actually show the following improvement of Theorem 1.17.

Theorem 1.18 *If* $f : \mathbb{R}^n \to [-\infty, \infty]$ *is an arbitrary function, then it follows for every* $\mathbf{x} \in \mathbb{R}^n$ *that*

$$f_{ec}(\mathbf{x}) = \sup_{\mathbf{a} \in \mathbb{R}^n} c_{\mathbf{a}}(\mathbf{a}^\top \mathbf{x})$$

with the function $c_{\mathbf{a}}$ *given in Definition 1.29.*

Proof. It follows for every $\mathbf{a}$ and $\mathbf{x} \in \mathbb{R}^n$ that $c_{\mathbf{a}}(\mathbf{a}^\top \mathbf{x}) \leq f(\mathbf{x})$. Since $c_{\mathbf{a}} \in \mathcal{C}_0$ this implies by Lemma 1.46 that the function $\mathbf{x} \to c_{\mathbf{a}}(\mathbf{a}^\top \mathbf{x})$ is evenly quasiconvex and so by Lemma 1.29 we obtain for every $\mathbf{x} \in \mathbb{R}^n$ that $f_{ec}(\mathbf{x}) \geq \sup_{\mathbf{a} \in \mathbb{R}^n} c_{\mathbf{a}}(\mathbf{a}^\top \mathbf{x})$. Suppose now by contradiction that $f_{ec}(\mathbf{x}_0) > \sup_{\mathbf{a} \in \mathbb{R}^n} c_{\mathbf{a}}(\mathbf{a}^\top \mathbf{x}_0)$ for some $\mathbf{x}_0$ and so there exists some $\gamma \in \mathbb{R}$ satisfying

$$f_{ec}(\mathbf{x}_0) > \gamma > \sup_{\mathbf{a} \in \mathbb{R}^n} c_{\mathbf{a}}(\mathbf{a}^\top \mathbf{x}_0). \tag{1.87}$$

If the set $L(f_{ec}, \gamma)$ is empty we obtain $f(\mathbf{x}) \geq f_{ec}(\mathbf{x}) > \gamma$ for every $\mathbf{x} \in \mathbb{R}^n$ and this implies $c_{\mathbf{a}}(\mathbf{a}^\top \mathbf{x}_0) \geq \gamma$ for every $\mathbf{a} \in \mathbb{R}^n$ contradicting relation (1.87). Therefore the set $L(f_{ec}, \gamma)$ is nonempty and since by Lemma 1.29 the function f_{ec} is evenly quasiconvex one can find a collection of vectors $(\mathbf{a}_i, b_i)_{i \in I}$ satisfying

$$L(f_{ec}, \gamma) = \cap_{i \in I} H^{<}(\mathbf{a}_i, b_i). \tag{1.88}$$

By relation (1.87) we know $\mathbf{x}_0$ does not belong to $L(f_{ec}, \gamma)$ and so by relation (1.88) there exists some $i \in I$ and a nonzero vector $\mathbf{a}_i$ satisfying $\mathbf{a}_i^\top \mathbf{x}_0 \geq b_i$. This implies using $f(\mathbf{x}) \geq f_{ec}(\mathbf{x})$ for every $\mathbf{x} \in \mathbb{R}^n$ and relation (1.88) that

$$\{\mathbf{y} \in \mathbb{R}^n : \mathbf{a}_i^\top \mathbf{y} \geq \mathbf{a}_i^\top \mathbf{x}_0\} \subseteq \{\mathbf{y} \in \mathbb{R}^n : f(\mathbf{y}) > \gamma\}$$

and so it follows that $c_{\mathbf{a}_i}(\mathbf{a}_i^\top \mathbf{x}_0) \geq \gamma$. This yields $\sup_{\mathbf{a} \in \mathbb{R}^n} c_{\mathbf{a}}(\mathbf{a}^\top \mathbf{x}_0) \geq c_{\mathbf{a}_i}(\mathbf{a}_i^\top \mathbf{x}_0) \geq \gamma$ contradicting relation (1.87). This shows the desired representation and our proof is completed. □

Also for l.s.c. quasiconvex functions one can show the following improvement of Theorem 1.16. Observe this formula is more complicated than the corresponding formula for evenly quasiconvex functions.

Theorem 1.19 *If* $f : \mathbb{R}^n \to [-\infty, \infty]$ *is an arbitrary function, then it follows for every* $\mathbf{x} \in \mathbb{R}^n$ *that*

$$f_{\overline{q}}(\mathbf{x}) = \overline{f_q}(\mathbf{x}) = \sup_{\mathbf{a} \in \mathbb{R}^n} \overline{c_{\mathbf{a}}}(\mathbf{a}^\top \mathbf{x}) = \sup_{\mathbf{a} \in \mathbb{R}^n} c_{\mathbf{a}}^{\Diamond}(\mathbf{a}^\top \mathbf{x})$$

with $\overline{c_{\mathbf{a}}}$ *denoting the l.s.c. hull of the function* $c_{\mathbf{a}}$ *and* $c_{\mathbf{a}}^{\Diamond}$ *listed in Definition 1.20.*

Proof. By Lemma 1.32 and relation 1.70 it is sufficient to show for every $\mathbf{x} \in \mathbb{R}^n$ that $\overline{f_q}(\mathbf{x}) = \sup_{\mathbf{a} \in \mathbb{R}^n} c_{\mathbf{a}}^{\Diamond}(\mathbf{a}^\top \mathbf{x})$. To verify this we first observe for every $\mathbf{a}$ and $\mathbf{x} \in \mathbb{R}^n$ that $c_{\mathbf{a}}(\mathbf{a}^\top \mathbf{x}) \leq f(\mathbf{x})$ and so we obtain $c_{\mathbf{a}}^{\Diamond}(\mathbf{a}^\top \mathbf{x}) \leq f(\mathbf{x})$ for every $\mathbf{x}$. By Lemma 1.32 the function $c_{\mathbf{a}}^{\Diamond} : \mathbb{R} \to [-\infty, \infty]$ is l.s.c. and nondecreasing and this implies by Lemma 1.46 that $\mathbf{x} \to c_{\mathbf{a}}^{\Diamond}(\mathbf{a}^\top \mathbf{x})$ is quasiconcex and l.s.c.. Therefore we obtain for every $\mathbf{x}$ that

$$\overline{f_q}(\mathbf{x}) \geq \sup_{\mathbf{a} \in \mathbb{R}^n} c_{\mathbf{a}}^{\Diamond}(\mathbf{a}^\top \mathbf{x}).$$

Suppose now by contradiction that $\overline{f_q}(\mathbf{x}_0) > \sup_{\mathbf{a} \in \mathbb{R}^n} c_{\mathbf{a}}^{\Diamond}(\mathbf{a}^\top \mathbf{x}_0)$ for some $\mathbf{x}_0$ and so there exists some $\gamma \in \mathbb{R}$ satisfying

$$\overline{f_q}(\mathbf{x}_0) > \gamma > \sup_{\mathbf{a} \in \mathbb{R}^n} c_{\mathbf{a}}^{\Diamond}(\mathbf{a}^\top \mathbf{x}_0). \tag{1.89}$$

If the set $L(\overline{f_q}, \gamma)$ is empty we obtain $f(\mathbf{x}) \geq \overline{f_q}(\mathbf{x}) > \gamma$ and we obtain as in Theorem 1.18 a contradiction with relation (1.89). Therefore, the closed convex set $L(\overline{f_q}, \gamma)$ is nonempty and since by relation (1.89) it holds that $\mathbf{x}_0$ does not belong to $L(\overline{f_q}, \gamma)$ there exist by Theorem 1.1 some nonzero vector $\mathbf{a}_0 \in \mathbb{R}^n$ and $\beta \in \mathbb{R}$ satisfying $\mathbf{a}_0^\top \mathbf{x} < \beta < \mathbf{a}_0^\top \mathbf{x}_0$ for every $\mathbf{x} \in L(\overline{f_q}, \gamma)$. Hence it follows for every $\mathbf{y}$ satisfying $\mathbf{a}_0^\top \mathbf{y} \geq \beta$ that $f(\mathbf{y}) \geq \overline{f_q}(\mathbf{y}) > \gamma$ and this yields $c_{\mathbf{a}_0}(\beta) \geq \gamma$. Using this observation we obtain

$$c_{\mathbf{a}_0}^{\Diamond}(\mathbf{a}_0^\top \mathbf{x}_0) = \sup_{s < \mathbf{a}_0^\top \mathbf{x}_0} c_{\mathbf{a}_0}(s) \geq c_{\mathbf{a}_0}(\beta) \geq \gamma$$

and this contradicts relation (1.89) completing the proof. □

It is also possible to show for every $\mathbf{a} \in \mathbb{R}^n$ that the function $c_{\mathbf{a}}^{\diamond}$ is actually the inverse of another function.

Lemma 1.48 *If $f : \mathbb{R}^n \to [-\infty, \infty]$ is a function with dom(f) nonempty and the function $h_{\mathbf{a}} : \mathbb{R} \to [-\infty, \infty], \mathbf{a} \in \mathbb{R}^n$ is given by $h_{\mathbf{a}}(\alpha) := \sup\{\mathbf{a}^\top \mathbf{y} : \mathbf{y} \in L(f, \alpha)\}$, then it follows for every $t \in \mathbb{R}$ that*

$$c_{\mathbf{a}}^{\diamond}(t) = \inf\{\alpha \in \mathbb{R} : h_{\mathbf{a}}(\alpha) \geq t\}.$$

Proof. Since dom(f) is nonempty, there exists some $\alpha \in \mathbb{R}$ satisfying $L(f, \alpha)$ is nonempty. If for some $\alpha_0 \in \mathbb{R}$ it follows that $h_{\mathbf{a}}(\alpha_0) \geq t$, then for every $s < t$ there exists some $\mathbf{y}_0$ satisfying $f(\mathbf{y}_0) \leq \alpha_0$ and $\mathbf{a}^\top \mathbf{y}_0 \geq s$. This implies $\alpha_0 \geq f(\mathbf{y}_0) \geq c_{\mathbf{a}}(s)$ and hence $\inf\{\alpha \in \mathbb{R} : h_{\mathbf{a}}(\alpha) \geq t\} \geq c_{\mathbf{a}}(s)$. Since $s < t$ we obtain $\inf\{\alpha \in \mathbb{R} : h_{\mathbf{a}}(\alpha) \geq t\} \geq \sup_{s<t} c_{\mathbf{a}}(s) = c_{\mathbf{a}}^{\diamond}(t)$ and to show equality we assume by contradiction that there exists some t_0 satisfying

$$\inf\{\alpha \in \mathbb{R} : h_{\mathbf{a}}(\alpha) \geq t_0\} > c_{\mathbf{a}}^{\diamond}(t_0).$$

If this holds one can find some α_0 satisfying $\alpha_0 > c_{\mathbf{a}}^{\diamond}(t_0)$ and $h_{\mathbf{a}}(\alpha_0) < t_0$. Hence there exists some $\epsilon > 0$ satisfying $\alpha_0 > c_{\mathbf{a}}^{\diamond}(t_0)$ and $h_{\mathbf{a}}(\alpha_0) < t_0 - \epsilon$. Since $h_{\mathbf{a}}(\alpha_0) < t_0 - \epsilon$ we obtain for every $\mathbf{y}$ satisfying $\mathbf{a}^\top \mathbf{y} \geq t_0 - \epsilon$ that $f(\mathbf{y}) > \alpha_0$. This implies $c_{\mathbf{a}}(t_0 - \epsilon) \geq \alpha_0$ and it follows $\alpha_0 > c_{\mathbf{a}}^{\diamond}(t_0) \geq c_{\mathbf{a}}(t_0 - \epsilon) \geq \alpha_0$. This is clearly a contradiction and the proof is completed. □

In case dom(f) is empty and so $f \equiv \infty$ and we use the well-known convention that $\sup\{\varnothing\} = -\infty$ and $\inf\{\varnothing\} = \infty$ then it is easy to verify that the above relation still holds. The next result first verified in [15] is an immediate consequence of Lemma 1.48 and Theorem 1.19.

Theorem 1.20 *If $f : \mathbb{R}^n \to [-\infty, \infty]$ is an arbitrary function, then it follows that*

$$\overline{f_q}(\mathbf{x}) = f_{\overline{q}}(\mathbf{x}) = \sup_{\mathbf{a} \in \mathbb{R}^n} \inf\{\alpha \in \mathbb{R} : \sup_{\mathbf{y} \in L(f,\,\alpha)} \mathbf{a}^\top \mathbf{y} \geq \mathbf{a}^\top \mathbf{x}\}$$

for every $\mathbf{x} \in \mathbb{R}^n$.

Actually the result in Theorem 1.18 and 1.19 can be seen as a generalization of the Fenchel-Moreau theorem for l.s.c. convex hulls. To show this we need to generalize the notion of conjugate and biconjugate functions used within convex analysis. Since we are dealing with extended real valued functions we use the convention that $(-\infty) + (+\infty) = (+\infty) + (-\infty) = -\infty$ and $-(-\infty) = \infty$.

Definition 1.30 *Let $\mathcal{C}$ be a nonempty collection of extended real valued univariate functions. For any function $f : \mathbb{R}^n \to [-\infty, \infty]$ and $c \in \mathcal{C}$ the function $f^c(\mathbf{a}) := \sup\{c(\mathbf{a}^\top \mathbf{x}) - f(\mathbf{x}) : \mathbf{x} \in \mathbb{R}^n\}$ is called the c-conjugate function of the function f. The function $f^{\mathcal{C}\mathcal{C}}(\mathbf{x}) := \sup\{c(\mathbf{x}^\top \mathbf{a}) - f^c(\mathbf{a}) : \mathbf{a} \in \mathbb{R}^n, c \in \mathcal{C}\}$ is called the bi-$\mathcal{C}$-conjugate function of f.*

By a similar proof as in Lemma 1.41 it is easy to give a geometrical interpretation of the biconjugate function.

Lemma 1.49 *For $\mathcal{C}$ a nonempty collection of extended real valued univariate functions and $f : \mathbb{R}^n \to [-\infty, \infty]$ an arbitrary function it follows that $(\mathbf{a}, r) \in epi(f^c)$ if and only if $a \in \mathcal{C}\mathcal{A}_f$ with $a(\mathbf{x}) = c(\mathbf{a}^\top \mathbf{x}) - r$ and $c \in \mathcal{C}$. Additionally, it holds that $f^{\mathcal{C}\mathcal{C}}(\mathbf{x}) = \sup\{a(\mathbf{x}) : a \in \mathcal{C}\mathcal{A}_f\}$ for every $\mathbf{x} \in \mathbb{R}^n$.*

Combining now Lemma 1.49 and Theorem 1.17 we immediately obtain for the sets $\mathcal{C}_i, i = 0, 1$ the following generalization of the Fenchel-Moreau theorem.

Theorem 1.21 *For any function $f : \mathbb{R}^n \to [-\infty, \infty]$ it follows that $f^{\mathcal{C}_0\mathcal{C}_0}(\mathbf{x}) = f_{ec}(\mathbf{x})$ and $f^{\mathcal{C}_1\mathcal{C}_1}(\mathbf{x}) = \overline{f_q}(\mathbf{x}) = f_{\overline{q}}(\mathbf{x})$ for every $\mathbf{x} \in \mathbb{R}^n$.*

Proof. By Lemma 1.49 we obtain $f^{\mathcal{C}_i\mathcal{C}_i}(\mathbf{x}) = \sup\{a(\mathbf{x}) : a \in \mathcal{C}_i\mathcal{A}_f\}, i = 0, 1$ and this shows by Theorem 1.17 the desired result. □

By Theorem 1.18, 1.19 and 1.21 we obtain the formulas

$$f^{\mathcal{C}_0\mathcal{C}_0}(\mathbf{x}) = f_{ec}(\mathbf{x}) = \sup_{\mathbf{a} \in \mathbb{R}^n} c_{\mathbf{a}}(\mathbf{a}^\top \mathbf{x})$$

and

$$f^{\mathcal{C}_1\mathcal{C}_1}(\mathbf{x}) = \overline{f_q}(\mathbf{x}) = f_{\overline{q}}(\mathbf{x}) = \sup_{\mathbf{a} \in \mathbb{R}^n} c_{\mathbf{a}}^{\Diamond}(\mathbf{a}^\top \mathbf{x}) \tag{1.90}$$

for every $\mathbf{x} \in \mathbb{R}^n$. Considering these formulas we now wonder whether it is possible to achieve the same result using a smaller set of extended real valued univariate functions.

Definition 1.31 *For any $r \in \mathbb{R}$ the function $c_r : \mathbb{R} \to [-\infty, \infty]$ is given by $c_r(t) = -\infty$ for $t < r$ and $c_r(t) = r$ for every $t \geq r$. The set $\mathcal{C}_r \subseteq \mathcal{C}_0$ consists now of all functions $c_r, r \in \mathbb{R}$, while the set $\overline{\mathcal{C}_r}$ consists of all functions $\overline{c_r}$, $r \in \mathbb{R}$ with $\overline{c_r}$ the l.s.c. hull of the function c_r.*

If $f : \mathbb{R}^n \to (-\infty, \infty]$ is an arbitrary function, then for $r \in \mathbb{R}$ and $\mathbf{a} \neq \mathbf{0}$ we obtain

$$f^{c_r}(\mathbf{a}) = \max\{-\infty, \sup\{r - f(\mathbf{y}) : \mathbf{a}^\top \mathbf{y} \geq r\}\} = r - c_{\mathbf{a}}(r) \tag{1.91}$$

with c_a defined in Definition 1.29. Moreover, for $\mathbf{a} = \mathbf{0}$ and $r \leq 0$, it follows that $f^{c_r}(\mathbf{0}) = \sup\{c_r(0) - f(\mathbf{y}) : \mathbf{y} \in \mathbb{R}^n\} = r - \inf\{f(\mathbf{y}) : \mathbf{y} \in \mathbb{R}^n\}$ and this shows

$$f^{c_r}(\mathbf{0}) = r - c_{\mathbf{0}}(r) = r - c_{\mathbf{0}}(0), r \leq 0. \tag{1.92}$$

Also for $r > 0$ it is easy to verify that $f^{c_r}(\mathbf{0}) = -\infty$ and so we have computed for every $r \in \mathbb{R}$ the c_r-conjugate function of the function f. To evaluate the $\overline{c_r}$-conjugate function of f we observe by Lemma 1.32 that $\overline{c_r}(t) = -\infty$ for every $t \leq r$ and $\overline{c_r}(t) = r$ for every $t > r$. Again considering $\mathbf{a} \neq \mathbf{0}$ it follows that

$$\begin{aligned} f^{\overline{c_r}}(\mathbf{a}) &= \max\{-\infty, \sup\{r - f(\mathbf{y}) : \mathbf{a}^\top \mathbf{y} > r\}\} \\ &= r - \inf\{f(\mathbf{y}) : \mathbf{a}^\top \mathbf{y} > r\}. \end{aligned} \tag{1.93}$$

Moreover, for $\mathbf{a} = 0$ and $r < 0$, we obtain that

$$\begin{aligned} f^{\overline{c_r}}(\mathbf{0}) &= \sup\{\overline{c_r}(0) - f(\mathbf{y}) : \mathbf{y} \in \mathbb{R}^n\} \\ &= r - \inf\{f(\mathbf{y}) : \mathbf{y} \in \mathbb{R}^n\} = r - c_{\mathbf{0}}(r), \end{aligned} \tag{1.94}$$

while for $r \geq 0$ it is easy to verify that $f^{\overline{c_r}}(\mathbf{0}) = -\infty$. Using the above computations we will first evaluate in the proof of Lemma 1.50 the bi-$\mathcal{C}_r$-conjugate function of a function $f : \mathbb{R}^n \to (-\infty, \infty]$, while in the proof of Lemma 1.51 the same computation will be carried out for a bi-$\overline{\mathcal{C}_r}$-conjugate function of the same function f.

Lemma 1.50 *For every* $\mathbf{x} \in \mathbb{R}^n$ *and* $f : \mathbb{R}^n \to (-\infty, \infty]$ *it follows for every* $\mathbf{x} \in \mathbb{R}^n$ *that*

$$f^{\mathcal{C}_r \mathcal{C}_r}(\mathbf{x}) = \sup_{\mathbf{a} \in \mathbb{R}^n} \inf\{f(\mathbf{y}) : \mathbf{a}^\top \mathbf{y} \geq \mathbf{a}^\top \mathbf{x}\} = f_{ec}(\mathbf{x}).$$

Proof. By relation (1.92) and $f^{c_r}(\mathbf{0}) = -\infty$ for every $r > 0$ we obtain using the convention $-\infty - (-\infty) = -\infty + \infty = -\infty$ that

$$\sup_{r \in \mathbb{R}}\{c_r(0) - f^{c_r}(\mathbf{0})\} = \sup_{r \leq 0} c_{\mathbf{0}}(0) = c_{\mathbf{0}}(0). \tag{1.95}$$

Also by relation (1.91) and $(-\infty) - (-\infty) = (-\infty) + \infty = -\infty$ it follows for every $\mathbf{x}$ that

$$\sup_{\mathbf{a} \neq \mathbf{0},\, r \in \mathbb{R}}\{c_r(\mathbf{a}^\top \mathbf{x}) - f^{c_r}(\mathbf{a})\} = \sup_{\mathbf{a} \neq \mathbf{0},\, r \leq \mathbf{a}^\top \mathbf{x},\, r \in \mathbb{R}} c_{\mathbf{a}}(r).$$

This shows, using $c_{\mathbf{a}}$ is nondecreasing for every $\mathbf{a} \neq \mathbf{0}$, that

$$\sup_{\mathbf{a} \neq \mathbf{0},\, r \in \mathbb{R}}\{c_r(\mathbf{a}^\top \mathbf{x}) - f^{c_r}(\mathbf{a})\} = \sup_{\mathbf{a} \neq \mathbf{0}} c_{\mathbf{a}}(\mathbf{a}^\top \mathbf{x}) \tag{1.96}$$

and so $f^{\mathcal{C}_r\mathcal{C}_r}(\mathbf{x}) = \sup_{\mathbf{a}\in\mathbb{R}^n} c_{\mathbf{a}}(\mathbf{a}^\top\mathbf{x})$ using relations (1.95) and (1.96). This shows the first equality and the second one is already listed in Theorem 1.18. □

The next result yields a similar result as Lemma 1.50 for a quasiconvex and l.s.c. function.

Lemma 1.51 *For every* $\mathbf{x} \in \mathbb{R}^n$ *and* $f : \mathbb{R}^n \to (-\infty, \infty]$ *it follows for every* $\mathbf{x} \in \mathbb{R}^n$ *that*

$$f^{\overline{\mathcal{C}_r}\,\overline{\mathcal{C}_r}}(\mathbf{x}) = \sup_{\mathbf{a}\in\mathbb{R}^n} \sup_{s<\mathbf{a}^\top\mathbf{x}} \inf\{f(\mathbf{y}) : \mathbf{a}^\top\mathbf{y} \geq s\} = \overline{f_q}(\mathbf{x}) = f_{\overline{q}}(\mathbf{x}).$$

Proof. By relation (1.93) and $f^{\overline{\mathcal{C}_r}}(\mathbf{0}) = -\infty$ for every $r \geq 0$ we obtain using $-\infty - (-\infty) = -\infty + \infty = -\infty$ that

$$\sup_{r\in\mathbb{R}}\{\overline{c_r}(0) - f^{\overline{\mathcal{C}_r}}(\mathbf{0})\} = \sup_{r<0} c_{\mathbf{0}}(r) = c_{\mathbf{0}}^{\diamond}(0). \tag{1.97}$$

Also by relation (1.92) and $(-\infty)-(-\infty) = (-\infty)+\infty = -\infty$ it follows with $h(\mathbf{x}) := \sup_{\mathbf{a}\neq\mathbf{0},\, r\in\mathbb{R}}\{\overline{c_r}(\mathbf{a}^\top\mathbf{x}) - f^{\overline{\mathcal{C}_r}}(\mathbf{a})\}$ that

$$h(\mathbf{x}) = \sup_{\mathbf{a}\neq\mathbf{0},\, r<\mathbf{a}^\top\mathbf{x},\, r\in\mathbb{R}} \inf\{f(\mathbf{y}) : \mathbf{a}^\top\mathbf{y} > r\}. \tag{1.98}$$

Since $\inf\{f(\mathbf{y}) : \mathbf{a}^\top\mathbf{y} > r\} \geq c_{\mathbf{a}}(r)$ for every $r \in \mathbb{R}$ and $\mathbf{a} \neq \mathbf{0}$ we obtain by relation (1.98) that

$$h(\mathbf{x}) \geq \sup_{\mathbf{a}\neq\mathbf{0},\, r<\mathbf{a}^\top\mathbf{x},\, r\in\mathbb{R}} c_{\mathbf{a}}(r) = \sup_{\mathbf{a}\neq\mathbf{0}} c_{\mathbf{a}}^{\diamond}(\mathbf{a}^\top\mathbf{x}). \tag{1.99}$$

Applying now relations (1.90), (1.97) and (1.99) it holds for every $\mathbf{x} \in \mathbb{R}^n$ that

$$f^{\mathcal{C}_r\mathcal{C}_r}(\mathbf{x}) \geq \sup_{\mathbf{a}\in\mathbb{R}^n} c_{\mathbf{a}}^{\diamond}(\mathbf{a}^\top\mathbf{x}) = \overline{f_q}(\mathbf{x}) = f^{\mathcal{C}_1\mathcal{C}_1}(\mathbf{x}). \tag{1.100}$$

Since $\overline{\mathcal{C}_r} \subseteq \mathcal{C}_1$ it follows that $f^{\mathcal{C}_1\mathcal{C}_1} \geq f^{\mathcal{C}_r\mathcal{C}_r}$ and this shows by relation (1.100) the desired result. □

In the last two lemmas we have shown that it is sufficient for any function f satisfying $f > -\infty$ to consider the class of $\mathcal{C}_r$-affine minorants and the class of $\overline{\mathcal{C}_r}$-affine minorants for approximating f_{ec}, respectively $\overline{f_q}$. This concludes the section on quasiconvex duality. In the next section we will discuss some important applications.

4. On applications of convex and quasiconvex analysis

In this section we will discuss different applications of the theory of convex and quasiconvex analysis. In Subsection 4.1 we consider applications to noncooperative game theory, while in Subsection 4.2 we discuss its applications to optimization problems and in particular to Lagrangian duality. Finally in Subsection 4.3 we will use the duality representation of evenly quasiconvex functions to show that every positively homogeneous evenly quasiconvex function satisfying $f(\mathbf{0}) = 0$ and $f > -\infty$ is actually the minimum of two positively homogeneous l.s.c. convex functions. This result was first verified by Crouzeix (cf. [15]) for a slightly smaller class of quasiconvex functions and serves as a very nice application of quasiconvex duality.

4.1 Minimax theorems and noncooperative game theory

To introduce the field of infinite antagonistic game theory (cf.[72]) we assume that the set of pure strategies of player 1 is given by some nonempty set $A \subseteq \mathbb{R}^n$, while the set of pure strategies of player 2 is given by $B \subseteq \mathbb{R}^m$. If player 1 chooses the pure strategy $\mathbf{a} \in A$ and player 2 chooses the pure strategy $\mathbf{b} \in B$, then player 2 has to pay to player 1 an amount $f(\mathbf{a}, \mathbf{b})$ with $f : A \times B \to [0, \infty]$ a given function. This function is called the payoff function and for simplicity this function is taken to be nonnegative. Since player 1 likes to gain as much profit as possible, but at the moment he does not know how to achieve this, he first decides to compute a lower bound on his profit. To compute this lower bound player 1 argues as follows : if he decides to choose action $\mathbf{a} \in A$, then it follows that he wins at least $\inf_{\mathbf{b}\in B} f(\mathbf{a}, \mathbf{b})$ irrespective of the action of player 2. Therefore a lower bound on the profit for player 1 is given by

$$r_* := \sup_{\mathbf{a}\in A} \inf_{\mathbf{b}\in B} f(\mathbf{a}, \mathbf{b}). \tag{1.101}$$

Similarly player 2 likes to minimize his losses but since he does not know how to achieve this he also decides to compute first an upper bound on his losses. To compute this upper bound player 2 argues as follows. If he decides to choose action $\mathbf{b}$ it follows that he loses at most $\sup_{\mathbf{a}\in A} f(\mathbf{a}, \mathbf{b})$ and this is independent of the action of player 1. Therefore an upper bound on his losses is given by

$$r^* := \inf_{\mathbf{b}\in B} \sup_{\mathbf{a}\in A} f(\mathbf{a}, \mathbf{b}). \tag{1.102}$$

Since the profit of player 1 is at least r_* and the losses of player 2 is at most r^* and the losses of player 2 are the profits of player 1 it follows directly that $r_* \leq r^*$. In general $r_* < r^*$, but under some properties on the action set and payoff function one can show that $r_* = r^*$. By the above inequality it follows immediately that $r_* = r^*$ for $r^* = -\infty$ and so we assume in the remainder of this section that $r^* > -\infty$. The equality $r_* = r^*$ is called a minimax result and if additionally inf and sup are attained an optimal strategy for both players can be easily derived. For player 1 it is possible to achieve at least a profit r_*, independent of the action of player 2, while for player 2 it is possible to achieve at most a loss r^* independent of the action of player 1. Since $r^* = r_* := v$ and both players have opposite interests, they will choose an action which achieves the value v and so player 1 will choose that action $\mathbf{a}_0 \in A$ satisfying

$$\inf_{\mathbf{b}\in B} f(\mathbf{a}_0, \mathbf{b}) = \max_{\mathbf{a}\in A} \inf_{\mathbf{b}\in B} f(\mathbf{a}, \mathbf{b}).$$

Moreover, player 2 will choose that strategy $\mathbf{b}_0 \in B$ satisfying

$$\sup_{\mathbf{a}\in A} f(\mathbf{a}, \mathbf{b}_0) = \min_{\mathbf{b}\in B} \sup_{\mathbf{a}\in A} f(\mathbf{a}, \mathbf{b}).$$

Since for $r_* = r^*$ and the additional assumption that the infimum and supremum are attained, it is clear how the optimal strategies should be chosen we will investigate in this subsection for which payoff functions and strategies the minimax result $r_* = r^*$ holds. Before discussing this, we give the following example for which this equality does not hold.

Example 1.14 Consider the continuous payoff function $f : [0,1] \times [0,1] \to [0,\infty)$ given by $f(a,b) = (a-b)^2$. For this function it holds for every $0 \leq a \leq 1$ that $\inf_{b\in[0,1]}(a-b)^2 = 0$ and so $r_* := \sup_{0\leq a\leq 1} \inf_{0\leq b\leq 1}(a-b)^2 = 0$. Moreover, it follows that $\sup_{0\leq a\leq 1}(a-b)^2 = (1-b)^2$ for every $0 \leq b < \frac{1}{2}$ and $\sup_{0\leq a\leq 1}(a-b)^2 = b^2$ for every $\frac{1}{2} \leq b \leq 1$. This shows $r^* := \inf_{0\leq b\leq 1} \sup_{0\leq a\leq 1}(a-b)^2 = \frac{1}{4}$ and so r_* does not equal r^*. For this example it is not obvious which strategies should be selected by the two players.

By extending the sets of the so-called pure strategies of each player it is possible to show under certain conditions that the extended game satisfies a minimax result. In the next definition we introduce the set of *mixed strategies.*

Definition 1.32 *For a nonempty set D of pure strategies and* $\mathbf{d} \in D$ *let* $\epsilon_{\mathbf{d}}$ *denote the one-point probability measure concentrated on the set* $\{\mathbf{d}\}$ *and denote by* $\mathcal{P}_D$ *the set of all probability measures on D with a finite support.*

Introducing the unit simplex $\Delta_k := \{\alpha : \sum_{i=1}^k \alpha_i = 1, \alpha_i \geq 0, 1 \leq i \leq k\}$ it follows by Definition 1.32 that λ belongs to the set $\mathcal{P}_D$ if and only if there exist some $k \in \mathbb{N}$ and set $\{\mathbf{d}_1, ..., \mathbf{d}_k\} \subseteq D$ consisting of different elements such that

$$\lambda = \sum_{i=1}^{k} \lambda_i \epsilon_{\mathbf{d}_i}, (\lambda_1, ..., \lambda_k) \in \Delta_k \text{ and } \lambda_i > 0.$$

Clearly the set $\mathcal{P}_D$ can seen as the convex hull of the set $\{\epsilon_{\mathbf{d}} : \mathbf{d} \in D\}$ and so it is convex. A game theoretic interpretation of a strategy $\lambda \in \mathcal{P}_D$ is now given by the following. If a player with pure strategy set D selects the mixed strategy $\lambda = \sum_{i=1}^{k} \lambda_i \epsilon_{\mathbf{d}_i} \in \mathcal{P}_D$, then with probability $\lambda_i, 1 \leq i \leq k$ this player will use the pure strategy $\mathbf{d}_i \in D$. By this interpretation it is clear that the set D of pure strategies can be identified with the set of one-point Borel probability measures $\{\epsilon_{\mathbf{d}} : \mathbf{d} \in D\}$. We now assume that player 1 uses the set $\mathcal{P}_A$ of mixed strategies and the same holds for player 2 using the set $\mathcal{P}_B$. This means that the payoff function f should be extended to a function $f_e : \mathcal{P}_A \times \mathcal{P}_B \to \mathbb{R}$ and this extension is given by

$$f_e(\lambda, \mu) := \sum_{i=1}^{k} \sum_{j=1}^{l} \lambda_i \mu_j f(\mathbf{a}_i, \mathbf{b}_j) \tag{1.103}$$

with $\lambda = \sum_{i=1}^{k} \lambda_i \epsilon_{\mathbf{a}_i} \in \mathcal{P}_A$ and $\mu = \sum_{j=1}^{l} \mu_j \epsilon_{\mathbf{b}_j} \in \mathcal{P}_B$. This extension represents the expected profit for player 1 or expected loss of player 2 if player 1 selects the mixed strategy $\lambda \in \mathcal{P}_A$ and player 2 selects the mixed strategy $\mu \in \mathcal{P}_B$. Without any conditions on the pure strategy sets A and B and the function f one can show the next result.

Lemma 1.52 *For any set A and B of pure strategies it follows that*

$$\inf_{\mu \in \mathcal{P}_B} \sup_{\lambda \in \mathcal{P}_A} f_e(\lambda, \mu) = \inf_{\mu \in \mathcal{P}_B} \sup_{\mathbf{a} \in A} f_e(\epsilon_{\mathbf{a}}, \mu)$$

and

$$\sup_{\lambda \in \mathcal{P}_A} \inf_{\mu \in \mathcal{P}_B} f_e(\lambda, \mu) = \sup_{\lambda \in \mathcal{P}_A} \inf_{\mathbf{b} \in B} f_e(\lambda, \epsilon_{\mathbf{b}}).$$

Proof. Since $\{\epsilon_{\mathbf{a}} : \mathbf{a} \in A\} \subseteq \mathcal{P}_A$ it follows that

$$\inf_{\mu \in \mathcal{P}_B} \sup_{\lambda \in \mathcal{P}_A} f_e(\lambda, \mu) \geq \inf_{\mu \in \mathcal{P}_B} \sup_{\mathbf{a} \in A} f_e(\epsilon_{\mathbf{a}}, \mu).$$

To verify the reverse inequality we observe for every mixed strategy $\mu \in \mathcal{P}_B, \lambda \in \mathcal{P}_A$ and relation (1.103) that $f_e(\lambda, \mu) \leq \sup_{\mathbf{a} \in A} f_e(\epsilon_{\mathbf{a}}, \mu)$. This implies

$$\sup_{\lambda \in \mathcal{P}_A} f_e(\lambda, \mu) \leq \sup_{\mathbf{a} \in A} f_e(\epsilon_{\mathbf{a}}, \mu)$$

and so the first formula is verified. The second formula can be shown by exactly the same argument. □

It is now possible to show that the extended game given by f_e and the mixed strategy sets $\mathcal{P}_A$ and $\mathcal{P}_B$ satisfies a minimax result under some topological conditions on the function f and the sets A and B of pure strategies. The next result was first given by Ville (cf. [70], [18], [72]) using a much more complicated proof. In the next alternative proof we only use the separation result for convex sets listed in Theorem 1.3 and the well-known result that a continuous function on a compact set is uniformly continuous (cf. [43]).

Theorem 1.22 *If the pure strategy sets* $A \subseteq \mathbb{R}^n$ *and* $B \subseteq \mathbb{R}^m$ *are compact and the function* $f : A \times B \to \mathbb{R}$ *is continuous, then it follows that*

$$\inf_{\mu \in \mathcal{P}_B} \sup_{\lambda \in \mathcal{P}_A} f_e(\lambda, \mu) = \sup_{\lambda \in \mathcal{P}_A} \inf_{\mu \in \mathcal{P}_B} f_e(\lambda, \mu).$$

Proof. It is easy to see that the inequality $\geq$ holds and so we only need to verify the reverse inequality. By Lemma 1.52 it is now sufficient to show that $\inf_{\mu \in \mathcal{P}_B} \sup_{\mathbf{a} \in A} f_e(\epsilon_{\mathbf{a}}, \mu) \leq \sup_{\lambda \in \mathcal{P}_A} \inf_{\mathbf{b} \in B} f_e(\lambda, \epsilon_{\mathbf{b}})$. By scaling we may assume that

$$\sup_{\lambda \in \mathcal{P}_A} \inf_{\mathbf{b} \in B} f_e(\lambda, \epsilon_{\mathbf{b}}) = 1 \tag{1.104}$$

and so need to show that

$$\inf_{\mu \in \mathcal{P}_B} \sup_{\mathbf{a} \in A} f_e(\epsilon_{\mathbf{a}}, \mu) \leq 1$$

Assume now by contradiction that there exists some $\gamma > 0$ satisfying

$$\sup_{\mathbf{a} \in A} f_e(\epsilon_{\mathbf{a}}, \mu) \geq 1 + \gamma \tag{1.105}$$

for every $\mu \in \mathcal{P}_B$. Since the function f is continuous on the compact set $A \times B$, it is well-known (cf. [64], [43]) that the function f is uniformly continuous on $A \times B$. Hence there exists some $\delta > 0$ such that for every $\mathbf{a}_1, \mathbf{a}_2 \in A$ satisfying $\|\mathbf{a}_1 - \mathbf{a}_2\| \leq \delta$ it follows that $\sup_{\mathbf{b} \in B} |f(\mathbf{a}_1, \mathbf{b}) - f(\mathbf{a}_2, \mathbf{b}| \leq \frac{\gamma}{2}$. This implies for every $\mathbf{a}_1, \mathbf{a}_2 \in A$ satisfying $\|\mathbf{a}_1 - \mathbf{a}_2\| \leq \delta$ that

$$\sup_{\mu \in \mathcal{P}_B} |f_e(\epsilon_{\mathbf{a}_1}, \mu) - f_e(\epsilon_{\mathbf{a}_2}, \mu)| \leq \frac{\gamma}{2}. \tag{1.106}$$

Since A is compact one can find a finite set $I \subseteq A$ satisfying $A \subseteq \cup_{\mathbf{a} \in I}(\mathbf{a} + \delta E)$ and this shows by relations (1.106) and (1.105) that

$$\max_{\mathbf{a} \in I} f_e(\epsilon_{\mathbf{a}}, \mu) \geq \sup_{\mathbf{a} \in A} f_e(\epsilon_{\mathbf{a}}, \mu) - \frac{\gamma}{2} \geq 1 + \frac{\gamma}{2} \tag{1.107}$$

for every $\mu \in \mathcal{P}_B$. Introducing the convex set V given by

$$V := co(\{(f_e(\epsilon_{\mathbf{a}}, \epsilon_{\mathbf{b}}))_{\mathbf{a} \in I}, \mathbf{b} \in B\}) \subseteq \mathbb{R}^{|I|}$$

it follows by the definition of V that $\mathbf{z}^\top = (z_1, ..., z_{|I|})$ belongs to V if and only if there exists some mixed strategy $\mu \in \mathcal{P}_B$ satisfying $\mathbf{z}^\top = (f_e(\epsilon_{\mathbf{a}}, \mu))_{\mathbf{a} \in I}$. This implies by relation (1.107) that

$$V \subseteq \{\mathbf{z} \in \mathbb{R}^{|I|} : \max_{1 \leq i \leq |I|} z_i \geq 1 + \frac{\gamma}{2}\}$$

and so the convex sets $\{\mathbf{z} \in \mathbb{R}^{|I|} : \max_{i \leq i \leq |I|} z_i < 1 + \frac{\gamma}{2}\}$ and V are disjoint. Applying now Theorem 1.3 one can find some mixed strategy $\lambda \in \mathcal{P}_I$ satisfying $\inf_{\mathbf{b} \in B} f_e(\lambda, \epsilon_{\mathbf{b}}) \geq 1 + \frac{\gamma}{2}$ and this contradicts relation (1.104). □

Actually the result in Theorem 1.22 holds under weaker topological conditions on the function f. However the proof of that result uses the Riesz representation theorem for the set of continuous functions on a compact Hausdorff space, the Banach-Alaoglu theorem and infinite dimensional separation (cf. [29]) and is beyond the scope of this chapter. The result listed in Theorem 1.22 is the most important result in infinite antagonistic game theory and fits within a chain of equivalent minimax theorems (cf. [28]). For one of these equivalent minimax results another alternative proof using also finite dimensional separation is given in [26]. Although not listed in [28], one result which also fits within this chain is the famous Sion's minimax theorem (cf. [66]) for quasiconcave-quasiconvex bifunctions.

Theorem 1.23 *If* $A \subseteq \mathbb{R}^n$ *is compact and convex,* $B \subseteq \mathbb{R}^m$ *is convex and the function* $f : A \times B \to \mathbb{R}$ *satisfies* $\mathbf{a} \to f(\mathbf{a}, \mathbf{b})$ *is quasiconcave and upper semicontinuous for every* $\mathbf{b} \in B$ *and* $\mathbf{b} \to f(\mathbf{a}, \mathbf{b})$ *is quasiconvex and lower semicontinuous for every* $\mathbf{a} \in A$*, then it follows that*

$$\max_{\mathbf{a} \in A} \inf_{\mathbf{b} \in B} f(\mathbf{a}, \mathbf{b}) = \inf_{b \in B} \max_{a \in A} f(\mathbf{a}, \mathbf{b}).$$

This result was proved using the Knaster-Kuratowski-Mazurkiewicz (KKM) lemma (cf. [74]). This lemma is the basis of fixed point theory and nonlinear functional analysis. It is also possible to give a more elementary proof of Sion's minimax theorem based on finite dimensional separation between convex sets. However, the most elementary proof of Sion's minimax theorem is given by an adaptation of the so-called level set method due to Joó (cf. [40], [41], [42]). This method first translates the minimax equality into an equivalent geometrical condition of a nonempty intersection of a collection of upper-level sets. Under the assumptions of Sion's minimax theorem it is now possible to verify this geometrical condition using compactness arguments and the well-known elementary topological result that every convex set is connected (cf. [36]). To start with our analysis we first introduce for every $\mathbf{a} \in A$ and $\mathbf{b} \in B$ the functions $f_{\mathbf{a}} : B \to \mathbb{R}$ and $f_{\mathbf{b}} : A \to \mathbb{R}$ defined by

$$f_{\mathbf{a}}(\mathbf{b}) = f_{\mathbf{b}}(\mathbf{a}) := f(\mathbf{a}, \mathbf{b}). \tag{1.108}$$

Also we introduce for every $r \in \mathbb{R}$ and $\mathbf{b} \in B$ the upper-level set $U(f_{\mathbf{b}}, r) \subseteq A$ given by

$$U(f_{\mathbf{b}}, r) := \{\mathbf{a} \in A : f_{\mathbf{b}}(\mathbf{a}) \geq r\}. \tag{1.109}$$

It is now easy to show the following result (cf. [41]).

Lemma 1.53 *It follows that* $r^* = r_*$ *if and only if* $\cap_{b \in B} U(f_{\mathbf{b}}, r) \neq \emptyset$ *for every* $r < r^*$.

Proof. If $r^* = r_* > -\infty$, then for every $r < r^*$ there exist by the definition of r_* some $\mathbf{a}_0 \in A$ satisfying $\inf_{\mathbf{b} \in B} f(\mathbf{a}_0, \mathbf{b}) > r$. This shows that $\mathbf{a}_0$ belongs to the intersection $\cap_{\mathbf{b} \in B} U(f_{\mathbf{b}}, r)$ and so $\cap_{\mathbf{b} \in B} U(f_{\mathbf{b}}, r)$ is nonempty. To verify the reverse implication it is sufficient to verify that $r_* \geq r^*$ or equivalently $r_* > r^* - \epsilon$ for every $\epsilon > 0$. Consider now $r := r^* - \epsilon$ for some $\epsilon > 0$. By our assumption it follows that the intersection $\cap_{\mathbf{b} \in B} U(f_{\mathbf{b}}, r)$ is nonempty and so there exists some $\mathbf{a}_0 \in A$ satisfying $\inf_{\mathbf{b} \in B} f(\mathbf{a}_0, \mathbf{b}) \geq r$. This implies that $r_* = \sup_{\mathbf{a} \in A} \inf_{\mathbf{b} \in B} f(\mathbf{a}, \mathbf{b}) \geq r$ and so the proof is completed. □

By Lemma 1.53 we need to show that $\cap_{\mathbf{b} \in B} U(f_{\mathbf{b}}, r) \neq \emptyset$ for every $r < r^*$. Before proving this result we consider an arbitrary finite set $\{\mathbf{b}_0, \mathbf{b}_1, ..., \mathbf{b}_k\} \subseteq B$ and introduce the affine mapping $p : [0,1] \to B$, given by

$$p(\lambda) = \lambda \mathbf{b}_0 + (1 - \lambda)\mathbf{b}_1, \tag{1.110}$$

and the set valued mapping $\Phi_r : [0,1] \to 2^A$, given by

$$\Phi_r(\lambda) = (\cap_{i=2}^k U(f_{\mathbf{b}_i}, r)) \cap U(f_{p(\lambda)}, r). \tag{1.111}$$

To verify the main result we need the following elementary lemma.

Lemma 1.54 *If the functions* $f_{\mathbf{a}} : B \to \mathbb{R}$ *are quasiconvex on the convex set* B *for every* $\mathbf{a} \in A$, *then it follows for every* $\lambda_0, \lambda_1 \in [0,1]$ *and* $0 < \alpha < 1$ *that*

$$\Phi_r(\alpha\lambda_0 + (1 - \alpha)\lambda_1) \subseteq \Phi_r(\lambda_0) \cup \Phi_r(\lambda_1)$$

for every $r \in \mathbb{R}$.

Proof. If the vector $\mathbf{a}$ belongs to $\Phi_r(\alpha\lambda_0 + (1-\alpha)\lambda_1)$, then by definition $\mathbf{a} \in \cap_{i=2}^k U(f_{\mathbf{b}_i}, r)$ and $f(\mathbf{a}, p(\alpha\lambda_0 + (1 - \alpha)\lambda_1)) \geq r$. This implies, using p is affine, that

$$f(\mathbf{a}, \alpha p(\lambda_0) + (1 - \alpha)p(\lambda_1)) \geq r \tag{1.112}$$

and by the quasiconvexity of the functions $f_{\mathbf{a}}$ we obtain by relation (1.112) that $\max\{f(\mathbf{a},p(\lambda_0)),f(\mathbf{a},p(\lambda_1))\} \geq r$. Hence it follows that $\mathbf{a}$ belongs to $\Phi_r(\lambda_0) \cup \Phi_r(\lambda_1)$ and the result is proved. □

In order to prove the next important lemma we denote by $\mathcal{F}(B)$ the set of all finite subsets of B.

Lemma 1.55 *If the functions $f_{\mathbf{b}} : A \to \mathbb{R}$ are quasiconcave and upper semicontinuous for every $\mathbf{b} \in B$ and the functions $f_{\mathbf{a}} : B \to \mathbb{R}$ are quasiconvex and lower semicontinuous for every $\mathbf{a} \in A$, then it follows for every J belonging to $\mathcal{F}(B)$ and $r < r^*$ that $\cap_{\mathbf{b}\in J} U(f_{\mathbf{b}}, r) \neq \emptyset$.*

Proof. If J is a subset of B consisting of one element the result clearly holds by the definition of r^* listed in relation (1.102). Suppose now for all sets J belonging to $\mathcal{F}(B)$ and consisting of at most k elements that

$$\cap_{b\in J} U(f_{\mathbf{b}}, r) \neq \emptyset \tag{1.113}$$

for every $r < r^*$. To prove the result for all sets J belonging to $\mathcal{F}(B)$ consisting of at most $k+1$ elements, we assume by contradiction that there exists some set $J = \{\mathbf{b}_0, ..., \mathbf{b}_k\} \subseteq B$ and some $r < r^*$ satisfying

$$\cap_{i=0}^{k} U(f_{\mathbf{b}_i}, r) = \emptyset. \tag{1.114}$$

Consider now for the points $\mathbf{b}_0$ and $\mathbf{b}_1$ the set valued mapping $\Phi_r : [0,1] \to 2^A$ given by relation (1.111). By our induction hypothesis listed in relation (1.113) and the assumption that the functions $f_{\mathbf{b}}, \mathbf{b} \in B$ are quasiconcave and upper semicontinuous we obtain that the sets $\Phi_r(\lambda)$ are nonempty, closed and convex for every $0 \leq \lambda \leq 1$. By relation (1.114) it follows that

$$\Phi_r(0) \cap \Phi_r(1) = \emptyset \tag{1.115}$$

and so the nonempty sets

$$S_i := \{0 \leq \lambda \leq 1 : \Phi_r(\lambda) \subseteq \Phi_r(i)\}, i = 0, 1$$

are disjoint and $S_0 \cup S_1 \subseteq [0,1]$. To show that $S_0 \cup S_1 = [0,1]$ consider for a given $0 \leq \lambda \leq 1$ the closed sets

$$A_i := \Phi_r(\lambda) \cap \Phi_r(i), i = 0, 1. \tag{1.116}$$

By Lemma 1.54 we obtain that $\Phi_r(\lambda) \subseteq \Phi_r(0) \cup \Phi_r(1)$ and so

$$A_0 \cup A_1 = \Phi_r(\lambda). \tag{1.117}$$

Also by relation (1.115) the sets A_0 and A_1 are disjoint and since $\Phi_r(\lambda)$ is convex and hence connected (cf. [36]) we obtain by relation (1.117) that either A_0 or A_1 is empty. This implies using again Lemma 1.54 that $\Phi_r(\lambda) \subseteq \Phi_r(0)$ or $\Phi_r(\lambda) \subseteq \Phi_r(1)$ and so $\lambda \in S_0 \cup S_1$. Hence we have shown that the sets S_0 and S_1 satisfy

$$S_0 \cap S_1 = \emptyset \text{ and } S_0 \cup S_1 = [0,1]. \tag{1.118}$$

We will now verify that the sets S_0 and S_1 are open in [0, 1] and to do so consider some $\lambda_0 \in S_0$ (a similar proof applies to S_1). Since $\Phi_r(\lambda_0)$ is nonempty for every $r < r^*$ it follows by the definition of $\Phi_r(\lambda_0)$ that

$$\sup_{\mathbf{a}\in A} \inf_{\mathbf{b}\in B_0} f(\mathbf{a}, \mathbf{b}) \geq r^* > r$$

with $B_0 := \{\mathbf{b}_2, ..., \mathbf{b}_k, \lambda_0 \mathbf{b}_0 + (1 - \lambda_0)\mathbf{b}_1\}$. This means that there exists some $\mathbf{a}_0 \in \cap_{i=2}^k U(f_{\mathbf{b}_i}, r)$ satisfying

$$f(\mathbf{a}_0, \lambda_0 \mathbf{b}_0 + (1 - \lambda_0)\mathbf{b}_1) > r \tag{1.119}$$

and by lower semicontinuity of the function $f_{\mathbf{a}_0}$ and relation (1.119) there exist some $\epsilon > 0$ such that

$$f(\mathbf{a}_0, \lambda \mathbf{b}_0 + (1 - \lambda)\mathbf{b}_1) > r$$

for every $\lambda \in \mathcal{N} := (\lambda_0 - \epsilon, \lambda_0 + \epsilon) \cap [0,1]$. Hence we obtain that $\mathbf{a}_0 \in \Phi_r(\lambda)$ for every $\lambda \in \mathcal{N}$ and since $\lambda_0 \in S_0$ this implies by relations (1.118) and (1.15) that $\Phi_r(\lambda) \subseteq \Phi_r(0)$ for every $\lambda \in \mathcal{N}$. Hence S_0 is an open set and since similarly S_1 is open we obtain by relation (1.118) and [0, 1] connected that either S_0 or S_1 is empty. This yields a contradiction with $S_i, i = 0, 1$ nonempty and so relation (1.114) cannot hold. □

It is now possible to give a proof of Sion's minimax result.

Proof. (Sion's minimax theorem). Since A is compact and $f_{\mathbf{b}}$ is upper semicontinuous we obtain that the set $U(f_{\mathbf{b}}, r)$ is compact. By the finite intersection property for compact sets we obtain by Lemma 1.55 that $\cap_{\mathbf{b}\in B} U(f_{\mathbf{b}}, r) \neq \emptyset$ for every $r < r^*$ and this shows by Lemma 1.53 that $r^* = r_*$. Since A compact and $f_{\mathbf{b}}$ upper semicontinuous and $f_{\mathbf{a}}$ lower semicontinuous it follows by a standard argument that we may replace sup by max. □

Actually we can also apply Sion's minimax theorem to prove Theorem 1.22. Looking at the proof of Theorem 1.22 we observe in relation (1.107) that

$$\inf_{\mu\in\mathcal{P}_B} \max_{\mathbf{a}\in I} f_e(\epsilon_{\mathbf{a}}, \mu) \geq 1 + \frac{\gamma}{2}$$

with I belonging to $\mathcal{F}(A)$. This shows by Lemma 1.52 that

$$\inf_{\mu\in\mathcal{P}_B} \max_{\lambda\in\mathcal{P}_I} f_e(\lambda, \mu) \geq 1 + \frac{\gamma}{2}. \tag{1.120}$$

To the expression in relation (1.120) we may now apply Sion's minimax theorem and so we obtain

$$\max_{\lambda \in \mathcal{P}_I} \inf_{\mu \in \mathcal{P}_B} f_e(\lambda, \mu) \geq 1 + \frac{\gamma}{2}$$

and in a similar way we obtain a contradiction with relation (1.104). In the next subsection we will consider applications of convex analysis to optimization theory.

4.2 Optimization theory and duality

In this subsection we will show how the tools of convex analysis can be used within optimization theory. In particular we introduce the dual of an optimization problem and derive some important properties of this dual problem. To start with a general introduction to optimization theory let $f : \mathbb{R}^n \to [-\infty, \infty]$ be an arbitrary function and consider the so-called primal optimization problem given by

$$v(P) := \inf\{f(\mathbf{x}) : \mathbf{x} \in \mathbb{R}^n\}. \tag{P}$$

In this optimization problem the infimum need not be attained. Since f represents an extended real valued function the above optimization problem also covers optimization problems with restrictions. Associate now with the function f a function $F : \mathbb{R}^n \times \mathbb{R}^m \to [-\infty, \infty]$ satisfying $F(\mathbf{x}, \mathbf{0}) = f(\mathbf{x})$ for every $\mathbf{x}$ and consider the so-called *perturbation function* $p : \mathbb{R}^m \to [-\infty, \infty]$ given by

$$p(\mathbf{y}) := \inf\{F(\mathbf{x}, \mathbf{y}) : \mathbf{x} \in \mathbb{R}^n\}. \tag{1.121}$$

It is easy to verify (remember the strict epigraph and the effective domain of a function are listed in relation (1.43) and (1.45)!) that

$$\widetilde{epi}(p) = A(\widetilde{epi}(F)) \text{ and } dom(p) = A(dom(F)) \tag{1.122}$$

with $A : \mathbb{R}^{n+m} \to \mathbb{R}^m$ the projection of $\mathbb{R}^{n+m}$ onto $\mathbb{R}^m$ given by $A(\mathbf{x}, \mathbf{y}) = \mathbf{y}$. Also by the definition of the function F we obtain that $p(\mathbf{0}) = v(P)$. In the next definition we introduce the dual of the optimization problem *(P)* (cf. [63]).

Definition 1.33 *The so-called dual problem of optimization problem (P) is given by*

$$v(D) := \sup\{-p^*(\mathbf{a}) : \mathbf{a} \in \mathbb{R}^m\} \tag{D}$$

with p^* the conjugate function of p listed in Definition 1.25.

By Definitions 1.33 and 1.25 it follows that $v(D) = p^{**}(\mathbf{0})$ and since $p^{**}(\mathbf{0}) \leq p(\mathbf{0})$ the inequality $v(D) \leq v(P)$ always holds. We are now interested under which conditions on the perturbation function p it follows that $v(D) = v(P)$. If $v(P) = -\infty$, then the inequality $v(D) \leq v(P)$ implies $v(D) = v(P) = -\infty$ and every $\mathbf{a} \in \mathbb{R}^m$ is an optimal solution of the dual problem *(D)*. Therefore we only need to consider $v(P) > -\infty$. Consider now the cases $v(P)$ is finite and $v(P) = \infty$. Observe the last case only happens if $\text{dom}(f)$ is empty. For $v(P)$ finite, one can now show the following result. This result is a direct consequence of Theorem 1.12 giving a dual characterization of a convex function (Fenchel-Moreau theorem) and Theorem 1.13.

Theorem 1.24 *If the function* $p : \mathbb{R}^m \to [-\infty, \infty]$ *is convex and* $p(\mathbf{0})$ *is finite, then it follows that*

$$v(P) = v(D) \Leftrightarrow \textit{the function } p \textit{ is l.s.c.at } \mathbf{0}.$$

Moreover, if $\mathbf{0}$ *belongs to* $ri(dom(p))$, *then the dual problem has an optimal solution and* $v(D) = v(P)$.

Proof. Since the function p is convex, l.s.c. at $\mathbf{0}$ and $p(\mathbf{0})$ finite it follows by relations (1.59) and (1.66) that $p_{\overline{c}}(\mathbf{0}) = \overline{p_c}(\mathbf{0}) = \overline{p}(\mathbf{0}) = p(\mathbf{0})$ is finite and this implies by Lemma 1.40 that $\mathcal{A}_p$ is nonempty. Therefore $\overline{p}(\mathbf{x}) > -\infty$ for every x and by the Fenchel-Moreau theorem (Theorem 1.12) we obtain $v(P) = p(\mathbf{0}) = \overline{p}(\mathbf{0}) = p^{**}(\mathbf{0}) = v(D)$. To prove the reverse implication we observe by Theorem 1.12 and $v(P) = v(D)$ is finite that $p(\mathbf{0}) = p^{**}(\mathbf{0}) = cl(p)(\mathbf{0})$ is finite. Hence it must follow by Definition 1.26 that $\overline{p}(\mathbf{0}) = p(\mathbf{0})$ and by relation (1.59) the function p is l.s.c. at $\mathbf{0}$. To show the second part it follows by Theorem 1.13 that $\partial p(\mathbf{0})$ is nonempty and by Lemma 1.43 it is now easy to verify that any $\mathbf{a}_0 \in \partial p(\mathbf{0})$ is an optimal solution of the dual problem. Moreover, by Lemma 1.36 we obtain that $p(\mathbf{0}) = \overline{p}(\mathbf{0})$ and we can apply the first part. □

Finally we consider the case $v(P) = \infty$. In general it does not hold even for p convex and l.s.c. in $\mathbf{0}$ that $v(P) = v(D)$. To show this we will discuss in Example 1.16 a linear programming problem satisfying $v(P) = \infty$ and $v(D) = -\infty$.

If $f : \mathbb{R}^n \to \mathbb{R}$ is some real valued function and $\mathbf{g} : \mathbb{R}^n \to \mathbb{R}^m$ a vector valued function represented by $\mathbf{g}(\mathbf{x}) := (g_1(\mathbf{x}), ..., g_m(\mathbf{x}))$, $g_i : \mathbb{R}^n \to \mathbb{R}$, then an important special case of optimization problem *(P)* is given by

$$\inf\{f(\mathbf{x}) : \mathbf{g}(\mathbf{x}) \in -K, \mathbf{x} \in D\} \tag{P_1}$$

with $K \subseteq \mathbb{R}^m$ a nonempty convex cone and $D \subseteq \mathbb{R}^n$ some nonempty set. The above optimization problem includes some important classes of optimization problems listed in the following example.

Example 1.15

1 If $f(\mathbf{x}) = \mathbf{c}^\top\mathbf{x}$ and $\mathbf{g}(\mathbf{x}) = A\mathbf{x} - \mathbf{b}$ with A some $m \times n$ matrix, $K = \{\mathbf{0}\} \subseteq \mathbb{R}^m$ and $D = \mathbb{R}^n_+$, then optimization problem (P_1) reduces to the so-called *linear programming problem* (cf. [4], [54], [19])

$$\inf\{\mathbf{c}^\top\mathbf{x} : A\mathbf{x} = \mathbf{b}, \mathbf{x} \geq \mathbf{0}\}.$$

2 If $f(\mathbf{x}) = \mathbf{c}^\top\mathbf{x}$ and $\mathbf{g}(\mathbf{x}) = A\mathbf{x} - \mathbf{b}$ with A some $m \times n$ matrix, $K = \{\mathbf{0}\} \subseteq \mathbb{R}^m$ and $D \subseteq \mathbb{R}^n$ is some closed convex cone, then optimization problem (P_1) reduces to a so-called *conic convex programming problem* (cf. [53]), given by

$$\inf\{\mathbf{c}^\top\mathbf{x} : \mathbf{A}\mathbf{x} = \mathbf{b}, \mathbf{x} \in D\}$$

3 If $m = n$ and $g(\mathbf{x}) = -\mathbf{x}$, then optimization problem (P_1) reduces to a so-called *generalized geometric programming problem* (cf. [57]), given by

$$\inf\{f(\mathbf{x}) : \mathbf{x} \in K \cap D\}.$$

4 If the nonempty convex cone $K \subseteq \mathbb{R}^m$ is given by $K = \mathbb{R}^p_+ \times \{\mathbf{0}\}$ with $\mathbf{0} \in \mathbb{R}^{m-p}, p \leq m$ and the set $D = \mathbb{R}^n$, then optimization problem (P_1) reduces to the classical *nonlinear programming problem* (cf. [3], [54], [19])

$$\inf\{f(\mathbf{x}) : g_i(\mathbf{x}) \leq 0, i = 1, ..., p, g_i(\mathbf{x}) = 0, p + 1 \leq i \leq n\}.$$

For optimization problem (P_1) the so-called *Lagrangian perturbation scheme* is used and this means that the function $F : \mathbb{R}^n \times \mathbb{R}^m \to [-\infty, \infty]$ is given by

$$F(\mathbf{x}, \mathbf{y}) = \begin{cases} f(\mathbf{x}) & \text{for } \mathbf{x} \in D \text{ and } \mathbf{g}(\mathbf{x}) \in -K + \mathbf{y} \\ \infty & \text{otherwise} \end{cases}.$$

For this specific choice of F we obtain by relation (1.121) that

$$p(\mathbf{y}) = \inf\{f(\mathbf{x}) : \mathbf{x} \in D, \mathbf{y} \in \mathbf{g}(\mathbf{x}) + K\}. \tag{1.123}$$

Using the representation of p, listed in relation (1.123), one can give a more detailed expression of the dual problem. Observe this dual problem is called the *Lagrangian dual problem.*

Lemma 1.56 *If the function* $\theta : K^0 \to [-\infty, \infty]$ *is given by* $\theta(\mathbf{a}) = \inf\{f(\mathbf{x}) - \mathbf{a}^\top\mathbf{g}(\mathbf{x}) : \mathbf{x} \in D\}$, *then the Lagrangian dual of optimization problem* (P_1) *equals*

$$v(LD) := \sup\{\theta(\mathbf{a}) : \mathbf{a} \in K^0\}. \tag{LD}$$

Proof. By the definition of the function p it follows for every $\mathbf{a} \in \mathbb{R}^n$ that

$$\begin{aligned} -p^*(\mathbf{a}) &= -\sup_{\mathbf{y}\in\mathbb{R}^m}\{\mathbf{a}^\top \mathbf{y} - \inf\{f(\mathbf{x}) : \mathbf{y} \in \mathbf{g}(\mathbf{x}) + K, \mathbf{x} \in D\}\} \\ &= -\sup_{\mathbf{y}\in\mathbb{R}^m} \sup\{\mathbf{a}^\top \mathbf{y} - f(\mathbf{x}) \colon \mathbf{y} \in \mathbf{g}(\mathbf{x}) + K, \mathbf{x} \in D\} \\ &= \inf\{f(\mathbf{x}) - \mathbf{a}^\top \mathbf{y} : \mathbf{y} \in \mathbf{g}(\mathbf{x}) + K, \mathbf{x} \in D\}. \end{aligned}$$

This shows

$$-p^*(\mathbf{a}) = \inf\{f(\mathbf{x}) - \mathbf{a}^\top(\mathbf{g}(\mathbf{x}) + \mathbf{k}) : \mathbf{k} \in K, \mathbf{x} \in D\}$$

and to simplify the above expression we first consider a vector $\mathbf{a}$ belonging to K^0. Since by definition $\mathbf{a}^\top \mathbf{k} \leq 0$ for every $\mathbf{k} \in K$ and $\mathbf{0} \in cl(K)$ this implies

$$-p^*(\mathbf{a}) = \inf\{f(\mathbf{x}) - \mathbf{a}^\top \mathbf{g}(\mathbf{x}) \colon \mathbf{x} \in D\} = \theta(\mathbf{a}).$$

Moreover, if the vector a does not belong to K^0 one can find some $\mathbf{k}_0 \in K$ satisfying $\mathbf{a}^\top \mathbf{k}_0 > 0$. Since $\alpha \mathbf{k}_0 \in K$ for every $\alpha > 0$ and the set D is not empty this yields $-p^*(\mathbf{a}) = -\infty$ and the desired result is verified. □

By Lemmas 1.24 and 1.56 the following result about the Lagrangian dual problem is easy to derive.

Theorem 1.25 *If the primal problem is represented by* (P_1) *and the vector valued function* $\mathbf{h} : \mathbb{R}^n \to \mathbb{R}^{m+1}$ *is given by* $\mathbf{h}(\mathbf{x}) := (\mathbf{g}(\mathbf{x}), f(\mathbf{x}))$ *and satisfies* $\mathbf{h}(D) + (K \times (0, \infty))$ *is convex and* $\mathbf{0} \in ri(\mathbf{g}(D) + K)$, *then it follows that* $\infty > v(P_1) = v(LD)$ *and the Lagrangian dual problem (LD) has an optimal solution.*

Proof. Since by assumption $\mathbf{0}$ belongs to $ri(\mathbf{g}(D) + K) \subseteq \mathbf{g}(D) + K$ we obtain that the feasible region of the optimization problem (P_1) is not empty and this shows $v(P_1) < \infty$. For $v(P_1) = -\infty$ the result follows immediately and so we only consider $v(P_1)$ is finite. To apply Theorem 1.24 we first need to verify whether the function p is convex. It is easy to check that

$$\widetilde{epi}(F) = \{(\mathbf{x}, \mathbf{y}, r) \in \mathbb{R}^{n+m+1} : \mathbf{y} \in \mathbf{g}(\mathbf{x}) + K, \mathbf{x} \in D \text{ and } r > f(\mathbf{x})\}$$

and this implies by relation (1.122) that $\widetilde{epi}(p) = \mathbf{h}(D) + (K \times (0, \infty))$. By assumption this set is convex and hence by Lemma 1.24 the perturbation function p is convex. Also by relation (1.122) we obtain $ri(dom(p)) =$

$ri(\mathbf{g}(D) + K)$ and applying Lemma 1.56 and Theorem 1.24 the desired result follows. □

The condition $\mathbf{0} \in ri(\mathbf{g}(D) + K)$ is known in the literature as the *generalized Slater condition.* Observe, if f is a convex function and $\mathbf{g}$ is a so-called K-convex vector valued function (cf. [73], [6]), then it follows that $\widetilde{epi}(F)$ is a convex set and hence also $\mathbf{h}(D)+(K \times (0,\infty))$ is convex. Also it is possible to prove related results under slightly weaker conditions (cf. [25],[24]). As shown by the next lemma the Lagrangian dual (LD) of a conic convex programming problem is again a conic convex programming problem. Due to the recent developments in interior point methods this class of optimization problems became very important (cf. [53]).

Lemma 1.57 *If the primal problem (P_1) is a conic convex programming problem given by*

$$\inf\{\mathbf{c}^\top \mathbf{x} : A\mathbf{x} = \mathbf{b}, \mathbf{x} \in D\}$$

with $D \subseteq \mathbb{R}^n$ some closed convex cone and there exists some $\mathbf{x}_0 \in ri(D)$ satisfying $A\mathbf{x}_0 = \mathbf{b}$, then it follows that

$$\infty > v(P_1) = v(LD) = \inf\{\mathbf{a}^\top \mathbf{b} : A^\top \mathbf{a} - \mathbf{s} = \mathbf{c}, \mathbf{s} \in D^0\}$$

and the last dual conic convex optimization problem has an optimal solution.

Proof. By part 2 of Example 1.15 we know that a conic convex programming problem is a special case of optimization problem (P_1) with $K = \{\mathbf{0}\} \subseteq \mathbb{R}^m$, the vector valued function $\mathbf{h}$, listed in Theorem 1.25, given by $\mathbf{h}(\mathbf{x}) = (A\mathbf{x}-\mathbf{b}, \mathbf{c}^\top \mathbf{x})$ and $D \subseteq \mathbb{R}^n$ a closed convex cone. Clearly for this choice the set $\mathbf{h}(D) + (\{\mathbf{0}\} \times (0,\infty))$ is convex. Moreover, by Lemma 1.18 the generalized Slater condition reduces to

$$\mathbf{0} \in ri(\mathbf{g}(D) + \{\mathbf{0}\}) = ri(A(D) - \mathbf{b}) = A(ri(D)) - \mathbf{b}$$

and by our assumption this condition is satisfied. Therefore the above result is an immediate consequence of Theorem 1.25 once we have evaluated $\theta(\mathbf{a})$ for $\mathbf{a} \in K^0 = \mathbb{R}^m$. Observe now that

$$\begin{aligned} \theta(\mathbf{a}) &= \inf\{\mathbf{c}^\top \mathbf{x} - \mathbf{a}^\top (A\mathbf{x} - \mathbf{b}) : \mathbf{x} \in D\} \\ &= \mathbf{a}^\top \mathbf{b} + \inf\{(\mathbf{c} - A^\top \mathbf{a})^\top \mathbf{x} : \mathbf{x} \in D\} \end{aligned} \tag{1.124}$$

and since

$$\inf\{(\mathbf{c} - A^\top \mathbf{a})^\top \mathbf{x} : \mathbf{x} \in D\} = \begin{cases} 0 & \text{for } A^\top \mathbf{a} - \mathbf{c} \in D^0 \\ -\infty & \text{otherwise} \end{cases}$$

the desired result follows by Theorem 1.25. □

Using Lemma 1.57 with $D = \mathbb{R}^n_+$ it follows that the Lagrangian dual of the linear programming problem $\inf\{\mathbf{c}^\top \mathbf{x} : A\mathbf{x} = \mathbf{b}, \mathbf{x} \geq \mathbf{0}\}$ is given by $\sup\{\mathbf{b}^\top \mathbf{a} : A^\top \mathbf{a} \leq \mathbf{c}\}$ and so this dual problem reduces to the ordinary dual listed in many text books (cf. [4]). Since the set $D = \mathbb{R}^n_+$ is a polyhedral convex cone (cf. [63]), the generalized Slater condition in Lemma 1.57 can be replaced by the condition that the feasible region of the linear programming problem is nonempty. Actually it can be shown for every polyhedral convex cone D that the associated conic convex programming problem reduces to a linear programming problem and so it is only useful to consider conic convex programming problems with a nonpolyhedral convex cone D. It is also possible to extend the above duality results for conic convex programming problems to a larger class of problems than the one having a generalized Slater point and for more details on this the reader is referred to [67]. To conclude this section we consider the following example of a linear programming problem satisfying $v(P) = \infty$ and $v(D) = -\infty$.

Example 1.16 Consider the linear programming problem

$$\inf\{-x_1 - x_2 : x_1 - x_2 \geq 1, -x_1 + x_2 \geq 1, \mathbf{x} \in \mathbb{R}^2_+\}.$$

Clearly this optimization problem has an empty feasible region and so $v(P_1) = \infty$. Penalizing the restrictions $x_1 - x_2 - 1 \geq 0$ and $-x_1 + x_2 - 1 \geq 0$ using the nonpositive Lagrangian multipliers a_1 and a_2 we obtain that the Lagrangian function $\theta : \mathbb{R}^2_+ \to [-\infty, \infty)$ is given by

$$\theta(\mathbf{a}) = \inf\{x_1(\lambda_1 - \lambda_2 - 1) + x_2(\lambda_2 - \lambda_1 - 1) : \mathbf{x} \in \mathbb{R}^2_+\}$$

Observe now for every $\mathbf{a} \in \mathbb{R}^2_+$ that

$$a_1 - a_2 - 1 \geq 0 \Rightarrow a_2 - a_1 - 1 \leq -2$$

and

$$a_2 - a_1 - 1 \geq 0 \Rightarrow a_1 - a_2 - 1 \leq -2$$

and by this observation it follows that $\theta(\mathbf{a}) = -\infty$ for every $\mathbf{a} \in \mathbb{R}^2_+$ or equivalently $v(D) = -\infty$.

One can also use the same Lagrangian perturbation scheme and the dual representation of an evenly quasiconvex function and the corresponding c_r-conjugate function to introduce the so-called surrogate dual. Due to limited space we will not discuss the properties of such a dual but refer the reader to the literature cited in [27]. This concludes our discussion on duality and optimization problems. In the next subsection we will consider the structure of positively homogeneous evenly quasiconvex functions.

4.3 Positively homogeneous evenly quasiconvex functions and dual representations

In this subsection the dual representation of an evenly quasiconvex function is used to show a remarkable property of a positively homogeneous evenly quasiconvex function. In [11] a similar property is also derived for a positively homogeneous quasiconvex function. As such the results in [11] apply to a larger class of functions, but are slightly weaker. Also the proof technique used in [11] is more direct and based on the geometrical aspects of convexity, whereas the approach used in this chapter is a natural consequence of the dual representation of an evenly quasiconvex function discussed in Subsection 3.4. To start with the dual approach we consider a positively homogeneous evenly quasiconvex function $f : \mathbb{R}^n \to (-\infty, \infty]$ satisfying $\mathbf{0} \in dom(f)$. Since f is positively homogeneous and $f(\mathbf{x}) > -\infty$ for every $\mathbf{x}$ we obtain by Lemma 1.25 that $\mathbf{0} \in dom(f)$ if and only if $f(\mathbf{0}) = 0$. Considering for every $\mathbf{a} \in \mathbb{R}^n$ the function $c_{\mathbf{a}} : \mathbb{R} \to [-\infty, \infty)$, given by

$$c_{\mathbf{a}}(t) := \inf\{f(\mathbf{y}) : \mathbf{a}^{\top}\mathbf{y} \geq t\}, \tag{1.125}$$

(see also Definition 1.29) it is easy to verify the next result.

Lemma 1.58 *If* $f : \mathbb{R}^n \to [-\infty, \infty]$ *is positively homogeneous, then for every* $\mathbf{a} \in \mathbb{R}^n$ *it follows that the function* $c_{\mathbf{a}} : \mathbb{R}^n \to [-\infty, \infty]$ *is positively homogeneous and nondecreasing.*

Proof. For any nonzero vector $\mathbf{a}$ it is obvious by relation (1.125) that the function $c_{\mathbf{a}}$ is nondecreasing. Also by Lemma 1.25 we obtain for every $\alpha > 0$ and $t \in \mathbb{R}$ that

$$c_{\mathbf{a}}(\alpha t) = \inf\{f(\alpha\mathbf{y}) : \mathbf{a}^{\top}\mathbf{y} \geq t\} = \alpha c_{\mathbf{a}}(t)$$

and so the result is verified for every nonzero $\mathbf{a}$. Moreover, for $\mathbf{a} = \mathbf{0}$, we obtain for every $\alpha > 0$ and $t \leq 0$ that

$$c_{\mathbf{0}}(\alpha t) = \inf\{f(\mathbf{y}) : \mathbf{y} \in \mathbb{R}^n\} = \inf\{f(\alpha\mathbf{y}) : \mathbf{y} \in \mathbb{R}^n\} = \alpha c_{\mathbf{0}}(t),$$

while for $\alpha > 0$ and $t > 0$ it follows using the convention $\inf\{\varnothing\} = \infty$, that $c_{\mathbf{0}}(\alpha t) = \inf\{\varnothing\} = \infty = \alpha c_{\mathbf{0}}(t)$. Trivially the function $c_{\mathbf{0}}$ is nondecreasing and the proof is completed. □

To analyze the behaviour of a positively homogeneous evenly quasiconvex function f satisfying $f(\mathbf{x}) > -\infty$ for every $\mathbf{x}$ and $\mathbf{0} \in dom(f)$, we first decompose this function. Using a slightly different decomposition as done by Crouzeix (cf. [15],[11]) we introduce the nonnegative function $f_+ : \mathbb{R}^n \to [0, \infty]$, given by

$$f_+(\mathbf{x}) := \begin{cases} 0 & \text{if } \mathbf{x} \in \widetilde{L}(f, 0) \\ f(\mathbf{x}) & \text{otherwise} \end{cases} . \tag{1.126}$$

with the strict lower level set $\widetilde{L}(f, 0)$ of the function f of level 0 listed in relation (1.52). Using now $f(\mathbf{x}) > -\infty$ for every $\mathbf{x}$ and $\mathbf{0} \in dom(f)$ we immediately obtain that $f_+(\mathbf{0}) = f(\mathbf{0}) = 0$. Moreover, the function $f_- : \mathbb{R}^n \to (-\infty, \infty]$ is given by

$$f_-(\mathbf{x}) = \begin{cases} f(\mathbf{x}) & \text{if } \mathbf{x} \in cl(\widetilde{L}(f, 0)) \\ \infty & \text{otherwise} \end{cases} . \tag{1.127}$$

To analyze the function f_- it is only interesting to consider positively homogeneous evenly quasiconvex functions f satisfying $\widetilde{L}(f, 0)$ is nonempty. If this holds, we obtain by Lemma 1.25 that $\widetilde{L}(f, 0)$ is a nonempty convex cone and since $\mathbf{0} \in cl(\widetilde{L}(f, 0))$ it follows that $f_-(\mathbf{0}) = f(\mathbf{0}) = 0$. Also for every $r \in \mathbb{R}$ we obtain that

$$\widetilde{L}(f_-, r) = cl(\widetilde{L}(f, 0)) \cap \widetilde{L}(f, r) \tag{1.128}$$

and this yields for $r = 0$ that $\widetilde{L}(f_-, 0) = \widetilde{L}(f, 0)$. By relation (1.127) we therefore obtain for $\widetilde{L}(f, 0)$ is not empty that

$$dom(f_-) \subseteq cl(\widetilde{L}(f, 0)) = cl(\widetilde{L}(f_-, 0)). \tag{1.129}$$

Since trivially $f_-(\mathbf{x}) \geq f(\mathbf{x})$ for every $\mathbf{x}$ it is easy to verify considering the cases $f(\mathbf{x}) \geq 0$ and $f(\mathbf{x}) < 0$ that

$$f(\mathbf{x}) = \min\{f_+(\mathbf{x}), f_-(\mathbf{x})\} \tag{1.130}$$

for every $\mathbf{x} \in \mathbb{R}^n$. For the functions f_+ and f_- one can now show the following result.

Lemma 1.59 *If $f : \mathbb{R}^n \to [-\infty, \infty]$ is a positively homogeneous evenly quasiconvex function, then the functions f_+ and f_- are positively homogeneous and evenly quasiconvex.*

Proof. Since f is positively homogeneous and evenly quasiconvex (and hence quasiconvex) we obtain by Lemma 1.25 and 1.27 that $\widetilde{L}(f, 0)$ is a (possibly empty) convex cone. This implies again by Lemma 1.25 that f_+ is positively homogeneous. To show that f_+ is evenly quasiconvex we observe that $L(f_+, r) = L(f, r)$ for every $r > 0$. Also by the definition of f_+ we obtain

$$L(f_+, 0) = \widetilde{L}(f, 0) \cup \{\mathbf{x} : \mathbf{x} \notin \widetilde{L}(f, 0) \text{ and } f(\mathbf{x}) \leq 0\} = L(f, 0).$$

This shows, using the fact that f_+ is a nonnegative function and f is evenly quasiconvex, that also f_+ is evenly quasiconvex. To verify the same result for f_- we observe, since $cl(\widetilde{L}(f, 0))$ is also a (possibly empty) convex cone, that f_- is positively homogeneous. Moreover, for every $r \in \mathbb{R}$ we know by relation (1.128) that $L(f_-, r) = cl(\widetilde{L}(f, 0)) \cap L(f, r)$ and applying Lemma 1.21 and f is evenly quasiconvex it follows that f_- is evenly quasiconvex. □

We will now apply the dual representation of an evenly quasiconvex function and show the following result for a nonnegative positively homogeneous evenly quasiconvex function f with $\mathbf{0} \in dom(f)$. A related result is also discussed in [11]. Recall that a function is called sublinear, if it is positively homogeneous and convex.

Lemma 1.60 *If $f : \mathbb{R}^n \to [-\infty, \infty]$ is a nonnegative positively homogeneous evenly quasiconvex function with $\mathbf{0} \in dom(f)$, then f is a nonnegative l.s.c. sublinear function.*

Proof. By the dual representation of an evenly quasiconvex function (see Theorem 1.18) we obtain that

$$f(\mathbf{x}) = f_{ec}(\mathbf{x}) = \sup_{\mathbf{a} \in \mathbb{R}^n} c_{\mathbf{a}}(\mathbf{a}^\top \mathbf{x}). \tag{1.131}$$

Since $f \geq 0$ it follows by the definition of $c_{\mathbf{a}}$ that $c_{\mathbf{a}}$ is a nonnegative function for every $\mathbf{a} \in \mathbb{R}^n$. Moreover, using $f(\mathbf{0}) = 0$ and $0 \leq c_{\mathbf{a}}(0) \leq f(\mathbf{0})$ we obtain $c_{\mathbf{a}}(0) = 0$. Also for $\mathbf{x} \in \mathbb{R}^n$ and $\mathbf{a} \in \mathbb{R}^n$ satisfying $\mathbf{a}^\top \mathbf{x} \leq 0$ it follows by the monotonicity of $c_{\mathbf{a}}$ that $0 \leq c_{\mathbf{a}}(\mathbf{a}^\top \mathbf{x}) \leq c_{\mathbf{a}}(0) = 0$ and this implies $c_{\mathbf{a}}(\mathbf{a}^\top \mathbf{x}) = 0$ for every $\mathbf{a}^\top \mathbf{x} \leq 0$. Moreover, for $\mathbf{a}^\top \mathbf{x} > 0$ we obtain by Lemma 1.58 that $c_{\mathbf{a}}(\mathbf{a}^\top \mathbf{x}) = r_{\mathbf{a}} \mathbf{a}^\top \mathbf{x}$ with $r_{\mathbf{a}} := c_{\mathbf{a}}(1) \geq 0$ and combining both observations yields

$$c_{\mathbf{a}}(\mathbf{a}^\top \mathbf{x}) = \max\{r_{\mathbf{a}} \mathbf{a}^\top \mathbf{x}, 0\}$$

for every $\mathbf{a} \in \mathbb{R}^n$. Applying now relation (1.131) yields

$$f(\mathbf{x}) = \sup_{\mathbf{a} \in \mathbb{R}^n} \max\{r_{\mathbf{a}} \mathbf{a}^\top \mathbf{x}, 0\} = \max\{\sup_{\mathbf{a} \in \mathbb{R}^n} r_{\mathbf{a}} \mathbf{a}^\top \mathbf{x}, 0\} \tag{1.132}$$

and since $\mathbf{x} \to \sup_{\mathbf{a} \in \mathbb{R}^n} r_{\mathbf{a}} \mathbf{a}^\top \mathbf{x}$ is a l.s.c. sublinear function the desired result follows by relation (1.132). □

Since by relation (1.126) we obtain that $f_+(\mathbf{x}) = \max\{f(\mathbf{x}), 0\}$ and for f positively homogeneous and evenly quasiconvex the function f_+

is also positively homogeneous and evenly quasiconvex (Lemma 1.59) we may apply Lemma 1.60 and so we obtain the result that f_+ is a nonnegative l.s.c. sublinear function in case the function f is positively homogeneous, evenly quasiconvex, $\mathbf{0} \in dom(f)$ and $f(\mathbf{x}) > -\infty$ for every $\mathbf{x}$. Finally we will show the following result for a positively homogeneous evenly quasiconvex function f satisfying $\widetilde{L}(f,0)$ nonempty, $dom(f) \subseteq cl(\widetilde{L}(f,0))$ and $f(\mathbf{x}) > -\infty$ for every $\mathbf{x}$..

Lemma 1.61 *If* $f : \mathbb{R}^n \to (-\infty, \infty]$ *is a positively homogeneous evenly quasiconvex function with* $\mathbf{0} \in dom(f) \subseteq cl(\widetilde{L}(f,0))$, *then* f *is a nonpositive l.s.c. sublinear function.*

Proof. By the dual representation of an evenly quasiconvex function we obtain (see Theorem 1.18) that

$$f(\mathbf{x}) = \sup_{\mathbf{a}\in\mathbb{R}^n} c_{\mathbf{a}}(\mathbf{a}^\top \mathbf{x}). \tag{1.133}$$

If the vector $\mathbf{a}$ does not belong to the polar cone $(\widetilde{L}(f,0))^0$, then there exists some $\mathbf{x}_0$ satisfying $f(\mathbf{x}_0) < 0$ and $r := \mathbf{a}^\top \mathbf{x}_0 > 0$. By Lemma 1.58 this yields for every $t > 0$ that

$$c_{\mathbf{a}}(t) = c_{\mathbf{a}}(tr^{-1}r) = tr^{-1}c_{\mathbf{a}}(r) \leq tr^{-1}f(\mathbf{x}_0) < 0$$

and so $c_{\mathbf{a}}(\infty) := \lim_{t\uparrow\infty} c_{\mathbf{a}}(t) = -\infty$. Since the function $c_{\mathbf{a}}$ is nondecreasing, this shows that $c_{\mathbf{a}}(t) = -\infty$ for every $t \in \mathbb{R}$ and using the fact that $f(\mathbf{x}) > -\infty$ for every x and relation (1.133) we obtain

$$f(\mathbf{x}) = \sup\{c_{\mathbf{a}}(\mathbf{a}^\top \mathbf{x}) : \mathbf{a} \in (\widetilde{L}(f,0))^0\}. \tag{1.134}$$

If the vector $\mathbf{a}$ belongs to $(\widetilde{L}(f,0))^0$ and $\mathbf{a}^\top \mathbf{x} \geq t > 0$ for some $\mathbf{x} \in \mathbb{R}^n$, then clearly $\mathbf{x}$ does not belong to $cl(\widetilde{L}(f,0))$. Since $dom(f) \subseteq cl(\widetilde{L}(f,0))$ this implies that $f(\mathbf{x}) = \infty$ and so we have shown for every $\mathbf{a}$ belonging to $(\widetilde{L}(f,0))^0$ that

$$c_{\mathbf{a}}(t) = \infty \text{ for every } t > 0. \tag{1.135}$$

To analyze $c_{\mathbf{a}}(t)$ for $\mathbf{a} \in (\widetilde{L}(f,0))^0$ and $t \leq 0$ we first assume that there exists some $\mathbf{x}_0$ satisfying $f(\mathbf{x}_0) < 0$ and $\mathbf{a}^\top \mathbf{x}_0 = 0$. By Lemma 1.58 it holds that $\alpha c_{\mathbf{a}}(0) = c_{\mathbf{a}}(0)$ for every $\alpha > 0$ and since $c_{\mathbf{a}}(0) \leq f(\mathbf{x}_0) < 0$ we obtain that $c_{\mathbf{a}}(0) = -\infty$. Hence it follows that $c_{\mathbf{a}}(t) \leq c_{\mathbf{a}}(0) = -\infty$ for every $t \leq 0$ and we have shown for every $\mathbf{a} \in (\widetilde{L}(f,0))^0$, for which there exists some $\mathbf{x}_0 \in \widetilde{L}(f,0)$ satisfying $\mathbf{a}^\top \mathbf{x}_0 = 0$, that

$$c_{\mathbf{a}}(t) = -\infty \text{ for every } t \leq 0 \tag{1.136}$$

Using again the fact that $f(\mathbf{x}) > -\infty$ for every $\mathbf{x}$ and relations (1.134), (1.135) and (1.136) yields

$$f(\mathbf{x}) = \sup\{c_{\mathbf{a}}(\mathbf{a}^\top \mathbf{x}) : \mathbf{a} \in D\} \tag{1.137}$$

for every $\mathbf{x} \in \mathbb{R}^n$ with

$$D := \{\mathbf{a} \in (\widetilde{L}(f,0))^0 : \mathbf{a}^\top \mathbf{y} < 0 \text{ for every } \mathbf{y} \in \widetilde{L}(f,0)\}.$$

We will now analyze the behaviour of $\mathbf{x} \to c_{\mathbf{a}}(\mathbf{a}^\top \mathbf{x})$ for an arbitrary $\mathbf{a}$ belonging to D. If $\mathbf{a}^\top \mathbf{x} > 0$ it follows by relation (1.135) that $c_{\mathbf{a}}(\mathbf{a}^\top \mathbf{x}) = \infty$. Also, if $\mathbf{a}^\top \mathbf{x} = 0$, then for every $\mathbf{y}$ satisfying $\mathbf{a}^\top \mathbf{y} \geq \mathbf{a}^\top \mathbf{x} = 0$ we obtain, using $\mathbf{a} \in D$, that $f(\mathbf{y}) \geq 0$ and since $\mathbf{0} \in dom(f)$ this implies $c_{\mathbf{a}}(\mathbf{a}^\top \mathbf{x}) = 0$. Finally, for $\mathbf{a} \in D$ and $\mathbf{a}^\top \mathbf{x} < 0$ it follows by Lemma 1.58 that $c_{\mathbf{a}}(\mathbf{a}^\top \mathbf{x}) = q_{\mathbf{a}} \mathbf{a}^\top \mathbf{x}$ with $q_{\mathbf{a}} := -c_{\mathbf{a}}(-1)$ and since $\widetilde{L}(f,0)$ is nonempty we obtain by relation (1.137) that $0 < q_{\mathbf{a}} \leq \infty$. Hence we have shown for every $\mathbf{a} \in D$ that

$$c_{\mathbf{a}}(\mathbf{a}^\top \mathbf{x}) = \begin{matrix} q_{\mathbf{a}} \mathbf{a}^\top \mathbf{x} & \text{if } \mathbf{a}^\top \mathbf{x} < 0 \\ 0 & \text{if } \mathbf{a}^\top \mathbf{x} = 0 \\ \infty & \text{if } \mathbf{a}^\top \mathbf{x} < 0 \end{matrix} \,. \tag{1.138}$$

Again by relation (1.137) and $f(\mathbf{x}) > -\infty$ for every $\mathbf{x}$ we obtain that the set $D_0 := \{\mathbf{a} \in D : 0 < q_{\mathbf{a}} < \infty\}$ is nonempty and by relations (1.138) and (1.137) this shows

$$f(\mathbf{x}) = \sup\{c_{\mathbf{a}}(\mathbf{a}^\top \mathbf{x}) : \mathbf{a} \in D_0\}. \tag{1.139}$$

Since for $\mathbf{a} \in D_0$ it follows that $-\infty < c_{\mathbf{a}}(\mathbf{a}^\top \mathbf{x}) = q_{\mathbf{a}} \mathbf{a}^\top \mathbf{x}$ for $\mathbf{a}^\top \mathbf{x} \leq 0$ and ∞ otherwise, this is clearly a l.s.c. sublinear function and by relation (1.139) the desired result follows. □

Since by Lemma 1.59 and relation (1,127) the function f_- satisfies the conditions of Lemma 1.61 for f a positively homogeneous evenly quasiconvex function with $\mathbf{0} \in dom(f)$ and $f(\mathbf{x}) > -\infty$ for every $\mathbf{x}$ it follows that f_- is a nonpositive l.s.c. sublinear function. Using relation (1.130) and Lemma 1.59 up to 1.61 the following remarkable result follows immediately.

Theorem 1.26 *If $f : \mathbb{R}^n \to (-\infty, \infty]$ is a positively homogeneous evenly quasiconvex function and $\mathbf{0} \in dom(f)$, then f can be written as the minimum of a nonpositive l.s.c. sublinear function and a nonnegative l.s.c. sublinear function.*

Proof. If $\widetilde{L}(f,0)$ is empty then f is a nonnegative function and the result follows by Lemma 1.60. Moreover, if $\widetilde{L}(f,0)$ is nonempty, then by relation (1.130) it follows that $f = \min(f_+, f_-)$ and applying the observations after Lemma 1.60 and 1.61 yields the desired result. □

By Theorem 1.26 every positively homogeneous evenly quasiconvex function f satisfying $f(\mathbf{x}) > -\infty$ for every x and $\mathbf{0} \in dom(f)$ must be the minimum of two l.s.c. sublinear functions and so it is also l.s.c.. By relation (1.77) these l.s.c. sublinear functions can be written as support functions. This is a rather remarkable result, which does not hold in general for evenly quasiconvex functions. As an example we mention the evenly quasiconvex function $\text{sign}(x)$ given by

$$sign(x) = -1 \text{ if } x < 0,\ sign(0) = 0 \text{ and } sign(x) = 1 \text{ if } x > 0$$

which is neither upper or lower semicontinuous at 0. To conclude this subsection we observe that Theorem 1.26 is an extension of the main result in Crouzeix (cf. [15]). For related results see also [14], [13], [12] and [11]. Introducing now the Dini upper directional derivative $\mathbf{d} \to f'_+(\mathbf{x}, \mathbf{d})$ given by

$$f'_+(\mathbf{x}, \mathbf{d}) := \limsup_{t \downarrow 0} t^{-1}(f(\mathbf{x} + t\mathbf{d}) - f(\mathbf{x}))$$

(cf. [30], [11]) it is possible to use the above so-called Crouzeix representation theorem for positively homogeneous quasiconvex functions to analyze the global behaviour of the function $\mathbf{d} \to f'_+(\mathbf{x}, \mathbf{d})$ for f quasiconvex (cf. [15], [44], [45], [33], [11]). This concludes our discussion of positively homogeneous evenly quasiconvex functions and dual representations. In the next section we mention some milestone papers and books within the long history of convex and quasiconvex analysis.

5. Some remarks on the history of convex and quasiconvex analysis

In this section[1] we will discuss the origin of the important notions used in convex and quasiconvex analysis. It seems that the field of convex geometry and convex bodies in two and three dimensional space was first studied systematically by H.Brunn (cf. [7], [8]) and Minkowski (cf. [51]). Brunn (cf. [9]) and Minkowski (cf. [52]) also proved the existence of support hyperplanes. Also at the end of 19*th* and the beginning of the 20*th* century Farkas showed in a series of papers (cf. [45], [61]) the alternative theorem for linear inequality systems and this result became known as Farkas lemma within linear programming. Although this result was listed with an incorrect proof in some of his earlier papers a

[1]The authors like to thank Prof. J. Kolumbán (Cluj) and Prof. S. Komlósi (Pecs) for pointing out some of the early developments.

correct proof of this result appeared in [20]. More fundamental ideas about the related field of necessary optimality conditions for nonlinear optimization subject to inequality constraints can be found in papers by Fourier, Cournot, Gauss, Ostrogradsky, and Hamel (cf. [61]). On the other hand, more early references related to the study of convex sets are listed in the reprinted version of the 1934 book of Bonnesen and Fenchel (cf. [5]), Fenchel (cf. [22]), Valentine (cf. [68]) and Varberg (cf. [62]). Also at the beginning of the 20th century convex functions were introduced by Jenssen (cf. [39]) and more than forty years later a thorough study of conjugate functions in $\mathbb{R}^n$ was initiated by Fenchel (cf. [21]). Although Mandelbrojt (cf. [48]) already introduced the conjugate function in $\mathbb{R}^n$ for $n = 1$ (cf. [69]), it was Fenchel, who first realized the importance of the conjugacy concept in convex analysis. Four years before the milestone paper of Fenchel, also the first book on convex functions written in French by Popoviciu (cf. [60]) was published. In the English scientific community the unpublished lecture notes by Fenchel (cf. [22]) were a long time the main source of references. This book served as the main inspiration for the classical book of Rockafellar (cf. [63]) as noted in its preface. Also in this preface it is mentioned that Prof. Tucker suggested the name convex analysis and this became the standard word for this field. The introduction of quasiconvex functions started later. Although in most of the literature de Finetti ([16]) is mentioned as being the first author introducing quasiconvex functions, these functions were already considered by von Neumann (cf. [71] and independently Popoviciu (cf. [59]). Actually von Neumann (cf. [71]) already proved in 1928 a minimax theorem on simplices for bifunctions which are quasiconcave in one variable and quasiconvex in the other variable. A generalization of this result was rediscovered by Sion (cf. [66]) 30 years later. For more details on the development of quasiconvex functions the reader is referred to [2]. To develop results for the surrogate dual concept developed by Glover (cf. [31]) an adhoc approach involving the $\overline{c_r}$-conjugate function was initiated by Greenberg and Pierskalla (cf. [32]). Their results were generalized and put into the proper framework of dual representations by Crouzeix in a series of milestone papers (cf. [12], [13], [15], [14]). In these papers Crouzeix focussed his attention on the dual representation of the l.s.c. hull of a quasiconvex function. Although Fenchel (cf. [23]) already introduced the concept of an evenly convex set the usefulness of this concept leading to a more symmetrical dual representation of an evenly quasiconvex function was discovered independently by Passy and Prisman (cf. [55]) and Martínez Legaz (cf. [50]). This concludes our short excursion, which is by no means complete, to the history of convex and quasiconvex analysis.

References

[1] Aubin, J.B., *Optima and Equilibria (An introduction to nonlinear analysis),* Graduate Texts in Mathematics, v. 140, Springer Verlag, Berlin, 1993.

[2] Avriel, M., Diewert, W.E., Schaible, S. and I. Zang, *Generalized Concavity,* in Mathematical Concepts and Methods in Science and Engineering, v. 36, Plenum Press, New York, 1988.

[3] Bazaara, M.S., Sherali, H.D. and C.N. Shetty, *Nonlinear Programming (Theory and Applications),* Wiley, New York, 1993.

[4] Bazaara, M.S., Jarvis, J.J. and H.D. Sherali, *Linear Programming and Network Flows (Second Edition),* Wiley, New York, 1990.

[5] Bonnesen, T. and W. Fenchel, *Theorie der konvexen Körper,* Springer Verlag, Berlin, 1974.

[6] Breckner, W.W. and G.Kassay, *A systematization of convexity concepts for sets and functions,* Journal of Convex Analysis, 4, 1997, 1-19.

[7] Brunn, H., *Über Ovale und Eiflächen,* Inauguraldissertation, University of München, 1887.

[8] Brunn, H., *Über Kurven ohne Wendepunkte,* Habilitationsschrift, University of München, 1889.

[9] Brunn, H., *Zur Theorie der Eigebiete,* Arch. Math. Phys. Ser.3, 17, 1910, 289-300.

[10] Choquet, G., *Lectures on Analysis: Volume 2 Representation Theory,* in Mathematics Lecture Note Series, W.A. Benjamin, London, 1976.

[11] Crouzeix, J.-P., *Continuity and differentiability of quasiconvex functions,* Chapter 3 in this volume.

[12] Crouzeix, J.-P., *Contributions a l' étude des fonctions quasiconvexes,* Université de Clermond-Ferrand 2, France, 1977.

[13] Crouzeix, J.-P., *Conditions for convexity of quasiconvex functions,* Mathematics of Operations Research, 5, 1980, 120-125.

[14] Crouzeix, J.-P., *A duality framework in quasiconvex programming,* in Generalized Concavity in Optimization and Economics, Schaible, S. and W.T. Ziemba, eds., Academic Press, 1981, 207-225.

[15] Crouzeix, J.-P., *Continuity and differentiability properties of quasiconvex functions on* $\mathbb{R}^n$, in Generalized Concavity in Optimization and Economics, Schaible, S. and W.T. Ziemba, eds., Academic Press, 1981, 109-130.

[16] De Finetti, B., *Sulle stratificazioni convesse,* Ann. Mat. Pura Appl., [4] 30, 1949, 173-183.

[17] Dudley, R.M., *Real Analysis and Probability,* Wadsworth and Brooks/Cole, Pacific Grove, 1989.

[18] Edwards, R.E., *Functional Analysis: Theory and Applications,* Holt, Rinehart and Winston, Chicago, 1965.

[19] Faigle, U., Kern, W. and G. Still, *Algorithmic Principles of Mathematical Programming,* in Kluwer texts in Mathematical Sciences, v. 24, Kluwer Academic Publishers, Dordrecht, 2002.

[20] Farkas, J., *Uber die Theorie der einfache Ungleichungen,* Journal für die Reine und Angewandte Mathematik, 124, 1901, 1-27.

[21] Fenchel, W., *On conjugate convex functions,* Canadian Journal of Mathematics, 1, 1949, 73-77.

[22] Fenchel, W., *Convex Cones, Sets and Functions,* Lecture notes, Princeton University, 1951.

[23] Fenchel, W., *A remark on convex sets and polarity,* Communication Seminar on Mathematics, University of Lund supplementary volume, 1952, 22-89.

[24] Frenk, J.B.G and G. Kassay, *On classes of generalized convex functions, Gordan-Farkas type theorems and Lagrangian duality,* Journal of Optimization Theory and Applications, [2] 102, 1999, 315-343.

[25] Frenk, J.B.G. and G. Kassay, *Lagrangian duality and cone convexlike functions,* Econometric Institute, Erasmus University Rotterdam, Tech. Rep. 2000-27/A, 2000.

[26] Frenk, J.B.G. and G. Kassay, *Minimax results and finite dimensional separation,* Journal of Optimization Theory and Applications, [2] 113, 2002, 409-421.

[27] Frenk, J.B.G., Dias, D.M.L. and J. Gromicho, *Duality theory for convex and quasiconvex Functions and its application to optimization,* in Generalized Convexity: Proceedings Pecs, Hungary, 1992, Komlósi, S., Rapcsák, T. and S. Schaible, eds., Lecture notes in Economics and Mathematics, 405, 1994, 153-171.

[28] Frenk, J.B.G., Kassay, G. and J. Kolumbán, *On equivalent results in minimax theory,* European Journal of Operational Research 157, 2004, 46-58.

[29] Frenk, J.B.G., Protassov, V. and G. Kassay, *On Borel probability measures and noncooperative game theory,* Report Series Econometric Institute, Erasmus University Rotterdam, Tech. Rep. EI-2002-32, 2002.

[30] Giorgio, G. and S. Komlósi, *Dini derivatives in optimization-Part I,* Rivista di Matematica per le Scienze Economiche e Sociali, [1] 15, 1992, 3-30.

[31] Glover, F., *Surrogate constraints,* Operations Research, 16, 1967, 741-749.

[32] Greenberg, H.J. and W.P. Pierskalla, *Quasiconjugate functions and surrogate duality,* Cahiers du Centre d'étude de Recherche Operationelle, [4] 15, 1973, 437-448.

[33] Gromicho, J., *Quasiconvex Optimization and Location Theory,* in Applied Optimization, v. 9, Kluwer Academic Publishers, Dordrecht, 1998.

[34] Hiriart-Urruty, J.B. and C. Lemaréchal, *Convex Analysis and Minimization Algorithms I,* in Grundlehren der Mathematischen Wissenschaften, v. 305, Springer Verlag, Berlin, 1993.

[35] Hiriart-Urruty, J.B. and C. Lemaréchal, *Convex Analysis and Minimization Algorithms II,* in Grundlehren der Mathematischen Wissenschaften, v. 306, Springer Verlag, Berlin, 1993.

[36] Hocking, J.G and G.S. Young, *Topology,* Addison Wesley, Reading, Massachusetts, 1961.

[37] Holmes, R.B., *Geometric Functional Analysis and its Applications,* Springer, New York, 1975.

[38] Jahn, J., *Mathematical Vector Optimization in Partially Ordered Spaces,* Peter Lang, Fankfurt am Main, 1986.

[39] Jensen, J.L.W.V., *Sur les functions convexes et les inégalités entre des valeurs moyennes,* Acta Math., 30, 1906, 175-193.

[40] Joó, I., *A simple proof for von Neumann's minimax theorem,* Acta Sci. Math. Szeged, 42, 1980, 91-94.

[41] Joó, I., *Note on my paper: A simple proof for von Neumann's minimax theorem,* Acta Math. Hung., [3-4] 44, 1984, 171-176.

[42] Joó, I., *On some convexities,* Acta Math.Hung., [1-2] 54, 1989, 163-172.

[43] Kolmogorov, A.N. and S.V. Fomin, *Introductory Real Analysis,* in Dover books on Mathematics, Dover Publications, New York, 1975.

[44] Komlósi, S., *Quasiconvex first-order approximations and Kuhn-Tucker type optimality conditions,* European Journal of Operational Research, 65, 1993, 327-335.

[45] Komlósi, S., *Farkas theorems for positively homogeneous quasiconvex functions,* Journal of Statistics and Management Systems, 5, 2002, 107-123.

[46] Kreyszig, E., *Introductory Functional Analysis with Applications,* Wiley, New York, 1978.

[47] Lancaster, P. and M. Tismenetsky, *The Theory of Matrices (Second Edition with Applications),* in Computer Science and Applied Mathematics, Academic Press, Orlando, 1985.

[48] Mandelbrojt, S., *Sur les fonctions convexes,* C.R. Acad. Sci. Paris, 209, 1939, 977-978.

[49] Martínez-Legaz, J.-E., *Generalized convex duality and its economic applications,* Chapter 6 in this volume.

[50] Martínez-Legaz, J.-E., *Un concepto generalizado deconjugacion, application a las funciones quasiconvexas,* M.Sc. Thesis, Universidad de Barcelona, 1981.

[51] Minkowski, H., *Algemeine Lehrsätze über die convexen Polyeder,* Nachr. Ges. Wiss. Göttingen Math. Phys. K1, 1897, 198-219.

[52] Minkowski, H., *Geometrie der Zahlen,* Teubner, Leipzig, 1910.

[53] Nesterov, Y. and A. Nemirovski, *Interior Point Polynomial Algorithms in Convex Programming,* in SIAM Studies in Applied Mathematics, SIAM, Philadelphia, 1994.

[54] Nocedal, J. and S.J. Wright, *Numerical Optimization,* in Springer series in Operations Research, Springer Verlag, New York, 1999.

[55] Passy, U. and E.Z. Prisman, *Conjugacy in quasi-convex programming,* Mathematical Programming, 30, 1984, 121-146.

[56] Penot, J.-P. and M. Volle, *On quasi-convex duality,* Mathematics of Operations Research, [4] 15, 1990, 597-624.

[57] Peterson, E.L., *The fundamental relations between geometric programming duality, parametric programming duality and ordinary Lagrangian duality,* Annals of Operations Research, 105, 2001, 109-153.

[58] Ponstein, J., *Approaches to the Theory of Optimization,* Cambridge University Press, Cambridge, 1980.

[59] Popoviciu, T., *Deux remarques sur les fonctions convexes,* Bull. Sc. Acad. Roumaine, 20, 1938, 45-49.

[60] Popoviciu, T., *Les Fonctions Convexes,* Hermann, Paris, 1945.

[61] Prekopa, A., *On the development of optimization theory,* American Mathematical Monthly, [7] 87, 1980, 527-542.

[62] Roberts, A.V. and D.E. Varberg, *Convex Functions,* Academic Press, New York, 1973.

[63] Rockafellar, R.T., *Convex Analysis,* in Princeton Mathematical Series, v. 28, Princeton University Press, Princeton, New Jersey, 1972.

[64] Rudin, W., *Principles of Mathematical Analysis,* McGraw-Hill, New York, 1976.

[65] Rudin, W., *Functional Analysis,* McGraw-Hill, New York, 1991.

[66] Sion, M., *On general minimax theorems,* Pacific Journal of Mathematics, 8, 1958, 171-176.

[67] Sturm, J.F., *Primal-Dual Interior Point Approach to Semidefinite Programming,* PhD. Thesis, Econometric Institute, Erasmus University Rotterdam, 1997.

[68] Valentine, F.A., *Convex Sets,* Mc-Graw Hill, San Francisco, 1964.

[69] van Tiel, J., *Convex Analysis: an Introductory Text,* Wiley, New York, 1984.

[70] Ville, J., *Sur la théorie générale des jeux au intervient l'habilité des jouers,* [2] (5), in Traité du calcul des probabilités et de ses applications, E. Borel and others, eds., Gauthier-Villars Cie, Paris, 1938, 105-113.

[71] Von Neumann, J., *Zur Theorie der Gesellschaftsspiele,* Math. Ann, 100, 1928, 295-320.

[72] Vorob'ev, N.N., *Game Theory: Lectures for Economists and Systems Scientists,* in Applications of Mathematics, v. 7, Springer Verlag, New York, 1977.

[73] Wolkowicz, H., *Some applications of optimization in matrix theory,* Linear Algebra and Its Applications, 40, 1981, 101-118.

[74] Yuang, G.X.-Z., *KKM Theory and Applications in Nonlinear Analysis,* Marcel Dekker, New York, 1999.