CONVEX AND QUASICONVEX ANALYSIS.

J.B.G.FRENK

Contents

1. Sets and Functions on finite dimensional linear spaces.

In this section we consider some basic properties of sets and functions on finite dimensional real linear spaces. In the first subsection we discuss the interplay between sets, hull operations on sets and their topological properties. Moreover in this section linear subspaces, affine and convex sets and cones are introduced. In subsection 2 extended real valued functions on finite dimensional real linear spaces and their properties will be considered. Finally in subsection 3 we will combine some of the results derived in subsection 1 and 2 and end up with an easy proof of the separation result of a convex set and some point outside this convex set. This separation result serves as the main tool within the field of convex and quasiconvex analysis. Although we only deal with finite dimensional linear spaces the basic ideas of the proofs can also be used to prove similar results in infinite dimensional linear spaces. At the same time we have tried to make the proofs as transparent and as simple as possible. Observe most of the material in this section can be found in Rockafellar (cf. [18]), Rudin (cf.[19]) and Hiriart-Urruty and Lemarechal (cf.[8]). For proofs of similar results in infinite dimensional linear (topological) spaces one should consult Chapter 2 and 3 of van Tiel (cf.[21]).

1.1. **Sets and hull operations.** Before introducing some well-known topological concepts in finite dimensional linear spaces we observe that our universe is always given by the n-dimensional Euclidean set R^n. In this set the usual componentswise addition of elements and scalar multiplication of a real number with an element is defined together with the Euclidean inproduct $\langle .,. \rangle : R^n \times R^n \to R$ given by

$$\langle \mathbf{x}, \mathbf{y} \rangle := \mathbf{x}^\intercal \mathbf{y} = \sum_{i=1}^{n} x_i y_i$$

for every $\mathbf{x}, \mathbf{y}$ belonging to R^n. The elements of R^n are called *vectors*[1] or *points*[2] and they are represented by boldfaced characters. The Euclidean norm $\| \mathbf{x} \|$ of the vector $\mathbf{x}$ is given by the nonnegative value

$$\| \mathbf{x} \| := \sqrt{\mathbf{x}^\intercal \mathbf{x}} \geq 0$$

and the set $E \subseteq R^n$ defined by

$$E := \{\mathbf{x} \in R^n : \| \mathbf{x} \| < 1\}.$$

is called the Euclidean unit ball. Moreover, for the sets $A, B \subseteq R^n$ and $\alpha, \beta \in R$ the *Minkowsky sum* $\alpha A + \beta B$ is given by

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

To define a so-called *topology* on the Euclidean space R^n we introduce the following definition taken from Rudin (cf.[19]).

[1] vectors
[2] points

Definition 1.1. *A point $\mathbf{x} \in R^n$ is called an interior point*[3] *of the set $S \subseteq R^n$ if there exists some $\epsilon > 0$ such that*

$$\mathbf{x} + \epsilon E \subseteq S.$$

The set $S \subseteq R^n$ is called open or an open set[4] *if every element of S is an interior point of S. Finally, the set $S \subseteq R^n$ is called closed or a closed set*[5] *if its complement S^c given by*

$$S^c := \{\mathbf{x} \in R^n : \mathbf{x} \notin S\}$$

is an open set.

By definition the empty set $\emptyset$ is open and it is easy to verify that the universe R^n is also an open set. The union $\cup_{i \in I} S_i$ of open sets $S_i, i \in I$ is again an open set and for any finite index set I the intersection $\cap_{i \in I} S_i$ is also open. Similar results hold for closed sets with union replaced by intersection and intersection by union. Since the union of open sets is again open and the empty set $\emptyset$ is open it is easy to construct the biggest open set (possibly empty) contained within the set S and this set is denoted by $\text{int}(S)$. Clearly it follows that

$$\text{int}(S) = \cup\{A : A \subseteq S \text{ and } A \text{ open}\}. \tag{1.1}$$

By the definition of $\text{int}(S)$ it is clear that S equals $\text{int}(S)$ if and only if the set S is open. Due to the intersection of closed sets is again closed and the set R^n is closed it follows that the so called *closure*[6] of S representing the smallest closed set containing S and denoted by $\text{cl}(S)$ is given by

$$\text{cl}(S) = \cap\{A : S \subseteq A \text{ and } A \text{ closed}\}. \tag{1.2}$$

By the definition of $\text{cl}(S)$ it is clear that S equals $\text{cl}(S)$ if and only if the set S is closed. Both constructions are examples of a so-called *hull operation*[7] applied to the set S and in these particular cases the first construction is called the *open hull operation*[8] while the second one is called the *closed hull operation*[9]. To relate the above hull constructions we observe by Definition 1.1 and relations (1.1) and (1.2) that

$$\text{cl}(S)^c = \cup\{B : B \subseteq S^c \text{ and } B \text{ open}\} = \text{int}(S^c) \tag{1.3}$$

To give a more convenient representation of the closure of S we need to introduce the next definition.

Definition 1.2. *A vector $\mathbf{x} \in R^n$ is called a limit point of the nonempty set $S \subseteq R^n$ if for every $\epsilon > 0$ the intersection of the set $\mathbf{x} + \epsilon E$ and S is*

---

[3] interior point
[4] open set
[5] closed set
[6] closure
[7] hull operation
[8] open hull operation
[9] closed hull operation

nonempty. Moreover, the nonempty set $S \subseteq R^n$ has a limit point if there exists some $\mathbf{x} \in R^n$ such that $\mathbf{x}$ is a limit point of the set S.

For any $\mathbf{x} \in R^n$ it is is immediately clear by Definition 1.2 that

$$\mathbf{x} \text{ limit point of } S \Leftrightarrow \mathbf{x} \in S + \epsilon E \text{ for every } \epsilon > 0. \tag{1.4}$$

Observe a point $\mathbf{x}$ is a limit point of the set S does not imply that this point also belongs to S. As an example we consider the set $S = \{\frac{1}{n} : n \in N\}$. Clearly for this set it follows that 0 is a limit point of S while 0 does not belong to this set. A more convenient characterization of the closure of a set and a closed set is listed by the following result.

Lemma 1.1. *For any nonempty set $S \subseteq R^n$ it follows that*

$$cl(S) = \cap_{\epsilon > 0}\{S + \epsilon E\}.$$

Moreover, the nonempty set $S \subseteq R^n$ is closed if and only if every limit point of the set S belongs to S.

Proof. By relation (1.3) we obtain

$$\mathbf{y} \in \mathrm{cl}(S) \Leftrightarrow \mathbf{y} \notin \mathrm{int}(S^c) \Leftrightarrow (\mathbf{y} + \epsilon E) \cap S \neq \emptyset \text{ for every } \epsilon > 0.$$

Hence it follows that

$$\mathbf{y} \in \mathrm{cl}(S) \Leftrightarrow \mathbf{y} \in S + \epsilon E \text{ for every } \epsilon > 0.$$

and this shows the first part. To check the second part we observe by relation (1.4) and the representation of the closure of a set verified in the first part of this lemma that

$$S \text{ closed } \Leftrightarrow S = \cap_{\epsilon > 0}\{S + \epsilon E\} \Leftrightarrow \text{limitpoint of } S \text{ belongs to } S.$$

and this shows the second part. ∎

A property related to open sets and very useful within finite dimensional optimization is given by compactness.

Definition 1.3. *An open cover of the nonempty set $S \subseteq R^n$ is a collection of open sets $S_i \subseteq R^n, i \in I$ satisfying*

$$S \subseteq \cup_{i \in I} S_i.$$

A nonempty set $S \subseteq R^n$ is called compact[10] *if every open cover of S contains a finite subcover.Moreover, a nonempty set $S \subseteq R^n$ is said to be sequentially compact*[11] *if every infinite subset of S has a limit point and this limit point belongs to S.*

Without proof we mention (cf.[19]) that compact sets are closed and bounded and closed subsets of compact sets are compact. To identify compact sets within R^n we mention the following important result (cf. [19]).

[10] compact

[11] sequentially compact

Lemma 1.2. *The nonempty set* $\Pi_{i=1}^{n}[a_i, b_i] \subseteq R^n$ *given by*

$$\Pi_{i=1}^{n}[a_i, b_i] := \{\mathbf{x} \in R^n : \mathbf{x} = (x_1, ..., x_n), a_i \leq x_i \leq b_i\}$$

is compact.

The most important consequence of Lemma 1.2 and the previous observations on compact sets is given by the following characterization of compact sets in R^n (cf.[19]).

Lemma 1.3. *If* $S \subseteq R^n$ *is a nonempty set then it follows that*

$$S \text{ closed and bounded} \Leftrightarrow S \text{ compact} \Leftrightarrow S \text{ sequentially compact.}$$

In infinite dimensional metric linear spaces it does not hold in general that any closed and bounded set is compact (cf.[2]). However, in these more general linear spaces one can show by a similar proof as used in R^n that compactness is equivalent to sequential compactness. Since in optimization we are dealing with *sequences generated by algorithms* sequential compactness is very important while closeness and boundedness are the most easy conditions to *check* for compactness. A useful and important observation of Lemma 1.3 is given by the following result of which the proof is taken from Kreyszig (cf.[2]). Before discussing this result we introduce the definition of linear independence.

Definition 1.4. *The vectors* $\mathbf{x}_1, ..., \mathbf{x}_k$ *are called linear independent*[12] *if*

$$\sum\nolimits_{i=1}^{k} \alpha_i \mathbf{x}_i = \mathbf{0} \Rightarrow \alpha_i = 0, 1 \leq i \leq k.$$

It is now possible to show the following result.

Lemma 1.4. *If* $\{\mathbf{x}_1, ..., \mathbf{x}_m\} \subseteq R^n$ *is a set of linear independent vectors then there exists some* $c > 0$ *such that for every* $\alpha_i \in R, 1 \leq i \leq m$ *it follows that*

$$\| \sum\nolimits_{i=1}^{m} \alpha_i \mathbf{x}_i \| \geq c \sum\nolimits_{i=1}^{m} |\alpha_i|.$$

Proof. We only need to show the above inequality for $s := \sum_{i=1}^{m} |\alpha_i| > 0$ and by normalizing we may assume without loss of generality that $s = 1$ or equivalently $\boldsymbol{\alpha} \in D$ with

$$D := \{\boldsymbol{\alpha} \in R^m : \sum\nolimits_{i=1}^{m} |\alpha_i| = 1\}.$$

Suppose by contradiction hat the above inequality does not hold. This implies that we can find a sequence $\boldsymbol{\alpha}^{(n)} \in D, n \in N$ satisfying

$$\boldsymbol{\alpha}^{(n)} = (\alpha_1^{(n)}, ..., \alpha_m^{(n)}) \text{ and } \lim_{n \uparrow \infty} \| \sum\nolimits_{i=1}^{m} \alpha_i^{(n)} \mathbf{x}_i \| = 0. \tag{1.5}$$

[12]linear independent

Clearly the set D is closed and bounded and hence by Lemma 1.3 the set D is sequentially compact. This means that the set $\{\boldsymbol{\alpha}^{(n)}; n \in N\} \subseteq D$ has a convergent subsequence $\{\boldsymbol{\alpha}^{(n)} : n \in N_0 \subseteq N\}$ with

$$\lim_{n \in N_0} \boldsymbol{\alpha}^{(n)} = \boldsymbol{\alpha}^{(\infty)} \in D$$

and so it follows by relation (1.5) that

$$0 = \lim_{n \in N_0 \uparrow \infty} \parallel \sum_{i=1}^{m} \alpha_i^{(n)} \mathbf{x}_i \parallel = \parallel \sum_{i=1}^{m} \alpha_i^{(\infty)} \mathbf{x}_i \parallel$$

By the independence of the set $\{\mathbf{x}_1, ..., \mathbf{x}_m\}$ we obtain that $\boldsymbol{\alpha}^{(\infty)} = \mathbf{0}$ contradicting $\boldsymbol{\alpha}^{(\infty)} \in D$. Hence our initial assumption is incorrect and we have shown the desired inequality. ∎

Until now we did not introduce sets with additional *algebraic* properties. The first sets to be considered are the main topic of study within *linear algebra* (cf.[18]).

Definition 1.5. *A nonempty set $L \subseteq R^n$ is called a linear subspace*[13] *if $\alpha L + \beta L \subseteq L$ for every real scalars α and β. Moreover, a nonempty set $M \subseteq R^n$ is called an affine set or affine* [14] *if $\alpha M + (1-\alpha)M \subseteq M$ for every real scalar α.*

It is immediately clear that a linear subspace is an affine set. The next result (cf.[18]) shows that linear subspaces and affine sets are closely related.

Lemma 1.5. *It follows that the nonempty set M is affine and $\mathbf{0} \in M$ if and only if M is a nonempty linear subspace.Moreover, for each nonempty affine set M there exists a unique linear subspace L_M satisfying $M = L_M + \mathbf{x}$ for any given $\mathbf{x} \in M$.*

Proof. To show the first part it is clear that any linear subspace contains the zero element and is an affine set. To verify that any affine set containing the zero element is a linear subspace we observe for every $\alpha \in R$ that

$$\alpha M = \alpha M + (1-\alpha)\mathbf{0} \subseteq M. \tag{1.6}$$

This implies for every $\alpha, \beta \in R$ that $\alpha M + \beta M \subseteq M + M$ and since by the definition of an affine set and relation (1.6) we obtain

$$M + M = 2(\frac{1}{2}M + \frac{1}{2}M) \subseteq 2M \subseteq M$$

it follows that $\alpha M + \beta M \subseteq M$. To verify the second part we observe for every $\mathbf{x} \in M$ that the set $M - \mathbf{x}$ is an affine set containing the zero element and hence by the first part a linear subspace. To prove the uniqueness let $M = L_1 + \mathbf{x}$ and $M = L_2 + \mathbf{y}$ with $\mathbf{x}, \mathbf{y} \in M$ and L_1, L_2 linear subspaces.

[13] linear subspace
[14] affine set

This implies $L_1 = L_2 + \mathbf{y} - \mathbf{x}$ and since the zero element belongs to L_1 we obtain that $\mathbf{x} - \mathbf{y}$ belongs to the linear subspace L_2. Hence it follows that

$$L_1 = L_2 + \mathbf{y} - \mathbf{x} \subseteq L_2 - L_2 = L_2$$

By a similar proof one can prove the reverse inclusion and this verifies the second part. ∎

Although in the next section we will study in detail vector valued mappings we mention in this section two classess of mappings which preserve respectively linear subspaces and affine sets.

Definition 1.6. *A mapping $A : R^n \to R^m$ is called linear or a linear mapping*[15] *if*

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

for every real scalars α, β and $\mathbf{x}, \mathbf{y} \in R^n$. Moreover, the mapping A is called affine or an affine mapping[16] *if*

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

for every real scalar α and $\mathbf{x}, \mathbf{y} \in R^n$.

It is easy to check that a linear mapping $A : R^n \to R^m$ is completely determined by $A\mathbf{e}_i$ with $\mathbf{e}_i$ denoting the ith unit vector in R^n and so there is a one-to-one correspondence with the set of $n \times m$ matrices. It is also easy to check that

$$A : R^n \to R^m \text{ is affine} \Leftrightarrow \text{the mapping } \mathbf{x} \to A(\mathbf{x}) - A(\mathbf{0}) \text{ is linear}$$

Moreover, by the definition of a linear mapping the set $A(L)$ given by

$$A(L) := \{A(\mathbf{x}) : \mathbf{x} \in L\}$$

is a linear subspace of R^m for L a linear subspace of R^n while the same holds for an affine mapping with linear replaced by affine.

To study hull operations on linear subspaces and affine sets we first observe that the intersection $\cap_{i \in I} L_i$ is again a linear subspace for any collection $L_i, i \in I$ of linear subspaces. Observe the same preservation result holds for affine sets. Since the set R^n is a linear subspace we can as before apply to any nonempty set $S \subseteq R^n$ the so-called *linear hull operation*[17] and construct the set

$$\text{lin}(S) := \cap\{L : S \subseteq L \text{ and } L \text{ a linear subspace}\}. \tag{1.7}$$

By the preservation of linear subspaces under intersection the above set is clearly the smallest linear subspace containing S and as one might expect this set is called the *linear hull generated by the set S*[18]. In case the set S has a finite number of elements the linear hull is called *finitely generated*[19].

[15] linear mapping
[16] affine mapping
[17] linear hull operation
[18] linear hull generated by S
[19] finitely generated linear hull

Similarly one can construct by means of the so-called *affine hull operation*[20] the smallest affine set containing S. This set denoted by $\text{aff}(S)$ is called the *affine hull generated by the set* S[21] and is given by

$$\text{aff}(S) := \cap\{M : S \subseteq M \text{ and } M \text{ an affine set}\}. \tag{1.8}$$

Since any linear subspace is an affine set it is clear that $\text{aff}(S) \subseteq \text{lin}(S)$ and again if the set S has a finite number of elements the affine hull is called *finitely generated.* To give an alternative representation of these sets we introduce the next definition.

Definition 1.7. *If $S \subseteq R^n$ is a nonempty set then a finite linear combination*[22] *of elements of the set S is given by $\sum_{i=1}^m \alpha_i \mathbf{x}_i$ with α_i real and $\{\mathbf{x}_1, ..., \mathbf{x}_m\} \subseteq S$. Moreover, a finite affine combination*[23] *of elements of the set S is given by $\sum_{i=1}^m \alpha_i \mathbf{x}_i$ with $\sum_{i=1}^n \alpha_i = 1$ and $\{\mathbf{x}_1, ..., \mathbf{x}_m\} \subseteq S$.*

It is now easy to give the following characterization of a linear subspace L and an affine set M. This immediately yields a representation for the linear and affine hull of a set S. Since the proof is trivial it is omitted.

Lemma 1.6. *A nonempty set $L \subseteq R^n$ is a linear subspace if and only if it contains all the finite linear combinations of elements of L. Moreover, the set lin(S) equals all finite linear combinations of the nonempty set $S \subseteq R^n$. A similar observation holds for affine sets with linear combination replaced by affine combination.*

By Lemma 1.6 it is clear that a linear and affine hull generated by a set S is given by

$$\text{lin}(S) = \cup_{m=1}^{\infty}\{\sum_{i=1}^{m} \alpha_i S : \alpha_i \text{ real}\} \tag{1.9}$$

and

$$\text{aff}(S) = \cup_{m=1}^{\infty}\{\sum_{i=1}^{m} \alpha_i S : \alpha_i \text{ real and } \sum_{i=1}^{m} \alpha_i = 1\} \tag{1.10}$$

and the above formulas are examples of a so-called *primal representation.* Using these formulas it is easy to show that there exists a close relation between a linear and affine hull. This is also to be expected by the second part of Lemma 1.5.

Lemma 1.7. *For any nonempty set $S \subseteq R^n$ and $\mathbf{x}_0$ belonging to aff(S) it follows that*

$$\textit{aff}(S) = \mathbf{x}_0 + \textit{lin}(S - \mathbf{x}_0).$$

[20] affine hull operation
[21] affine hull generated by S
[22] finite linear combination
[23] finite affine combination

Proof. For any $\mathbf{x}_0$ belonging to $\text{aff}(S)$ and $\alpha_i, 1 \leq i \leq m$ satisfying $\sum_{i=1}^m \alpha_i = 1$ it follows by relation (1.9) that

$$\sum_{i=1}^m \alpha_i S = \sum_{i=1}^m \alpha_i (S - \mathbf{x}_0) + \mathbf{x}_0 \subseteq \text{lin}(S - \mathbf{x}_0) + \mathbf{x}_0.$$

This implies by relation (1.10) that

$$\text{aff}(S) \subseteq \mathbf{x}_0 + \text{lin}(S - \mathbf{x}_0).$$

Moreover, since $\text{aff}(S) - \mathbf{x}_0$ is an affine set containing the zero element and $S - \mathbf{x}_0$ we obtain by Lemma 1.5 that $\text{aff}(S) - \mathbf{x}_0$ is a linear subspace containing $S - \mathbf{x}_0$ and so

$$\text{lin}(S - \mathbf{x}_0) \subseteq \text{aff}(S) - \mathbf{x}_0,$$

This implies $\text{lin}(S - \mathbf{x}_0) + \mathbf{x}_0 \subseteq \text{lin}(S)$ and so the desired representation is verified. ∎

The next result shows that the affine hull of the cartesian product of the sets S_1 and S_2 equals the cartesian product of the affine hulls.

Lemma 1.8. *If $S_1 \subseteq R^n$ and $S_2 \subseteq R^m$ are nonempty sets then it follows that*

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

Proof. By the representation for the affine hull of a set given by relation (1.10) it follows that

$$\text{aff}(S_1 \times S_2) \subseteq \text{aff}(S_1) \times \text{aff}(S_2)$$

Moreover, if $(\mathbf{x}, \mathbf{y})$ belongs to $\text{aff}(S_1) \times S_2$ then again by relation (1.10) one can find some points $\mathbf{x}_i \in S_1, 1 \leq i \leq m$ satisfying

$$\mathbf{x} = \sum_{i=1}^m \alpha_i \mathbf{x}_i \text{ and } \sum_{i=1}^m \alpha_i = 1.$$

This implies

$$\begin{pmatrix} \mathbf{x} \\ \mathbf{y} \end{pmatrix} = \sum_{i=1}^m \alpha_i \begin{pmatrix} \mathbf{x}_i \\ \mathbf{y} \end{pmatrix} \in \text{aff}(S_1 \times S_2)$$

and so it follows that

$$\text{aff}(S_1) \times S_2 \subseteq \text{aff}(S_1 \times S_2) \tag{1.11}$$

Similarly for $(\mathbf{x}, \mathbf{y})$ belonging to $\text{aff}(S_1) \times \text{aff}(S_2)$ one can find points $\mathbf{y}_i \in S_2, 1 \leq i \leq m$ satisfying

$$\mathbf{y} = \sum_{i=1}^m \alpha_i \mathbf{y}_i \text{ and } \sum_{i=1}^m \alpha_i = 1.$$

and this implies by relation (1.11) that

$$\begin{pmatrix} \mathbf{x} \\ \mathbf{y} \end{pmatrix} = \sum_{i=1}^m \alpha_i \begin{pmatrix} \mathbf{x} \\ \mathbf{y}_i \end{pmatrix} \in \sum_{i=1}^m \alpha_i \text{aff}(S_1 \times S_2) = \text{aff}(S_1 \times S_2).$$

Hence we have shown that

$$\text{aff}(S_1) \times \text{aff}(S_2) \subseteq \text{aff}(S_1 \times S_2)$$

and combining this with relation (1.11) the desired result follows. ∎

By Lemma 1.6 it is easy to show for any affine mapping $A : R^n \to R^m$ that

$$A(\text{aff}(S)) = \text{aff}(A(S)) \tag{1.12}$$

and this relation in combination with Lemma 1.8 yields the following rule for the affine hull for the affine hull of the sum of sets.

Lemma 1.9. *If the set $S_1, S_2 \subseteq R^n$ are nonempty and α, β some scalars then it follows that*

$$\text{aff}(\alpha S_1 + \beta S_2) = \alpha\, \text{aff}(S_1) + \beta\, \text{aff}(S_2).$$

Proof. Introduce the linear mapping $A : R^{2n} \to R^n$ given by $A(\mathbf{x}, \mathbf{y}) = \alpha\mathbf{x} + \beta\mathbf{y}$. Applying this mapping to relation (1.12) with S replaced by $S_1 \times S_2$ it follows by Lemma 1.8 that

$$\begin{aligned} \text{aff}(\alpha S_1 + \beta S_2) &= \text{aff}(A(S_1 \times S_2)) = A(\text{aff}(S_1 \times S_2)) \\ &= A(\text{aff}(S_1) \times \text{aff}(S_2)) = \alpha\text{aff}(S_1) + \beta\text{aff}(S_2) \end{aligned}$$

and this shows the desired result. ∎

An improvement of Lemma 1.6 in finite dimensional linear spaces is given by the observation that any linear subspace (affine set) can be represented by the set of finite linear (affine) combinations of a *finite* and *fixed* subset $S \subseteq R^n$. Due to this finite representation it can be shown that linear subspaces and affine sets are closed. Before presenting this improvement we introduce the following definition.

Definition 1.8. *The vectors $\mathbf{x}_0, ..., \mathbf{x}_k$ are called affinely independent*[24] *if the vectors $\mathbf{x}_1 - \mathbf{x}_0, ..., \mathbf{x}_k - \mathbf{x}_0$ are linear independent.*

To explain the name linear and affine independent we observe that the vectors $\mathbf{x}_1, ..., \mathbf{x}_k$ are linear independent if and only if any $\mathbf{x}$ belonging to $\text{lin}(\{\mathbf{x}_1, ..., \mathbf{x}_k\})$ can be represented as a *unique linear combination* of the vectors $\mathbf{x}_1, ..., \mathbf{x}_k$. An immediate consequence of the next result shows that a similar observation also holds for affinely independent vectors with unique linear combination replaced by unique affine combination.

Lemma 1.10. *The vectors $\mathbf{x}_0, ..., \mathbf{x}_k$ are affinely independent if and only if the system*

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

has a unique solution and this is given by $\alpha_i = 0$ for every $0 \leq i \leq k$.

[24] affinely independent

Proof. For a given set of affine independent vectors $\mathbf{x}_0, ..., \mathbf{x}_k$ we consider the system

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

Clearly for $\alpha_0, ..., \alpha_k$ satisfying the above system we obtain that

$$\sum_{i=1}^{k} \alpha_i (\mathbf{x}_i - \mathbf{x}_0) = \mathbf{0} \text{ and } \sum_{i=0}^{k} \alpha_i = 0.$$

and this shows by Definition 1.8 applied to the first relation and $\sum_{i=0}^{k} \alpha_i = 0$ that $\alpha_i = 0$ for every $0 \leq i \leq k$. To verify the reverse inclusion we consider for a given set of vectors $\mathbf{x}_0, ..., \mathbf{x}_k$ the system

$$\sum_{i=1}^{k} \alpha_i (\mathbf{x}_i - \mathbf{x}_0) = \mathbf{0}.$$

or equivalently the system

$$-\sum_{i=1}^{k} \alpha_i \mathbf{x}_0 + \sum_{i=1}^{k} \alpha_i \mathbf{x}_i = \mathbf{0}$$

By our assumption the only solution of the above system is given by $\alpha_i = 0, 0 \leq i \leq k$ and so the vectors $\mathbf{x}_1 - \mathbf{x}_0, ..., \mathbf{x}_k - \mathbf{x}_0$ are linear independent. ∎

It is now possible to show the following improvement of Lemma 1.6.

Lemma 1.11. *For any nonempty linear subspace $L \subseteq R^n$ there exists a set of linear independent vectors $\mathbf{x}_1, ..., \mathbf{x}_k$, $k \leq n$ such that*

$$L = lin(\{\mathbf{x}_1, ..., \mathbf{x}_k\})$$

and for any nonempty affine set $M \subseteq R^n$ there exists a set of affinely independent vectors $\mathbf{x}_0, ..., \mathbf{x}_k, k \leq n$ satisfying

$$M = aff(\{\mathbf{x}_0, ..., \mathbf{x}_k\}).$$

Proof. It is well known from linear algebra (cf.[17]) that any nonempty linear subspace $L \subseteq R^n$ is generated by a finite set of at most n linear independent vectors. Moreover, by Lemma 1.5 and the first part of this lemma it follows for any affine set $M \subseteq R^n$ and $\mathbf{x}_0 \in M$ that

$$M - \mathbf{x}_0 = \text{lin}(\{\mathbf{x}_1 - \mathbf{x}_0, ..., \mathbf{x}_k - \mathbf{x}_0\})$$

with the set of vectors $\mathbf{x}_1 - \mathbf{x}_0, ..., \mathbf{x}_k - \mathbf{x}_0, k \leq n$ linear independent or equivalently $\mathbf{x}_i, 0 \leq i \leq k$ affine independent. Applying now Lemma 1.7 implies that

$$M = \mathbf{x}_0 + \text{lin}(\{\mathbf{x}_1 - \mathbf{x}_0, ..., \mathbf{x}_k - \mathbf{x}_0\}) = \text{aff}(\{\mathbf{x}_0, ..., \mathbf{x}_k\})$$

and this shows the desired result. ∎

In case the linear subspace L is represented by the linear hull of k linear independent vectors the dimension $\dim(L)$ is given by k. By the definition of linear independence any $\mathbf{x}$ belonging to L can be written as a unique linear combination of the linear independent vectors $\mathbf{x}_1, ..., \mathbf{x}_k$, while at the same time this implies that $\dim(L)$ is well defined. Moreover, the dimension

$\dim(M)$ of an affine set M is by definition the dimension of the unique subspace L_M parallel to M. By Lemma 1.10 it is also clear that any $\mathbf{x}$ belonging to M can be written as a unique affine combination of affine independent vectors $\mathbf{x}_0, ..., \mathbf{x}_k$. Finally we observe for completeness that the *dimension* $\dim(S)$ *of an arbitrary set* S[25] is given by $\dim(\text{aff}(S))$. An immediate and important consequence of the finite representation of any affine set is given by the next result.

Lemma 1.12. *Any nonempty affine set $M \subseteq R^n$ is closed.*

Proof. Since any affine set is a translation of a linear subspace it is sufficient to prove that a linear subspace is closed. By Lemma 1.11 it follows for a given linear subspace $L \subseteq R^n$ that there exists a set of linear independent vectors $\mathbf{x}_1, .., \mathbf{x}_k, k \leq n$ satisfying $L = \text{lin}(\{\mathbf{x}_1, .., \mathbf{x}_k\})$ and to prove that this set is closed we consider some sequence $\{\mathbf{y}_n : n \in N\} \subseteq L$ given by

$$\mathbf{y}_n = \sum_{i=1}^k \alpha_i^{(n)} \mathbf{x}_i \text{ and satisfying } \lim_{n \uparrow \infty} \mathbf{y}_n = \mathbf{y}_\infty. \tag{1.13}$$

By lemma 1.1 it is now sufficient to show that $\mathbf{y}_\infty$ belongs to L. Since the set $\{\mathbf{x}_1, .., \mathbf{x}_k\}$ consists of linear independent vectors we can apply Lemma 1.4 and so by relation (1.13) there exist some $c_0 > 0$ and $c > 0$ such that

$$c_0 \geq \| \mathbf{y}_n \| \geq c \sum_{i=1}^k |\alpha_i^{(n)}|$$

for every $n \in N$. Hence by Lemma 1.3 the sequence $\boldsymbol{\alpha}^{(n)}$ is contained in a compact set and has therefore a convergent subsequence with limit $\boldsymbol{\alpha}^{(\infty)}$. This implies by relation (1.13) that

$$\mathbf{y}_\infty = \sum_{i=1}^m \alpha_i^{(\infty)} \mathbf{x}_i$$

and this show that $\mathbf{y}_\infty$ belongs to L. ∎

Since by Lemma 1.12 any affine or linear hull of a given nonempty set $S \subseteq R^n$ is closed we obtain by the definition of the hull operation that

$$\text{cl}(S) \subseteq \text{aff}(S) \subseteq \text{lin}(S)$$

and this yields by the monotonicity of the hull operation that

$$\text{aff}(\text{cl}((S)) = \text{aff}(S) \text{ and } \text{lin}(\text{cl}(S)) = \text{lin}(S). \tag{1.14}$$

In contrast to the primal representation of a linear subspace or affine set given by Lemma 1.11 we can also give a so-called *dual representation*[26] 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 characterisation can be seen as a "improvement " of the hull operation given by relations (1.7) and (1.8).

[25] dimension of arbitrary set S

[26] dual representation

Definition 1.9. *If $S \subseteq R^n$ is some nonempty set then the nonempty set $S^\perp \subseteq R^n$ given by*

$$S^\perp = \{\mathbf{x}^* \in R^n : \mathbf{x}^\intercal \mathbf{x}^* = 0 \text{ for every } \mathbf{x} \in S\}$$

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 nonempty linear subspace. Moreover, a basic result (cf.[17]) in linear algebra is given by the following.

Lemma 1.13. *For any nonempty linear subspace L its so-called biorthogonal complement $(L^\perp)^\perp$ equals L. Moreover, every $\mathbf{x} \in R^n$ can be uniquely decomposed as the sum of an element from L and $L^\perp$ and these elements are respectively the orthogonal projection of $\mathbf{x}$ on L and $L^\perp$. Finally $n = \dim(L) + \dim(L^\perp)$.*

By the above lemma a so-called dual representation of any linear hull $\text{lin}(S)$ can be given using the following procedure. It is easy to verify for any $S_1 \subseteq S_2$ that $S_2^\perp \subseteq S_1^\perp$ and so $(S_1^\perp)^\perp \subseteq (S_2^\perp)^\perp$. Since $S \subseteq \text{lin}(S)$ this yields by Lemma 1.13 that

$$(S^\perp)^\perp \subseteq (\text{lin}(S)^\perp)^\perp = \text{lin}(S).$$

Moreover, $(S^\perp)^\perp$ is clearly a linear subspace containing S and hence by the definition of a linear hull we obtain the dual representation

$$(S^\perp)^\perp = \text{lin}(S). \tag{1.15}$$

For affine hulls it follows by Lemma 1.7 and relation (1.15) that

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

for every $\mathbf{x}_0$ belonging to $\text{aff}(S)$. Since it is easy to verify that $\boldsymbol{\lambda}^\top(\mathbf{x}_1 - \mathbf{x}_0) = 0$ for every $\boldsymbol{\lambda}$ belonging to $(S - \mathbf{x}_0)^\perp$ and $\mathbf{x}_1 \in \text{aff}(S)$ we obtain that $(S - \mathbf{x}_0)^\perp \subseteq (S - \mathbf{x}_1)^\perp$ for every $\mathbf{x}_0, \mathbf{x}_1 \in \text{aff}(S)$ and by a similar argument the reverse inclusion also holds. Hence it follows that

$$(S - \mathbf{x}_0)^\perp = (S - \mathbf{x}_1)^\perp$$

for every $\mathbf{x}_0, \mathbf{x}_1$ belonging to $\text{aff}(S)$ and so a dual representation of the affine hull of a set S is given by

$$\text{aff}(S) = \mathbf{x}_0 + ((S - \mathbf{x}_1)^\perp)^\perp \tag{1.16}$$

for every $\mathbf{x}_0, \mathbf{x}_1 \in \text{aff}(S)$. For arbitrary affine sets a consequence of the dual representation (1.16) is given by the following lemma and this shows again that this characterization is a representation from "outside" the set.

Lemma 1.14. *The set $M \subseteq R^n$ is a nonempty affine set of dimension $n - m \leq n$ if and only if there exists some $m \times n$ matrix A of rank m and some $\mathbf{d} \in R^m$ such that*

$$M = \{\mathbf{x} \in R^n : A\mathbf{x} = \mathbf{d}\}.$$

A similar result holds for any linear subspace L with $\mathbf{d} = \mathbf{0}$.

Proof. We only show that any linear subspace has the above representation. The remaining parts can be easily verified. Clearly by Lemma 1.13 the linear subspace $L^{\perp}$ has dimension m and so by Lemma 1.11 the linear subspace $L^{\perp}$ equals $\text{lin}(\{\mathbf{x}_1,...,\mathbf{x}_m\})$ with $\mathbf{x}_1,...,\mathbf{x}_m$ a set of m linearly independent vectors. Taking the matrix A consisting of the columns $\mathbf{x}_1,...,\mathbf{x}_m$ it follows due to relation (1.15) that

$$L = (L^{\perp})^{\perp} = \{\mathbf{x} \in R^n : A\mathbf{x} = \mathbf{0}\}$$

and this shows the desired result. ∎

The above result concludes our discussion of linear subspaces and affine sets and we continue now with sets which are the main topic of study within the field of convex and quasiconvex analysis. These sets are introduced in the next definition.

Definition 1.10. *A nonempty set $C \subseteq R^n$ is called a convex set*[27] *or convex if $\alpha C+(1-\alpha)C \subseteq C$ for every $0 < \alpha < 1$. Moreover, a nonempty set $K \subseteq R^n$ is called a cone if $\alpha K \subseteq K$ for every $\alpha > 0$.*

Observe an affine set is clearly a convex set but not every convex set is an affine set and hence convex analysis is an extension of linear algebra. Also it is easy to verify for any affine mapping $A : R^n \to R^m$ that the set $A(C)$ is convex for any nonempty convex set $C \subseteq R^n$ and that the set $A(K)$ is a cone for $K \subseteq R^n$ a nonempty cone and $A : R^n \to R^m$ a linear mapping. By a similar proof as used in Lemma 1.5 one can also show for every cone K that

$$K \text{ convex} \Leftrightarrow K + K \subseteq K. \tag{1.17}$$

To relate convex sets to convex cones[28] we observe for $R_+ := [0, \infty)$ and any nonempty set $S \subseteq R^n$ that

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

is a cone. This implies by relation (1.17) that the set $R_+(C \times \{1\})$ is a convex cone for any nonempty convex set $C \subseteq R^n$ It is now clear for any set $S \subseteq R^n$ that

$$R_+(S \times \{1\}) \cap (R^n \times \{1\}) = S \times \{1\} \tag{1.18}$$

and so any convex set C can be seen as an intersection of the convex cone $R_+(C \times \{1\})$ and the affine set $R^n \times \{1\}$. This shows that convex sets are closely related to convex cones and by relation (1.18) 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. We also mention that it is easy to verify that the set $A(C)$ is convex for any convex set C and $A : R^n \to R^m$ an affine mapping, while the set $A(K)$ is a convex cone for K a convex cone and A a linear mapping. To

[27] convex set

[28] convex cone

introduce an important class of convex sets we consider the affine mappings $A : R^n \to R$ given by

$$A(\mathbf{x}) = \mathbf{a}^\intercal \mathbf{x} + b$$

with $\mathbf{a} \in R^n$ and $b \in R$. The set $H^{<}(\mathbf{a}, b)$ represented by

$$H^{<}(\mathbf{a}, b) = \{\mathbf{x} \in R^n : \mathbf{a}^\intercal \mathbf{x} < b\} \tag{1.19}$$

is called a halfspace[29] and clearly this halfspace is an open convex set. Moreover, the set

$$H^{\leq}(\mathbf{a}, b) = \{\mathbf{x} \in R^n : \mathbf{a}^\intercal \mathbf{x} \leq b\} \tag{1.20}$$

is also called a halfspace and this set is clearly a closed convex set. An important subclass of convex sets is now given by the following definition (cf.[22]).

Definition 1.11. *A set $C_e \subseteq R^n$ is called evenly convex or an evenly convex set if C_e is the intersection of a collection of open halfspaces or $C_e = R^n$.*

Since any closed halfspace $H^{\leq}(\mathbf{a}, b)$ can be obtained by intersecting the open halfspaces $H^{<}(\mathbf{a}, b + \frac{1}{n}), n \geq 1$ it follows that any closed halfspace is evenly convex. In Section 3 we will show that any closed and open convex set is evenly convex. However, there exist convex sets which are not evenly convex and an example of a convex set which is not evenly convex is shown at page 10 of Gromicho (cf. [9]).

Since the intersection $\cap_{i \in I} C_i$ is again convex for any collection $C_i, i \in I$ of convex sets and R^n is a convex set we can apply to any nonempty set S the so-called *convex hull operation*[30] and this results in the set

$$\text{co}(S) := \cap \{C : S \subseteq C \text{ and } C \text{ convex }\}. \tag{1.21}$$

By the preservation of convexity under intersection the set $\text{co}(S)$ is clearly the smallest convex set containing S and as one might expect this set is called the *convex hull generated by* S[31]. Moreover, if S is a finite set then the convex hull $\text{co}(S)$ is called *finitely generated.* Since by definition evenly convex sets are also closed under intersection we can similarly apply to any nonempty set S the so-called *evenly convex hull operation*[32] and this yields the set

$$\text{eco}(S) := \cap \{C_e : S \subseteq C_e \text{ and } C_e \text{ evenly convex }\}. \tag{1.22}$$

This set is called the *evenly convex hull generated by* S[33] and by the above observations it is the smallest evenly convex set containing S. Since any evenly convex set is convex it follows that in general $\text{co}(S) \subseteq \text{eco}(S)$. By the so-called *canonic hull operation*[34] one can also construct the smallest

[29] halfspace
[30] convex hull operation
[31] convex hull generated by S
[32] evenly convex hull operation
[33] evenly convex hull generated by S
[34] canonic hull operation

convex cone containing S and the smallest convex cone containing $S \cup \{\mathbf{0}\}$. The last set is given by

$$\text{cone}(S) := \cap\{K : S \cup \{0\} \subseteq K \text{ and } K \text{ convex cone}\}. \tag{1.23}$$

and unfortunately this set is called the *convex cone generated by* S[35] (cf.[18]). Clearly the set cone(S) is in general not equal to the smallest 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 1.12. *If $S \subseteq R^n$ is a nonempty set then a finite (strict) canonical combination of elements of the set S is given by $\sum_{i=1}^m \alpha_i \mathbf{x}_i$ with α_i (positive) nonnegative and $\{\mathbf{x}_1, ..., \mathbf{x}_n\} \subseteq S$. Moreover, a finite convex combination of elements of the set S is given by $\sum_{i=1}^m \alpha_i \mathbf{x}_i$ with α_i nonnegative satisfying $\sum_{i=1}^n \alpha_i = 1$ and $\{\mathbf{x}_1, ..., \mathbf{x}_n\} \subseteq S$.*

It is now easy to give the following so-called *primal representation* of a convex set C and a convex cone K.

Lemma 1.15. *A nonempty set $C \subseteq R^n$ is convex if and only if it contains all finite convex combinations of elements of C. Moreover, a cone $K \subseteq R^n$ is convex if and only if it contains all finite strict canonical combinations of the set K.*

Proof. Clearly if S contains all finite convex combinations then in particular $\alpha S + (1-\alpha)S \subseteq S$ for every $0 < \alpha < 1$ and so S is convex. To prove the reverse implication let S be convex and assume any convex combination of k elements of S belongs to S. We will now show that this also holds for any convex combination of $k+1$ elements of S. Introduce therefore the positive scalars $\alpha_i, i = 1, .., k+1$ satisfying $\sum_{i=1}^{k+1} \alpha_i = 1$. If $\beta_i := \frac{\alpha_i}{1-\alpha_{k+1}} > 0$ for $i = 1, ..., k$ then clearly the identity

$$\sum_{i=1}^{k+1} \alpha_i S = (1-\alpha_{k+1}) \sum_{i=1}^{k} \beta_i S + \alpha_{k+1} S. \tag{1.24}$$

holds. Since $\sum_{i=1}^k \beta_i = 1$ this yields by our induction assumption that $\sum_{i=1}^k \beta_i S \subseteq S$ and hence by the convexity of S and relation (1.24) we obtain that any convex combination of $k+1$ elements belong to S. This shows the first part and to verify the second part one can apply a similar proof. ∎

An immediate consequence of the above lemma is given by the next result.

Lemma 1.16. *For any set $S \subseteq R^n$ it follows that the set co(S) equals all finite convex combinations of the set S. Moreover, the set cone(S) equals all finite canonical combinations of the set S while the smallest convex cone containing S equals all finite strict canonical combinations of the set S.*

Proof. We only give a proof of the first part since the other parts can be proved similarly. If V denotes the set of all finite convex combinations of the set S then by Lemma 1.15 the set V is a convex set containing S and

[35]convex cone generated by S

hence $\text{co}(S) \subseteq V$. Since $\text{co}(S)$ is convex we can apply again Lemma 1.15 and so $\text{co}(S)$ contains all its finite convex combinations. In particular it contains all the finite convex combinations of the set S and hence $V \subseteq \text{co}(S)$. ∎

By Lemma 1.16 it is clear that

$$\text{cone}(S) = \cup_{m=1}^{\infty}\{\sum_{i=1}^{m} \alpha_i S : \alpha_i \geq 0\} \tag{1.25}$$

while

$$\text{co}(S) = \cup_{m=1}^{\infty}\{\sum_{i=1}^{m} \alpha_i S : \sum_{i=1}^{m} \alpha_i = 1, \alpha_i > 0\}. \tag{1.26}$$

We observe that the above representations are the "convex equivalences" of the representation for $\text{lin}(S)$ and $\text{aff}(S)$ given by relations (1.9) and (1.10). Moreover, to relate the above representations it is easy to see that

$$\text{cone}(S) = R_+(\text{co}(S)). \tag{1.27}$$

The above characterizations can also be seen as *primal representations.* Starting with the study of convex sets and cones and reconsidering the finite representation discussed in Lemma 1.11 of affine sets and linear subspaces we might now wonder whether a convex cone K containing $\mathbf{0}$ can always be seen as a set of finite canonical combinations of a finite and fixed set $S \subseteq R^n$.

Example 1.1. *Contrary to linear subspaces it is not true that any convex cone containing* $\mathbf{0}$ *is generated by a finite set. An example is given by the so-called* ice-cream *(convex and closed)* cone K *represented by*

$$K := \{(\mathbf{x}, t) : \| \mathbf{x} \| \leq t\} \subseteq R^{n+1}$$

A similar observation holds for convex sets as shown by the Euclidean unit ball E.

Despite this negative result it is possible in finite dimensional linear spaces to improve for convex cones and convex sets the representation given by relations (1.25) and (1.26). In the next result it is shown that any element belonging to $\text{cone}(S)$ can be written as a canonical combination of at most n linear independent vectors belonging to S. This is called *Caratheodory's theorem* for convex cones. Using this result and relation (1.18) a related result holds for convex sets and in this case linear independent is replaced by affine independent and at most n is replaced by at most $n+1$. Clearly this result is weaker than the corresponding result for affine sets and linear subspaces since for affine sets and linear subspaces we can take for any element belonging to these sets the *same finite set*.

Lemma 1.17. *For any nonempty set* $S \subseteq R^n$ *and any* $\mathbf{x} \in \text{cone}(S)$ *there exists a set of linear independent vectors* $\mathbf{x}_1, ..., \mathbf{x}_k, k \leq n$ *belonging to* S *such that* $\mathbf{x}$ *can be written as a finite canonical combination of these vectors or equivalently*

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

Moreover, for any $\mathbf{x} \in co(S)$ *there exists a set of affinely independent vectors* $\mathbf{x}_0, ..., \mathbf{x}_k, k \leq n+1$ *belonging to* S *such that* $\mathbf{x}$ *can be written as a finite convex combination of these vectors or equivalently*

$$\mathbf{x} = \sum\nolimits_{i=1}^{k} \alpha_i \mathbf{x}_i, \ \alpha_i \geq 0 \ and \ \sum\nolimits_{i=1}^{k} \alpha_i = 1.$$

Proof. To show the first part we consider an arbitrary $\mathbf{x}$ belonging to cone(S). Clearly for $\mathbf{x} = \mathbf{0}$ the result holds and so $\mathbf{x}$ should be nonzero. Applying now relation (1.25) one can find a finite set $\{\mathbf{x}_1, ..., \mathbf{x}_m\} \subseteq S$ such that

$$\mathbf{y} = \sum\nolimits_{i=1}^{m} \alpha_i \mathbf{x}_i, \ \alpha_i \text{ positive}$$

If the set of vectors $\mathbf{x}_1, ..., \mathbf{x}_m$ are linear independent then clearly $m \leq n$ and we are done. Therefore let the set of vectors $\mathbf{x}_1, ..., \mathbf{x}_m$ be linear dependent and by this assumption one can find some nonzero sequence $\{\beta_i : 1 \leq i \leq m\}$ satisfying

$$\sum\nolimits_{i=1}^{m} \beta_i \mathbf{x}_i = \mathbf{0}. \tag{1.28}$$

If some of the scalars β_i are positive introduce

$$\epsilon := \min\{\frac{\alpha_i}{\beta_i} : \beta_i > 0\} \text{ and } i^* := \arg\min\{\frac{\alpha_i}{\beta_i} : \beta_i > 0\}.$$

and this implies by relation (1.28) that

$$\mathbf{y} = \sum\nolimits_{i=1, i \neq i^*}^{m} (\alpha_i - \epsilon \beta_i) \mathbf{x}_i. \tag{1.29}$$

For all β_i negative we replace the minus sign by a plus sign in the above relation. By relation (1.29) we obtain that $\mathbf{y}$ can be written as a canonical combination of at most $m-1$ vectors and by applying the same procedure until all the vectors in the canonical sum are linearly independent we obtain the desired result. Observe this can only happen for $m \leq n$. To show the second part it follows for any $\mathbf{x} \in \text{co}(S)$ that

$$(\mathbf{x}, 1) \in \text{co}(S) \times \{1\} \subseteq R_+(\text{co}(S) \times \{1\}).$$

Applying now relation (1.18) and observing that $R_+(\text{co}(S) \times \{1\}) \subseteq R^{n+1}$ is the convex cone generated by $S \times \{1\}$ we obtain by the first part that there exists a set $(\mathbf{x}_0, 1), ...(\mathbf{x}_k, 1) \in R^{n+1}, k \leq n+1$ of linear independent vectors such that

$$\mathbf{x} = \sum\nolimits_{i=0}^{k} \alpha_i \mathbf{x}_i, \ \alpha_i \geq 0 \text{ and } 1 = \sum\nolimits_{i=0}^{k} \alpha_i.$$

Since the set $(\mathbf{x}_0, 1), ...(\mathbf{x}_k, 1) \in R^{n+1}, k \leq n$ are linear independent this implies that the system

$$\sum_{i=0}^{k} \lambda_i \binom{\mathbf{x}_i}{1} = \mathbf{0}$$

has the unique solution $\lambda_i = 0$ and so by Lemma 1.10 we obtain that the vectors $\mathbf{x}_0, ..., \mathbf{x}_k$ are affinely independent. ∎

By the above result it seems that convex cones (convex sets) are generalizations of linear subspaces (affine sets). Unfortunately opposed to linear subspaces it is not true that any convex cone is closed. The same holds for convex sets. It will be shown for *closed* convex sets and *closed* convex cones that it is relatively easy to give a *dual representation* of those sets and this is the main reason why we like to identify which classes of convex sets and convex cones are closed. Since affine sets can always be generated by a *finite set of affine independent vectors* (and this guarantees that affine sets are closed) and we know by Example 1.1 that this is not true for convex sets one might now wonder which property replacing finiteness should be imposed on S to guarantee that $co(S)$ is closed. Looking at the following counterexample it is not sufficient to impose that the generator S is a closed set and this implies that we need a stronger property.

Example 1.2. *If the set $S \subseteq R^2$ is given by the closed set*

$$S = \mathbf{0} \cup \{(x,1) : x \geq 0\}$$

then it follows by Lemma 1.16 that

$$co(S) = \{(x_1, x_2) : 0 < x_2 \leq 1, x_1 \geq 0\} \cup \{\mathbf{0}\}$$

and this convex set 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 by Lemma 1.3 compact one can show that co(S) is indeed closed and even compact. At the same time this 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*[36] Δ_{n+1} in R^{n+1} defined by

$$\Delta_{n+1} := \{\boldsymbol{\alpha} : \sum\nolimits_{i=1}^{n+1} \alpha_i = 1 \text{ and } \alpha_i \geq 0\} \subseteq R^{n+1}. \tag{1.30}$$

By Lemma 1.17 it follows that

$$\text{co}(S) = f(\Delta_{n+1} \times S^{n+1}) \tag{1.31}$$

with S^m denoting the m-fold Cartesian product of the set $S \subseteq R^n$ and the function f is given by

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

A related observation holds for convex cones and using the above observations one can now show the following result.

Lemma 1.18. *If the nonempty set $S \subseteq R^n$ 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.*

[36]unit simplex

Proof. The unit simplex Δ_{n+1} is closed and bounded and hence by Lemma 1.3 compact. Since S is compact it is easy to show again by Lemma 1.3 that the Cartesian product $\Delta_{n+1} \times S^{n+1}$ is compact. Since it is well-known (cf.[19]) that $h(A)$ is compact for A is compact and $h : R^m \to R^p$ a continuous vector valued function we obtain by the continuity of the function f and relation () that co(S) is compact. This proves the first part and to verify the second part we first observe by relation (1.27) that cone(S)=R_+(co(S)) and so it is sufficient to show that the set R_+(co(S)) is closed. Consider now an arbitrary sequence $t_n\mathbf{x}_n, n \in N$ belonging to R_+(co(S)) and satisfying

$$\lim_{n\uparrow\infty} t_n\mathbf{x}_n = \mathbf{y}. \tag{1.32}$$

By the continuity of the Euclidean norm relation (1.32) implies that

$$\lim_{n\uparrow\infty} t_n \parallel \mathbf{x}_n \parallel = \lim_{n\uparrow\infty} \parallel t_n\mathbf{x}_n \parallel = \parallel \mathbf{y} \parallel . \tag{1.33}$$

Moreover, by the first part the set co(S) is compact and so we can find by Lemma 1.3 a subsequence $N_0 \subseteq N$ satisfying

$$\lim_{n\in N_0\uparrow\infty} \mathbf{x}_n = \mathbf{x}_\infty \in \text{co}(S). \tag{1.34}$$

Since the zero element does not belong to co(S) this implies

$$\lim_{n\in N_0\uparrow\infty} \|\mathbf{x}_n\| = \|\mathbf{x}_\infty\| > 0.$$

and hence by (1.33) we obtain

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

This means that the sequence $t_n, n \in N_0$ is convergent to a finite number t_∞ and this implies by relations (1.32) and (1.34) that

$$\mathbf{y} = t_\infty\mathbf{x}_\infty \in R_+(\text{co}(S))$$

showing the desired result. ∎

One may wonder whether for S compact and $\mathbf{0} \in$ co(S) the set cone(S) is still closed. As shown by the following counterexample this does not hold.

Example 1.3. *If the set $S \subseteq R^2$ is given by the compact set*

$$S = \{(x_1, x_2) : (x_1 - 1)^2 + x_2^2 \leq 1\}$$

then clearly $\mathbf{0} \in S$ and by relation (1.27) it follows that

$$cone(S) = \{(x_1, x_2) : x_1 > 0\} \cup \{\mathbf{0}\}$$

Observe now that the set cone(S) is not closed. This shows that the condition $\mathbf{0} \notin S$ is necessary in Lemma 1.18.

An immediate consequence of Caratheodory's theorem and Lemma 1.18 is given by the next result for convex cones generated by some nonempty set S.

Lemma 1.19. *If the set $S \subseteq R^n$ contains a finite number of elements then the set cone(S) is closed.*

Proof. For any finite set S we consider the finite set V given by

$$V = \{S_i : S_i \subseteq S \text{ and } S_i \text{ consists of linear independent vectors}\}$$

By the first part of Lemma 1.17 it follows for any $\mathbf{x}$ belonging to cone(S) that there exists some $S_i \in V$ such that $\mathbf{x} \in \text{cone}(S_i)$ and this yields

$$\text{cone}(S) = \cup_{S_i \in V} \text{cone}(S_i) \tag{1.35}$$

Since S_i consists of a finite set of linear independent vectors it follows that S_i is compact and 0 does not belong to co(S) and so by the second part of Lemma 1.18 we obtain that cone(S_i) is closed for every S_i belonging to the finite set V. Hence it follows by relation (1.35) that the set cone(S) is closed and this shows the desired result. ∎

A generalization of the orthogonality relation for linear subspaces is given by the polarity relation for convex cones.

Definition 1.13. *If $K \subseteq R^n$ is a nonempty convex cone then the set K^0 given by*

$$K^0 := \{\mathbf{x}^* \in R^n : \mathbf{x}^\intercal \mathbf{x}^* \leq 0 \text{ for every } \mathbf{x} \in K\}$$

is called the polar cone[37] *of K.*

In case L is a linear subspace it is easy to verify that $L^0 = L^\perp$ and so the *polar operator* applied to a linear subspace reduces to the *orthogonal operator.* Moreover, it is also easy to verify that the nonempty set K^0 is a closed convex cone. Without proof we now mention for K a closed convex cone that $(K^0)^0 = K$ and this so-called bipolar result enables us to give a dual representation for closed convex sets and closed convex cones. We only mention this result to make clear that in convex analysis we are actually trying to generalize the orthogonality relation applied to linear subspaces and this enables us to obtain also for closed convex sets and closed convex cones dual representations. Due to this it is hopefully clear that in convex analysis one is interested in closed convex cones and closed convex sets. We will continue with these dual representations in Section 1.3 after a proof of the bipolar result. To be able to prove the strongest possible results for convex sets in finite dimensional spaces we also need to introduce the definition of a relative interior point. This generalizes the notion of an interior point given by Definition 1.1.

Definition 1.14. *A vector $\mathbf{x} \in R^n$ is called a relative interior point*[38] *of the set $S \subseteq R^n$ if $\mathbf{x}$ belongs to aff(S) and there exists some $\epsilon > 0$ such that*

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

[37] polar cone
[38] relative interior point

A set $S \subseteq R^n$ is called regular or a regular set[39] *if the set $ri(S)$ with*

$$ri(S) := \{\mathbf{x} \in R^n : \textit{the vector } \mathbf{x} \textit{ is a relative interior point of } S\}$$

is nonempty. Moreover, the set $S \subseteq R^n$ is called relatively open or a relatively open set[40] *if S equals $ri(S)$.*

As shown by the next example it is quite natural to assume that $\mathbf{x}$ belongs to aff(S). This assumption implies by the second part of the definition of a relative interior point that $\mathbf{x}$ belongs to S.

Example 1.4. *Consider the set $S \subseteq R^2$ given by*

$$S = \{0\} \times [1, -1]$$

and let $\mathbf{x} = (1, 0)$. Clearly the set $aff(S)$ is given by $\{0\} \times R$ and for $\epsilon = 1$ it follows that

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

If in the definition of a relative interior point one would delete 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 would not be natural.

By the above definition it is clear for $S \subseteq R^n$ full dimensional or equivalently aff$(S) = R^n$ 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 manifold is relatively open. Moreover, since by Lemma 1.12 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 1.5. *Consider the convex sets $C_1 = \{0\}$ and $C_2 = [0, 1]$. For these sets it follows that $ri(C_1) = \{0\}$ and $ri(C_2) = (0, 1)$ and so $ri(C_1) \nsubseteq ri(C_2)$. Moreover, it follows that $aff(C_1) \neq aff(C_2)$.*

To guarantee that the relative interior operator is monotone when applied to the sets $S_1 \subseteq S_2 \subseteq R^n$ we need to impose the additional condition that aff$(S_1) =$ aff(S_2). If this holds it is easy to check

$$S_1 \subseteq S_2 \Rightarrow \mathrm{ri}(S_1) \subseteq \mathrm{ri}(S_2). \tag{1.36}$$

By the above observation it is important to known which different sets cannot be distinguished by the affine operator. The next result shows that this holds for the sets S, cl(S), co(S) and cl(co(S)).

Lemma 1.20. *It follows for every nonempty set $S \subseteq R^n$ that*

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

[39] regular set
[40] relatively open set

Proof. Since the affine operator is monotone and $S \subseteq \text{co}(S) \subseteq \text{cl}(\text{co}(S))$ and $S \subseteq \text{cl}(S) \subseteq \text{cl}(\text{co}(S))$ for any nonempty set S it is sufficient to verify that $\text{aff}(S) = \text{aff}(\text{cl}(\text{co}(S)))$. By Lemma 1.12 the set $\text{aff}(S)$ containing S is closed and since $\text{aff}(S)$ is also convex and $\text{cl}(\text{co}(S))$ is the smallest closed convex set containing S it follows that

$$\text{cl}(\text{co}(S)) \subseteq \text{aff}(S).$$

This yields by the monotonicity of the affine operator that

$$\text{aff}(\text{cl}(\text{co}(S))) \subseteq \text{aff}(\text{aff}(S)) = \text{aff}(S)$$

Again by the monotonicity of the affine operator it follows that $\text{aff}(S) \subseteq \text{aff}(\text{cl}(\text{co}(S)))$ and this verifies the desired result. ∎

By relation (1.36) and Lemma 1.20 it follows immediately that

$$\text{ri}(S) \subseteq \text{ri}(\text{cl}(S)) \subseteq \text{ri}(\text{cl}(\text{co}(S))) \text{ and } \text{ri}(S) \subseteq \text{ri}(\text{co}(S)) \tag{1.37}$$

for arbitrary sets $S \subseteq R^n$. Moreover, since in Lemma 1.8 it is shown that

$$\text{aff}(S_1 \times S_2) = \text{aff}(S_1) \times \text{aff}(S_2)$$

for any nonempty sets $S_1 \subseteq R^n$ and $S_2 \subseteq R^m$and $E_{m+n} \subseteq E_m \times E_n$ with E_n denoting the n-dimensional Euclidean unit ball it is easy to verify that

$$\text{ri}(S_1 \times S_2) = \text{ri}(S_1) \times \text{ri}(S_2). \tag{1.38}$$

An alternative definition of a relative interior point which is needed to show that the relative interior operator is invariant when applied to a relatively open set is given by the next lemma.

Lemma 1.21. *If the set $S \subseteq R^n$ is regular then the vector $\mathbf{x}$ is a relative interior point of the set S if and only $\mathbf{x}$ belongs to aff(S) and there exists some $\epsilon > 0$ such that*

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

Proof. Since by assumption

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

with $\text{ri}(S)$ nonempty and $\mathbf{x}$ belongs to $\text{aff}(S)$ it follows immediately that $\mathbf{x} \in \text{ri}(S)$. To verify the reverse implication let $\mathbf{x}$ be a relative interior point of the set S. This means that the point $\mathbf{x}$ belongs to $\text{aff}(S)$ and there exists some $\epsilon > 0$ such that

$$(\mathbf{x} + \epsilon E) \cap \text{aff}(S) \subseteq S. \tag{1.39}$$

Since $\mathbf{x}$ belongs to $\text{aff}(S)$ we obtain that the intersection $(\mathbf{x} + \delta E) \cap \text{aff}(S)$ is nonempty for every $\delta > 0$. Consider now some arbitrary point $\mathbf{y}$ belonging to $(\mathbf{x} + \frac{\epsilon}{2}E) \cap \text{aff}(S)$. For this point $\mathbf{y}$ it follows that

$$\mathbf{y} + \frac{\epsilon}{2}E \subseteq \mathbf{x} + \frac{\epsilon}{2}E + \frac{\epsilon}{2}E = \mathbf{x} + \epsilon E$$

and applying relation (1.39) it is clear that

$$(1.40) \qquad (\mathbf{y}+\frac{\epsilon}{2}E) \cap \text{aff}(S) \subseteq (\mathbf{x}+\epsilon E) \cap \text{aff}(S) \subseteq S$$

Due to $\mathbf{y}$ also belongs to $\text{aff}(S)$ we obtain by relation (1.40) that $\mathbf{y}$ is a relative interior point of S and since $\mathbf{y}$ is arbitrary chosen this implies

$$(\mathbf{x}+\frac{\epsilon}{2}E) \cap \text{aff}(S) \subseteq \text{ri}(S)$$

and we have shown the desired result. ∎

In case we consider the rational numbers Q it is clear that the set Q is not regular and so there exist sets which are not regular. The next result shows that for regular sets $S \subseteq R^n$ the affine hull operation cannot distinguish the sets $\text{ri}(S)$ and S and so this lemma can be seen as an extension of Lemma 1.20.

Lemma 1.22. *If the set $S \subseteq R^n$ is regular then it follows that*

$$\textit{aff}(\textit{ri}(S)) = \textit{aff}(S).$$

Proof. By the monotonicity of the affine hull operator it is clear that

$$\text{aff}(\text{ri}(S)) \subseteq \text{aff}(S).$$

To show the converse inclusion we consider some $\mathbf{x} \in S$. Since the set S is regular one can find some $\mathbf{y} \in \text{ri}(S) \subseteq S$ and so by Lemma 1.21 there exists some $\epsilon > 0$ satisfying

$$(1.41) \qquad (\mathbf{y}+\epsilon E) \cap \text{aff}(S) \subseteq \text{ri}(S).$$

Clearly the set $[\mathbf{y},\mathbf{x}] := \{(1-\alpha)\mathbf{y}+\alpha\mathbf{x} : 0 \leq \alpha \leq 1\}$ belongs to $\text{co}(S) \subseteq \text{aff}(S)$ and this implies by relation (1.41) that

$$(\mathbf{y}+\epsilon E) \cap [\mathbf{y},\mathbf{x}] \subseteq \text{ri}(S).$$

This means that the halfline starting in $\mathbf{y}$ and passing through $\mathbf{x}_1 \in (\mathbf{y}+\epsilon E) \cap [\mathbf{y},\mathbf{x}]$ is contained in $\text{aff}(\text{ri}(S))$ and contains $\mathbf{x}$. Hence $\mathbf{x}$ belongs to $\text{aff}(\text{ri}(S))$ and we have shown that $S \subseteq \text{aff}(\text{ri}(S))$ This yields that $\text{aff}(S) \subseteq \text{aff}(\text{ri}(S))$ and the lemma has been verified. ∎

An immediate consequence of Lemma 1.22 and 1.21 is given by the observation that for any regular set $S \subseteq R^n$ it follows that $\mathbf{x}$ relative interior point of S if and only if $\mathbf{x}$ belongs to $\text{aff}(\text{ri}(S))$ and there exists some $\epsilon > 0$ satisfying $(\mathbf{x}+\epsilon E) \cap$ $\text{aff}(\text{ri}(S)) \subseteq \text{ri}(S)$. This implies for every regular set $S \subseteq R^n$ that

$$\text{ri}(\text{ri}(S)) = \text{ri}(S).$$

and since by definition $\text{ri}(\emptyset) = \emptyset$ implying that the above result is also correct for any nonregular set we obtain for any set S that

$$(1.42) \qquad \text{ri}(\text{ri}(S)) = \text{ri}(S).$$

Keeping in mind the close relationship between affine hulls and convex sets and the observation that affine manifolds are regular (in fact $\text{ri}(M) = M$!)

we might wonder whether convex sets are regular. This result indeed holds and to show this we introduce the class of convex hulls generated by a finite set of affinely independent vectors. Clearly these sets mostly "resemble" affine hulls.

Definition 1.15. *A k-simplex $S \subseteq R^n$ is the convex hull of $k+1$ affinely independent vectors $\mathbf{x}_0, ..., \mathbf{x}_k$ or equivalently $S = co(\{\mathbf{x}_0, ..., \mathbf{x}_k\})$.*

Since by Lemma 1.20 it follows that

$$\text{aff}(\text{co}(\{\mathbf{x}_0, ..., \mathbf{x}_k\})) = \text{aff}(\{\mathbf{x}_0, ..., \mathbf{x}_k\})$$

we obtain by the definition of the dimension of a set and Lemma 1.7 that for any k-simplex $S \subseteq R^n$ the dimension $\dim(S)$ is given by

$$\dim(S) = \dim(\text{lin}(\{\mathbf{x}_1 - \mathbf{x}_0, ..., \mathbf{x}_k - \mathbf{x}_0\})) = k \leq n. \tag{1.43}$$

If we do not want to stress the dimension we also refer to a k-simplex as a simplex. Observe the already encountered unit simplex $\Delta_{n+1} \subseteq R^{n+1}$ given by relation (1.30) has the representation

$$\Delta_{n+1} = \text{co}(\{\mathbf{e}_1, ..., \mathbf{e}_{n+1}\})$$

with $\mathbf{e}_i \in R^{n+1}$ denoting the ith unit vector,$1 \leq i \leq n+1$ and so this set is actually a n-simplex in R^{n+1}. For simplices the next result is geometrically obvious and so we will not give a proof of this result.

Lemma 1.23. *Any k-simplex $S \subseteq R^n$ given by $S = co(\{\mathbf{x}_0, ..., \mathbf{x}_k\})$ is regular and the set $ri(S)$ has the representation*

$$ri(S) = \{\sum_{i=0}^{k} \alpha_i \mathbf{x}_i : \sum_{i=0}^{k} \alpha_i = 1, \alpha_i > 0\}.$$

To show that any convex set C is regular we need to prove the following result.

Lemma 1.24. *For every nonempty convex set $C \subseteq R^n$ it follows that there exist a simplex $S_{\max}$ such that $S_{\max} \subseteq C$ and $aff(C) = aff(S_{\max})$.*

Proof. Since $C \subseteq R^n$ is nonempty it clearly contains the 0-simplex $\text{co}(\{\mathbf{x}\})$ for any $\mathbf{x} \in C$ and by relation (1.43) it will never contain a $(n+1)$-simplex. Hence it follows that

$$k_{\max} := \max\{k : \text{there exists a } k\text{-simplex } S \subseteq C\}$$

is well defined. For the selected $k_{\max}$-simplex $S_{\max} \subseteq C$ given by

$$S_{\max} = \text{co}(\{\mathbf{x}_0, ..., \mathbf{x}_{k_{\max}}\})$$

it clearly follows that $\text{aff}(S_{\max}) \subseteq \text{aff}(C)$. To show that the inclusion

$$\text{aff}(C) \subseteq \text{aff}(S_{\max})$$

holds it is sufficient to verify that $C \subseteq \text{aff}(S_{\max})$. To prove this we assume by contradiction that there exists some $\mathbf{x} \in C$ with the property that

$$\mathbf{x} \notin \text{aff}(S_{\max}) = \text{aff}(\{\mathbf{x}_0, ..., \mathbf{x}_{k_{\max}}\}). \tag{1.44}$$

If additionally the vectors $\mathbf{x}_0 - \mathbf{x}, ..., \mathbf{x}_{k_{\max}} - \mathbf{x}$ are linear dependent it must follow by definition that the vectors $\mathbf{x}, \mathbf{x}_0, ..., \mathbf{x}_{k_{\max}}$ are affinely dependent and this implies by Lemma 1.10 that the system

$$\beta\mathbf{x} + \sum_{i=0}^{k_{\max}} \alpha_i \mathbf{x}_i = \mathbf{0} \text{ and } \beta + \sum_{i=0}^{k_{\max}} \alpha_i = 0. \tag{1.45}$$

has a nonzero solution $(\beta^*, \alpha_1^*, ..., \alpha_{k_{\max}}^*)$. Since by asssumption the vectors $\mathbf{x}_0, ..., \mathbf{x}_{k_{\max}}$ are affinely independent it follows by contradiction and relation (1.45) that the scalar β^* is nonzero and hence we obtain that

$$\mathbf{x} = -\frac{1}{\beta} \sum_{i=0}^{k_{\max}} \alpha_i \mathbf{x}_i \text{ and } -\frac{1}{\beta} \sum_{i=0}^{k_{\max}} \alpha_i = 1.$$

Therefore the vector $\mathbf{x}$ belongs to $\text{aff}(\{\mathbf{x}_0, ..., \mathbf{x}_{k_{\max}}\})$ and this contradicts relation (1.44). Hence it must follow that the vectors $\mathbf{x}_0 - \mathbf{x}, ..., \mathbf{x}_{k_{\max}} - \mathbf{x}$ are linear independent and this yields that the simplex

$$S = \text{co}(\{\mathbf{x}, \mathbf{x}_0, ..., \mathbf{x}_{k_{\max}}\}) \subseteq C$$

has dimension $k_{\max} + 1$. Again we obtain a contradiction and so it follows that $C \subseteq \text{aff}(S_{\max})$ showing the desired result. ∎

The next important existence result is an immediate consequence of Lemma 1.23 and 1.24.

Lemma 1.25. *Every nonempty convex set $C \subseteq R^n$ is regular.*

We will now list some important properties of relative interiors. To start with this we first verify the following technical result.

Lemma 1.26. *If $S_1, S_2 \subseteq R^n$ are nonempty sets then it follows for every $0 < \alpha < 1$ that*

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

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 $\mathbf{y} \in \text{aff}(S_1)$. It is now necessary to verify that $\mathbf{x}_2$ belongs to $S_2 \cap \text{aff}(S_1)$. By the definition of $\mathbf{y}$ and $0 < \alpha < 1$ we obtain that

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

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

Applying now Lemma 1.26 the next important result for convex sets can be shown. This result will play an important role in the proof of the subsequent results.

Lemma 1.27. *If $C \subseteq R^n$ 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

$$\alpha \mathrm{cl}(C) + (1-\alpha)\mathbf{x}_2 \subseteq \mathrm{ri}(C)$$

for any fixed $\mathbf{x}_2 \in \mathrm{ri}(C)$ and $0 < \alpha < 1$. Clearly this set belongs to $\mathrm{aff}(C)$ and since $\mathbf{x}_2$ belongs to $\mathrm{ri}(C) \subseteq C$ there exists some $\epsilon > 0$ satisfying

$$(\mathbf{x}_2 + \frac{(1+\alpha)\epsilon}{1-\alpha}E) \cap \mathrm{aff}(C) \subseteq C. \tag{1.46}$$

Moreover, by Lemma 1.1 we know that

$$\mathrm{cl}(C) \subseteq C + \epsilon E.$$

and this implies

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

Hence by Lemma 1.26 it follows that

$$\begin{aligned} & (\alpha \mathrm{cl}(C) + (1-\alpha)\mathbf{x}_2 + \epsilon E) \cap \mathrm{aff}(C) \\ \subseteq \quad & \alpha C + (1-\alpha)\left((\mathbf{x}_2 + \frac{1+\alpha}{1-\alpha}\epsilon E) \cap \mathrm{aff}(C)\right) \end{aligned}$$

and this yields by relation (1.46) and the convexity of the set C that

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

Hence we have verified that

$$\alpha \mathrm{cl}(C) + (1-\alpha)\mathbf{x}_2 \subseteq \mathrm{ri}(C)$$

and this shows the result. ∎

By Lemma 1.25 and 1.27 it follows immediately for any nonempty convex set C that the set $\mathrm{ri}(C)$ is nonempty and convex. Also by Lemma 1.1 it is easy to verify that $\mathrm{cl}(C)$ is a convex set. An easy and important consequence of Lemma 1.27 is given by the observation that the relative interior operator cannot distinguish the convex sets C and $\mathrm{cl}(C)$. A similar observation applies to the closure operator applied to the convex sets $\mathrm{ri}(C)$ and C.

Lemma 1.28. *If $C \subseteq R^n$ 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 relation we only need to check that $\mathrm{cl}(C) \subseteq \mathrm{cl}(\mathrm{ri}(C))$. To verify this we consider some $\mathbf{x} \in \mathrm{cl}(C)$ and since $\mathrm{ri}(C)$ is nonempty we select some $\mathbf{y}$ belonging to $\mathrm{ri}(C)$. By Lemma 1.27 the half-open line segment $[\mathbf{y}, \mathbf{x})$ belongs to $\mathrm{ri}(C)$ and this implies that the vector $\mathbf{x}$ belongs to $\mathrm{cl}(\mathrm{ri}(C)$. Hence we have shown that

$$\mathrm{cl}(C) \subseteq \mathrm{cl}(\mathrm{ri}(C))$$

and the first equality is verified. To prove the second relation it follows immediately by relation (1.36) that $\mathrm{ri}(C) \subseteq \mathrm{ri}(\mathrm{cl}(C))$. To verify the inclusion

$\text{ri}(\text{cl}(C)) \subseteq \text{ri}(C)$ consider some arbitrary $\mathbf{x}$ belonging to $\text{ri}(\text{cl}(C))$ and so one can find some $\epsilon > 0$ satisfying

$$(\mathbf{x} + \epsilon E) \cap \text{aff}(\text{cl}(C)) \subseteq \text{cl}(C). \tag{1.47}$$

Moreover, due to $\text{ri}(C)$ is nonempty one can find some $\mathbf{y}$ belonging to $\text{ri}(C)$ and for this specific $\mathbf{y}$ construct the line $M := \{(1-t)\mathbf{x}+t\mathbf{y} : t \in R\}$ through the points $\mathbf{x}$ and $\mathbf{y}$. Since $\mathbf{x}$ belongs to $\text{ri}(\text{cl}(C)) \subseteq \text{cl}(C)$ and $\mathbf{y}$ belongs to $\text{ri}(C) \subseteq \text{cl}(C)$ it follows that $M \subseteq \text{aff}(\text{cl}(C))$ and so by relation (1.47) there exists some $\mu < 0$ satisfying

$$\mathbf{y}_1 := (1-\mu)\mathbf{x} + \mu\mathbf{y} \in \text{cl}(C).$$

By the definition of $\mathbf{y}_1$ it follows that

$$\mathbf{x} = \frac{1}{1-\mu}\mathbf{y}_1 - \frac{\mu}{1-\mu}\mathbf{y} \tag{1.48}$$

and since $\mathbf{y}_1$ belongs to $\text{cl}(C)$ and $\mathbf{y}$ belongs to $\text{ri}(C)$ this yields by Lemma 1.27 and relation (1.48) that $\mathbf{x} \in \text{ri}(C)$. Hence we have shown that

$$\text{ri}(\text{cl}(C)) \subseteq \text{ri}(C)$$

and this proves the second equality. ∎

In the above lemma one might wonder whether the convexity of the set C is necessary. In the following example we present a nonconvex regular set S with $\text{ri}(S)$ and $\text{cl}(S)$ convex and this set does not satisfy the result of Lemma 1.28.

Example 1.6. *Let $S \subseteq R$ be given by the set $S := [0,1] \cup ((1,2] \cap Q)$. This set is clearly not convex and ri$(S) = (0,1)$ while cl$(S) = [0,2]$. Moreover, by this observation we obtain immediately that ri(cl$(S)) \neq$ ri(S) and cl(ri$(S)) \neq$ cl(S).*

Looking at Example 1.6 it is possible to slightly weaken the assumption in Lemma 1.28 that C is a nonempty convex set.

Definition 1.16. *A nonempty set $S \subseteq R^n$ is called almost convex*[41] *if the set cl(S) is convex and ri(cl$(S)) \subseteq S$.*

It is now possible to prove the following version of Lemma 1.28 for almost convex sets. This result also serves as an alternative definition of an almost convex set.

Lemma 1.29. *For any nonempty set $S \subseteq R^n$ it follows that*

$$\begin{aligned} S \text{ is almost convex} &\Leftrightarrow \text{cl}(S) \text{ convex and } \text{ri}(\text{cl}(S)) = \text{ri}(S) \\ &\Leftrightarrow \text{ri}(S) \text{ convex and } \text{cl}(\text{ri}(S)) = \text{cl}(S). \end{aligned}$$

[41] almost convex set

Proof. If the nonempty set S is almost convex it follows by Lemma 1.25 that $\mathrm{cl}(S)$ is a regular set and so by Lemma 1.22 we obtain

$$\mathrm{aff}(\mathrm{ri}(\mathrm{cl}(S))) = \mathrm{aff}(\mathrm{cl}(S)) = \mathrm{aff}(S).$$

Since by definition $\mathrm{ri}(\mathrm{cl}(S)) \subseteq S$ this implies by relations (1.42) and (1.36) that

$$\mathrm{ri}(\mathrm{cl}(S)) = \mathrm{ri}(\mathrm{ri}(\mathrm{cl}(S))) \subseteq \mathrm{ri}(S)$$

and using relation (1.37) we obtain $\mathrm{ri}(\mathrm{cl}(S)) = \mathrm{ri}(S)$. To verify the second if-implication we observe that $\mathrm{cl}(S)$ convex implies $\mathrm{ri}(\mathrm{cl}(S))$ is convex and since $\mathrm{ri}(\mathrm{cl}(S)) = \mathrm{ri}(S)$ this yields that the set $\mathrm{ri}(S)$ is convex. Applying now Lemma 1.28 to the convex set $\mathrm{cl}(S)$ it follows

$$\mathrm{cl}(\mathrm{cl}(S)) = \mathrm{cl}(\mathrm{ri}(\mathrm{cl}(S))).$$

and using $\mathrm{ri}(\mathrm{cl}(S)) = \mathrm{ri}(S)$ we obtain

$$\mathrm{cl}(S) = \mathrm{cl}(\mathrm{cl}(S)) = \mathrm{cl}(\mathrm{ri}(\mathrm{cl}(S)) = \mathrm{cl}(\mathrm{ri}(S)).$$

To complete the proof we still need to show that $\mathrm{ri}(S)$ convex and $\mathrm{cl}(\mathrm{ri}(S)) = \mathrm{cl}(S)$ implies that the set S is almost convex. Since $\mathrm{ri}(S)$ is convex we obtain that $\mathrm{cl}(\mathrm{ri}(S))$ is convex and this yields using $\mathrm{cl}(\mathrm{ri}(S)) = \mathrm{cl}(S)$ that $\mathrm{cl}(S)$ is convex. Moreover, applying first Lemma 1.28 to the convex set $\mathrm{ri}(S)$ and relation (1.42) it follows that

$$\mathrm{ri}(\mathrm{cl}(\mathrm{ri}(S))) = \mathrm{ri}(\mathrm{ri}(S)) = \mathrm{ri}(S).$$

This implies using $\mathrm{cl}(S) = \mathrm{cl}(\mathrm{ri}(S))$ that

$$\mathrm{ri}(\mathrm{cl}(S)) = \mathrm{ri}(\mathrm{cl}(\mathrm{ri}(S))) = \mathrm{ri}(S) \subseteq S$$

and hence we have verified that the set S is almost convex. ∎

By Lemma 1.28 and 1.29 a convex set is almost convex and any nonempty almost convex set is regular. The following example presents a nonconvex set which is almost convex.

Example 1.7. *Let S denote a hypercube with some of the edges partly deleted. As an example we take $S = [0,1] \times [0,1] \backslash \{(1, x_2) : \frac{3}{4} < x_2 < 1\}$. Clearly this set is not convex but it is certainly almost convex.*

Applying now Lemma 1.29 and relation (1.42) it is possible to show the following improvement of Lemma 1.27.

Lemma 1.30. *If $S \subseteq R^n$ is a nonempty almost convex set then it follows for every $0 \leq \alpha < 1$ that*

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

Proof. Since S is a nonempty almost convex set we obtain by Lemma 1.29 and relation (1.42) that

$$\alpha \mathrm{cl}(S) + (1-\alpha)\mathrm{ri}(S) = \alpha \mathrm{cl}(\mathrm{ri}(S)) + (1-\alpha)\mathrm{ri}(\mathrm{ri}(S)) \tag{1.49}$$

for every $0 \leq \alpha < 1$. Due to $\text{ri}(S)$ is a nonempty convex set we may apply Lemma 1.27 to relation (1.49) and this implies

$$\alpha \text{cl}(S) + (1-\alpha)\text{ri}(S) \in \text{ri}(\text{ri}(S))$$

Using now relation (1.42) yields the desired result. ∎

Since for an almost convex set S it follows that $\text{ri}(\text{cl}(S)) = \text{ri}(S)$ it is possible to give an equivalent definition of a relative interior point of an almost convex set. This result is very useful in the proof of the weak separation theorem.

Lemma 1.31. *If the nonempty set $S \subseteq R^n$ is almost convex then if follows that $\mathbf{x} \in ri(S)$ if and only if $\mathbf{x} \in aff(S)$ and there exists some $\epsilon > 0$ such that*

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

An easy and important consequence of Lemma 1.29 is given by the following result.

Lemma 1.32. *If $S_1, S_2 \subseteq R^n$ are nonempty almost convex sets then it follows that*

$$cl(S_1) = cl(S_2) \Leftrightarrow ri(S_1) = ri(S_2) \Leftrightarrow ri(S_1) \subseteq S_2 \subseteq cl(S_1).$$

Proof. Since $\text{cl}(S_1) = \text{cl}(S_2)$ we obtain by Lemma 1.29 that

$$\text{ri}(S_1) = \text{ri}(\text{cl}(S_1)) = \text{ri}(\text{cl}(S_2)) = \text{ri}(S_2).$$

Moreover, if $\text{ri}(S_1) = \text{ri}(S_2)$ it follows immediately that $\text{ri}(S_1) \subseteq S_2$. Again by Lemma 1.29 we obtain

$$S_2 \subseteq \text{cl}(S_2) = \text{cl}(\text{ri}(S_2)) = \text{cl}(\text{ri}(S_1)) = \text{cl}(S_1).$$

To verify the last implication we observe in case $\text{ri}(S_1) \subseteq S_2 \subseteq \text{cl}(S_1)$ that again by Lemma 1.29

$$\text{cl}(S_1) = \text{cl}(\text{ri}(S_1)) \subseteq \text{cl}(S_2) \subseteq \text{cl}(S_1)$$

and this shows the desired result. ∎

We will now give a primal representation of the relative interior of an almost convex set S. This result will be used in the proof of the behaviour of the relative interior operator under affine mappings.

Lemma 1.33. *If $S \subseteq R^n$ is a nonempty almost convex set then it follows that*

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

Proof. We first verify the inclusion

$$\mathrm{ri}(S) \subseteq \{\mathbf{x} \in S : \forall_{\mathbf{y} \in \mathrm{cl}(S)}\ \exists_{\mu<0} \text{ such that } (1-\mu)\mathbf{x} + \mu\mathbf{y} \in S\}.$$

Let $\mathbf{x} \in \mathrm{ri}(S) \subseteq \mathrm{cl}(S)$ and consider an arbitrary $\mathbf{y} \in \mathrm{cl}(S)$. Clearly for any scalar t the vector $(1-t)\mathbf{x}+t\mathbf{y}$ belongs to $\mathrm{aff}(S)$. Also, since $\mathbf{x} \in \mathrm{ri}(S)$ there exists some $\epsilon > 0$ satisfying

$$(\mathbf{x} + \epsilon E) \cap \mathrm{aff}(S) \subseteq S \tag{1.50}$$

and so one can find some $\mu < 0$ such that

$$(1-\mu)\mathbf{x} + \mu\mathbf{y} \in \mathbf{x} + \epsilon E.$$

Applying now relation (1.50) we obtain that $(1-\mu)\mathbf{x} + \mu\mathbf{y}$ belongs to S and this shows the desired inclusion. To verify the desired result it is now sufficient to prove the inclusion

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

Consider now an arbitrary $\mathbf{x}$ belonging to the first set. By Lemma 1.29 we know that the set $\mathrm{ri}(S)$ is nonempty and so by our assumption there exists for a given $\mathbf{y} \in \mathrm{ri}(S) \subseteq \mathrm{cl}(S)$ some $\mu < 0$ satisfying

$$\mathbf{y}_1 := (1-\mu)\mathbf{x} + \mu\mathbf{y} \in S.$$

This yields by the definition of $\mathbf{y}_1$ that

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

and since $\mathbf{y} \in \mathrm{ri}(S)$ and $\mathbf{y}_1 \in S$ it follows by Lemma 1.30 that $\mathbf{x} \in \mathrm{ri}(S)$ showing the desired result. ∎

The above result is equivalent to the geometrically obvious fact that for S an almost convex set and any $\mathbf{x} \in \mathrm{ri}(S)$ and $\mathbf{y} \in S$ the line segment $[\mathbf{y}, \mathbf{x}]$ can be extended beyond $\mathbf{x}$ without leaving S. Also, by relation (1.42) and Lemma 1.29 another primal representation of $\mathrm{ri}(S)$ with S an almost convex set is given by

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

Since affine mappings preserve convexity it is also of interest to know how the relative interior operator behaves under affine mappings. This is discussed in the next result.

Lemma 1.34. *If $A : R^n \to R^m$ is an affine mapping and $C \subseteq R^n$ is a nonempty convex set then it follows that*

$$A(ri(C)) = ri(A(C)).$$

Moreover, if $C \subseteq R^m$ is a convex set satisfying

$$A^{-1}(ri(C)) := \{\mathbf{x} \in R^n : A(\mathbf{x}) \in ri(C)\}$$

is nonempty then

$$ri(A^{-1}(C)) = A^{-1}(ri(C)).$$

Proof. To verify the first equality we first observe by the continuity of the mapping A that for any set $S \subseteq R^n$ it must follow that

$$A(\mathrm{cl}(S)) \subseteq \mathrm{cl}(A(S)).$$

This shows in combination with Lemma 1.28 that

$$A(C) \subseteq A(\mathrm{cl}(C)) = A(\mathrm{cl}(\mathrm{ri}(C))) \subseteq \mathrm{cl}(A(\mathrm{ri}(C))) \subseteq \mathrm{cl}(A(C)$$

and taking the closure at both sides implies

$$\mathrm{cl}(A(C)) = \mathrm{cl}(A(\mathrm{ri}(C))) \tag{1.51}$$

Since $\mathrm{ri}(C)$ is a convex set it follows that the set $A(\mathrm{ri}(C))$ is convex and this implies by Lemma 1.32 with $S_1 := A(C)$ and $S_2 := A(\mathrm{ri}(C))$ and relation (1.51) that also

$$\mathrm{ri}(A(C)) = \mathrm{ri}(A(\mathrm{ri}(C)) \subseteq A(\mathrm{ri}(C)).$$

To prove the reverse inclusion

$$A(\mathrm{ri}(C)) \subseteq \mathrm{ri}(A(C))$$

consider an arbitrary $A(\mathbf{y})$ with $\mathbf{y} \in \mathrm{ri}(C)$. Since by Lemma 1.25 the set $\mathrm{ri}(A(C)) \subseteq A(C)$ is nonempty one can find some $\mathbf{y}_1 \in C$ satisfying $A(\mathbf{y}_1) \in \mathrm{ri}(A(C))$ and using this point $\mathbf{y}_1$ construct the line connecting $\mathbf{y}$ and $\mathbf{y}_1$. Since $\mathbf{y} \in \mathrm{ri}(C)$ and $\mathbf{y}_1 \in C$ there exists by Lemma 1.33 some $\mu < 0$ satisfying

$$\mathbf{y}_2 := (1-\mu)\mathbf{y} + \mu\mathbf{y}_1 \in C$$

Hence it follows that

$$\mathbf{y} = \frac{1}{1-\mu}\mathbf{y}_2 - \frac{\mu}{1-\mu}\mathbf{y}_1$$

and so

$$A(\mathbf{y}) = \frac{1}{1-\mu}A(\mathbf{y}_2) - \frac{\mu}{1-\mu}A(\mathbf{y}_1)$$

with $A(\mathbf{y}_2) \in A(C)$ and $A(\mathbf{y}_1) \in \mathrm{ri}(A(C))$. This implies by Lemma 1.27 that $A(\mathbf{y}) \in \mathrm{ri}(A(C))$ and this shows the first result. The other result can be proved similarly and so we omit it. ∎

Before showing the next result for almost convex sets we observe it is easy to verify that

$$A(\mathrm{cl}(S)) \subseteq \mathrm{cl}(A(S))$$

for any affine mapping $A : R^n \to R^m$ and $S \subseteq R^n$ an arbitrary nonempty set. Taking now closures at both sides it follows by the monotonicity of the closure operator that $\mathrm{cl}(A(\mathrm{cl}(S)) \subseteq \mathrm{cl}(\mathrm{cl}(A(S)))$ $=\mathrm{cl}(A(S))$ and since trivially $\mathrm{cl}(A(S)) \subseteq \mathrm{cl}(A(\mathrm{cl}(S))$ we obtain the equality

$$\mathrm{cl}(A(\mathrm{cl}(S))) = \mathrm{cl}(A(S)). \tag{1.52}$$

Using lemma 1.34 it is now possible to prove the following composition result for almost convex sets.

Lemma 1.35. *If $A : R^n \to R^m$ is an affine mapping and $S \subseteq R^n$ is a nonempty almost convex set then it follows that the set $A(S)$ is almost convex and*

$$A(ri(S)) = ri(A(S)).$$

Proof. To show that the set $A(S)$ is almost convex for S a nonempty almost convex set we observe that $\text{cl}(S)$ is convex and hence by relation (1.52) the set $\text{cl}(A(S))$ is convex. Moreover, by the same relation and Lemma 1.28 applied to the set $A(\text{cl}(S))$ it follows that

$$\text{ri}(\text{cl}(A(S))) = \text{ri}(\text{cl}(A(\text{cl}(S)))) = \text{ri}(A(\text{cl}(S))).$$

Since by Lemma 1.34 and $\text{cl}(S)$ is a convex set we know that $\text{ri}(A(\text{cl}(S))) = A(\text{ri}(\text{cl}(S)))$ this implies in combination with Lemma 1.29 that

$$\text{ri}(\text{cl}(A(S))) = A(\text{ri}(\text{cl}(S))) = \text{ri}(A(\text{cl}(S))) = A(\text{ri}(S)). \tag{1.53}$$

Applying now Lemma 1.29 shows that the set $A(S)$ is almost convex. To check the second part we observe by Lemma 1.29 and relation (1.53) that

$$\text{ri}(A(S)) = \text{ri}(\text{cl}(A(S)) = A(\text{ri}(S))$$

and this shows the desired result. ∎

By Lemma 1.35 we obtain for any nonempty almost convex cone $K \subseteq R^n$ and every $\alpha > 0$ that

$$\alpha\text{ri}(K) = \text{ri}(\alpha K) \subseteq \text{ri}(K) \tag{1.54}$$

and this implies by Lemma 1.27 that $\text{ri}(K)$ is a nonempty convex cone for K an almost convex cone. Applying relation (1.54) and Lemma 1.30 we also obtain for any nonempty almost convex cone K that

$$\text{cl}(K) + \text{ri}(K) = 2(\frac{1}{2}\text{cl}(K) + \frac{1}{2}\text{ri}(K)) \subseteq 2\text{ri}(K) \subseteq \text{ri}(K). \tag{1.55}$$

An immediate consequence of Lemma 1.35 and relation (1.38) is given by the following result.

Lemma 1.36. *If the nonempty sets $S_i \subseteq R^n, i = 1, 2$ are almost convex and α, β some scalars then it follows that*

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

Proof. Introduce the linear mapping $A : R^{2n} \to R^n$ given by $A(\mathbf{x}, \mathbf{y}) = \alpha\mathbf{x} + \beta\mathbf{y}$. Applying this mapping to Lemma 1.35 with S replaced by $S_1 \times S_2$ and using relation (1.38) it follows that

$$\begin{aligned} \text{ri}(\alpha S_1 + \beta S_2) &= \text{ri}(A(S_1 \times S_2)) = A(\text{ri}(S_1 \times S_2)) \\ &= A(\text{ri}(S_1) \times \text{ri}(S_2)) = \alpha\text{ri}(S_1) + \beta\text{ri}(S_2) \end{aligned}$$

and this shows the desired result. ∎

We might now wonder whether the nonempty intersection of almost convex sets is again almost convex. In the next example we show that in general this is not the case.

Example 1.8. *Let the sets S_1 and $S_2 \subseteq R^2$ be given by*

$$S_1 := [1,2] \times [0,1] \text{ and } S_2 := [0,1] \times [0,1] \backslash \{(1,x_2) : \frac{1}{4} < x_2 < \frac{3}{4}\}.$$

Both sets are almost convex and their nonempty intersection is given by the set

$$S_1 \cap S_2 = \{(1,x_2) : 0 \leq x_2 \leq \frac{1}{4} \text{ or } \frac{3}{4} \leq x_2 \leq 1\}.$$

which is not almost convex.

In the next lemma we introduce an additional condition which guarantees that the intersection of almost convex sets is again almost convex. At the same time it shows how almost convex sets and their closures and relative interiors behave under intersections.

Lemma 1.37. *If the sets $S_1, S_2 \subseteq R^n$ are almost convex and $ri(S_1) \cap ri(S_2)$ is nonempty then the set $S_1 \cap S_2$ is almost convex. Moreover, it follows that*

$$cl(S_1 \cap S_2) = cl(S_1) \cap cl(S_2) \text{ and } ri(S_1 \cap S_2) = ri(S_1) \cap ri(S_2).$$

Proof. We first show that $\text{cl}(S_1 \cap S_2) = \text{cl}(S_1) \cap \text{cl}(S_2)$. Since it is clear that $\text{cl}(S_1 \cap S_2) \subseteq \text{cl}(S_i)$ for every $i = 1,2$ we obtain

$$\text{cl}(S_1 \cap S_2) \subseteq \text{cl}(S_1) \cap \text{cl}(C_2)$$

To verify the other inclusion let $\mathbf{x} \in \text{cl}(S_1) \cap \text{cl}(S_2)$ and consider some $\mathbf{y} \in \text{ri}(S_1) \cap \text{ri}(S_2)$. For every $i = 1,2$ it follows by Lemma 1.30 that the half-open linesegment $[\mathbf{y}, \mathbf{x})$ belongs to $\text{ri}(S_i)$ for $i = 1,2$ and hence $[\mathbf{y}, \mathbf{x})$ belongs to $\text{ri}(S_1) \cap \text{ri}(S_2)$.This implies

$$\mathbf{x} \in \text{cl}(\text{ri}(S_1) \cap \text{ri}(S_2)) \subseteq \text{cl}(S_1 \cap S_2)$$

and the first equality is proved. To verify that the intersection $S_1 \cap S_2$ is almost convex we observe by the previous part and $\text{cl}(S_i)$ convex that $\text{cl}(S_1 \cap S_2)$ is a nonempty convex set. Moreover, since $\text{ri}(S_i), i = 1,2$ is also a nonempty convex set and $\text{ri}(\text{ri}(S_i)) = \text{ri}(S_i)$ it follows by the previous equality and Lemma 1.29 that

$$\text{(1.56)} \qquad \begin{aligned} \text{cl}(S_1 \cap S_2) &= \text{cl}(S_1) \cap \text{cl}(S_2) = \text{cl}(\text{ri}(S_1)) \cap \text{cl}(\text{ri}(S_2)) \\ &= \text{cl}(\text{ri}(S_1) \cap \text{ri}(S_2)) \end{aligned}$$

Since the intersection $\text{ri}(S_1) \cap \text{ri}(S_2)$ is convex for $S_i, i = 1,2$ almost convex this implies by Lemma 1.28 that

$$\text{ri}(\text{cl}(S_1 \cap S_2)) = \text{ri}(\text{cl}(\text{ri}(S_1) \cap \text{ri}(S_2))) = \text{ri}(\text{ri}(S_1) \cap \text{ri}(S_2))$$

and so this set is contained in $S_1 \cap S_2$. This shows by Lemma 1.29 that the set $S_1 \cap S_2$ is nonempty and almost convex. To verify that $\text{ri}(S_1 \cap S_2) = \text{ri}(S_1) \cap \text{ri}(S_2)$ we obtain by relation (1.56) and Lemma 1.32 with S_1 replaced by $S_1 \cap S_2$ and S_2 by $\text{ri}(S_1) \cap \text{ri}(S_2)$ that

$$\text{ri}(S_1 \cap S_2) \subseteq \text{ri}(S_1) \cap \text{ri}(S_2).$$

To verify the reverse inclusion let $\mathbf{x}$ belong to $\mathrm{ri}(S_1)\cap\mathrm{ri}(S_2)$. Since we know that $S_1\cap S_2$ is a nonempty almost convex set it follows by Lemma 1.31 that the set $\mathrm{ri}(S_1\cap S_2)$ is nonempty. Consider now an arbitrary $\mathbf{y}\in\mathrm{ri}(S_1\cap S_2)\subseteq S_1\cap S_2$. Using this vector $\mathbf{y}$ and $\mathbf{x}\in\mathrm{ri}(S_i), i=1,2$ it follows by Lemma 1.33 that there exists some $\mu<0$ such that

$$\mathbf{y}_1 := (1-\mu)\mathbf{x}+\mu\mathbf{y}\in S_i, i=1,2,$$

and so $\mathbf{y}_1$ belongs to $S_1\cap S_2$. By the definition of the vector $\mathbf{y}_1$ we obtain that

$$\mathbf{x}=\frac{1}{1-\mu}\mathbf{y}_1-\frac{\mu}{1-\mu}\mathbf{y}.$$

and since $\mathbf{y}\in\mathrm{ri}(S_1\cap S_2)$ and $\mathbf{y}_1\in S_1\cap S_2$ it follows by Lemma 1.32 that the point $\mathbf{x}$ belongs to $\mathrm{ri}(S_1\cap S_2)$. Hence we have shown that

$$\mathrm{ri}(S_1)\cap\mathrm{ri}(S_2)\subseteq\mathrm{ri}(S_1\cap S_2)$$

and this proves the desired result. ∎

Observe by a similar proof as in Lemma 1.37 one can verify for S_i, $i\in I$ almost convex and $\cap_{i\in I}\mathrm{ri}(S_i)$ is nonempty that

$$\mathrm{cl}(\cap_{i\in I}S_i)=\cap_{i\in I}\mathrm{cl}(S_i).$$

Moreover, if the set I is finite, it also follows that

$$\mathrm{ri}\,(\cap_{i\in I}S_i)=\cap_{i\in I}\mathrm{ri}(S_i).$$

Moreover, the proof breaks down for the last case if I is not finite. Similarly it is necessary to assume that $\mathrm{ri}(S_1)\cap\mathrm{ri}(S_2)$.is nonempty and all this is shown by means of the following counterexample for convex sets C_i.

Example 1.9.

1. As a counterexample we mention for $C_\alpha\subseteq R$ given by

   $$C_\alpha=[0,1+\alpha],\alpha>0$$

   that

   $$\mathrm{ri}(\cap_{\alpha>0}C_\alpha)=\mathrm{ri}([0,1])=(0,1)$$

   Moreover, since for each $\alpha>0$ it follows that $\mathrm{ri}(C_\alpha)=(0,1+\alpha)$ we obtain that $\cap_{\alpha>0}ri(C_\alpha)=(0,1]$.
2. To show that $\mathrm{ri}(C_1)\cap\mathrm{ri}(C_2)$ should be nonempty in Lemma 1.37 we consider the following example. Let

   $$C_1=\{\mathbf{x}\in R^2:x_1>0,x_2>0\}\cup\{\mathbf{0}\}$$

   and

   $$C_2=\{\mathbf{x}\in R^2:x_2=0\}$$

   Clearly we obtain that

   $$\mathrm{ri}(C_1)=\{\mathbf{x}:x_1>0,x_2>0\}\text{ and }\mathrm{ri}(C_2)=C_2$$

and for these two sets

$$\text{ri}(C_1) \cap \text{ri}(C_2) = \varnothing \text{ and } \text{ri}(C_1 \cap C_2) \neq \text{ri}(C_1) \cap \text{ri}(C_2).$$

Also it is easy to see that $\text{cl}(C_1 \cap C_2) \neq \text{cl}(C_1) \cap \text{cl}(C_2)$.

This last example concludes our discussion of sets and hull operations. In the next section we will discuss in detail functions, their relations with sets and hull operations.

1.2. **Functions and Hull operations.** In this section we will discuss the interplay between extended real valued functions and sets. To start with this discussion let $f : R^n \to [-\infty, \infty]$ be an extended real valued function and associate with f its so-called *epigraph*[42] $\text{epi}(f)$ given by

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

A related set is given by the strict *epigraph* $\text{epi}_{\text{S}}(f)$ given by

$$\text{epi}_{\text{S}}(f) := \{\{(\mathbf{x}, r) \in R^{n+1} : f(\mathbf{x}) < r\} \subseteq R^{n+1}.$$

Since in the previous section we considered sets and their properties it would be advantageous to relate functions to sets and use the properties of sets to derive properties of functions. Within convex and quasiconvex analysis there are essentially two ways to do this. The first way to achieve this is to observe that another primal representation of the function f is given by the obvious relation (cf.[18])

$$f(\mathbf{x}) = \inf\{r : (\mathbf{x}, r) \in \text{epi}(f)\}. \tag{1.57}$$

Observe by definition we set $\inf\{\varnothing\} = \infty$ and this does only happen if the vector $\mathbf{x}$ does not belong to the so-called *effective domain*[43] $\text{dom}(f)$ of the function f given by

$$\text{dom}(f) := \{\mathbf{x} \in R^n : f(\mathbf{x}) < \infty\}.$$

Moreover, for $\text{dom}(f)$ nonempty we obtain that $\text{dom}(f) = A(\text{epi}(f))$ with A the projection of R^{n+1} onto R^n given by $A(\mathbf{x}, r) = \mathbf{x}$. The representation of the function f given by relation (1.57) is especially useful in the study of convex functions since convexity of a function f means by definition that the epigraph $\text{epi}(f)$ of f is a convex set. In quasiconvex analysis another representation is useful. To introduce this we define the so-called *lower-level set*[44] $L(f, r), r \in R$ of a function f given by

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

A related set is given by the *strict lower-level set*[45] $L_{\text{S}}(f, r)$ given by

$$L_{\text{S}}(f, r) := \{\mathbf{x} \in R^n : f(\mathbf{x}) < r\}.$$

[42] epigraph of a function
[43] effective domain of a function f
[44] lower-level set of a function f
[45] strict lower-level set of a function f

As observed by Crouzeix (cf.[11]) another primal represention of the function f is given by

$$f(\mathbf{x}) = \inf\{r : \mathbf{x} \in L(f,r)\}. \tag{1.58}$$

This representation is useful in the study of quasiconvex functions since quasiconvexity of a function f means by definition that the lower-level sets are convex. It is easy to verify (make a picture!) for every $r \in R$ that

$$\text{epi}(f) \cap (R^n \times \{r\}) = L(f,r) \times \{r\}. \tag{1.59}$$

and this relation immediately shows that a convex function is also a quasiconvex function. Before discussing hull operations on functions we introduce the class of lower semicontinuous functions.

Definition 1.17. *Let $f : R^n \to [-\infty, \infty]$ be some extended real valued function. The function f is called lower semicontinuous at $\mathbf{x} \in R^n$ if*

$$\liminf_{\mathbf{y} \to \mathbf{x}} f(\mathbf{y}) = f(\mathbf{x})$$

with liminf $_{\mathbf{y}\to\mathbf{x}} f(\mathbf{y})$ given by

$$\lim_{\epsilon > 0} \inf\{f(\mathbf{y}) : \mathbf{y} \in \mathbf{x} + \epsilon E\} := \sup_{\epsilon > 0} \inf\{f(\mathbf{y}) : \mathbf{y} \in \mathbf{x} + \epsilon E\}. \tag{1.60}$$

Moreover, the function $f : R^n \to [-\infty, \infty]$ is called upper semicontinuous at $\mathbf{x} \in R^n$ 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 $\mathbf{x}$. Finally the function $f : R^n \to [-\infty, \infty]$ is called lower semicontinuous[46] *(upper semicontinuous)*[47] *if f is lower semicontinuous (upper semicontinuous) at every $\mathbf{x} \in R^n$ and it is called continuous* [48] *if it is both upper semicontinuous and lower semicontinuous.*

Observe we sometimes abbrevate lower semicontinuous to l.s.c. To relate the above definition of liminf to the liminf of a sequence we observe that the familiar liminf of a sequence is defined by

$$\liminf_{k \uparrow \infty} f(\mathbf{y}_k) := \lim_{n \uparrow \infty} \inf_{k \geq n} f(\mathbf{y}_k)$$

and using this definition one can easily show the following result.

Lemma 1.38. *The function $f : R^n \to [-\infty, \infty]$ is lower semicontinuous at $\mathbf{x} \in R^n$ if and only if for every sequence $\{\mathbf{y}_k : k \in N\}$ with limit $\mathbf{x} \in R^n$ it follows that*

$$\liminf_{k \uparrow \infty} f(\mathbf{y}_k) \geq f(\mathbf{x}).$$

[46] lower semicontinuous function
[47] upper semicontinuous function
[48] continuous function

Proof. Clearly for every sequence $\{\mathbf{y}_k : k \in N\}$ with limit $\mathbf{x}$ and f lower semicontinuous at $\mathbf{x}$ we obtain that

$$f(\mathbf{x}) = \liminf_{\mathbf{y}\to\mathbf{x}} f(\mathbf{y}) \leq \liminf_{k\uparrow\infty} f(\mathbf{y}_k).$$

To show the reverse implication it follows by the definition of liminf as given by relation (1.60) that there exists some sequence $\{\mathbf{y}_k : k \in N\}$ with limit $\mathbf{x} \in R^n$ satisfying

$$\liminf_{\mathbf{y}\to\mathbf{x}} f(\mathbf{y}) = \liminf_{k\uparrow\infty} f(\mathbf{y}_k)$$

and this yields by our assumption that

$$\liminf_{\mathbf{y}\to\mathbf{x}} f(\mathbf{y}) \geq f(\mathbf{x}).$$

By relation (1.60) it is clear that

$$\liminf_{\mathbf{y}\to\mathbf{x}} f(\mathbf{y}) \leq f(\mathbf{x})$$

and this shows the result. ∎

The following result gives an important characterisation of lower semicontinuity using the *lower-level set* or the *epigraph* of a function (cf.[18]),([1])).

Theorem 1.39. *If $f : R^n \to [-\infty, \infty]$ is an extended real valued function then it follows that*

$$f \text{ l.s.c.} \Leftrightarrow \text{epi}(f) \text{ closed} \Leftrightarrow L(f,r) \text{ closed for every } r \in R.$$

Proof. Clearly the above conditions are equivalent if the function f is identically ∞ and so we assume that there exists some $\mathbf{x} \in R^n$ satisfying $f(\mathbf{x}) < \infty$ or equivalently dom(f) is nonempty. To prove that f is lower semicontinuous implies epi(f) is closed we need to check that the set epi(f) satisfies the second part of Lemma 1.1. Consider therefore an arbitrary sequence $\{(\mathbf{x}_k, r_k) : k \in N\} \subseteq \text{epi}(f)$ converging to $(\mathbf{x}, r) \in R^{n+1}$. Since by definition $f(\mathbf{x}_k) \leq r_k$ this implies by the lower semicontinuity of the function f and Lemma 1.38 that

$$\infty > r = \lim_{k\uparrow\infty} r_k \geq \liminf_{k\uparrow\infty} f(\mathbf{x}_k) \geq f(\mathbf{x})$$

and this shows that $(\mathbf{x}, r)$ belongs to epi(f). By Lemma 1.1 it follows now that epi(f) is closed. To verify epi(f) closed implies that $L(f,r)$ is closed for every $r \in R$ we obtain by relation (1.59) and $R^n \times \{r\}$ is a closed subset of R^{n+1} that the set $L(f,r) \times \{r\}$ is closed and this implies that $L(f,r)$ is closed. Finally we need to check that $L(f,r)$ closed for every $r \in R$ implies that f is lower semicontinuous. By the definition of liminf as given by relation (1.60) it follows with

$$\beta := \liminf_{\mathbf{y}\to\mathbf{x}} f(\mathbf{y}) \tag{1.61}$$

that $\beta \leq f(\mathbf{x})$ and so we need to prove that $\beta \geq f(\mathbf{x})$. Without loss of generality we may now assume that $f(\mathbf{x}) > -\infty$ and $-\infty \leq \beta < \infty$. Suppose now by contradiction that $\beta < f(\mathbf{x})$. Since by assumption $f(\mathbf{x}) > -\infty$ and $-\infty \leq \beta < \infty$ there exists some finite constant c satisfying

$$\beta < c < f(\mathbf{x}). \tag{1.62}$$

This implies by relation (1.61) that one can find a sequence $\mathbf{y}_k, k \in N$ with limit $\mathbf{x} \in R^n$ and $f(\mathbf{y}_k) \leq c$. Hence it follows that $\mathbf{y}_k$ belongs to $L(f, c)$ and since by assumption $L(f, c)$ is closed we obtain by Lemma 1.1 that $\mathbf{x} \in \mathcal{L}_f(c)$ or equivalently $f(\mathbf{x}) \leq c$. This contradicts relation (1.62) and so it follows that f is lower semicontinuous at $\mathbf{x}$ for any $\mathbf{x} \in R^n$. ∎

Since lower semicontinuity is an important property to solve minimization problems it is useful to know under which operations lower semicontinuity is preserved. It is easy to verify for a collection of functions $f_i, i \in I$ that

$$\text{epi}(\sup_{i \in I} f_i) = \cap_{i \in I} \text{epi}(f_i) \tag{1.63}$$

and since the intersection of closed sets is again a closed set this implies by Theorem 1.39 that the function $\sup_{i \in I} f_i$ is lower semicontinuous if each function f_i is lower semicontinuous. It is also easy to see that

$$I \text{ finite } \Rightarrow \text{epi}(\min_{i \in I} f_i) = \cup_{i \in I} \text{epi}(f_i) \tag{1.64}$$

and this shows since the finite union of closed sets is closed again that the function $\min_{i \in I} f_i$ is lower semicontinuous if each f_i is lower semicontinuous. Finally we observe for functions $f_i, i = 1, 2$ with $\text{liminf}_{\mathbf{y}_k \to \mathbf{x}} f_i(\mathbf{y}_k) > -\infty$ that

$$\lim \inf_{\mathbf{y}_k \to \mathbf{x}} (\alpha f_1 + \beta f_2)(\mathbf{y}_k) \geq \alpha \lim \inf_{\mathbf{y}_k \to \mathbf{x}} f_1(\mathbf{y}_k) + \beta \lim \inf_{\mathbf{y}_k \to \mathbf{x}} f_2(\mathbf{y}_k)$$

for every $\alpha, \beta > 0$. This shows by Lemma 1.38 that every strict canonical combination of the lower semicontinuous functions $f_i, i = 1, 2$ with $f_i > -\infty$ is again lower semicontinuous. Observe the assumption $f_i > -\infty$ is included to avoid undefined expressions like $\infty - \infty$. We will now show by means of the next result known as the *Weierstrass-Lebesgue Theorem* why lower semicontinuity in combination with compactness is a useful property for minimization problems.

Theorem 1.40. *If the function $f : R^n \to [-\infty, \infty]$ is lower semicontinuous and $S \subseteq R^n$ is a nonempty compact set then it follows that the optimization problem*

$$\inf\{f(\mathbf{x}) : \mathbf{x} \in S\} \tag{P}$$

has an optimal solution.

Proof. If we denote the optimal objective value of optimization problem (P) by $v(P)$ then clearly the result holds whenever there exists some $\mathbf{x} \in S$ satisfying $f(\mathbf{x}) = -\infty$ and so without loss of generality we may assume that $f(\mathbf{x}) > -\infty$ for every $\mathbf{x} \in S$. By Theorem 1.39 the decreasing sequence of upper-level sets

$$U_n := \{\mathbf{x} \in R^n : f(\mathbf{x}) > n\}, n \in Z$$

are open and since $f > -\infty$ on S the collection $\{U_n : n \in Z\}$ forms an open cover of S. By the compactness of S there exists a finite subcover and since $U_{n+1} \subseteq U_n$ for every $n \in Z$ this implies that one can find some $m \in Z$ satisfying $S \subseteq U_m$. Hence the function f is uniformly bounded from below on S and therefore $v(P) > -\infty$. If we assume by contradiction that $f(\mathbf{x}) > v(P)$ for every $\mathbf{x} \in S$ then clearly the collection $\{U_{v(P)+\frac{1}{n}} : n \in N\}$ of open sets is an open cover of S and again by the compactness of S there exists a finite subcover. Since $U_{v(P)+\frac{1}{n}} \subseteq U_{v(P)+\frac{1}{n+1}}$ for every $n \in N$ this implies that there exists some $m \in N$ satisfying $S \subseteq U_{v(P)+\frac{1}{m}}$ and so we obtain a contradiction with the definition of $v(P)$. Hence it must follow that there exists some $\mathbf{x} \in S$ with $f(\mathbf{x}) = v(P)$ and this shows the result. ∎

An application of Theorems 1.39 and 1.40 is given by the next useful preservation property for lower semicontinuous functions.

Lemma 1.41. *If the function $F : R^{m+n} \to [-\infty, \infty]$ is lower semicontinuous and $S \subseteq R^m$ is a nonempty compact set that it follows that the function $p : R^n \to [-\infty, \infty]$ given by*

$$p(\mathbf{y}) = \inf\{F(\mathbf{x}, \mathbf{y}) : \mathbf{x} \in S\}$$

is a lower semicontinuous function.

Proof. Let $r \in R$ and consider a sequence $\{\mathbf{y}_n : n \in N\} \subseteq L(p, r)$ satisfying $\lim_{n\to\infty} \mathbf{y}_n = \mathbf{y}_\infty$. By Theorem 1.40 there exist for every $\mathbf{y}_n$ some $\mathbf{x}_n \in S$ satisfying

$$F(\mathbf{x}_n, \mathbf{y}_n) = p(\mathbf{y}_n) \le r.$$

Since S is compact one can find by Lemma 1.3 a subsequence $\mathbf{x}_n, n \in N_0$ converging to $\mathbf{x}_\infty \in S$ and this implies by the lower semicontinuity of F and Lemma 1.38 that

$$r \ge \lim \inf_{n \in N_0 \uparrow \infty} p(\mathbf{y}_n) = \lim \inf_{n \in N_0 \uparrow \infty} F(\mathbf{x}_n, \mathbf{y}_n) \ge F(\mathbf{x}_\infty, \mathbf{y}_\infty).$$

Hence it follows that $r \ge p(\mathbf{y}_\infty)$ and this shows that $L(p, r)$ is closed. Applying now Theorem 1.39 yields the desired result. ∎

This concludes our discussion of lower semicontinuous functions. We will now introduce the definition of a convex and almost convex function.

Definition 1.18. *The function $f : R^n \to [-\infty, \infty]$ is called convex or a convex function*[49] *if epi(f) is a convex set and it is called almost convex or an almost convex function if epi(f) is an almost convex set. Moreover, the function $f : R^n \to [-\infty, \infty]$ is called positively homogeneous or a positively homogeneous function*[50] *if epi(f) is a cone.*

Using the definition of a cone and an epigraph it is easy to prove the following result.

[49] convex function

[50] positively homogeneous function

Lemma 1.42. *The function $f : R^n \to [-\infty, \infty]$ is positively homogeneous if and only if $f(\alpha\mathbf{x}) = \alpha f(\mathbf{x})$ for every $\mathbf{x} \in R^n$ and $\alpha > 0$.*

Proof. Since epi(f) is a cone we obtain for every $\mathbf{x}$ belonging to R^n and satisfying $f(\mathbf{x}) = -\infty$ that $(\alpha\mathbf{x}, \alpha r)$ belongs to epi(f) for every $r \in R$ and this implies by the definition of an epigraph that $f(\alpha\mathbf{x}) = -\infty = \alpha f(\mathbf{x})$. Moreover, if $\mathbf{x}$ belongs to dom(f) and $f(\mathbf{x}) > -\infty$ we obtain that $(\alpha\mathbf{x}, \alpha f(\mathbf{x}))$ belongs to epi(f) and hence we obtain $f(\alpha\mathbf{x}) \leq \alpha f(\mathbf{x}) < \infty$. At the same time it follows using $(\alpha\mathbf{x}, f(\alpha\mathbf{x})$ belongs to epi(f) that also $(\mathbf{x}, \alpha^{-1} f(\alpha\mathbf{x}))$ belongs to epi(f) and so $f(\mathbf{x}) \leq \alpha^{-1} f(\alpha\mathbf{x})$ or equivalently $f(\alpha\mathbf{x}) \geq \alpha f(\mathbf{x})$. This verifies $f(\alpha\mathbf{x}) = \alpha f(\mathbf{x})$ for $f(\mathbf{x})$ finite and since it is easy to show for $f(\mathbf{x}) = \infty$ that $f(\alpha\mathbf{x}) = \infty$ we have verified the above equality. To prove the reverse implication is trivial and so we omit it. ∎

Again by the special structure of an epigraph an equivalent definition of a convex function is given in the next result.

Lemma 1.43. *A function $f : R^n \to [-\infty, \infty]$ is convex $\Leftrightarrow$ epi$_S(f)$ is a convex set.*

Proof. To show that epi$_S(f)$ is a convex set whenever f is a convex function let $(\mathbf{x}_i, r_i), i = 1, 2$ belong to epi$_s(f)$. This implies that there exists some constants β_i satisfying

$$f(\mathbf{x}_i) \leq \beta_i < r_i, i = 1, 2. \tag{1.65}$$

Since epi(f) is a convex set we obtain by relation (1.65) that

$$(\alpha\mathbf{x}_1 + (1-\alpha)\mathbf{x}_2, \alpha\beta_1 + (1-\alpha)\beta_2) \in \text{epi}(f)$$

for every $0 < \alpha < 1$ and so by applying again relation (1.65) it follows that

$$f(\alpha\mathbf{x}_1 + (1-\alpha)\mathbf{x}_2) \leq \alpha\beta_1 + (1-\alpha)\beta_2 < \alpha r_1 + (1-\alpha) r_2$$

Hence epi$_S(f)$ is a convex set and to prove the reverse implication consider $(\mathbf{x}_i, r_i) \in$ epi(f). Clearly for every $\epsilon > 0$ it follows that $(\mathbf{x}_i, r_i + \epsilon) \in$ epi$_S(f)$. Hence by the convexity of the set epi$_S(f)$ we obtain that

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

for every $0 < \alpha < 1$ and letting $\epsilon \downarrow 0$ yields

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

Hence epi(f) is a convex set and the result is verified. ∎

An equivalent representation of Lemma 1.43 is given by the observation (cf.[18]) that a function $f : R^n \to [-\infty, \infty]$ is convex if and only if

$$f(\alpha\mathbf{x}_1 + (1-\alpha)\mathbf{x}_2) < \alpha\mu_1 + (1-\alpha)\mu_2 \tag{1.66}$$

whenever $f(\mathbf{x}_i) < \mu_i \in R$. In case we know additionally that $f > -\infty$ we obtain by relation (1.57) that f is convex if and only if

$$f(\alpha\mathbf{x}_1 + (1-\alpha)\mathbf{x}_2) \leq \alpha f(\mathbf{x}_1) + (1-\alpha) f(\mathbf{x}_2) \tag{1.67}$$

and so we recover the more familiar definition of a convex function. In case we are considering a function $f > -\infty$ and in relation (1.67) the inequality sign can be replaced by a strict inequality sign for every $\mathbf{x}_1 \neq \mathbf{x}_2$ then the function f is called strictly convex or a *strictly convex function*[51]. Necessarily this function must have a nonempty effective domain $\text{dom}(f)$ and a function with $f > \infty$ and $\text{dom}(f)$ nonempty is called *proper*[52].

As for lower semicontinuous functions one is interested under which operations convexity is preserved. Applying relation (1.63) and using that the intersection of convex sets preserves convexity it follows immediately that the function $\sup_{i \in I} f_i$ is convex if f_i is convex for every $i \in I$. Moreover, by relation (1.66) we obtain that any strict canonical combination of the convex functions $f_i, i = 1, 2$ with f_1 proper is again convex. Finally we consider the function $p : R^m \to [-\infty, \infty]$ defined by

$$p(\mathbf{y}) = \inf\{F(\mathbf{x}, \mathbf{y}) : \mathbf{x} \in C\} \tag{1.68}$$

with $F : R^{n+m} \to [-\infty, \infty]$ an extended real valued function and $C \subseteq R^n$ a nonempty convex set. For this function it follows that

$$\text{epi}_S(p) = \{(\mathbf{y}, r) \in R^{m+1} : \exists_{\mathbf{x} \in C} \text{ satisfying } (\mathbf{x}, \mathbf{y}, r) \in \text{epi}_S(F)\}.$$

This implies with $A : R^{n+m+1} \to R^{m+1}$ denoting the projection of R^{n+m+1} onto R^{m+1} represented by $A(\mathbf{x}, \mathbf{y}, r) := (\mathbf{y}, r)$ that

$$\text{epi}_S(p) = A(\text{epi}_S(F) \cap (C \times R^{m+1})). \tag{1.69}$$

By the preservation of convexity under linear transformations it follows immediately using relation (1.69) and Lemma 1.43 that the function p is convex if the function F is convex and C is a nonempty convex set. Clearly this condition is sufficient. As shown by the following example the above function p plays a prominent pole within finite dimensional optimization theory in the construction of the so-called Lagrangian dual problem.

Example 1.10. *In optimization theory one studies the following continuous or discrete optimization problem* (P) *given by*

$$(P) \qquad \inf\{f_0(\mathbf{x}) : \mathbf{f}(\mathbf{x}) \in -K, \mathbf{x} \in D\}.$$

with $f_0 : R \to R$ *some function and* $\mathbf{f} : R^n \to R^m$ *some vector-valued function. Moreover, the set* $K \subseteq R^m$ *is a nonempty convex cone and* $D \subseteq R^n$ *some continuous or discrete set. The above optimization problem covers a lot of special and well studied cases. First of all we mention nonlinear programming problems*[53] *with equality and inequality constraints (cf.*[14]*) given by*

$$\inf\{f_0(\mathbf{x}) : f_i(\mathbf{x}) \leq 0, 1 \leq i \leq p \text{ and } f_i(\mathbf{x}) = 0,\ p+1 \leq i \leq m\}$$

[51] strictly convex function
[52] proper function
[53] nonlinear programming problems

with $p \leq m$. Special cases of nonlinear programming problems are fractional programming problems[54] *(cf.[13]) where the objective function f_0 is given by the ration of two functions and geometric programming problems. Other special cases are linear programming problems*[55] *(cf. [15]) given by*

$$\inf\{\mathbf{c}^\intercal \mathbf{x} : \mathbf{x} - \mathbf{b} \in L, \mathbf{x} \geq \mathbf{0}\}$$

and replacing the cone R^n_+ by the convex cone K we obtain the so-called conic convex programming problems[56] *(cf.[24]) given by*

$$\inf\{\mathbf{c}^\intercal \mathbf{x} : \mathbf{x} - \mathbf{b} \in L, \mathbf{x} \in -K\}.$$

In both linear and conic convex programming problems the function $\mathbf{f}$ is represented by the affine mapping $\mathbf{f} : R^n \to R^n$ with $\mathbf{f}(\mathbf{x}) = \mathbf{x} - \mathbf{b}$, while by Lemma 1.5 and 1.14 a more familiar representation of linear programming problems is given by

$$\inf\{\mathbf{c}^\intercal \mathbf{x} : A\mathbf{x} = \mathbf{d}, \mathbf{x} \geq \mathbf{0}\}$$

Another important class of optimization problems are integer linear programming problems [57]*(cf.[5]) given by*

$$\inf\{\mathbf{c}^\intercal \mathbf{x} : \mathbf{x} - \mathbf{b} \in L, \mathbf{x} \in Z^n_+\}.$$

In the construction of so-called primal dual algorithms to solve some of the above optimization problems the so-called Lagrangian dual characterisation of the primal problem (P) plays an important role. To construct the Lagrangian dual of (P) a perturbed optimization problem

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

is introduced with the perturbation function[58] *$F : R^{n+m} \to [-\infty, \infty]$ given by*

$$F(\mathbf{x}, \mathbf{y}) = f_0(\mathbf{x}) \text{ for } \mathbf{x} \in D \text{ and } \mathbf{f}(\mathbf{x}) \in -K + \mathbf{y}$$

and ∞ otherwise. It is easy to check that

$$epi_S(F) = \{(\mathbf{x}, \mathbf{y}, r) \in R^{n+m+1} : \mathbf{y} \in \mathbf{f}(\mathbf{x}) + K, \mathbf{x} \in D \text{ and } r > f_0(\mathbf{x})\}.$$

and this implies with A denoting the projection of R^{n+m+1} onto R^{m+1} that the set $A(epi_S(F))$ is given by

$$\{(\mathbf{y}, r) : \exists_{\mathbf{x} \in D}\ \mathbf{y} \in \mathbf{f}(\mathbf{x}) + K \text{ and } r > f_0(\mathbf{x})\} = \mathbf{F}(D) + K \times (0, \infty)$$

with $\mathbf{F} : R^n \to R^{m+1}$ denoting the vector valued function

$$\mathbf{F}(\mathbf{x}) := (\mathbf{f}(\mathbf{x}), f_0(\mathbf{x})).$$

[54] fractional programming problems
[55] linear programming problems
[56] conic convex programming problems
[57] integer linear programming problems
[58] perturbation function

By relation (1.69) this yields

$$epi_S(p) = \mathbf{F}(D) + (K \times [0,\infty)). \tag{1.70}$$

and applying now Lemma (1.43) it follows that

$$p \text{ convex } \Leftrightarrow \text{the set } \mathbf{F}(D) + (K \times (0,\infty)) \text{ is convex.} \tag{1.71}$$

In case we use the primal representation of an arbitrary function f given by relation (1.57) and the different hull operations on a set it is easy to introduce the different so-called hull functions of f. This is achieved by applying one of the hull operations to the set $\text{epi}(f)$ and then define the associated hull function by means of relation (1.57) with $\text{epi}(f)$ replaced by its hull operation. The first hull function constructed this way is given in the next definition.

Definition 1.19. *For any function $f : R^n \to [-\infty,\infty]$ the function $\overline{f} : R^n \to [-\infty,\infty]$ given by*

$$\overline{f}(\mathbf{x}) = \inf\{r : (\mathbf{x}, r) \in cl(epi(f))\} \tag{1.72}$$

is called the lower semicontinuous hull function[59] *of the function f.*

For the lower semicontinuous hull $\overline{f}$ of an arbitrary function f with $\text{dom}(f)$ nonempty the following result holds for its effective domain.

Lemma 1.44. *For any function $f : R^n \to [-\infty,\infty]$ with $dom(f)$ nonempty we obtain*

$$dom(f) \subseteq dom(\overline{f}) \subseteq cl(dom(f)).$$

Moreover, if additionally f is almost convex then it follows

$$ri(dom(f)) = ri(dom(\overline{f}))$$

Proof. To prove the first inclusion we observe for any $\mathbf{x}$ belonging to $\text{dom}(f)$ and $\text{dom}(f) = A(\text{epi}(f))$ with A denoting the projection of R^{n+1} onto R^n that there exists some $r_1 \in R$ with $(\mathbf{x}, r_1)$ belonging to $\text{epi}(f) \subseteq \text{cl}(\text{epi}(f))$ and this shows by relation (1.72) that $\overline{f}(\mathbf{x}) \leq r_1$ or equivalently $\mathbf{x}$ belongs to $\text{dom}(\overline{f})$. Moreover, if $\mathbf{x}$ belongs to $\text{dom}(\overline{f})$ there exists some $r_1 \in R$ satisfying $\overline{f}(\mathbf{x}) \leq r_1$ and this implies by relation (1.72) that $(\mathbf{x}, r_1 + 1) \in \text{cl}(\text{epi}(f))$. Hence the point $\mathbf{x}$ belongs to $A(\text{cl}(\text{epi}(f))) \subseteq \text{cl}(A(\text{epi}(f)))$ and this shows the desired result. To verify the second result we observe by Lemma 1.37 that the set $\text{dom}(f) = A(\text{epi}(f))$ is almost convex and this implies by the first part and Lemma 1.29 that $\text{ri}(\text{dom}(f)) = \text{ri}(\text{dom}(\overline{f}))$. ∎

The next result is the "function equivalence" of the construction of the closure of a nonempty set by means of the closed hull operation.

Lemma 1.45. *For any function $f : R^n \to [-\infty,\infty]$ the lower semicontinuous hull function $\overline{f}$ given by relation (1.72) is the greatest lower semicontinuous function majorized by f and the epigraph of this function equals $cl(epi(f))$.*

[59] lower semicontinuous hull function of f

Proof. Since clearly $\text{cl}(\text{cl}(\text{epi}(f))) = \text{cl}(\text{epi}(f))$ we obtain by relation (1.72) that

$$(\mathbf{x}, r) \in \text{epi}(\overline{f}) \Leftrightarrow \forall_{\epsilon>0}\ (\mathbf{x}, r+\epsilon) \in \text{cl}(\text{epi}(f)) \Leftrightarrow (\mathbf{x}, r) \in \text{cl}(\text{epi}(f))$$

and so its epigraph equals $\text{cl}(\text{epi}(f))$. Hence by Theorem 1.39 we obtain that $\overline{f}$ is lower semicontinous. To show for any lower semicontinuous function $h \leq f$ that $h \leq \overline{f}$ we observe for $h \leq f$ that $\text{epi}(f) \subseteq \text{epi}(h)$ and by Theorem 1.39 this yields

$$\text{cl}(\text{epi}(f)) \subseteq \text{cl}(\text{epi}(h)) = \text{epi}(h).$$

Using now the definition of $\overline{f}$ as presented by relation (1.72) it follows that $h \leq \overline{f}$ and this shows the desired result. ∎

By Lemma 1.45 and the definition of an almost convex function it follows immediately that

$$f \text{ almost convex } \Leftrightarrow \overline{f} \text{ is convex and } \text{ri}(\text{epi}(\overline{f})) \subseteq \text{epi}(f). \tag{1.73}$$

Moreover, by the preservation of lower semicontinuity under the sup operation it is also immediately clear by Lemma 1.45 that

$$\overline{f} = \sup\{h : h \leq f \text{ and } h : R^n \to [-\infty, \infty] \text{ l.s.c.}\}. \tag{1.74}$$

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

Lemma 1.46. *For any function $f : R^n \to [-\infty, \infty]$ and $\mathbf{x} \in R^n$ it follows that*

$$\overline{f}(\mathbf{x}) = \liminf_{\mathbf{y} \to \mathbf{x}} f(\mathbf{y}).$$

Proof. Since $\overline{f} \leq f$ and $\overline{f}$ lower semicontinuous 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 r and $\epsilon > 0$ satisfying

$$f(\mathbf{x} + \mathbf{y}) > r > \overline{f}(\mathbf{x})$$

for every $\mathbf{y} \in \epsilon E$. This implies that the open set $(\mathbf{x} + \epsilon E) \times (-\infty, r)$ containing the point $(\mathbf{x}, \overline{f}(\mathbf{x}))$ has a nonempty intersection with $\text{epi}(f)$. However, by Lemma 1.45 it follows that $(\mathbf{x}, \overline{f}(\mathbf{x}))$ belongs to $\text{cl}(\text{epi}(f))$ and so by Lemma 1.1 every open set containing $(\mathbf{x}, \overline{f}(\mathbf{x}))$ must have a nonempty intersection with $\text{epi}(f)$. Hence we obtain a contradiction and so the result is proved. ∎

By Lemma 1.46 and Definition 1.17. it is now clear that

$$f \text{ lower semicontinuous at } \mathbf{x} \Leftrightarrow \overline{f}(\mathbf{x}) = f(\mathbf{x}). \tag{1.75}$$

To improve the above result for almost convex functions f we need to give a representation of the relative interior of the epigraph of an almost convex function.

Lemma 1.47. *If $f : R^n \to [-\infty, \infty]$ is an almost convex function with dom(f) nonempty then the set ri(epi(f)) is nonempty and*

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

Proof. Since dom(f) is nonempty it follows that the convex set epi(f) is nonempty and hence by Lemma 1.29 the set ri(epi(f)) is nonempty. If $A : R^{n+1} \to R^n$ is given by $A((\mathbf{x}, r)) = \mathbf{x}$ then we obtain by Lemma 1.35 that

$$\text{ri}(\text{dom}(f)) = \text{ri}(A(\text{epi}(f)) = A(\text{ri}(\text{epi}(f)) \tag{1.76}$$

Consider now an arbitrary $(\mathbf{x}, r)$ satisfying $\mathbf{x} \in \text{ri}(\text{dom}(f))$ and $f(\mathbf{x}) < r$. By relation (1.76) it must follow that

$$(\{\mathbf{x}\} \times R) \cap \text{ri}(\text{epi}(f)) \neq \varnothing$$

and since the affine manifold $\{\mathbf{x}\} \times R$ is relatively open we may apply Lemma 1.37. Hence the intersection $(\{\mathbf{x}\} \times R) \cap$ ri(epi(f)) equals

$$\text{ri}((\{\mathbf{x}\} \times R) \cap \text{epi}(f)) = \text{ri}([f(\mathbf{x}), \infty)) = (f(\mathbf{x}), \infty). \tag{1.77}$$

and so we obtain that $(\mathbf{x}, r) \in \text{ri}(\text{epi}(f))$. To show the reverse inclusion consider some $(\mathbf{x}, r) \in \text{ri}(\text{epi}(f))$. By relation (1.76) clearly $\mathbf{x} \in \text{ri}(\text{dom}(f))$ and so by relation (1.77) it must follow that $f(\mathbf{x}) < r$. ∎

In case f is a almost convex function with dom(f) nonempty the result of Lemma 1.46 can be improved as follows.

Lemma 1.48. *If $f : R^n \to [-\infty, \infty]$ is an almost convex function with dom(f) nonempty then it follows for every $\mathbf{x}_1 \in$ ri(dom(f)) that*

$$\overline{f}(\mathbf{x}) = \lim_{t \downarrow 0} f(\mathbf{x} + t(\mathbf{x}_1 - \mathbf{x})).$$

Moreover, if $\mathbf{x} \in$ ri(dom(f)) then it follows that the function f is lower semicontinuous at $\mathbf{x}$ or equivalently

$$\overline{f}(\mathbf{x}) = \liminf_{\mathbf{y} \to \mathbf{x}} f(\mathbf{y}) = f(\mathbf{x}).$$

Proof. By Lemma 1.46 we obtain that

$$\overline{f}(\mathbf{x}) = \liminf_{\mathbf{y} \to \mathbf{x}} f(\mathbf{y}) \leq \liminf_{t \downarrow 0} f(\mathbf{x} + t(\mathbf{x}_1 - \mathbf{x})). \tag{1.78}$$

If $\overline{f}(\mathbf{x}) = \infty$ then the result holds by the previous inequality and so we assume that $\overline{f}(\mathbf{x}) < \infty$. Since by assumption $\mathbf{x}_1 \in \text{ri}(\text{dom}(f))$ we obtain by

Lemma 1.47 that $(\mathbf{x}_1, r_1)$ belongs to $\text{ri}(\text{epi}(f))$ for every $r_1 > f(\mathbf{x}_1)$ and due to $(\mathbf{x}, \overline{f}(\mathbf{x}) \in \text{epi}(\overline{f}) = \text{cl}(\text{epi}(f))$ this implies by Lemma 1.30 that

$$(t\mathbf{x}_1 + (1-t)\mathbf{x}, tr_1 + (1-t)\overline{f}(\mathbf{x})) \in \text{epi}(f)$$

for every $0 < t < 1$. Hence it follows that

$$f(\mathbf{x} + t(\mathbf{x}_1 - \mathbf{x})) = f(t\mathbf{x}_1 + (1-t)\mathbf{x}) \leq tr_1 + (1-t)\overline{f}(\mathbf{x})$$

and by this inequality we obtain

$$\limsup_{t \downarrow 0} f(\mathbf{x} + t(\mathbf{x}_1 - \mathbf{x})) \leq \overline{f}(\mathbf{x}).$$

Combining this inequality with relation (1.78) yields the first equality. To prove the second equality we observe by Lemma 1.44 that $\text{ri}(\text{dom}(f)) = \text{ri}(\text{dom}(f))$ and this implies by Lemma 1.47 applied to $\overline{f}$ and Lemma 1.29 that

$$\{(\mathbf{x}, r) : r > \overline{f}(\mathbf{x}), \, \mathbf{x} \in \text{ri}(\text{dom}(f))\} = \text{ri}(\text{epi}(f)) \subseteq \text{epi}(f).$$

Consider now an arbitrary $\mathbf{x}$ belonging to $\text{ri}(\text{dom}(f))$. By the above inclusion we obtain for every $\epsilon > 0$ that $(\mathbf{x}, \overline{f}(\mathbf{x}) + \epsilon)$ belongs to $\text{epi}(f)$ and this shows $\overline{f}(\mathbf{x}) + \epsilon \geq f(\mathbf{x})$. Hence by letting $\epsilon \downarrow 0$ it follows that $\overline{f}(\mathbf{x}) \geq f(\mathbf{x})$ and since by Lemma 1.45 we know that $\overline{f}(\mathbf{x}) \leq f(\mathbf{x})$ the second part using also Lemma 1.46 is proved. ∎

Finally we observe for any function $f : R^n \to [-\infty, \infty]$ that

$$\text{epi}(\overline{p}) = \text{cl}(\text{epi}(f)) = \text{cl}(\text{epi}_S(f))$$

and this shows for the function p presented in Example 1.10 that

$$\overline{p} \text{ convex} \Leftrightarrow \text{cl}(\mathbf{F}(D) + (K \times (0, \infty))) \text{ is convex.} \tag{1.79}$$

Another important hull function related to the representation (1.57) is given in the the following definition.

Definition 1.20. *For any function $f : R^n \to [-\infty, \infty]$ the function $co(f) : R^n \to [-\infty, \infty]$ given by*

$$co(f)(\mathbf{x}) := \inf\{r : (\mathbf{x}, r) \in co(epi(f))\} \tag{1.80}$$

is called the convex hull function[60] *of the function f.*

For the convex hull $\text{co}(f)$ of an arbitrary function f with $\text{dom}(f)$ nonempty it follows that its effective domain has the following representation.

Lemma 1.49. *For any function $f : R^n \to [-\infty, \infty]$ with $dom(f)$ nonempty it follows that*

$$dom(co(f)) = co(dom(f)).$$

[60]convex hull function of f

Proof. If the point $\mathbf{x}$ belongs to the set co(dom(f)) then by relation (1.26) there exists some points $\mathbf{x}_i \in \text{dom}(f), 1 \leq i \leq m$ such that $\mathbf{x} = \sum_{i=1}^m \alpha_i \mathbf{x}_i$ with $\alpha_i > 0$ and $\sum_{i=1}^m \alpha_i = 1$. Hence it follows that for every $1 \leq i \leq m$ that the vectors $(\mathbf{x}_i, f(\mathbf{x}_i))$ belong to epi(f) and this shows that the vector $(\sum_{i=1}^m \alpha_i \mathbf{x}_i, \sum_{i=1}^m \alpha_i f(\mathbf{x}_i))$ belongs to co(epi(f)). Since $f(\mathbf{x}_i) < \infty$ we obtain that $\sum_{i=1}^m \alpha_i f(\mathbf{x}_i) < \infty$ and hence by relation (1.80) it follows that $\text{co}(f)(\mathbf{x}) \leq \sum_{i=1}^m \alpha_i f(\mathbf{x}_i)$. By this observation we know that $\mathbf{x}$ belongs to dom(co(f)) and to verify the reverse inclusion we only need to observe for $\mathbf{x}$ belonging to dom(co(f)) that by relation (1.80) it follows that $(\mathbf{x}, \text{co}(f)(\mathbf{x}) + 1)$ belongs to co(epi(f)). Hence we obtain that $\mathbf{x} \in A(\text{co}(\text{epi}(f)) = \text{co}(A(\text{epi}(f)) = \text{co}(\text{dom}(f))$ with A denoting the projection of R^{n+1} onto R^n and this shows the desired result. ∎

The following result is also easy to prove.

Lemma 1.50. *For any function $f : R^n \to [-\infty, \infty]$ the convex hull function co(f) given by relation (1.80) is the greatest convex function majorized by f and the strict epigraph of this function is given by co(epi(f)).*

Proof. By relation (1.80) it follows that

$$(\mathbf{x}, r) \in \text{epi}_S(\text{co}(f)) \Leftrightarrow (\mathbf{x}, r) \in \text{co}(\text{epi}(f)).$$

This verifies the representation for the strict epigraph and by Lemma 1.43 the function co(f) is convex. Moreover, by a similar proof as used in the second part of Lemma 1.45 it is easy to show for any convex function $h \leq f$ that $h \leq \text{co}(f)$ and this shows the result. ∎

Again by the preservation of convexity under the sup operation and Lemma 1.50 it follows that

$$\text{co}(f) = \sup\{h : h \leq f \text{ and } h : R^n \to [-\infty, \infty] \text{ is convex}\}. \tag{1.81}$$

Combining the closure and convex hull operation by observing that the intersection of closed convex sets is again a closed convex set we finally obtain the most important hull function within the field of convex functions.

Definition 1.21. *For any function $f : R^n \to [-\infty, \infty]$ the function $\overline{co(f)} : R^n \to [-\infty, \infty]$ given by*

$$\overline{co(f)}(\mathbf{x}) = \{r : (\mathbf{x}, r) \in cl(co(epi(f)))\} \tag{1.82}$$

is called the lower semicontinuous convex hull function[61] *of the function f.*

By a similar proof as Lemma 1.45 and 1.44 it is easy to verify the following result.

Lemma 1.51. *For any function $f : R^n \to [-\infty, \infty]$ the lower semicontinuous convex hull function $\overline{co(f)}$ given by relation (1.82) is the greatest lower semicontinuous convex function majorized by f and the epigraph of*

[61] lower semicontinuous convex hull function of f

this function is given by $cl(co(epi(f)))$. *Moreover, for dom* (f) *nonempty it follows that*

$$dom(co(f)) \subseteq dom(\overline{co(f)}) \subseteq cl(dom(co(f))).$$

By the preservation of closed convex sets under intersection and Lemma 1.51 we obtain the representation

$$\overline{\text{co}(f)} = \sup\{h : h \leq f \text{ and } h : R \to [-\infty, \infty] \text{ convex and l.s.c.}\} \tag{1.83}$$

To relate the different hull functions based on relation (1.57) it follows by relations (1.74), (1.81) and (1.82) that

$$\overline{\text{co}(f)} \leq \text{co}(f) \leq f \text{ and } \overline{\text{co}(f)} \leq \overline{f} \leq f. \tag{1.84}$$

One might now wonder under which conditions the different hull functions coincide.Clearly for any function $f : R^n \to [-\infty, \infty]$ it follows by Lemma 1.50 that f is convex if and only if $\text{co}(f) = f$. Also it is easy to verify that f is lower semicontinuous if and only if $f = \overline{f}$. Some other cases are now considered in the next lemma.

Lemma 1.52. *For any function* $f : R^n \to [-\infty, \infty]$ *it follows* $\overline{f}$ *is convex if and only if* $\overline{\text{co}(f)} = \overline{f}$. *Moreover, the function* f *convex and lower semicontinuous if and only if* $\overline{\text{co}(f)} = f$.

Proof. For any function f it follows by relation (1.84) that $\overline{\text{co}(f)} \leq \overline{f}$. Since by our assumption $\overline{f}$ is convex and by Lemma 1.45 lower semicontinuous and majorized by f we obtain by Lemma 1.51 that $\overline{f} \leq \overline{\text{co}(f)}$. and this yields $\overline{\text{co}(f)} = \overline{f}$. Due to to $\overline{\text{co}(f)}$ is a convex function the reverse implication follows immediately. To prove the second equivalence relation we observe since f is convex that $\overline{f}$ is convex and so by the first part we obtain $\overline{\text{co}(f)} = \overline{f}$. Due to f lower semicontinuous we know $\overline{f} = f$ and this shows $\overline{\text{co}(f)} = f$. The proof of the reverse implication is trivial and so we omit it. ∎

This concludes our discussion of hull functions based on relation (1.57). We will now consider hull functions based on relation (1.58). However, before discussing these hull functions it is necessary to introduce a quasiconvex function.

Definition 1.22. *The function* $f : R^n \to [-\infty, \infty]$ *is called quasiconvex if the lower-level sets* $L(f, r)$ *for every* $r \in R$ *are convex. Moreover, the function* f *is called evenly quasiconvex if the lower level sets* $L(f, r)$ *are evenly quasiconvex.*

By a similar proof as in Lemma 1.43 one can show that

$$f \text{ quasiconvex } \Leftrightarrow L_{\text{S}}(f, r) \text{ is convex for every } r \in R. \tag{1.85}$$

Also it is easy to verify (cf.[13]) that a function $f : R^n \to [-\infty, \infty]$ is quasiconvex if and only if

$$f(\alpha \mathbf{x}_1 + (1 - \alpha)\mathbf{x}_2) \leq \max\{f(\mathbf{x}_1), f(\mathbf{x}_2)\}. \tag{1.86}$$

As for lower semicontinuous and convex functions one is interested under which operations quasiconvexity is preserved. It is easy to verify for a collection of functions $f_i, i \in I$ that

$$L(\sup_{i\in I} f_i, r) = \cap_{i\in I} L(f_i, r) \tag{1.87}$$

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

Example 1.11. *Consider the functions $f_i : R \to R, i = 1, 2$ given by*

$$f_1(x) = x$$

and

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

Both functions are quasiconvex but it is easy to verify by means of a picture that the sum of both functions is not quasiconvex.

If we consider as before the function $p : R^m \to [-\infty, \infty]$ given by relation (1.68) then it follows that

$$L_S(p, r) = \{\mathbf{y} \in R^m : \exists_{\mathbf{x}\in C} \text{ satisfying } (\mathbf{x}, \mathbf{y}) \in L(F, r)\}.$$

This implies with $A : R^{n+m} \to R^m$ denoting the projection of R^{n+m} onto R^m represented by $A(\mathbf{x}, \mathbf{y}) = \mathbf{y}$ that

$$L_S(p, r) = A(L(F, r) \cap (C \times R^n)). \tag{1.88}$$

By relation (1.88) it follows immediately that the function p is quasiconvex if the function F is quasiconvex and $\mathcal{C} \subseteq R^n$ is a nonempty convex set.

Example 1.12. *Considering the same function p and F as in Example 1.10 it follows that*

$$L_S(F, r) = \{(\mathbf{x}, \mathbf{y}) : f_0(\mathbf{x}) < r, \mathbf{x} \in D \text{ and } \mathbf{y} \in \mathbf{f}(\mathbf{x}) + K\}$$

This implies with $A : R^{n+m} \to R^m$ denoting the projection of R^{n+m} onto R^m that the set $A(L_S(F, r))$ equals

$$\{\mathbf{y} : \mathbf{y} \in \mathbf{f}(\mathbf{x}) + K \text{ and } f_0(\mathbf{x}) < r, \mathbf{x} \in D\} = \mathbf{f}(L_S(f_0, r) \cap D) + K$$

and so by relation (1.88) we obtain

$$L_S(p, r) = \mathbf{f}(L_S(f_0, r) \cap D) + K$$

By this observation we obtain that the function p is quasiconvex if and only if the set $\mathbf{f}(L_S(f_0, r) \cap D) + K$ is convex for every $r \in R$.

We will now introduce the differerent hull functions based on relation (1.58).

Definition 1.23. *For any function* $f : R^n \to [-\infty, \infty]$ *the function* $\overline{f} : R^n \to [-\infty, \infty]$ *given by*

$$\overline{f}(\mathbf{x}) = \inf\{r : \mathbf{x} \in cl(L(f, r))\} \tag{1.89}$$

is called the lower semicontinuous hull function of the function f.

Observe the above function is denoted similarly as the function introduced in Definition 1.19. Moreover, it has the same name since we did not specify with respect to which representation (relation 1.57 or relation(1.58)!) the closed hull operation is taken. However, this does not make any difference due to the following result (cf.[11]).

Lemma 1.53. *For any function* $f : R^n \to [-\infty, \infty]$ *the lower semicontinuous hull function* $\overline{f}$ *given by relation (1.89) is the greatest lower semicontinuous function majorized by* f. *Moreover, it follows that*

$$L(\overline{f}, r) = \cap_{\beta > r} cl(L(f, \beta)).$$

Proof. It is easy to verify by relation 1.89 that

$$L(\overline{f}, r) = \cap_{\beta > r}\{\mathbf{x} \in R^n : \overline{f}(\mathbf{x}) < \beta\} = \cap_{\beta > r}\text{cl}(L(f, \beta)) \tag{1.90}$$

Since the intersection of closed sets is again closed this yields by Theorem 1.39 that the function $\overline{f}$ is lower semicontinuous. To show that $\overline{f}$ is the greatest lower semicontinuous function majorized by f consider some lower semicontinuous function $h \leq f$. This implies that $L(f, r) \subseteq L(h, r)$ for every $r \in R$ and by the lower semicontinuity of h and Theorem 1.39 we obtain

$$\text{cl}(L(f, r)) \subseteq \text{cl}(L(h, r)) = L(h, r) \tag{1.91}$$

Applying now relations (1.90) and (1.91) it follows

$$L(\overline{f}, r) = \cap_{\beta > r}\text{cl}(L(f, \beta)) \subseteq \cap_{\beta > r} L(h, \beta) = L(h, r)$$

for every $r \in R$ and this shows $h \leq \overline{f}$. Hence the function $\overline{f}$ is the greatest lower semicontinuous function majorized by f and so the desired result is verified. ∎

Another hull function to be considered is given by the next definition (cf. [11]).

Definition 1.24. *For any function* $f : R^n \to [-\infty, \infty]$ *the function* $qc(f) : R^n \to [-\infty, \infty]$ *given by*

$$qc(f)(\mathbf{x}) = \inf\{r : \mathbf{x} \in co(L(f, r))\} \tag{1.92}$$

is called the quasiconvex hull function[62] *of the function* f.

The next result (cf.[11]) can be checked similarly as Lemma 1.53.

[62] quasiconvex hull function of f

Lemma 1.54. *For any function $f : R^n \to [-\infty, \infty]$ the quasiconvex hull function $qc(f)$ given by relation (1.92) is the greatest quasiconvex function majorized by f. Moreover, it follows that*

$$L(qc(f), r) = \cap_{\beta > r} co(L(f, \beta)).$$

As before it is clear that

$$\mathrm{qc}(f) = \sup\{h : h \leq f \text{ and } h : R^n \to [-\infty, \infty] \text{ quasiconvex}\}.$$

We now consider a hull function based on closed convex sets.(cf.[11]).

Definition 1.25. *For any function $f : R^n \to [-\infty, \infty]$ the function $\overline{qc(f)} : R^n \to [-\infty, \infty]$ given by*

$$\overline{qc(f)}(\mathbf{x}) = \inf\{r : \mathbf{x} \in cl(co(L(f, r)))\} \tag{1.93}$$

is called the lower semicontinuous quasiconvex hull function [63] *of the function f.*

Similarly as Lemma 1.53 one can show the following result.

Lemma 1.55. *For any function $f : R^n \to [-\infty, \infty]$ the lower semicontinuous quasiconvex hull function $\overline{qc(f)}$ given by relation (1.93) is the greatest lower semicontinuous quasiconvex function majorized by f. Moreover, it follows that*

$$L(\overline{qc(f)}, r) = \cap_{\beta > r} cl(co(L(f, \beta))).$$

As before it follows that

$$\overline{\mathrm{qc}(f)} = \sup\{h : h \leq f \text{ and } h : R^n \to [-\infty, \infty] \text{ quasiconvex and l.s.c}\}.$$

Finally we consider a hull function based on evenly convex sets. It will turn out that this function plays an important role in quasiconvex duality.

Definition 1.26. *For any function $f : R^n \to [-\infty, \infty]$ the function $eqc(f) : R^n \to [-\infty, \infty]$ given by*

$$eqc(f)(\mathbf{x}) = \inf\{r : \mathbf{x} \in eco(L(f, r))\} \tag{1.94}$$

is called the evenly quasiconvex hull function [64] *of the function f.*

Again one can verify the following result.

Lemma 1.56. *For any function $f : R^n \to [-\infty, \infty]$ the evenly quasiconvex hull function $eqc(f)$ given by relation (1.94) is the greatest evenly quasiconvex function majorized by f. Moreover, it follows that*

$$L(eqc(f), r) = \cap_{\beta > r} eco(L(f, \beta)))$$

[63] lower semicontinuous quasiconvex hull function of f

[64] evenly quasiconvex hull function

As before we obtain that

$$\mathrm{eqc}(f) = \sup\{h : h \leq f \text{ and } h : R^n \to [-\infty, \infty] \text{ evenly quasiconvex}\}.$$

Since we will prove in the next section that every closed convex set is evenly convex we finally remark that

$$\overline{\mathrm{qc}(f)} \leq \mathrm{eqc}(f) \leq \mathrm{qc}(f) \leq f$$

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. Penot and Volle (cf.[16]) discusses these hull operations and the relations with quasiconvex duality in linear topological vector spaces and actually in these notes their approach is translated to finite dimensional spaces thereby slightly generalizing the mile-stone papers of Crouzeix (cf.[11]). Observe in finite dimensional spaces one can show stronger results than in infinite dimensional spaces (think of relative interior and separation results to be discussed!) and so in finite dimensional spaces one has an additional structure which needs to be used. To be able to improve some of the representations of hull functions by means of a so-called dual representation and use these improved dual representation to define duality in optimization problems we first need to derive a (weak) and strong separation result between a (closed) convex set and a point outside this set. These separation results are the most important results within convex and quasiconvex analysis and the next section will be completely devoted to this topic.

1.3. **First order conditions and separation.** Since the well known separation result between a closed convex set and a point outside this set is a direct consequence of the first order conditions of the so-called minimum norm problem we first need to introduce the definition of a directional derivative. Although for convex functions directional derivatives always exist this is not the case for more general functions like quasiconvex functions. Therefore, in order to be as complete as possible, we need to introduce a general notion of a directional derivative which exists for any arbitrary proper function $f : R^n \to (-\infty, \infty]$. This means that we have to introduce the so-called upper and lower directional Dini derivatives (cf.[4]) and to avoid undefined combinations like $\infty - \infty$ we only assume in the next definition that $\mathbf{x}$ belongs to dom (f) or equivalently $f(\mathbf{x}) < \infty$.

Definition 1.27. *If $f : R^n \to (-\infty, \infty]$ is an arbitrary proper function with $\mathbf{x}$ belonging to dom(f) then the upper directional Dini derivative*[65] *of f at $\mathbf{x}$ in the direction $\mathbf{d}$ is given by*

$$D^+f(\mathbf{x}, \mathbf{d}) := \limsup_{t \downarrow 0} \frac{f(\mathbf{x} + t\mathbf{d}) - f(\mathbf{x})}{t},$$

[65] upper directional Dini derivative

while the lower directional Dini derivative[66] *of* f *at* $\mathbf{x}$ *in this direction is given by*

$$D_+ f(\mathbf{x}, \mathbf{d}) := \liminf_{t \downarrow 0} \frac{f(\mathbf{x} + t\mathbf{d}) - f(\mathbf{x})}{t}$$

Moreover, if $D^+ f(\mathbf{x}, \mathbf{d})$ *equals* $D_+ f(\mathbf{x}, \mathbf{d})$ *then the directional derivative*[67] $Df(\mathbf{x}, \mathbf{d})$ *of* f *at* $\mathbf{x}$ *in the direction* $\mathbf{d}$ *exists and this directional derivative is given by*

$$Df(\mathbf{x}, \mathbf{d}) := \lim_{t \downarrow 0} \frac{f(\mathbf{x} + t\mathbf{d}) - f(\mathbf{x})}{t}$$

Finally, if the directional derivative of f *at* $\mathbf{x}$ *in every direction* $\mathbf{d}$ *exists and for every* $\mathbf{d}$ *it follows that* $Df(\mathbf{x}, \mathbf{d}) = \mathbf{c}^\intercal \mathbf{d}$ *for some vector* $\mathbf{c}$ *then the function* f *is said to be Gâteaux differentiable*[68] *at* $\mathbf{x}$. *Moreover, if there exists some vector* $\mathbf{c}$ *such that*

$$\lim_{\|\mathbf{h}\| \to 0} \frac{f(\mathbf{x} + \mathbf{h}) - f(\mathbf{x}) - \mathbf{c}^\intercal \mathbf{h}}{\| \mathbf{h} \|} = 0$$

then the function is called Fréchet differentiable[69] *at* $\mathbf{x}$.

As already observed the directional derivative might not exist for arbitrary functions. As an example we consider the function $f : R \to R$ satisfying $f(x) = 1$ for $x \in Q$ and $f(x) = 0$ for $x \notin Q$. On the other hand, the upper and lower directional Dini derivative of f at $\mathbf{x}$ in the direction $\mathbf{d}$ with $f(\mathbf{x})$ finite always exist and these directional derivatives satisfy

$$-\infty \leq D_+ f(\mathbf{x}, \mathbf{d}) \leq D^+ f(\mathbf{x}, \mathbf{d}) \leq \infty$$

Also it is easy to show that

$$D_+ f(\mathbf{x}, \alpha\mathbf{d}) = \alpha D_+ f(\mathbf{x}, \mathbf{d}) \text{ and } D^+ f(\mathbf{x}, \alpha\mathbf{d}) = \alpha D^+ f(\mathbf{x}, \mathbf{d})$$

for every $\alpha > 0$ and so by Lemma 1.42 both functions are positively homogeneous[70]. Finally we observe that any Fréchet differentiable function at $\mathbf{x}$ is Gâteaux differentiable at $\mathbf{x}$ and in both cases the vector $\mathbf{c}$ equals the gradient $\nabla f(\mathbf{x})$. For convex functions the following important result is easy to prove.

Lemma 1.57. *If* $f : R^n \to (-\infty, \infty]$ *is a proper convex function with* $\mathbf{x}$ *belonging to dom*(f) *then it follows that the directional derivative of* f *at* $\mathbf{x}$ *in every direction* $\mathbf{d}$ *exists and*

$$Df(\mathbf{x}, \mathbf{d}) = \inf_{t > 0} \frac{f(\mathbf{x} + t\mathbf{d}) - f(\mathbf{x})}{t}.$$

[66] lower directional Dini derivative
[67] directional derivative
[68] Gateaux differentiable function
[69] Frechet differentiable function
[70] positively homogeneous function

Moreover, the function $\mathbf{d} \to Df(\mathbf{x},\mathbf{d})$ is positively homogeneous and convex.

Proof. In case $f(\mathbf{x}) = -\infty$ it follows immediately that $Df(\mathbf{x},\mathbf{d}) = \infty$ and the above representation trivially holds. Therefore we only consider $f(\mathbf{x})$ finite and for any direction $\mathbf{d} \in R^n$ introduce the function $h : [0,\infty) \to (-\infty,\infty]$ given by

$$h(t) := f(\mathbf{x} + t\mathbf{d}) - f(\mathbf{x}).$$

Since the function f is a proper convex function and $f(\mathbf{x})$ finite we obtain that $h > -\infty$ and convex. This implies by relation (1.67) and $h(0) = 0$ that $h(\alpha t) \leq \alpha h(t)$ for every $0 < \alpha < 1$ and $t > 0$ and so the function

$$t \to f_t(\mathbf{d}) := \frac{f(\mathbf{x} + t\mathbf{d}) - f(\mathbf{x})}{t}$$

is nondecreasing. This implies that

$$Df(\mathbf{x},\mathbf{d}) = \lim_{t\downarrow 0} f_t(\mathbf{d}) = \inf_{t>0} f_t(\mathbf{d})$$

and this shows the first part. To verify the second part we only need to prove that the function $\mathbf{d} \to Df(\mathbf{x},\mathbf{d})$ is convex. Observe now for fixed $t > 0$ and $0 < \alpha < 1$ that by the convexity of f we obtain

$$\begin{aligned} f_t(\alpha\mathbf{d}_1 + (1-\alpha)\mathbf{d}_2) &= \frac{f(\alpha(\mathbf{x} + t\mathbf{d}_1) + (1-\alpha)(\mathbf{x} + t\mathbf{d}_2)) - f(\mathbf{x})}{t} \\ &\leq \alpha f_t(\mathbf{d}_1) + (1-\alpha) f_t(\mathbf{d}_2) \end{aligned}$$

and since $f_t(\mathbf{d})$ belongs to $(-\infty,\infty]$ this shows by relation (1.67) that $\text{epi}(f_t)$ is a convex set. Since $f_t \leq f_s$ for every $s \geq t$ it follows that $\text{epi}(f_t) \subseteq \text{epi}(f_s)$ and so we obtain that the set

$$\text{epi}(Df(\mathbf{x},.)) = \cup_{t>0}\text{epi}(f_t)$$

is a convex set. ∎

For quasiconvex functions f with $f(\mathbf{x})$ finite the directional derivative might not exist. However, as shown by Crouzeix (cf.[11]) one can show the following result for this class of functions.

Lemma 1.58. *If $f : R^n \to (-\infty,\infty]$ is a quasiconvex function with $f(\mathbf{x})$ finite then it follows that the function $\mathbf{d} \to D^+f(\mathbf{x},\mathbf{d})$ is quasiconvex and positively homogeneous.*

Proof. We only need to verify that the function $\mathbf{d} \to D^+f(\mathbf{x},\mathbf{d})$ is quasiconvex since it is already shown that this function is positively homogeneous. By definition we know that

$$D^+f(\mathbf{x},\mathbf{d}) = \limsup_{t\downarrow 0} f_t(\mathbf{d})$$

Since f is quasiconvex it follows for fixed $t > 0$ that

$$\begin{aligned} f_t(\alpha \mathbf{d}_1 + (1-\alpha)\mathbf{d}_2) &= \frac{f(\alpha(\mathbf{x}+t\mathbf{d}_1) + (1-\alpha)(\mathbf{x}+t\mathbf{d}_2) - f(\mathbf{x})}{t} \\ &\leq \frac{\max\{f(\mathbf{x}+t\mathbf{d}_1), f(\mathbf{x}+t\mathbf{d}_2)\} - f(\mathbf{x})}{t} \end{aligned}$$

for every $0 < \alpha < 1$ and this shows for every $t > 0$ that

$$f_t(\alpha \mathbf{d}_1 + (1-\alpha)\mathbf{d}_2) \leq \max\{f_t(\mathbf{d}_1), f_t(\mathbf{d}_2)\} \tag{1.95}$$

By relation (1.95) and using contradiction it is easy to verify that

$$\begin{aligned} D^+ f(\mathbf{x}, \alpha \mathbf{d}_1 + (1-\alpha)\mathbf{d}_2) &= \limsup_{t \downarrow 0} f_t(\alpha \mathbf{d}_1 + (1-\alpha)\mathbf{d}_2) \\ &\leq \max\{\limsup_{t\downarrow 0} f_t(\mathbf{d}_1), \limsup_{t\downarrow 0} f_t(\mathbf{d}_2)\} \end{aligned}$$

and this shows by relation (1.86) that the upper directional Dini derivative of f at $\mathbf{x}$ is quasiconvex. ∎

In case it is also known that a positively homogeneous quasiconvex function is nonnegative and lower semicontinuous then it is possible to show by means of a duality representation that this function is convex (cf.[10]). For the moment we only mention this result which will be proved in the next section. An easy and well-known consequence of Lemma 1.57 are the so-called first order conditions for a convex program.

Definition 1.28. *A vector $\mathbf{x}_{opt}$ is called an optimal solution*[71] *of an optimization problem*

$$\inf\{f(\mathbf{x}) : \mathbf{x} \in S\}$$

if and only if $\mathbf{x}_{opt}$ belongs op S and $f(\mathbf{x}_{opt}) \leq f(\mathbf{x})$ for every $\mathbf{x} \in S$. Moreover, the above optimization problem is called a convex program[72] *if the function $f : R^n \to [-\infty, \infty]$ is convex and $S \subseteq R^n$ a nonempty convex set.*

In case $f(\mathbf{x}) = -\infty$ for some $\mathbf{x} \in S$ or $\mathrm{dom}(f) \cap S$ is empty then the optimization problem is not well behaving since an optimal solution can be found immediately. Therefore it is only interesting to study optimization problems with $f > -\infty$ and $\mathrm{dom}(f) \cap S$ nonempty. Observe the above optimization problem is the same as the optimization problem

$$\inf\{f_1(\mathbf{x}) : \mathbf{x} \in R^n\}$$

with $f_1(\mathbf{x}) = f(\mathbf{x})$ whenever $\mathbf{x} \in S$ and ∞ otherwise and for this problem the above assumptions hold if and only if the function f_1 is proper. Such optimization problems are therefore called for simplicity *proper*[73].

[71] optimal solution of optimization problem
[72] convex program
[73] proper optimization problem

Lemma 1.59. *If the optimization problem* $\inf\{f(\mathbf{x}) : \mathbf{x} \in C\}$ *is a proper convex program then it follows that*

$$\mathbf{x}_{opt} \textit{ optimal } \Leftrightarrow \forall_{\mathbf{x}\in C}\ Df(\mathbf{x}_{opt}, \mathbf{x}-\mathbf{x}_{opt}) \geq 0 \textit{ and } \mathbf{x}_{opt} \in C.$$

Proof. If $\mathbf{x}_{\mathrm{opt}}$ belonging to C satisfies $Df(\mathbf{x}_{\mathrm{opt}}, \mathbf{x}-\mathbf{x}_{\mathrm{opt}}) \geq 0$ for every $\mathbf{x} \in C$ then necessarily $f(\mathbf{x}_{\mathrm{opt}}) < \infty$ and hence by Lemma 1.57 we obtain for every $\mathbf{x} \in C$ that

$$f(\mathbf{x}) - f(\mathbf{x}_{\mathrm{opt}}) \geq Df(\mathbf{x}_{\mathrm{opt}}, \mathbf{x}-\mathbf{x}_{\mathrm{opt}}) \geq 0$$

and this proves that $\mathbf{x}_{\mathrm{opt}}$ is an optimal solution. To show the reverse implication it follows since we are dealing with a proper convex program and $\mathbf{x}_{\mathrm{opt}}$ optimal that $f(\mathbf{x}_{\mathrm{opt}}) < \infty$. Moreover, by the convexity of the set C and $\mathbf{x}_{\mathrm{opt}}$ optimal we obtain for every $\mathbf{x} \in C$ and $0 < t < 1$ that the vector $\mathbf{x}_{\mathrm{opt}} + t(\mathbf{x}-\mathbf{x}_{\mathrm{opt}})$ belongs to C and $f(\mathbf{x}_{\mathrm{opt}} + t(\mathbf{x}-\mathbf{x}_{\mathrm{opt}})) \geq f(\mathbf{x}_{\mathrm{opt}})$. This shows by Lemma 1.57 the desired result. ∎

Clearly for an arbitrary proper optimization problem with $C \subseteq R^n$ a nonempty convex set we obtain by a similar argument that

$$\mathbf{x}_{\mathrm{opt}} \text{ optimal } \Rightarrow \forall_{\mathbf{x}\in C}\ D_+f(\mathbf{x}_{\mathrm{opt}}, \mathbf{x}-\mathbf{x}_{\mathrm{opt}}) \geq 0 \text{ and } \mathbf{x}_{\mathrm{opt}} \in C. \tag{1.96}$$

Clearly if the objective function is Gâteaux differentiable we can replace in the above statements

$$D_+f(\mathbf{x}_{\mathrm{opt}}, \mathbf{x}-\mathbf{x}_{\mathrm{opt}}) \geq 0 \text{ for every } \mathbf{x} \in C$$

by

$$\nabla f(\mathbf{x}_{\mathrm{opt}})^{\intercal}(\mathbf{x}-\mathbf{x}_{\mathrm{opt}}) \geq 0 \text{ for every } \mathbf{x} \in C.$$

Therefore from a *computational point of view* the above condition seems to be easier to check if additionally f is Gâteaux differentiable. Observe that the reverse implication in relation (1.96) does not hold in general and so one has introduced in the literature (cf.[23],[3]) the class of pseudoconvex functions. The next definition is taken from Diewert (cf.[23]).

Definition 1.29. *A proper function* $f : R^n \to (-\infty, \infty]$ *is called pseudoconvex on the convex set* $C \subseteq R^n$ *if and only if*

$$f(\mathbf{x}_1) < f(\mathbf{x}_2) \textit{ for } \mathbf{x}_i \in C \Rightarrow D_+f(\mathbf{x}_2, \mathbf{x}_1-\mathbf{x}_2) < 0.$$

In optimization theory the direction $\mathbf{x}_1 - \mathbf{x}_2$ satisfying $D_+f(\mathbf{x}_2, \mathbf{x}_1-\mathbf{x}_2) < 0$ is called a *strict descent direction*[74]. An immediate consequence of the above definition is given by the following result (cf.[3]).

Lemma 1.60. *If the optimization problem* $\inf\{f(\mathbf{x}) : \mathbf{x} \in C\}$ *is a proper optimization problem and* f *is pseudoconvex on the convex set* C *then it follows that*

$$\mathbf{x}_{opt} \textit{ optimal } \Leftrightarrow \forall_{\mathbf{x}\in C}\ D_+f(\mathbf{x}_{opt}, \mathbf{x}-\mathbf{x}_{opt}) \geq 0 \textit{ and } \mathbf{x}_{opt} \in C.$$

[74]strict descent direction

Proof. The implication $\Rightarrow$ is obvious by relation (1.96). To show the reverse implication suppose by contradiction that $\mathbf{x}_{\text{opt}}$ is not optimal and so there exists some $\mathbf{x} \in C$ satisfying $f(\mathbf{x}) < f(\mathbf{x}_{\text{opt}})$ Since f is pseudoconvex on C this yields that $D_+ f(\mathbf{x}_{\text{opt}}, \mathbf{x} - \mathbf{x}_{\text{opt}}) < 0$ and hence our assumption does not hold. ∎

A nice characterisation of arbitrary proper pseudoconvex functions seems not to be possible (cf.[3]). Moreover, generalizing convexity and still knowing that the statement of Lemma 1.60 holds was the main reason to introduce the set of pseudoconvex functions on the convex set C. In case these functions are also Gâteaux differentiable the class of pseudoconvex functions is studied in detail by Komlosi (cf.[20]) and Crouzeix.(cf.[12]) To start with the most simplest proper convex optimization problem we observe that the function $\mathbf{x} \rightarrow \| \mathbf{x} \|$ is positively homogeneous. Moreover, for every real valued t it follows that

$$0 \leq \| \mathbf{x} - t\mathbf{y} \|^2 = \| \mathbf{x} \|^2 + t^2 \| \mathbf{y} \|^2 - 2t\mathbf{x}^\intercal \mathbf{y} \tag{1.97}$$

This implies $\lim_{t\uparrow\infty} \| \mathbf{x} - t\mathbf{y} \|^2 = \infty$ for every $\mathbf{y} \neq \mathbf{0}$ and hence by the Weierstrass-Lesbesgue theorem (Theorem 1.40) an optimal solution of the optimization problem

$$\inf\{\| \mathbf{x} - t\mathbf{y} \|^2 : t \in R\}$$

exist. This implies by relation (1.96) and the Gâteaux differentiability that an optimal solution must be equal to $\mathbf{x}^\intercal \mathbf{y} \| \mathbf{y} \|^{-2}$ and substituting this into relation (1.97) we obtain the well-known *Cauchy-Schwartz inequality* given by

$$(\mathbf{x}^\intercal \mathbf{y})^2 \leq \| \mathbf{x} \|^2 \| \mathbf{y} \|^2 \tag{1.98}$$

for every $\mathbf{x}, \mathbf{y} \in R^n$. By relation (1.98) it is easy to verify that the triangle inequality

$$\| \mathbf{x} + \mathbf{y} \| \leq \| \mathbf{x} \| + \| \mathbf{y} \| \tag{1.99}$$

holds and since the function $\mathbf{x} \rightarrow \| \mathbf{x} \|$ is positively homogeneous it follows by relation (1.99) that this function is finite valued and convex. Since for any increasing convex function $g : [0, \infty) \rightarrow R$ it is easy to check by the definition of convexity that the function $\mathbf{x} \rightarrow g(\| \mathbf{x} \|)$ is also convex we also obtain that the function $\mathbf{x} \rightarrow \| \mathbf{x} \|^2$ is convex. One can now consider the so-called *minimum norm problem*[75] given by

$$\mathrm{d}_C(\mathbf{y}) := \inf\{\frac{1}{2} \| \mathbf{y} - \mathbf{x} \|^2 : \mathbf{x} \in C\}. \tag{P_{minnorm}}$$

with $C \subseteq R^n$ a proper closed nonempty convex set and this is one of the most simplest proper convex programming problems. The vector $\mathbf{y}$ in optimization problem (P_{minnorm}) serves as a perturbation parameter since it is

[75] minimum norm problem

easy to check that

$$\mathrm{d}_C(\mathbf{y}) = \inf\{\frac{1}{2} \parallel \mathbf{x} \parallel^2 : \mathbf{x} \in C - \mathbf{y}\} = d_{C-\mathbf{y}}(\mathbf{0}).$$

By a standard application of the Weierstrass-Lebesgue theorem (Theorem 1.40) the minimum norm problem (P_{minnorm}) has an optimal solution. To show that this optimal solution is unique we observe for any $\mathbf{y}_1, \mathbf{y}_2$ belonging to R^n that

$$\frac{1}{2} \parallel \mathbf{y}_1 + \mathbf{y}_2 \parallel^2 + \frac{1}{2} \parallel \mathbf{y}_1 - \mathbf{y}_2 \parallel^2 = \parallel \mathbf{y}_1 \parallel^2 + \parallel \mathbf{y}_2 \parallel^2 . \tag{1.100}$$

For every $\mathbf{x}_1 \neq \mathbf{x}_2$ belonging to C it follows by relation (1.100) with $\mathbf{y}_i$ replaced by $\mathbf{y} - \mathbf{x}_i$ for $i = 1, 2$ that

$$\frac{1}{2}\|\mathbf{y} - \frac{1}{2}(\mathbf{x}_1 + \mathbf{x}_2)\|^2 < \frac{1}{4}\|\mathbf{y} - \mathbf{x}_1\|^2 + \frac{1}{4}\|\mathbf{y} - \mathbf{x}_2\|^2$$

and so for $\mathbf{x}_i, i = 1, 2$ different optimal solutions of the optimization problem (P_{minnorm}) we obtain

$$\frac{1}{2}\|\mathbf{y} - \frac{1}{2}(\mathbf{x}_1 + \mathbf{x}_2)\|^2 < \mathrm{d}_C(\mathbf{y}).$$

Since the set C is convex and hence $\frac{1}{2}(\mathbf{x}_1 + \mathbf{x}_2)$ belongs to C the objective function evaluated at this point has a lower objective value as the objective value $\mathrm{d}_C(\mathbf{y})$ of the optimal solution and so it cannot happen that there are two different optimal solutions. Therefore the optimal solution is unique and for simplicity this unique optimal solution is denoted by $\mathrm{p}_C(\mathbf{y})$. Moreover, if the function $f : R^n \to R$ is given by

$$f(\mathbf{x}) = \frac{1}{2}\|\mathbf{y} - \mathbf{x}\|^2$$

then it follows that the directional derivative $Df(\mathrm{p}_C(\mathbf{y}), \mathbf{d})$ in every direction $\mathbf{d}$ exists and this directional derivative equals

$$\frac{1}{2} \lim_{t \downarrow 0} \frac{\|\mathbf{y} - (\mathbf{x} + t\mathbf{d})\|^2 - \|\mathbf{y} - \mathbf{x}\|^2}{t} = (\mathrm{p}_C(\mathbf{y}) - \mathbf{y})^\top \mathbf{d} \tag{1.101}$$

The next result is one of the most important results within convex analysis and is an easy consequence of Lemma 1.59 and relation (1.101).

Lemma 1.61. *For any $\mathbf{y} \in R^n$ and $C \subseteq R^n$ a closed convex set it follows that*

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

Proof. Since $\mathrm{p}_C(\mathbf{y}) \in C$ denotes the unique optimal solution of the minimum norm problem and by relation (1.101) the directional derivative of the function $f(\mathbf{x}) = \frac{1}{2} \parallel \mathbf{y} - \mathbf{x} \parallel^2$ at $\mathrm{p}_C(\mathbf{y})$ in the feasible direction $\mathbf{x} - \mathrm{p}_C(\mathbf{y})$ for any $\mathbf{x} \in C$ is given by

$$Df(\mathrm{p}_C(\mathbf{y}), \mathbf{x} - \mathrm{p}_C(\mathbf{y})) = (\mathrm{p}_C(\mathbf{y}) - \mathbf{y})^\intercal (\mathbf{x} - \mathrm{p}_C(\mathbf{y}))$$

we obtain by Lemma 1.59 the desired result. To prove the reverse implication we consider some $\mathbf{z} \in C$ satisfying $(\mathbf{z} - \mathbf{y})^\intercal (\mathbf{x} - \mathbf{z}) \geq 0$ for every $\mathbf{x} \in C$. Again by Lemma 1.59 the vector $\mathbf{z}$ must be an optimal solution of the proper convex minimum norm problem and since we know that this problem has a unique optimal solution it must follow that $\mathbf{z} = \mathrm{p}_C(\mathbf{y})$. ∎

By the above first order conditions one can also verify that the vector valued function $\mathbf{y} \to \mathrm{p}_C(\mathbf{y})$ is a so-called contraction mapping[76]

Lemma 1.62. *For $C \subseteq R^n$ a proper closed nonempty convex set it follows that*

$$\|p_C(\mathbf{y}_1) - p_C(\mathbf{y}_1)\| \leq \|\mathbf{y}_1 - \mathbf{y}_2\|$$

for every $\mathbf{y}_1$ and $\mathbf{y}_2$ belonging to R^n.

Proof. By elementary calculations we obtain that

$$\|\mathrm{p}_C(\mathbf{y}_1) - \mathrm{p}_C(\mathbf{y}_2)\|^2 = (\mathrm{p}_C(\mathbf{y}_1) - \mathrm{p}_C(\mathbf{y}_2))^\top (\mathbf{y}_1 - \mathbf{y}_2) - \mathrm{I}_0 - \mathrm{I}_1$$

with

$$\mathrm{I}_0 := (\mathrm{p}_C(\mathbf{y}_1) - \mathbf{y}_1)^\top (\mathrm{p}_C(\mathbf{y}_2) - \mathrm{p}_C(\mathbf{y}_1))$$

and

$$\mathrm{I}_1 := (\mathrm{p}_C(\mathbf{y}_2) - \mathbf{y}_2)^\top (\mathrm{p}_C(\mathbf{y}_1) - \mathrm{p}_C(\mathbf{y}_2))$$

Since $\mathrm{p}_C(\mathbf{y}_1)$ and $\mathrm{p}_C(\mathbf{y}_2)$ belong to C we obtain by the first order conditions of Lemma 1.61 applied to $\mathrm{p}_C(\mathbf{y}_1)$, respectively $\mathrm{p}_C(\mathbf{y}_2)$ that the values I_0 and I_1 are nonnegative and so the inequality

$$\|\mathrm{p}_C(\mathbf{y}_1) - \mathrm{p}_C(\mathbf{y}_2)\|^2 \leq (\mathrm{p}_C(\mathbf{y}_1) - \mathrm{p}_C(\mathbf{y}_2))^\top (\mathbf{y}_1 - \mathbf{y}_2) \tag{1.102}$$

holds. Applying now the Cauchy-Schwartz inequality given by relation (1.98) to the last part of relation (1.102) yields

$$\|\mathrm{p}_C(\mathbf{y}_1) - \mathrm{p}_C(\mathbf{y}_2)\|^2 \leq \|\mathrm{p}_C(\mathbf{y}_1) - \mathrm{p}_C(\mathbf{y}_2)\| \|\mathbf{y}_1 - \mathbf{y}_2\|$$

and this shows the desired result. ∎

In case $K \subseteq R^n$is a closed convex cone we can improve the result of the previous lemma. Remember that the polar cone K^0 is given by

$$K^0 = \{\mathbf{x}^* : \mathbf{x}^\intercal \mathbf{x}^* \leq 0 \text{ for every } \mathbf{x} \in K\}.$$

and $\mathbf{x}_1 \perp \mathbf{x}_2$ if and only if $\mathbf{x}_1^\intercal \mathbf{x}_2 = 0$. The next result is due to Moreau.

Lemma 1.63. *For any $\mathbf{y} \in R^n$ and $K \subseteq R^n$ a closed convex cone it follows that*

$$\mathbf{z} = p_K(\mathbf{y}) \Leftrightarrow \ \mathbf{y} - \mathbf{z} \in K^0, \ \mathbf{z} \in K \textit{ and } \mathbf{y} - \mathbf{z} \perp \mathbf{z}.$$

[76] contraction mapping

Proof. If $\mathbf{z} = p_K(\mathbf{y})$ we obtain by Lemma 1.61 that

$$(p_K(\mathbf{y}) - \mathbf{y})^\intercal(\mathbf{x} - p_K(\mathbf{y})) \geq 0 \tag{1.103}$$

for every $\mathbf{x} \in K$. Since $p_K(\mathbf{y})$ belongs to K also $\alpha p_K(\mathbf{y})$ belongs to K for every $\alpha > 0$ and this implies by relation (1.103) that

$$(\alpha - 1)(p_K(\mathbf{y}) - \mathbf{y})^\intercal p_K(\mathbf{y}) \geq 0$$

for every $\alpha > 0$. Hence it follows that

$$(p_K(\mathbf{y}) - \mathbf{y})^\intercal p_K(\mathbf{y}) = 0 \tag{1.104}$$

and substituting relation (1.104) into relation (1.103) we obtain that

$$(p_K(\mathbf{y}) - \mathbf{y})^\intercal \mathbf{x} \geq 0$$

for every $\mathbf{x} \in K$ or equivalently $\mathbf{y} - p_K(\mathbf{y})$ belongs to K^0. To prove the reverse implication we observe for $\mathbf{z} \in K$ satisfying $\mathbf{y} - \mathbf{z} \in K^0$ and $\mathbf{y} - \mathbf{z} \perp \mathbf{z}$ that

$$(\mathbf{z} - \mathbf{y})^\intercal(\mathbf{x} - \mathbf{z}) \geq 0$$

for every $\mathbf{x} \in K$. This implies by Lemma 1.61 that $\mathbf{z} = p_K(\mathbf{y})$ and this proves the desired result. ∎

Specializing Lemma 1.61 for an affine set it is easy to show the following result.

Lemma 1.64. *For any $\mathbf{y} \in R^n$ and $M \subseteq R^n$ an affine set it follows that*

$$\mathbf{z} = p_M(\mathbf{y}) \Leftrightarrow \forall_{\mathbf{x} \in M} \ (\mathbf{z} - \mathbf{y})^\intercal(\mathbf{x} - \mathbf{z}) = 0 \text{ and } \mathbf{z} \ \in M.$$

Proof. Since $p_M(\mathbf{y})$ belongs to M and M is an affine set we obtain by Lemma 1.61 that

$$(p_M(\mathbf{y}) - \mathbf{y})^\intercal(\mathbf{x} - p_M(\mathbf{y})) \geq 0 \tag{1.105}$$

for every $\mathbf{x} \in M$. At the same time, since M is an affine manifold and $p_M(\mathbf{y})$ belongs to M we obtain that $2p_M(\mathbf{y}) - \mathbf{x}$ belongs to M for every $\mathbf{x} \in M$ and this implies by relation (1.105) with $\mathbf{x}$ replaced by $2p_M(\mathbf{y}) - \mathbf{x}$ that

$$(p_M(\mathbf{y}) - \mathbf{y})^\intercal(p_M(\mathbf{y}) - \mathbf{x}) \geq 0$$

for every $\mathbf{x} \in M$. Combining the two inequalities yields

$$(p_M(\mathbf{y}) - \mathbf{y})^\intercal(p_M(\mathbf{y}) - \mathbf{x}) = 0$$

and this shows the desired result. The reverse implication is now a direct consequence of Lemma 1.61. ∎

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.

Theorem 1.65. *If $S \subseteq R^n$ is a nonempty almost convex set and* $\mathbf{y}$ *does not belong to the set cl(S) then there exists some nonzero vector* $\mathbf{y}^* \in R^n$ *and* $\epsilon > 0$ *such that*

$$\mathbf{y}^{*\top}\mathbf{x} \geq \mathbf{y}^{*\top}\mathbf{y} + \epsilon$$

for every $\mathbf{x}$ *belonging to cl(S). In particular the vector* $\mathbf{y}^*$ *can be chosen equal to* $p_{cl(S)}(\mathbf{y}) - \mathbf{y}$.

Proof. Since by Lemma 1.29 the vector $\mathbf{y}$ does not belong to the closed convex set $\mathrm{cl}(S)$ it follows that the vector $\mathrm{p}_{\mathrm{cl}(S)}(\mathbf{y}) - \mathbf{y}$ belonging to $\mathrm{cl}(S) - \mathbf{y}$ is nonzero and so the scalar $\epsilon := \| \mathrm{p}_{\mathrm{cl}(S)}(\mathbf{y}) - \mathbf{y} \|^2$ is positive. Moreover, by Lemma 1.61 we obtain for every $\mathbf{x} \in \mathrm{cl}(S)$ that

$$\begin{aligned} & (\mathrm{p}_{\mathrm{cl}(S)}(\mathbf{y}) - \mathbf{y})^\intercal \mathbf{x} - \|\mathrm{p}_{\mathrm{cl}(S)}(\mathbf{y}) - \mathbf{y}\|^2 - (\mathrm{p}_{\mathrm{cl}(S)}(\mathbf{y}) - \mathbf{y})^\intercal \mathbf{y} \\ = \; & (\mathrm{p}_{\mathrm{cl}(S)}(\mathbf{y}) - \mathbf{y})^\intercal (\mathbf{x} - \mathrm{p}_{\mathrm{cl}(S)}(\mathbf{y})) \geq 0 \end{aligned}$$

and reordering this inequality yields

$$(\mathrm{p}_{\mathrm{cl}(S)}(\mathbf{y}) - \mathbf{y})^\intercal \mathbf{x} \geq (\mathrm{p}_{\mathrm{cl}(S)}(\mathbf{y}) - \mathbf{y})^\intercal \mathbf{y} + \epsilon$$

for every $\mathbf{x}$ belonging to $\mathrm{cl}(S)$. ∎

In the above result the condition that S is almost convex is too strong. Actually we only need that the set $\mathrm{cl}(S)$ is a convex set. However, we listed the almost convexity condition in order to be able to compare the above result with a weaker separation result to be discussed later. Observe now that the nonzero vector $\mathbf{y}^* \in \mathrm{cl}(S) - \mathbf{y}$ is called the *normal vector*[77] of the *separating hyperplane*[78]

$$H^{=}(\mathbf{a}, a) := \{\mathbf{x} \in R^n : \mathbf{a}^\top \mathbf{x} = a\}, \mathbf{a} = \mathbf{y}^* \text{ and } a = \mathbf{y}^{*\intercal}\mathbf{y} + \frac{\epsilon}{2}$$

and this hyperplane strongly separates the closed convex set $\mathrm{cl}(S)$ and $\mathbf{y}$. Due to this terminology the set $\mathrm{cl}(S)$ and $\mathbf{y}$ are said to be *strongly separated*[79] by the hyperplane $H^{=}(\mathbf{a}, a)$. Without loss of generality we may take as a normal vector of the hyperplane the vector $\mathbf{y}^* \parallel \mathbf{y}^* \|^{-1}$ and this vector has norm 1 and clearly belongs to $\mathrm{cone}(\mathrm{cl}(S) - \mathbf{y})$. Before discussing the "weak" form of the separation result in finite dimensional spaces we will consider some implications of the above strong separation result.

Definition 1.30. *If $S \subseteq R^n$ is some nonempty set then the function* $\sigma_S : R^n \to (-\infty, \infty]$ *given by*

$$\sigma_S(\mathbf{s}) := \sup\{\mathbf{s}^\intercal \mathbf{x} : \mathbf{x} \in S\}$$

is called the support function [80] *of the set S.*

[77] normal vector
[78] separating hyperplane
[79] strong separation
[80] support function

It is easy to see that any support function σ_S with $S \subseteq R^n$ a nonempty set is a lower semicontinuous proper convex function which is also positively homogeneous and satisfies $\sigma_S(\mathbf{0}) = 0$. The next result shows that a support function cannot distinguish between the set S and $\text{cl}(\text{co}(S))$.

Lemma 1.66. *For any nonempty set $S \subseteq R^n$ it follows that the support function σ_S of the set S coincides with the support function $\sigma_{cl(co(S))}$ of the set $cl(co(S))$.*

Proof. To prove the above result we first observe that clearly $\sigma_S \leq \sigma_{\text{cl}(\text{co}(S))}$. It is now sufficient to verify that $\sigma_S \geq \sigma_{\text{co}(S)} \geq \sigma_{\text{cl}(\text{co}(S))}$. To start with the first inequality consider an arbitrary vector $\mathbf{x}$ belonging to the set $\text{co}(S)$. By relation (1.26) there exist some vectors $\mathbf{x}_i \in S, 1 \leq i \leq m$ such that $\mathbf{x} = \sum_{i=1}^m \alpha_i \mathbf{x}_i$ with α_i positive and $\sum_{i=1}^m \alpha_i = 1$. Since $\sigma_S(\mathbf{s}) \geq \mathbf{s}^\top \mathbf{x}_i$ for every $1 \leq i \leq m$ and $\mathbf{s} \in R^n$ it follows that $\sigma_S(\mathbf{s}) \geq \sum_{i=1}^n \alpha_i \mathbf{s}^\top \mathbf{x}_i = \mathbf{s}^\top \mathbf{x}$ and this shows using $\mathbf{x} \in \text{co}(S)$ is arbitrary that $\sigma_S \geq \sigma_{\text{co}(S)}$. To verify the second inequality it is sufficient to check that $\sigma_S \geq \sigma_{\text{cl}(S)}$ for any set S and to prove this consider an arbitrary $\mathbf{x}$ belonging to $\text{cl}(S)$. By Lemma 1.1 there exists a sequence $\mathbf{x}_n \in S, n \in N$ with limit $\mathbf{x}$ and this shows by the continuity of a linear mapping that $\sigma_S(\mathbf{s}) \geq \lim_{n\uparrow\infty} \mathbf{s}^\top \mathbf{x}_n = \mathbf{s}^\top \mathbf{x}$. As before it follows since $\mathbf{x} \in \text{co}(S)$ is arbitrary that $\sigma_S \geq \sigma_{\text{cl}(S)}$ and this proves the result. ∎

A reformulation of Theorem 1.65 in terms of the support function of the closed convex set C is given by the following result.

Theorem 1.67. *If $S \subseteq R^n$ is a proper nonempty almost convex set then it follows that*

$$\mathbf{x}_0 \in cl(S) \Leftrightarrow \mathbf{s}^\intercal \mathbf{x}_0 \leq \sigma_{cl(S)}(\mathbf{s}) \text{ for every } \mathbf{s} \in R^n.$$

Proof. Clearly $\mathbf{x}_0 \in \text{cl}(S)$ implies that $\mathbf{s}^\intercal \mathbf{x} \leq \sigma_{\text{cl}(S)}(\mathbf{s})$ for every $\mathbf{s}$ belonging to R^n. To show the reverse implication let $\mathbf{s}^\intercal \mathbf{x}_0 \leq \sigma_{\text{cl}(S)}(\mathbf{s})$ for every $\mathbf{s} \in R^n$ and suppose by contradiction that $\mathbf{x}_0$ does not belong to the set $\text{cl}(S)$. By Lemma1.29 it follows that $\text{cl}(S)$ is convex and by Theorem 1.65 we obtain that there exists some nonzero vector $\mathbf{x}_0^* \in R^n$ and $\epsilon > 0$ satisfying $\mathbf{x}_0^{*\top} \mathbf{x} \geq \mathbf{x}_0^{*\top} \mathbf{x}_0 + \epsilon$ for every $\mathbf{x}$ belonging to $\text{cl}(S)$. This shows by the definition of a support function that

$$\sigma_{\text{cl}(S)}(-\mathbf{x}_0^*) \leq -\mathbf{x}_0^{*\top} \mathbf{x}_0 - \epsilon < -\mathbf{x}_0^{*\top} \mathbf{x}_0$$

contradicting our initial assumption and so it must follow that $\mathbf{x}_0$ belongs to $\text{cl}(S)$. ∎

The above result can be seen as a dual representation of a closed nonempty convex set. Also by Lemma 1.66 one can replace in Theorem 1.67 the support functtion $\sigma_{\text{cl}(S)}$ by σ_S. An immediate consequence of Theorem 1.67 is given by the next observation.

Theorem 1.68. *For any nonempty sets $S_1, S_2 \subseteq R^n$ it follows that*

$$\sigma_{S_1} \leq \sigma_{S_2} \Leftrightarrow cl(co(S_1) \subseteq cl(co(S_2)).$$

Proof. If $\text{cl}(\text{co}(S_1) \subseteq \text{cl}(\text{co}(S_2))$ we obtain by Lemma 1.66 that

$$\sigma_{S_1} = \sigma_{\text{cl}(\text{co}(S_1))} \leq \sigma_{\text{cl}(\text{co}(S_2))} = \sigma_{S_2}.$$

To verify the reverse implication assume that $\sigma_{S_1} \leq \sigma_{S_2}$and assume by contradiction that there exists some $\mathbf{x}_0 \in \text{cl}(\text{co}(S_1)$ which does not belong to $\text{cl}(\text{co}(S_2))$. By Theorem 1.67 and Lemma 1.66 this implies that there exists some $\mathbf{s}_0 \in R^n$ satisfying

$$\sigma_{S_2}(\mathbf{s}_0) = \sigma_{\text{cl}(\text{co}(S_2))}(\mathbf{s}_0) < \mathbf{s}_0^\top \mathbf{x}_0 \leq \sigma_{\text{cl}(\text{co}(S_1))}(\mathbf{s}_0) = \sigma_{S_1}(\mathbf{s}_0)$$

and this contradict our initial assumption. Hence it must follow that $\text{cl}(\text{co}(S_1) \subseteq \text{cl}(\text{co}(S_2))$ and this shows the desired result. ∎

Theorem 1.68 can be used to derive composition rules for subgradients of convex functions. To introduce the next result remember that in Definition 1.13 a polar cone K^0 is introduced and applying this operation twice the bipolar cone K^{00} is given by

$$K^{00} := (K^0)^0 = \{\mathbf{x} \in R^n : \mathbf{x}^\intercal \mathbf{x}^* \leq 0 \text{ for every } \mathbf{x}^* \in K^0\}.$$

Again by Theorem 1.67 it is easy to derive an important dual representation for closed convex cones. This result is known as the *bipolar theorem* and generalizes the biorthogonality relation for linear subspaces as discussed in Lemma 1.13.

Theorem 1.69. *If $K \subseteq R^n$ is a nonempty convex cone then it follows that*

$$cl(K) = K^{00}.$$

Proof. Since the set K is a convex cone we obtain that $\text{cl}(K)$ is a closed convex cone and this implies by the definition of a support function that $\sigma_{\text{cl}(K)}(\mathbf{s}) = 0$ for every $\mathbf{s} \in K^0$ and ∞ otherwise. Using this representation and applying Theorem 1.67 it follows that

$$\begin{aligned} \mathbf{x} \quad &\in \quad \text{cl}(K) \Leftrightarrow \mathbf{s}^\intercal \mathbf{x} \leq \sigma_{\text{cl}(K)}(\mathbf{s}) \text{ for every } \mathbf{s} \in R^n \\ &\Leftrightarrow \quad \mathbf{s}^\intercal \mathbf{x} \leq 0 \text{ for every } \mathbf{s} \in K^0 \Leftrightarrow \mathbf{x} \in K^{00} \end{aligned}$$

and this shows the desired result. ∎

It is also possible to give a dual representation of the nonempty relative interior $\text{ri}(K)$ of a convex cone K. To prove this result we first observe for any nonempty linear subspace L that $\text{p}_L(\mathbf{x}^*)$ denotes the orthogonal projection of the vector $\mathbf{x}^*$ on L. By Lemma 1.13 or Lemma 1.64 we know that any $\mathbf{x}^* \in R^n$ can be uniquely written as the sum of an element of L and of $L^\perp$ and this decomposition is given by

$$\mathbf{x}^* = \text{p}_L(\mathbf{x}^*) + \text{p}_{L^\perp}(\mathbf{x}^*) \tag{1.106}$$

Taking $L = K^{\perp}$ and $L^{\perp} = (K^{\perp})^{\perp}$ in relation (1.106) it is clear for every $\mathbf{x} \in K$ with K a proper convex cone that $\mathbf{x}^{*\top}\mathbf{x} = \mathrm{p}_{(K^{\perp})^{\perp}}(\mathbf{x}^*)^{\top}\mathbf{x}$ and this implies

$$\mathbf{x}^* \in K^0 \Leftrightarrow \mathrm{p}_{(K^{\perp})^{\perp}}(\mathbf{x}^*) \in K^0. \tag{1.107}$$

To prove the dual representation of $\mathrm{ri}(K)$ we need to verify the following auxiliary result.

Lemma 1.70. *For any nonempty convex cone $K \subseteq R^m$ it follows that*

$$K^0 = K^{\perp} \Leftrightarrow K \textit{ a linear subspace} \Leftrightarrow K^0 \cap (K^{\perp})^{\perp}\backslash\{0\} \textit{ is empty.}$$

Proof. We first show that $K^0 = K^{\perp}$ implies that K is a linear subspace. Since K is a convex cone and $K^{\perp}$ is a linear subspace it follows by Theorem 1.69 and $K^0 = K^{\perp}$ that

$$\mathrm{cl}(K) = (K^0)^0 = (K^{\perp})^0 = (K^{\perp})^{\perp}$$

and so $\mathrm{cl}(K)$ is a linear subspace. By the convexity of the set K we obtain by Lemma 1.28 that $\mathrm{ri}(\mathrm{cl}(K)) = \mathrm{ri}(K)$ and since the linear subspace $\mathrm{cl}(K)$ is relatively open this implies

$$\mathrm{cl}(K) = \mathrm{ri}(\mathrm{cl}(K)) = \mathrm{ri}(K) \subseteq K.$$

Trivially $K \subseteq \mathrm{ri}(K)$ and this yields that K equals $\mathrm{cl}(K)$ and is therefore a linear subspace. To show the result that K is a linear subspace implies $K^0 \cap (K^{\perp})^{\perp}\backslash\{0\}$ is empty it follows by our assumption that $K^0 = K^{\perp}$. It is easy to verify that $K^{\perp} \cap (K^{\perp})^{\perp} = \{0\}$ and so $K^0 \cap (K^{\perp})^{\perp}\backslash\{0\}$ is empty. To prove the last implication we assume by contradiction that $K^0\backslash K^{\perp}$ is nonempty. This implies by relations (1.105) and (1.106) that

$$\mathrm{p}_{(K^{\perp})^{\perp}}(\mathbf{x}^*) \neq 0 \text{ and } \mathrm{p}_{(K^{\perp})^{\perp}}(\mathbf{x}^*) \in K^0 \cap (K^{\perp})^{\perp}$$

for every $\mathbf{x}^* \in K^0\backslash K^{\perp}$ and so the vector $\mathrm{p}_{(K^{\perp})^{\perp}}(\mathbf{x}^*)$ belongs to the set $K^0 \cap (K^{\perp})^{\perp}\backslash\{0\}$. This contradicts our assumption and the lemma is proved. ∎

As shown by the following example the convexity of the cone is needed in the above result.

Example 1.13. *Consider the nonconvex cone $K = \{0\} \times (R\backslash\{0\}) \subseteq R^2$. For this cone it is easy to verify that $K^0 = K^{\perp} = R \times \{0\}$ and since 0 does not belong to K it follows that K is not a linear subspace.*

It is now possible to prove the following dual representation of $\mathrm{ri}(K)$ for any convex cone K.

Theorem 1.71. *For any nonempty convex cone $K \subseteq R^m$ it follows that*

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

Proof. In case K is a linear subspace it follows that $\mathrm{ri}(K) = K$ and by Lemma 1.70 we obtain that $K = (K^{\perp})^{\perp}$ and the second condition always holds. The reverse implication is also a direct consequence of Lemma 1.70 and so we only need to prove the result for K not a linear subspace. To prove the implication $\Rightarrow$ we first observe that $\mathrm{ri}(K) \subseteq \mathrm{aff}(K)$ and due to $0 \in \mathrm{aff}(K)$ we obtain by relation (1.16) that $\mathrm{aff}(K) = (K^{\perp})^{\perp}$. Consider now an arbitrary $\mathbf{x}^*$ belonging to $K^0 \cap (K^{\perp})^{\perp} \backslash \{0\}$ and since by assumption $\mathbf{x} \in \mathrm{ri}(K)$ there exists some $\epsilon > 0$ satisfying $\mathbf{x}+\epsilon\mathbf{x}^* \in K$. Due to $\mathbf{x}^* \in K^0 \backslash \{0\}$ this implies that

$$\mathbf{x}^{*\top}\mathbf{x} = \mathbf{x}^{*\top}(\mathbf{x}+\epsilon\mathbf{x}^*) - \epsilon\|\mathbf{x}^*\|^2 < 0$$

and we have shown the desired result. To verify the reverse implication consider some point $\mathbf{x}$ satisfying

$$\mathbf{x} \in (K^{\perp})^{\perp} \text{ and } \mathbf{x}^{*\top}\mathbf{x} < 0 \text{ for every } \mathbf{x}^* \in K^0 \cap (K^{\perp})^{\perp} \backslash \{0\}. \tag{1.108}$$

By the first part and $\mathrm{ri}(K)$ is nonempty every point $\mathbf{x} \in \mathrm{ri}(K)$ satisfies relation (1.108). Introducing the optimization problem

$$s(\mathbf{x}) := \sup\{\mathbf{x}^{*\top}\mathbf{x} : \mathbf{x}^* \in K^0 \cap (K^{\perp})^{\perp} \text{ and } \|\mathbf{x}^*\| = 1\}$$

we obtain by the compactness of the feasible region and the continuity of the objective function that by the Weierstrass-Lebesgue theorem (Theorem 1.40) an optimal solution exists and this implies by relation (1.108) that $s(\mathbf{x}) < 0$. If it can be shown that

$$(x - s(\mathbf{x})E) \cap \mathrm{aff}(K^{00}) \subseteq K^{00} \tag{1.109}$$

it follows by Theorem 1.69 and $\mathbf{x} \in (K^{\perp})^{\perp} = ((K^{00})^{\perp})^{\perp}$ that

$$\mathbf{x} \in \mathrm{ri}(K^{00}) = \mathrm{ri}(\mathrm{cl}(K)) = \mathrm{ri}(K)$$

and hence the desired result is proved. By this observation it is sufficient to verify relation (1.109) and to do this consider an arbitrary point $\mathbf{x}-s(\mathbf{x})\mathbf{y}$ belonging to $(\mathbf{x} - s(\mathbf{x})E) \cap \mathrm{aff}(K^{00})$. For this point we need to check that it belongs to K^{00} or equivalently

$$\boldsymbol{\lambda}^{\top}(\mathbf{x}-s(\mathbf{x})\mathbf{y}) \leq 0 \text{ for every } \boldsymbol{\lambda} \in K^0.$$

Due to $\mathbf{x}-s(\mathbf{x})\mathbf{y} \in \mathrm{aff}(K^{00}) = (K^{\perp})^{\perp}$ it follows for every $\boldsymbol{\lambda} \in K^0$ that

$$\begin{aligned} \boldsymbol{\lambda}^{\top}(\mathbf{x} - s(\mathbf{x})\mathbf{y}) &= \mathrm{p}_{(K^{\perp})^{\perp}}(\boldsymbol{\lambda})^{\top}(\mathbf{x} - s(\mathbf{x})\mathbf{y}) \\ &= \mathbf{x}^{\top}\mathrm{p}_{(K^{\perp})^{\perp}}(\boldsymbol{\lambda}) - s(\mathbf{x})\mathbf{y}^{\top}\mathrm{p}_{(K^{\perp})^{\perp}}(\boldsymbol{\lambda}). \end{aligned} \tag{1.110}$$

Applying the Cauchy-Schwartz inequality given by relation (1.98) to the last part and using $\mathbf{y} \in E$ we obtain

$$\mathbf{y}^{\top}\mathrm{p}_{(K^{\perp})^{\perp}}(\boldsymbol{\lambda}) \leq \|\mathbf{y}\|\|\mathrm{p}_{(K^{\perp})^{\perp}}(\boldsymbol{\lambda})\| \leq \|\mathrm{p}_{(K^{\perp})^{\perp}}(\boldsymbol{\lambda})\|$$

This yields by relation (1.110) that

$$\boldsymbol{\lambda}^{\top}(\mathbf{x}-s(\mathbf{x})\mathbf{y}) \leq \mathbf{x}^{\top}\mathrm{p}_{(K^{\perp})^{\perp}}(\boldsymbol{\lambda}) - s(\mathbf{x})\|\mathrm{p}_{(K^{\perp})^{\perp}}(\boldsymbol{\lambda})\| \tag{1.111}$$

and so for $p_{(K^{\perp})^{\perp}}(\boldsymbol{\lambda}) = 0$ relation (1.111) reduces to

$$\boldsymbol{\lambda}^{\top}(\mathbf{x}-s(\mathbf{x})\mathbf{y}) \leq 0.$$

To verify that the above inequality also holds for $p_{(K^{\perp})^{\perp}}(\boldsymbol{\lambda}) \neq 0$ we observe by relation (1.107) that

$$p_{(K^{\perp})^{\perp}}(\boldsymbol{\lambda}) \neq 0 \Leftrightarrow p_{(K^{\perp})^{\perp}}(\boldsymbol{\lambda}) \in K^0 \cap (K^{\perp})^{\perp} \backslash \{0\}$$

By the definition of $s(\mathbf{x})$ and relation (1.111) this implies

$$\boldsymbol{\lambda}^{\top}(\mathbf{x}-s(\mathbf{x})\mathbf{y}) \leq \mathbf{x}^{\top} p_{(K^{\perp})^{\perp}}(\boldsymbol{\lambda}) - s(\mathbf{x}) \| p_{(K^{\perp})^{\perp}}(\boldsymbol{\lambda}) \| \leq 0$$

and hence also for this case the result follows. ∎

The strong separation result of Theorem 1.69 can be used to prove the following "weaker" separation result valid under a weaker condition on the point $\mathbf{y}$. In this weaker form we assume that the vector $\mathbf{y}$ does not belong to $\text{ri}(S)$. By Theorem 1.65 it is clear that we may assume without loss of generality that $\mathbf{y}$ belongs to the *relative boundary*[81] $\text{rbd}(S) := \text{cl}(S) \backslash \text{ri}(S)$ of the almost convex set $S \subseteq R^n$.

Theorem 1.72. *If $S \subseteq R^n$ is a nonempty almost convex set and $\mathbf{y}$ belongs to $rbd(S)$ then there exists some nonzero vector $\mathbf{y}^*$ belonging to the unique linear subspace $L_{aff(S)}$ parallel to the set $aff(S)$ satisfying $\mathbf{y}^{*\intercal}\mathbf{x} \geq \mathbf{y}^{*\intercal}\mathbf{y}$ for every $\mathbf{x} \in S$. Moreover, for the same vector $\mathbf{y}^*$ there exists some $\mathbf{x}_0 \in S$ such that $\mathbf{y}^{*\intercal}\mathbf{x}_0 > \mathbf{y}^{*\intercal}\mathbf{y}$.*

Proof. By Lemma 1.29 it follows that the sets $\text{ri}(S) \subseteq S$ and $\text{cl}(S)$ are nonempty convex sets. Also by Lemma 1.29 we obtain that $\text{ri}(S) = \text{ri}(\text{cl}(S))$ and so $\mathbf{y} \in \text{rbd}(S)$ does not belong to $\text{ri}(\text{cl}(S))$.Consider now for every $n \in N$ the set $(\mathbf{y} + n^{-1}E) \cap \text{aff}(S)$. Due to $\mathbf{y}$ does not belong to $\text{ri}(\text{cl}(S))$ we obtain by the definition of the relative interior that there exists some vector $\mathbf{y}_n$ satisfying

$$\mathbf{y}_n \notin \text{cl}(S) \text{ and } \mathbf{y}_n \in (\mathbf{y} + n^{-1}E) \cap \text{aff}(S) \tag{1.112}$$

By Lemma 1.29 the set $\text{cl}(S)$ is a closed convex set and so it follows by relation 1.112 and Theorem 1.69 that one can find some vector $\mathbf{y}_n^* \in R^n$ such that

$$\| \mathbf{y}_n^* \| = 1 \text{ , } \mathbf{y}_n^* \in \text{cone}(\text{cl}(S) - \mathbf{y}_n) \text{ and } \mathbf{y}_n^{*\intercal}\mathbf{x} \geq \mathbf{y}_n^{*\intercal}\mathbf{y}_n \tag{1.113}$$

for every $\mathbf{x} \in \text{cl}(S)$. Since $\mathbf{y}_n$ belongs to $\text{aff}(S)$ this implies by relation (1.16) that

$$\| \mathbf{y}_n^* \| = 1 \text{ , } \mathbf{y}_n^* \in L_{\text{aff}(S)} \text{ and } \mathbf{y}_n^{*\intercal}\mathbf{x} \geq \mathbf{y}_n^{*\intercal}\mathbf{y}_n. \tag{1.114}$$

By relation (1.114) the sequence $\{\mathbf{y}_n^* : n \in N\}$ belongs to a compact set and so by Lemma 1.3 there exist a convergent subsequence $\{\mathbf{y}_n^* : n \in N_0\}$ with

$$\lim_{n \in N_0 \to \infty} \mathbf{y}_n^* = \mathbf{y}^*. \tag{1.115}$$

[81] relative boundary

This implies by relations (1.112),(1.114) and (1.115) that

$$\mathbf{y}^{*\intercal}\mathbf{x} = \lim_{n\in N_0\to\infty} \mathbf{y}_n^{*\intercal}\mathbf{x} \geq \lim_{n\in N_0\to\infty} \mathbf{y}_n^{*\top}\mathbf{y}_n = \mathbf{y}^{*\intercal}\mathbf{y} \tag{1.116}$$

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

$$\mathbf{y}^* \in L_{\text{aff}(S)} \text{ and } \| \mathbf{y}^* \|= 1. \tag{1.117}$$

Suppose now that there does not exist a $\mathbf{x}_0 \in S$ satisfying $\mathbf{a}^\intercal\mathbf{x}_0 > \mathbf{a}^\intercal\mathbf{y}$. By relation (1.116) this implies that $\mathbf{a}^\intercal(\mathbf{x}-\mathbf{y}) = 0$ for every $\mathbf{x} \in S$ and since $\mathbf{y}$ belongs to $\text{cl}(S) \subseteq \text{aff}(S)$ we obtain that y

$$\mathbf{a}^\intercal\mathbf{z} = 0 \tag{1.118}$$

for every $\mathbf{z}$ belonging to $L_{\text{aff}(S)}$. Since by relation (1.117) the vector $\mathbf{y}^*$ belongs to $L_{\text{aff}(S)}$ we obtain by relation (1.118) that $\| \mathbf{y}^* \|^2= 0$ and this yields a contradiction with $\| \mathbf{y}^* \|= 1$. Hence it must follow that there exists some $\mathbf{x}_0 \in S$ satisfying $\mathbf{y}^{*\intercal}\mathbf{x}_0 > \mathbf{y}^{*\top}\mathbf{y}$ and this proves the desired result. ∎

The separation of Theorem 1.72 is called a *proper separation*[82] between the set S and $\mathbf{y}$. Observe by Theorems 1.65 and 1.72 it is always possible to separate a convex set C and a point.$\mathbf{y}$ outside C. If we know additionally that this point $\mathbf{y}$ does not belong to $\text{cl}(C)$ then strong separation holds while for $\mathbf{y}$ belonging to $\text{cl}(C)$ and not to C we have proper separation. The above separation results are the corner stones of convex and quasiconvex analysis. An easy consequence of these results is given by the observation that closed convex sets and relatively open convex sets are evenly convex. Remember the definition of an evenly convex set is presented in Definition 1.11 and this subset of convex sets plays an important role in duality theory for quasiconvex functions.

Lemma 1.73. *If the proper convex set $C \subseteq R^n$ is closed or relatively open then C is evenly convex.*

Proof. Since the set $C \subseteq R^n$ is proper there exists some $\mathbf{y} \notin C$ and this implies by Theorem 1.65 and C closed and convex that the set C and $\mathbf{y}$ can be strongly separated. Hence there exist some $\mathbf{a} \in R^n$ and $b \in R$ satisfying $C \subseteq H^<(\mathbf{a}, b)$ with $H^<(\mathbf{a}, b)$ denoting the open halfspace given by

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

and this shows that the set $\mathcal{H}_C^<$ of all open halfspaces containing the set C is nonempty. By the definition of the set $\mathcal{H}_C^<$ it is clear that

$$C \subseteq \cap\{H^<(\mathbf{a}, b) : H^<(\mathbf{a}, b) \in \mathcal{H}_C^<\}$$

and applying again Theorem 1.65 one can show by contradiction that

$$C = \cap\{H^<(\mathbf{a}, b) : H^<(\mathbf{a}, b) \in \mathcal{H}_C^<\}$$

Hence the closed convex set C is evenly convex and this shows the first part. To show that any relatively open convex set C is evenly convex we observe

[82] proper separation

since any closed convex set is evenly convex and $C \subseteq \mathrm{cl}(C)$ that any element of the nonempty set $\mathcal{H}^{<}_{\mathrm{cl}(C)}$ also belongs to the set $\mathcal{H}^{<}_{C}$ of all open halfspaces containing the set C and this shows that

$$C \subseteq \cap\{H^{<}(\mathbf{a}, b) : H^{<}(\mathbf{a}, b) \in \mathcal{H}^{<}_{C}\}$$

To verify that C equals $\cap\{H^{<}(\mathbf{a}, b) : H^{<}(\mathbf{a}, b) \in \mathcal{H}^{<}_{C}\}$ and hence is evenly convex we assume by contradiction that there exists some

$$\mathbf{y} \in \cap\{H^{<}(\mathbf{a}, b) : H^{<}(\mathbf{a}, b) \in \mathcal{H}^{<}_{C}\} \text{ and } \mathbf{y} \notin C. \tag{1.119}$$

If $\mathbf{y}$ does not belong to $\mathrm{cl}(C)$ then again by Theorem 1.65 one can find some open halfspace containing C and not containing $\mathbf{y}$ and this contradicts relation (1.119). Moreover, if we consider the other case that $\mathbf{y}$ belongs to $\mathrm{rbd}(C)$ we obtain by relation (1.119), Theorem 1.72 and C relatively open that one can find some nonzero $\mathbf{y}^* \in L_{\mathrm{aff}(S)}$ satisfying

$$\mathbf{y}^{*\top}\mathbf{x} \geq \mathbf{y}^{*\top}\mathbf{y} \text{ for every } \mathbf{x} \in C \tag{1.120}$$

Since C is relatively open and $\mathbf{y}^* \in L_{\mathrm{aff}(S)}$ there exists for every $\mathbf{x} \in C$ some $\epsilon > 0$ satisfying $\mathbf{x} - \epsilon\mathbf{y}^*$ belongs to C and this yields by relation (1.120) applied to $\mathbf{x} - \epsilon\mathbf{y}^*$ that

$$\mathbf{y}^{*\top}\mathbf{x} = \mathbf{y}^{*\top}(\mathbf{x} - \epsilon\mathbf{y}^*) + \epsilon\|\mathbf{y}^*\| > \mathbf{y}^{*\top}\mathbf{y}$$

Hence we can find an open halfspace containing C which does not contain $\mathbf{y}$ and again by relation (1.119) we obtain a contradiction. This shows the second part and we are done. ∎

Without proof we now mention the following result for evenly convex sets (cf. [16]).

Lemma 1.74. *If $S \subseteq R^n$ is an evenly convex set and its complement S^c is also convex then it follows that S is either empty or R^n or an open or closed halfspace.*

In the nex section we will use these results to derive dual representations for convex and quasiconvex functions.

2. Dual representations and conjugation.

In the first part of this section we will consider in detail properties of convex functions which can be derived using the above strong and weak separation results. In particular we will discuss a dual representation of a lower semicontinuous proper convex function based on relation (1.95). Moreover, in the second subsection we will discuss similar properties of quasiconvex functions and in particular we derive a dual representation of an evenly quasiconvex function.

2.1. **Dual representations and conjugation for convex functions.** In this section we consider extended real valued convex functions $f : R^n \to [-\infty, \infty]$ for which there exists some element in its domain R^n with a finite function value. At first sight these functions seem complicated and as always in mathematics one tries to approximate these complicated functions by simpler functions. For convex functions these simpler functions are given by the so-called affine minorants.

Definition 2.1. *For any function $f : R \to (-\infty, \infty]$ the affine function $a : R^n \to R$ given by $a(\mathbf{x}) = \mathbf{a}^\intercal \mathbf{x} + \alpha$ with $\mathbf{a} \in R^n$ and $\alpha \in R$ is called an affine minorant*[83] *of the function f if*

$$f(\mathbf{x}) \geq a(\mathbf{x})$$

for every $\mathbf{x}$ belonging to 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 2.1. *For any function $f : R^n \to [-\infty, \infty]$ it follows that*

$$\mathcal{A}_f = \mathcal{A}_{co(f)} = \mathcal{A}_{\overline{co(f)}}$$

Proof. We only give a proof of the above result for $\mathcal{A}_f$ nonempty. Since $\overline{\text{co}(f)} \leq \text{co}(f) \leq f$ it follows immediately that

$$\mathcal{A}_{\overline{\text{co}(f)}} \subseteq \mathcal{A}_{\text{co}(f)} \subseteq \mathcal{A}_f.$$

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 (1.83) that $a \leq \overline{\text{co}(f)}$ and hence the afiine function a belongs to $\mathcal{A}_{\overline{\text{co}(f)}}$. ∎

Since an affine function is always finite valued the set of affine minorants of a function f is empty if there exists some $\mathbf{x} \in R^n$ satisfying $f(\mathbf{x}) = -\infty$ and so it is necessary to consider functions $f : R^n \to (-\infty, \infty]$. In the next theorem we give a necessary and sufficient condition for the set $\mathcal{A}_f$ of affine minorants of the function f to be nonempty.

Theorem 2.2. *If $f : R^n \to [-\infty, \infty]$ is an arbitrary function then it follows that*

$$\mathcal{A}_f \textit{ nonempty} \Leftrightarrow co(f)(\mathbf{x}) > -\infty \textit{ for every } \mathbf{x} \in R^n.$$

Proof. If the set $\mathcal{A}_f$ is nonempty then for a given a belonging to $\mathcal{A}_f$ we obtain by definition that $a(\mathbf{x}) = \mathbf{a}^\top \mathbf{x} + \alpha$ and $a(\mathbf{x}) \leq f(\mathbf{x})$ for every $\mathbf{x} \in R^n$. Since the function a is a convex function majorized by f this implies by relation (1.81) that

$$\text{co}(f)(\mathbf{x}) \geq \mathbf{a}^\top \mathbf{x} + \alpha$$

[83] affine minorant of the function f

and this shows the desired result. To show the reverse implication we consider some f satisfying $\mathrm{co}(f)(\mathbf{x}) > -\infty$ for every $\mathbf{x} \in R^n$. In case $\mathrm{dom}(\mathrm{co}(f))$ is empty it follows that $\mathrm{co}(f)$ is identically ∞ and hence f is identically ∞ and so trivially $\mathcal{A}_f$ is nonempty. Therefore consider the case that $\mathrm{dom}(\mathrm{co}(f))$ is nonempy and since by Lemma 1.49 this is a convex set it must follow by Lemma 1.25 that $\mathrm{ri}(\mathrm{dom}(\mathrm{co}(f)))$ is nonempty. Since by construction $\mathrm{co}(f) > -\infty$ is a convex function it follows by Lemma 1.48 for $\mathbf{x}_0 \in \mathrm{ri}(\mathrm{dom}(\mathrm{co}(f)))$ that

$$-\infty < \mathrm{co}(f)(\mathbf{x}_0) = \overline{\mathrm{co}(f)}(\mathbf{x}_0) < \infty. \tag{2.1}$$

This implies that the point $(\mathbf{x}_0, \mathrm{co}(f)(\mathbf{x}_0) - \epsilon)$ with $\epsilon > 0$ does not belong to the set $\mathrm{epi}(\overline{\mathrm{co}(f)})$ and by Lemma 1.51 this set equals the nonempty closed convex set $\mathrm{cl}(\mathrm{co}(\mathrm{epi}(f)))$. Applying now Theorem 1.65 there exists some nonzero vector $(\mathbf{x}_0^*, \beta^*)$ satisfying

$$\mathbf{x}_0^{*\top}\mathbf{x} + \beta^* r > \mathbf{x}_0^{*\top}\mathbf{x}_0 + \beta^*(\mathrm{co}(f)(\mathbf{x}_0) - \epsilon) \tag{2.2}$$

for every $(\mathbf{x}, r) \in \mathrm{epi}(\overline{\mathrm{co}(f)})$. Since by relation (2.1) the real valued vector $(\mathbf{x}_0, \mathrm{co}(f)(\mathbf{x}_0))$ belongs to $\mathrm{epi}(\overline{\mathrm{co}(f)})$ we obtain by relation (2.2) that

$$\mathbf{x}_0^{*\top}\mathbf{x}_0 + \beta^* \mathrm{co}(f)(\mathbf{x}_0) > \mathbf{x}_0^{*\top}\mathbf{x}_0 + \beta^*(\mathrm{co}(f)(\mathbf{x}_0) - \epsilon)$$

and this yields $\beta\epsilon > 0$. Due to $\epsilon > 0$ we obtain $\beta > 0$. Also by Lemma 1.49 it follows for every $\mathbf{x} \in \mathrm{co}(\mathrm{dom}(f))$ that the vector $(\mathbf{x}, r)$ with $r = \mathrm{co}(f)(\mathbf{x})$ belongs to $\mathrm{epi}(\mathrm{co}(f)) \subseteq \mathrm{epi}(\overline{\mathrm{co}(f)})$ and this implies dividing the inequality in relation (2.2) by $\beta > 0$ that

$$\mathrm{co}(f)(\mathbf{x}) \geq \mathrm{co}(f)(\mathbf{x}_0) - \frac{1}{\beta^*}\mathbf{x}_0^{*\top}(\mathbf{x} - \mathbf{x}_0) - \epsilon \tag{2.3}$$

for every $\mathbf{x} \in \mathrm{co}(\mathrm{dom}(f))$. Since $f \geq \mathrm{co}(f)$ this shows

$$f(\mathbf{x}) \geq \mathrm{co}(f)(\mathbf{x}) \geq \mathrm{co}(f)(\mathbf{x}_0) - \frac{1}{\beta^*}\mathbf{x}_0^{*\top}(\mathbf{x} - \mathbf{x}_0) - \epsilon \tag{2.4}$$

for every $\mathbf{x} \in \mathrm{dom}(f) \subseteq \mathrm{co}(\mathrm{dom}(f))$ and hence the set $\mathcal{A}_f$ of affine monorants of f is nonempty. ∎

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

Example 2.1. *For the concave function $f : R \to R$ given by $f(x) = -x^2$ it is easy to verify that $co(epi(f)) = R^2$ and $f > -\infty$. Hence we obtain that $\mathcal{A}_{co(f)}$ is empty and this yields by relation (2.5) that $\mathcal{A}_f$ is empty.*

In some cases it is difficult to check $\mathrm{co}(f) > -\infty$ and so we derive in the next lemma a sufficient condition.

Lemma 2.3. *If $f : R \to [-\infty, \infty]$ is an arbirary function and there exists some $\mathbf{y}$ satisfying $\overline{co(f)}(\mathbf{y})$ is finite then the set $\mathcal{A}_f$ is nonempty and $co(f)$ is proper.*

Proof. We may copy the same proof as used in Theorem 2.2 replacing everywhere $\text{co}(f)$ by $\overline{\text{co}(f)}$ and this proves that the set $\mathcal{A}_f$ is nonempty and hence $\text{co}(f) > -\infty$. To show that $\text{co}(f)$ is proper we still need to verify that $\text{dom}(\text{co}(f))$ is nonempty. Assume therefore by contradiction that $\text{co}(f)$ is identically ∞. Hence it must follow that f is identically ∞ and so also $\overline{\text{co}(f)}$ is identically ∞ . This contradicts our assumption that $\overline{\text{co}(f)}(\mathbf{y})$ is finite and so it must follow that $\text{co}(f)$ is proper. ∎

By Lemma 1.48 a sufficient condition to guarantee that there exists $\mathbf{y}$ satisfying $-\infty < \overline{\text{co}(f)}(\mathbf{y}) < \infty$ is the existence of some $\mathbf{y} \in \text{ri}(\text{dom}(\text{co}(f))$ satisfying $\text{co}(f)(\mathbf{y}) > -\infty$. In order to prove the next result known as Minkowski's theorem we introduce the the following set of convex functions.

Definition 2.2. *The function $f : R^n \to [-\infty, \infty]$ belongs to the set $\overline{\text{conv}(R^n)}$ if f is convex and lower semicontinuous and there exists some $\mathbf{y} \in R^n$ with $f(\mathbf{y})$ finite.*

It is now possible to prove the following important result.

Theorem 2.4. *If $f : R^n \to [-\infty, \infty]$ denotes some function with dom(f) nonempty then it follows that*

$$f \in \overline{\text{conv}(R^n)} \Leftrightarrow f(\mathbf{x}) = \sup\{a(\mathbf{x}) : a \in \mathcal{A}_f\} \text{ and } \mathcal{A}_f \text{ nonempty.}$$

Proof. If the function $f : R^n \to [-\infty, \infty]$ with $\text{dom}(f)$ nonempty has the representation

$$f(\mathbf{x}) = \sup\{a(\mathbf{x}) : a \in \mathcal{A}_f\} \text{ and } \mathcal{A}_f \text{ nonempty}$$

then clearly f is lower semicontinuous and convex. At the same time $f > -\infty$ and since $\text{dom}(f)$ is nonempty this yields that there exists some $\mathbf{y}$ with $f(\mathbf{y})$ finite and this shows that f belongs to the set $\overline{\text{conv}(R^n)}$. To prove the reverse implication we observe since f belongs to $\overline{\text{conv}(R^n)}$ that f equals $\overline{\text{co}(f)}$ and this shows by Lemma 2.3 that the set $\mathcal{A}_f$ is nonempty and hence

$$f(\mathbf{x}) \geq \sup\{a(\mathbf{x}) : a \in \mathcal{A}_f\}.$$

Suppose now by contradiction that there exists some $\mathbf{x}_0 \in R^n$ satisfying

$$f(\mathbf{x}_0) > \sup\{a(\mathbf{x}_0) : a \in \mathcal{A}_f\}. \tag{2.5}$$

If this holds there exists some finite γ such that

$$f(\mathbf{x}_0) > \gamma > \sup\{a(\mathbf{x}_0) : a \in \mathcal{A}_f\} \tag{2.6}$$

and this yields that the vector $(\mathbf{x}_0, \gamma)$ does not belong to the closed and convex epigraph $\text{epi}(f)$. By Theorem 1.65 one can now find some vector $(\mathbf{x}_0^*, \beta^*)$ and $\epsilon > 0$ satisfying

$$\mathbf{x}_0^{*\top}\mathbf{x} + \beta^* r \geq \mathbf{x}_0^{*\top}\mathbf{x}_0 + \beta^*\gamma + \epsilon \tag{2.7}$$

for every $(\mathbf{x}, r)$ belonging to the nonempty set $\text{epi}(f)$. Since for every $h > 0$ the vector $(\mathbf{x}, r + h)$ belongs to $\text{epi}(f)$ for $(\mathbf{x}, r) \in \text{epi}(f)$ it must follow by relation (2.7) that $\beta^* \geq 0$. If it happens that $f(\mathbf{x}_0) < \infty$ then we know that

$(\mathbf{x}_0, f(\mathbf{x}_0))$ belongs to epi(f) and this implies again by relation (2.7) that $\beta^*(f(\mathbf{x}_0) - \beta) > 0$. Due to $f(\mathbf{x}_0) - \beta > 0$ this yields $\beta^* > 0$ and so we obtain by relation (2.7) that

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

for every $\mathbf{x}$ belonging to dom(f). Hence we have found some $a \in \mathcal{A}_f$ satisfying $a(\mathbf{x}_0) = \beta$ and this contradicts relation (2.6). If $f(\mathbf{x}_0) = \infty$ and in relation (2.7) the scalar β^* is positive then by a similar proof we obtain a contradiction and so we consider the last case $f(\mathbf{x}_0) = \infty$ and $\beta^* = 0$. Hence it follows by relation (2.7) that

$$-\mathbf{x}_0^{*\top}(\mathbf{x} - \mathbf{x}_0) + \epsilon \leq 0 \tag{2.8}$$

for every $\mathbf{x}$ belonging to dom(f). Consider now the affine function $a_0 : R^n \to R$ given by

$$a_0(\mathbf{x}) = -\mathbf{x}_0^{*\top}(\mathbf{x} - \mathbf{x}_0) + \frac{1}{2}\epsilon$$

It is clear that $a_0(\mathbf{x}_0) > 0$ and by relation (2.8) we obtain

$$a_0(\mathbf{x}) < -\mathbf{x}_0^{*\top}(\mathbf{x} - \mathbf{x}_0) + \epsilon \leq 0 \tag{2.9}$$

for every $\mathbf{x} \in$ dom(f). Since the set $\mathcal{A}_f$ is nonempty consider now an arbitrary function a belonging to this set. By relation (2.9) it follows for every $\lambda > 0$ that the affine function

$$a(\mathbf{x}) + \lambda a_0(\mathbf{x}) \leq a(\mathbf{x}) \leq f(\mathbf{x}) \tag{2.10}$$

for every $\mathbf{x} \in$ dom(f) and so this affine function is an affine minorant of f . By relation (2.6) and $a_0(\mathbf{x}_0) > 0$ we obtain that scalar

$$\lambda_0 := \frac{\beta - a(\mathbf{x}_0)}{a_0(\mathbf{x}_0)} > 0$$

and this shows that

$$a(\mathbf{x}_0) + \lambda_0 a_0(\mathbf{x}_0) = \beta \tag{2.11}$$

Hence by relations (2.10) and (2.11) we obtain a contradiction with relation (2.6) and this shows the desired result. ∎

An immediate consequence of Minkowski's theorem and Lemma 2.1 is listed in the next result.

Lemma 2.5. *If $f : R^n \to [-\infty, \infty]$ is an arbitrary function satisfying $-\infty < \overline{co(f)}(\mathbf{y}) < \infty$ for some $\mathbf{y}$ then it follows that $\mathcal{A}_f$ is nonempty and*

$$\overline{co(f)}(\mathbf{x}) = \sup\{a(\mathbf{x}) : a \in \mathcal{A}_f\}.$$

In Theorem 2.4 we only guarantee that any function $f \in \overline{\text{conv}(R^n)}$ can be approximated from below by affine functions. However, it is sometimes useful to derive an approximation formula in terms of the origianl function

f. This formula was first constructed in its general form by Fenchel and it has an easy geometrical interpretation.

Definition 2.3. *Let $f : R^n \to [-\infty, \infty]$ be an arbitrary function. The conjugate function* [84] *$f^* : R^n \to [-\infty, \infty]$ of the function f or the Legendre-Young-Fenchel transform f^* of f is given by*

$$f^*(\mathbf{x}^*) = \sup\{\mathbf{x}^{*\top}\mathbf{x} - f(\mathbf{x}) : \mathbf{x} \in R^n\}.$$

Morever the biconjugate function [85] *$f^{**} : R^n \to [-\infty, \infty]$ of the function f is given by*

$$f^{**}(\mathbf{x}) := (f^*)^*(\mathbf{x}) := \sup\{\mathbf{x}^\top\mathbf{x}^* - f^*(\mathbf{x}^*) : \mathbf{x}^* \in R^n\}$$

By the above definition it is immediately clear that conjugate function f^* is convex and lower semicontinuous. Moreover, if the function $f : R^n \to [-\infty, \infty]$ 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 2.6. *If the set $\mathcal{A}_f$ of affine minorants of the function f is nonempty then it follows that*

$$(\mathbf{x}^*, r) \in epi(f^*) \Leftrightarrow a \in \mathcal{A}_f \text{ with } a(\mathbf{x}) = \mathbf{x}^{*\top}\mathbf{x} - r$$

and

$$f^{**}(\mathbf{x}) = \sup\{a(\mathbf{x}) : a \in \mathcal{A}_f\}.$$

Proof. To verify the first equivalence relation we observe for the affine minorant $a(\mathbf{x}) = \mathbf{x}^{*\top}\mathbf{x} - r \leq f(\mathbf{x})$ for every $\mathbf{x} \in R^n$ that

$$r \geq f^*(\mathbf{x}^*) = \sup\{\mathbf{x}^{*\top}\mathbf{x} - f(\mathbf{x}) : \mathbf{x} \in R^n\}$$

or equivalently $(\mathbf{x}^*, r)$ belonging to $\text{epi}(f^*)$. Moreover, if $(\mathbf{x}^*, r) \in R^{n+1}$ belongs to $\text{epi}(f^*)$ we obtain by the definition of an epigraph that $r \geq f^*(\mathbf{x}^*)$ and this implies for every $\mathbf{x} \in R^n$ that

$$a(\mathbf{x}) = \mathbf{x}^{*\top}\mathbf{x} - r \leq f(\mathbf{x}).$$

To prove the second equality we observe by definition that

$$f^{**}(\mathbf{x}) = \sup\{\mathbf{x}^\top\mathbf{x}^* - r : (\mathbf{x}^*, r) \in \text{epi}(f^*)\}.$$

Since by the first part $(\mathbf{x}^*, r) \in \text{epi}(f^*)$ if and only if $a(\mathbf{x}) = \mathbf{x}^{*\top}\mathbf{x} - r$ is an affine minorant of the function f this shows that

$$f^{**}(\mathbf{x}) = \sup\{a(\mathbf{x}) : a \in \mathcal{A}_f\}$$

and hence the second equality is verified. ∎

To prove one of the most important theorems in convex analysis we need to introduce the next definition.

[84] conjugate function of f

[85] biconjugate function

Definition 2.4. *If $f : R^n \to [-\infty, \infty]$ is an arbitrary function then the closure $cl(f) : R^n \to [-\infty, \infty]$ of the function f is given by*

$$cl(f) = \overline{f} \text{ if } \overline{f} > -\infty \text{ and } cl(f) = -\infty \text{ otherwise.}$$

Clearly the function $\mathrm{cl}(f)$ is lower semicontinuous and satisfies $\mathrm{cl}(f) \leq \overline{f}$. The next result is well-known and is known as the Fenchel-Moreau theorem.

Theorem 2.7. *For any function $f : R^n \to [-\infty, \infty]$ it follows that*

$$f^{**} = cl(co(f)).$$

Proof. If there exists some $\mathbf{x}_0$ satisfying $\overline{\mathrm{co}(f)}(\mathbf{x}_0) = -\infty$ then it follows in case $f^*(\mathbf{x}_0^*) < \infty$ for some $\mathbf{x}_0^*$ that there exists some finite α satisfying $\alpha \geq \mathbf{x}_0^{*\top}\mathbf{x} - f(\mathbf{x})$ for every $\mathbf{x} \in R^n$. This shows that the function $\mathbf{x} \to \mathbf{x}_0^{*\top}\mathbf{x} - \alpha$ is an affine minorant of the function f and so we obtain by relation (1.83) that $\overline{\mathrm{co}(f)}(\mathbf{x}_0) > -\infty$. This yields a contradiction and therefore f^* is identically ∞. This implies f^{**} is identically $-\infty$ and by the definition of the closure we obtain $f^{**} = \mathrm{cl}(\mathrm{co}(f))$. In case $\overline{\mathrm{co}(f)} > -\infty$ and there exists some $\mathbf{x}_0$ satisfying $\overline{\mathrm{co}(f)}(\mathbf{x}_0)$ is finite the result follows by Lemma 2.5 and Lemma 2.6. Finally if $\overline{\mathrm{co}(f)}$ is identically ∞ then clearly f is identically ∞ and so $f^*(\mathbf{x}) = -\infty$ and $f^{**}(\mathbf{x}) = \infty$ for every $\mathbf{x} \in R^n$. ∎

An important consequence of the Fenchel-Moreau theorem is given by the next result.

Theorem 2.8. *If the lower semicontinuous hull $\overline{f}$ of the function $f : R^n \to [-\infty, \infty]$ is convex then it follows that*

$$f^{**}(\mathbf{x}) = \overline{f}(\mathbf{x}) = \lim \inf_{\mathbf{y} \to \mathbf{x}} f(\mathbf{y})$$

for every $\mathbf{x}$ belonging to $dom(\overline{f})$.

Proof. It is well-known by Lemma 1.46 that $\overline{f}(\mathbf{x}) = \lim\inf_{\mathbf{y} \to \mathbf{x}} f(\mathbf{y})$. Moreover, since $\overline{f}$ is convex we obtain that $\overline{f}$ is a lower semicontinous convex function majorized by f and this shows by Lemma 1.51 that $\overline{f} \leq \overline{\mathrm{co}(f)}$. Trivially $\overline{\mathrm{co}(f)} \leq \overline{f}$ and hence we obtain that $\overline{\mathrm{co}(f)} = \overline{f}$. Since by assumption $\mathbf{x}$ belongs to $\mathrm{dom}(\overline{f})$ it follows that either $\overline{\mathrm{co}(f)}(\mathbf{x})$ is finite or $\overline{\mathrm{co}(f)}(\mathbf{x}) = -\infty$. If it happens that $\overline{\mathrm{co}(f)}(\mathbf{x}) = -\infty$ then $\mathrm{cl}(\mathrm{co}(f))$ is identically $-\infty$ and by the Fenchel-Moreau theorem we obtain that

$$f^{**}(\mathbf{x}) = -\infty = \overline{\mathrm{co}(f)}(\mathbf{x}) = \overline{f}(\mathbf{x}).$$

Also, if $\overline{\mathrm{co}(f)}(\mathbf{x})$ is finite then it follows that $\overline{\mathrm{co}(f)}$ is proper and by the definition of the closure it follows that $\mathrm{cl}(\mathrm{co}(f)) = \overline{\mathrm{co}(f)} = \overline{f}$. Applying now the Fenchel-Moreau theorem yields the desired result. ∎

Looking back at the proof of Theorem 2.2 we observe that the inequality in relation (2.3) is so important that it has been given a special name.

Definition 2.5. *For any function* $f : R^n \to (-\infty, \infty]$, $\mathbf{x}_0 \in R$ *and* $\epsilon > 0$ *the subset of* R^n *consisting of those vectors* $\mathbf{x}_0^*$ *satisfying*

$$f(\mathbf{x}) \geq f(\mathbf{x}_0) + \mathbf{x}_0^{*\top}(\mathbf{x} - \mathbf{x}_0) - \epsilon$$

for every $\mathbf{x} \in R^n$ *is called the* $\epsilon-$*subgradient set*[86] *of the function* f *at the point* $\mathbf{x}_0$. *This set is denoted by* $\partial_\epsilon f(\mathbf{x}_0)$ *and its elements are called* $\epsilon-$*subgradients*[87]. *Moreover, the set* $dom(\partial_\epsilon f)$ *is given by*

$$dom(\partial_\epsilon f) := \{\mathbf{x} \in R^n : \partial_\epsilon f(\mathbf{x}) \neq \emptyset\}$$

By Theorem 2.4 it follows that any convex function f: $R^n \to [-\infty, \infty]$ with $f(\mathbf{x}_0) > -\infty$ for some $\mathbf{x}_0 \in \text{ri}(\text{dom}(f))$ has a γ-subgradient at $\mathbf{x}_0$ for any $\gamma > 0$. Actually by using the stronger proper separation result one can show the following improvement for any $\mathbf{x}_0$ belonging to $\text{ri}(\text{dom}(f))$.

Theorem 2.9. *If* $f : R^n \to [-\infty, \infty]$ *is a convex function with* $f(\mathbf{x}_0)$ *finite for some* $\mathbf{x}_0$ *belonging to* $ri(dom(f))$ *then there exists some* $\mathbf{x}_0^*$ *belonging to the linear subspace* $L_{aff(dom(f))}$ *parallel to* $aff(dom(f))$ *satisfying*

$$f(\mathbf{x}) \geq f(\mathbf{x}_0) + \mathbf{x}_0^{*\top}(\mathbf{x} - \mathbf{x}_0)$$

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

Proof. Consider $\mathbf{x}_0 \in \text{ri}(\text{dom}(f))$ satisfying $f(\mathbf{x}_0)$ is finite. Since by Lemma 1.48 we obtain that $\overline{f}(\mathbf{x}_0) = f(\mathbf{x}_0)$ it follows by Lemma 2.5 that the convex function f is proper and by Lemma 1.47 this yields

$$(\mathbf{x}_0, f(\mathbf{x}_0)) \text{ belongs to } \text{rbd}(\text{epi}(f))$$

Since $\text{epi}(f)$ is a nonempty convex set we may apply Theorem 1.72 and so there exists some nonzero vector $(\mathbf{x}_0^*, \beta^*)$ belonging to the linear subspace $L_{\text{aff}(\text{epi}(f)}$ parallel to $\text{aff}(\text{epi}(f))$ satisfying

$$\mathbf{x}_0^{*\top}\mathbf{x} + \beta^* r \geq \mathbf{x}_0^{*\top}\mathbf{x}_0 + \beta^* f(\mathbf{x}_0) \tag{2.12}$$

for every $(\mathbf{x}, r)$ belonging to $\text{epi}(f)$. Since

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

and so $L_{\text{aff}(\text{epi}(f))} = L_{\text{aff}(\text{dom}(f))} \times R$ we obtain that $\mathbf{x}_0^*$ belongs to the linear subspace $L_{\text{aff}(\text{dom}(f))}$ parallel to $\text{aff}(\text{dom}(f))$ and so for every $t > 0$ it follows that

$$\mathbf{x}_0 + t\mathbf{x}_0^* \in \text{aff}(\text{dom}(f)). \tag{2.13}$$

Since we additionally know that $\mathbf{x}_0$ belongs to $\text{ri}(\text{dom}(f))$ there exists using relation (2.13) some $\epsilon > 0$ satisfying $\mathbf{x}_0 - \epsilon\mathbf{x}_0^* \in \text{dom}(f)$ and this implies by relation (2.12) for $r = f(\mathbf{x}_0 - \epsilon\mathbf{x}_0^*)$ that

$$-\epsilon \parallel \mathbf{x}_0^* \parallel^2 + \beta f(\mathbf{x}_0 - \epsilon\mathbf{x}_0^*) \geq \beta f(\mathbf{x}_0) \tag{2.14}$$

[86] γ-subgradient set of f at x_0

[87] $\gamma-$subgradients

Clearly by relation (2.12) it must follow that $\beta \geq 0$. To check that $\beta > 0$ we assume by contradiction that $\beta = 0$ and this yields by relation (2.14) that $\mathbf{x}_0^* = 0$. Hence the normal vector $(\mathbf{x}_0^*, \beta^*)$ must be the zero vector and this contradicts $(\mathbf{x}_0^*, \beta^*)$ nonzero. Dividing the expression in relation (2.12) by $\beta^* > 0$ yields for $r = f(\mathbf{x}), \mathbf{x} \in \text{dom}(f)$ that

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

and this shows with $\boldsymbol{\gamma}_0^* = -\frac{1}{\beta}\mathbf{x}_0^*$ belonging to the linear subspace $\mathcal{L}_{\text{aff}(\text{dom}(f))}$ parallel to aff(dom(f)) that the result holds. ∎

Observe the vector $\boldsymbol{\gamma}_0^* \in L_{\text{aff}(\text{dom}(f))}$ in the above theorem has a special name and as one might expect this is called a subgradient.

Definition 2.6. *For any function $f : R^n \to (-\infty, \infty]$ and $\mathbf{x}_0 \in R$ the subset of R^n consisting of those vectors $\mathbf{x}_0^*$ satisfying*

$$f(\mathbf{x}) \geq f(\mathbf{x}_0) + \mathbf{x}_0^{*\intercal}(\mathbf{x} - \mathbf{x}_0)$$

for every $\mathbf{x} \in R^n$ is called the subgradient set[88] *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*[89]*. Moreover, the set dom(∂f) is given by*

$$dom(\partial f) := \{\mathbf{x} \in R^n : \partial f(\mathbf{x}) \neq \emptyset\}$$

The next result can now be easily verified.

Lemma 2.10. *If $f : R^n \to [-\infty, \infty]$ is a convex function with $f(\mathbf{x}_0)$ finite for some $\mathbf{x}_0 \in ri(dom(f))$ then it follows that*

$$ri(dom(f)) \subseteq dom(\partial f) \subseteq dom(f)$$

Moreover, the set $\partial f(\mathbf{x})$ is closed and convex.

Proof. By Theorem 2.9 it follows immediately that ri(dom(f)) $\subseteq$ dom(∂f). To verify the other inclusion we know that dom(f) is nonempty and so let $\mathbf{y} \in \text{dom}(f)$ and consider an arbitrary $\mathbf{x} \in \text{dom}(\partial f)$. By the definition of the subgradient it follows with $\mathbf{x}^*$ belonging to $\partial f(\mathbf{x})$ that

$$\infty > f(\mathbf{y}) \geq f(\mathbf{x}) + \mathbf{x}^{*\intercal}(\mathbf{y} - \mathbf{x})$$

and this shows that $\mathbf{x}$ belongs to dom(f). The last part of the above lemma is esay to verify and so we omit it. ∎

By the following example it is shown that a subgradient might not exist for $\mathbf{x}$ belonging to the relative boundary of dom(f).

Example 2.2. *Consider the convex function $f : R \to (-\infty, \infty]$ given by*

$$f(x) = -\sqrt{x} \text{ for } x \geq 0 \text{ and } f(x) = \infty \text{ otherwise.}$$

Clearly 0 *belongs to the relative boundary of dom(f) but $\partial f(0)$ is empty.*

[88] subgradient set of f at x_0
[89] subgradients

In case the subgradient of a function at a certain point exists one can also prove the following relation between f and f^*.

Lemma 2.11. *For $f: R^n \to [-\infty, \infty]$ be a proper function it follows that*

$$\mathbf{x}_0^* \in \partial f(\mathbf{x}_0) \Leftrightarrow f(\mathbf{x}_0) + f^*(\mathbf{x}_0) = \mathbf{x}_0^\top \mathbf{x}_0^*.$$

Proof. If $\mathbf{x}_0^* \in \partial f(\mathbf{x}_0)$ it follows by definition that $f(\mathbf{x}) \geq f(\mathbf{x}_0) + \mathbf{x}_0^{*\top}(\mathbf{x} - \mathbf{x}_0)$. In case $f(\mathbf{x}_0) = \infty$ this implies that $f(\mathbf{x}) = \infty$ for every $\mathbf{x} \in R^n$ and this contradict that f is proper. Therefore it must follow that $f(\mathbf{x}_0)$ is finite and so we obtain for every $\mathbf{x} \in R^n$ that

$$\mathbf{x}_0^{*\top}\mathbf{x}_0 - f(\mathbf{x}_0) \geq \mathbf{x}^\top \mathbf{x}_0^* - f(\mathbf{x})$$

This yields by the definition of the conjugate function that

$$\mathbf{x}_0^{*\top}\mathbf{x}_0 - f(\mathbf{x}_0) \geq f^*(\mathbf{x}_0^*)$$

and since trivially $f^*(\mathbf{x}_0) \leq \mathbf{x}_0^{*\top}\mathbf{x}_0 - f(\mathbf{x}_0)$ the desired equality follows. To show the reverse iimplication we observe for every $\mathbf{x} \in R^n$ that

$$\mathbf{x}_0^{*\top}\mathbf{x} - f(\mathbf{x}) \leq f^*(\mathbf{x}_0^*) = \mathbf{x}_0^\top \mathbf{x}_0^* - f(\mathbf{x}_0).$$

Since f is proper this shows for every $\mathbf{x}$ belonging to $\text{dom}(f)$ that

$$f(\mathbf{x}) \geq f(\mathbf{x}_0) + \mathbf{x}_0^{*\top}(\mathbf{x} - \mathbf{x}_0)$$

and this shows $\mathbf{x}_0^*$ belongs to $\partial f(\mathbf{x}_0)$. ∎

A direct consequence of the above result is given by the following improvement of the Fenchel-Moreau theorem.

Lemma 2.12. *For $f: R^n \to (-\infty, \infty]$ a proper function it follows for every $\mathbf{x}_0$ belonging to dom(∂f) and $\mathbf{x}_0^* \in \partial f(\mathbf{x}_0)$ that*

$$f^{**}(\mathbf{x}_0) = \mathbf{x}_0^{*\top}\mathbf{x}_0 - f^*(\mathbf{x}_0^*) = f(\mathbf{x}_0)$$

Proof. It is easy to verify that $f^{**}(\mathbf{x}_0) \leq f(\mathbf{x}_0)$ and this shows by Lemma 2.11 that

$$f(\mathbf{x}_0) \geq f^{**}(\mathbf{x}_0) \geq \mathbf{x}_0^{*\top}\mathbf{x}_0 - f^*(\mathbf{x}_0^*) = f(\mathbf{x}_0).$$

for every $\mathbf{x}_0^*$ belonging to $\partial f(\mathbf{x}_0)$. ∎

Convex functions have remarkable continuity properties. Before mentioning the main result we introduce the next definition.

Definition 2.7. *A function $f: R^n \to R$ is called Lipschitz continuous*[90] *on the set $S \subseteq R^n$ if there exists some finite constant L (the so-called Lipschitz constant) satisfying*

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

for every $\mathbf{x}_1, \mathbf{x}_2$ belonging to S.

Without proof we now list the following result (cf.[8]).

[90] Lipschitz continuous function

Lemma 2.13. *If $f : R^n \to [-\infty, \infty]$ is a proper convex function with $f(\mathbf{y})$ finite for some $\mathbf{y} \in R^n$ then it follows that the function f is continuous on $ri(dom(f))$ and Lipschitz continuous on every compact subset of $ri(dom(f))$.*

Finally at the end of this section we will consider the general duality framework for optimization problems. To do so we consider for a given function $f : R^m \to [-\infty, \infty]$ the optimization problem

$$\inf\{f(\mathbf{x}) : \mathbf{x} \in R^m\}.$$

Since f represents an extended real valued function the above optimization problem also covers constraint optimization problems. Associate with the function f a function $F : R^m \times R^n \to [-\infty, \infty]$ satisfying $F(\mathbf{x}, \mathbf{0}) = f(\mathbf{x})$ and consider the so-called perturbation function $p : R^n \to [-\infty, \infty]$ defined by

$$p(\mathbf{y}) = \inf\{F(\mathbf{x}, \mathbf{y}) : \mathbf{x} \in R^m\}$$

By the definition of the function F we obtain that

$$p(\mathbf{0}) = \inf\{f(\mathbf{x}) : \mathbf{x} \in R^m\}.$$

and an example of such a function and the corresponding perturbation function is listed in Example 1.10. The dual problem of the above primal problem is now given by

$$p^{**}(\mathbf{0}) = \sup\{-p^*(\mathbf{y}) : \mathbf{y} \in R^n\}.$$

Clearly by the Fenchel-Moreau theorem we obtain that $p^{**}(\mathbf{0}) = \mathrm{cl}(p)(\mathbf{0})$ and this yields immediately for p a proper convex function that

$$p^{**}(\mathbf{0}) = p(\mathbf{0}) \Leftrightarrow p \text{ is lower semicontinuous in } \mathbf{0}$$

Moreover, if $\mathbf{0}$ belongs to $\mathrm{ri}(\mathrm{dom}(p))$ we obtain that the subgradient set of p at $\mathbf{0}$ is nonempty and by Lemma 2.12 any element from this subgradient set is an optimal solution of the dual problem. This concludes our discussion of the general framework of duality theory for optimization programs. For a detailed analysis of the Lagrangian perturbation function of Example 1.10 and its relation with the so-called cone convexlike vector valued functions the reader is referred to [6].

2.2. **Dual representations and conjugation for quasiconvex functions.** In this section we primarily study duality results for the class of evenly quasiconvex functions. Most of the the results of this section can be found in [16]. Unfortunately in this paper no geometrical interpretation of the results are given and for a clear geometrical interpretation the reader should consult [7]. In these papers 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 quasi-affine functions. As in convex analysis the used approximations and the generalized bi-conjugate functions have a clear geometrical interpretation (cf.[7]). Before introducing these simpler quasi-affine functions we denote by $\mathcal{G}_0$ the class of extended real valued and nondecreasing functions $c : R \to [-\infty, \infty]$. Moreover, the subclass $\mathcal{G}_1 \subseteq \mathcal{G}_0$ denotes the class of extended real valued lower semicontinuous and nondecreasing functions on R.

Definition 2.8. *A function $a_c : R^n \to [-\infty, \infty]$ is called a $\mathcal{G}_i$-affine function with $i = 0, 1$ if*

$$a_c(\mathbf{x}) = c(\mathbf{a}^\top \mathbf{x}) + \alpha$$

for some c belonging to $\mathcal{G}_i$, $\mathbf{a} \in R^n$ and $\alpha \in R$. The function a_c is called a $\mathcal{G}_i$-affine minorant of the the function $f : R^n \to [-\infty, \infty]$ if a_c is a $\mathcal{G}_i$-affine function and $a_c(\mathbf{x}) \leq f(\mathbf{x})$ for every $\mathbf{x} \in R^n$. Moreover, the set of $\mathcal{G}_i$-affine minorants of the function f is denoted by $\mathcal{G}_i\mathcal{A}_f$.

Since the function $c : R^n \to [-\infty, \infty]$ with c identically $-\infty$ belongs to the set $\mathcal{G}_1 \subseteq \mathcal{G}_0$ it follows immediately that the set of $\mathcal{G}_1$-affine minorants and the set of $\mathcal{G}_0$-affine minorants of any function $f : R^n \to [-\infty, \infty]$ is nonempty. This is a major difference with the set of affine minorants of a function f since this set might be empty. Observe in Theorem 2.2 we showed that this set is nonempty if and only if $\operatorname{co}(f) > -\infty$. We now show that a $\mathcal{G}_0$-affine function, respectively a $\mathcal{G}_1$-affine function is an evenly quasiconvex function, respectively a lower semicontinuous quasiconvex function.

Lemma 2.14. *If the function $a_c : R^n \to [-\infty, \infty]$ is $\mathcal{G}_0$-affine then it follows that a_c is evenly quasiconvex. Moreover, if the function a_c is $\mathcal{G}_1$-affine then it follows that a_c is lower semicontinuous and quasiconvex.*

Proof. To show that any $\mathcal{G}_0$-affine function is evenly quasiconvex we observe for every $r \in R$ that there exists some $c \in \mathcal{G}_0, \mathbf{a} \in R^n$ and $\alpha \in R$ such that $L(a_c, r) := \{\mathbf{x} \in R^n : c(\mathbf{a}^\top \mathbf{x}) \leq r\}$ and this shows

$$L(a_c, r) = \{\mathbf{x} \in R^n : \mathbf{a}^\top \mathbf{x} \in L(c, r)\} \tag{2.15}$$

with $L(c, r)$ denoting the lower level set of the function c of level r. Since c is nondecreasing it follows that the lower level set $L(c, r)$ for any $r \in R$ represents a (possibly empty) interval and if this interval is nonempty it has the form $(-\infty, \beta_r)$ or $(-\infty, \beta_r]$ with $\beta_r := \sup\{t : c(t) \leq r\}$. This implies by relation (2.15) that $L(a_c, r)$ is either empty, a closed or open halfspace and so we obtain that this set is evenly quasiconvex. To prove that any $\mathcal{G}_1$-affine function is a lower semicontinuous quasiconvex function we observe by the lower semicontinuity of c and Theorem 1.39 that the lower level set $L(c, r)$ is either empty or a closed set represented by $(-\infty, \beta_r)$ and this shows that the lower level set $L(a_c, r)$ is either empty or a closed halfspace. Applying now Theorem 1.39 yields the desired result. ∎

By Lemma 2.14 and Lemmas 1.54, 1.55 and 1.56 one can show similarly as in Lemma 2.1 the following result and so the proof of this result is omitted.

Lemma 2.15. *For any function* $f : R^n \to [-\infty, \infty]$ *it follows that*

$$\mathcal{G}_0\mathcal{A}_f = \mathcal{G}_0\mathcal{A}_{eqc(f)} \text{ and } \mathcal{G}_1\mathcal{A}_f = \mathcal{G}_1\mathcal{A}_{qc(f)} = \mathcal{G}_1\mathcal{A}_{\overline{qc(f)}}.$$

We will now show a generalization of Minkowsky's theorem (Theorem 2.4) valid for evenly quasiconvex and lower semicontinuous quasiconvex functions.

Theorem 2.16. *If* $f : R^n \to [-\infty, \infty]$ *is an arbitrary function then*

$$eqc(f)(\mathbf{x}) = \sup\{a_c(\mathbf{x}) : a_c \in \mathcal{G}_0\mathcal{A}_f\}.$$

Moreover, it follows that

$$\overline{qc(f)}(\mathbf{x}) = \sup\{a_c(\mathbf{x}) : a_c \in \mathcal{G}_1\mathcal{A}_f\}.$$

Proof. We only give a proof of the first formula since the proof of the second formula is similar and can be found in [16]. Since the set $\mathcal{G}_0\mathcal{A}_f$ is nonempty we obtain by Lemma 2.15 that

$$\text{eqc}(f)(\mathbf{x}) \geq \sup\{a_c(\mathbf{x}) : a_c \in \mathcal{G}_0\mathcal{A}_f\}.$$

If for some $\mathbf{x}_0$ it follows that $\text{eqc}(f)(\mathbf{x}_0) = -\infty$ the result follows immediately and so we assume that $\text{eqc}(f)(\mathbf{x}_0) > -\infty$. Suppose by contadiction that

$$\text{eqc}(f)(\mathbf{x}_0) > \sup\{a_c(\mathbf{x}_0) : a_c \in \mathcal{G}_0\mathcal{A}_f\}.$$

and let β be some finite number satisfying

$$\text{eqc}(f)(\mathbf{x}_0) > \beta > \sup\{a_c(\mathbf{x}_0) : a_c \in \mathcal{G}_0\mathcal{A}_f\}. \tag{2.16}$$

If it happens that $L(\text{eqc}(f), \beta)$ is empty it follows by relation (2.16) that $f(\mathbf{x}) \geq \text{eqc}(f)(\mathbf{x}) > \beta$ for every $\mathbf{x} \in R^n$ and so taking a_c identically equal to β we obtain that a_c is a $\mathcal{G}_0$-affine minorant of the function f and this contradicts relation (2.16). Suppose therefore that $L(\text{eqc}(f), \beta)$ is nonempty and since by Lemma 1.56 the function $\text{eqc}(f)$ is evenly quasiconvex one can find a collection of vectors $(\mathbf{a}_i, \alpha_i)_{i\in I}$ such that

$$L(\text{eqc}(f), \beta) = \cap_{i\in I} H^{<}(\mathbf{a}_i, \alpha_i) \tag{2.17}$$

By relation (2.16) it follows that the vector $\mathbf{x}_0$ does not belong to $L(\text{eqc}(f), \beta)$ and so by relation (2.17) there exists some $i_0 \in I$ satisfying

$$\mathbf{a}_{i_0}^\top \mathbf{x}_0 \geq \alpha_{i_0}. \tag{2.18}$$

Introduce now the increasing function $c : R^n \to [-\infty, \infty]$ represented by

$$c(t) = \inf\{f(\mathbf{y}) : \mathbf{y}^\top \mathbf{a}_{i_0} \geq t \text{ and } \mathbf{y} \in R^n\}$$

and consider the $\mathcal{G}_0$-affine function $a_c : R^n \to [-\infty, \infty]$ given by

$$a_c(\mathbf{x}) := c(\mathbf{a}_{i_0}^\top \mathbf{x}) = \inf\{f(\mathbf{y}) : \mathbf{y}^\top \mathbf{a}_{i_0} \geq \mathbf{x}^\top \mathbf{a}_{i_0}\}.$$

It is clear by the definition of a_c that $a_c(\mathbf{x}) \leq f(\mathbf{x})$ for every $\mathbf{x}$ belonging to R^n and so a_c is a $\mathcal{G}_0$-affine minorant of the function f. Moreover, by relation (2.18) we obtain for every $\mathbf{y}$ satisfying $\mathbf{y}^\top \mathbf{a}_{i_0} \geq \mathbf{x}_0^\top \mathbf{a}_{i_0}$ that $\mathbf{y}^\top \mathbf{a}_{i_0} \geq \alpha_{i_0}$

and this shows by relation (2.17) that $\mathbf{y}$ does not belong $L(\text{eqc}(f),\beta)$. This implies

$$\mathbf{y}^\top \mathbf{a}_{i_0} \geq \mathbf{x}_0^\top \mathbf{a}_{i_0} \Rightarrow f(\mathbf{y}) \geq \text{eqc}(f)(\mathbf{y}) \geq \beta$$

and so we obtain $a_c(\mathbf{x}_0) \geq \beta$. This yields a contradiction with relation (2.16) and the result is verified. ∎

By the above observation it is clear that for lower semicontinuous quasiconvex functions and evenly quasiconvex functions the affine functions for convex functions are replaced by respectively $\mathcal{G}_1$-affine and $\mathcal{G}_0$-affine functions. Using these functions one can also generalize the conjugate and biconjugate functions used within convex analysis.

Definition 2.9. *If $f : R^n \to [-\infty,\infty]$ is an arbitrary function and $c \in \mathcal{G}_i$ then the c-conjugate function $f^c : R^n \to [-\infty,\infty]$ of the function f is given by*

$$f^c(\mathbf{x}^*) = \sup\{c(\mathbf{x}^{*\top}\mathbf{x}) - f(\mathbf{x}) : \mathbf{x} \in R^n\}.$$

Moreover the bi-$\mathcal{G}_i$-conjugare function $f^{\mathcal{G}_i\mathcal{G}_i} : R^n \to [-\infty,\infty]$ of the function f is given by

$$f^{\mathcal{G}_i\mathcal{G}_i}(\mathbf{x}) := \sup\{c(\mathbf{x}^\top\mathbf{x}^*) - f^c(\mathbf{x}^*) : \mathbf{x}^* \in R^n, c \in \mathcal{G}_i\}.$$

By a similar proof as in Lemma 2.6 it is easy to give a geometrical interpretation of the biconjugate function.

Lemma 2.17. *For any function $f : R^n \to [-\infty,\infty]$ it follows that*

$$(\mathbf{x}^*, r) \in epi(f^c) \Leftrightarrow a_c \in \mathcal{G}_i\mathcal{A}_f \text{ with } a_c(\mathbf{x}) = c(\mathbf{x}^{*\top}\mathbf{x}) - r$$

and

$$f^{\mathcal{G}_i\mathcal{G}_i}(\mathbf{x}) = \sup\{a_c(\mathbf{x}) : a_c \in \mathcal{G}_i\mathcal{A}_f\}.$$

Combining now Lemma 2.17 and Theorem 2.16 we immediately obtain the following generalization of the Fenchel-Moreau theorem.

Theorem 2.18. *For any function $f : R^n \to [-\infty,\infty]$ it follows that $f^{\mathcal{G}_0\mathcal{G}_0} = eqc(f)$ and $f^{\mathcal{G}_1\mathcal{G}_1} = \overline{qc(f)}$*

Although the above formula yields a dual representation of an evenly quasiconvex hull and a lower semicontinuous convex hull the above formula is of no use since the definition of the bi-$\mathcal{G}_i$-conjugate function is still very complicated. Therefore we wonder whether it is possible to simplify these formulas. Indeed this is possible and this simplification is listed without proof in the following theorem for the case of the bi-$\mathcal{G}_0$ conjugate function. For this case the bi-conjugate function simplifies most and the proof of this result is purely algebraic and can be found in [16].

Theorem 2.19. *For any function $f : R^n \to [-\infty,\infty]$ it follows that*

$$f^{\mathcal{G}_0\mathcal{G}_0}(\mathbf{x}) = \sup_{\mathbf{x}^* \in R^n} \inf\{f(\mathbf{y}) : \mathbf{y}^\top\mathbf{x}^* \geq \mathbf{x}^\top\mathbf{x}^*\}$$

Combining now Theorem 2.19 and 2.18 we immediately obtain the following dual representation of the evenly quasiconvex hull eqc(f) of an arbitrary function f.

Theorem 2.20. *For any function $f : R^n \to [-\infty, \infty]$ it follows that*

$$eqc(f)(\mathbf{x}) = \sup_{\mathbf{x}^* \in R^n} \inf\{f(\mathbf{y}) : \mathbf{y}^\top \mathbf{x}^* \geq \mathbf{x}^\top \mathbf{x}^*\}$$

Actually the above result can also be proved directly copying the proof of the generalized Minkowsky theorem (Theorem 2.16). However, although it is mathematically easier to prove the above result directly we loose the geometrical interpretation of the above result. This concludes this short section on duality results for quasiconvex functions. Observe we did not introduce the concept of the c-subgradient but this can be done similarly as for convex functions. It is the intention of the author if time permits to write a book on convex and quasiconvex analysis and besides the theoretical results show how the above results can be used to build basic algorithmic procedures for convex and quasiconvex optimization problems.

References

[1] J.P. Aubin. *Optima and Equilibra (An introduction to nonlinear analysis)*, volume 140 of *Graduate Texts in Mathematics*. Springer Verlag, Berlin, 1993.

[2] E.Kreyszig. *Introductory Functional Analysis with Applications*. Wiley, New York, 1978.

[3] G.Giorgi and S.Komlosi. Dini derivatives in optimization-part II. *Rivista di matematica per le scienze economiche e sociali*, 15(2):3–23, 1992.

[4] G.Giorgio and S.Komlosi. Dini derivatives in optimization-part i. *Rivista di matematica per le scienze economiche e sociali*, 15(1):3–30, 1992.

[5] L.A.Wolsey G.L.Nemhauser. *Integer and Combinatorial Optimization*. Wiley, New York, 1988.

[6] J.B.G.Frenk and G.Kassay. Lagrangian duality and cone convexlike functions. Technical report, Econometric Institute, Erasmus university., 1999.

[7] D.M.L.Dias J.B.G.Frenk and J.Gromicho. Duality theory for convex and quasiconvex functions and its application to optimization. In T.Rapcsak S.Komlosi and S.Schaible, editors, *Generalized Convexity :Proceedings Pecs,Hungary,1992*, pages 153–171. Springer Verlag, Lecture notes in Economics and Mathematics 405, 1994.

[8] C.Lemarechal J.B.Hiriart Urruty. *Convex Analysis and Minimization Algorithms*, volume 1. Springer Verlag, 1993.

[9] J.Gromicho. *Quasiconvex Optimization and Location Theory*. Kluwer Academic publishers, Dordrecht, 1998.

[10] J.P.Crouzeix. Conditions for convexity of quasiconvex functions. *Mathematics of Operations Research*, 5(1):120–125, 1980.

[11] J.P.Crouzeix. Continuity and differentiability properties of quasiconvex functions. In S.Schaible and W.T.Ziemba, editors, *Generalized Concavity in Optimization and Economics*, pages 109–130. Academic Press, 1981.

[12] J.P.Crouzeix. *Generalized Convexity, Generalized Monotonicity*, volume 27 of *Nonconvex Optimization and its Applications*, chapter 10, pages 237–253. Kluwer Academic publishers, 1998.

[13] S.Schaible M.Avriel, W.E.Diewert and I.Zang. *Generalized Concavity*. Plenum Press, New York, 1988.

[14] C.M.Shetty M.S.Bazaraa, H.D.Sherali. *Nonlinear Programming (Theory and Applications)*. Wiley, New York, 1993.

[15] H.D.Sherali M.S.Bazaraa, J.J.Jarvis. *Linear Programming and Network Flows*. Wiley, New York, 1990.

[16] J.P. Penot and M.Volle. On quasi-convex duality. *Mathematics of Operations Research*, 15(4):597–624, 1990.

[17] P.Lancaster and M.Tismenetsky. *The Theory of Matrices*. Computer Science and Applied Mathematics. Academic Press, 1985.

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

[19] W. Rudin. Principles of mathematical analysis. *McGraw-Hill*, New York, 1976.

[20] S.Komlosi. On pseudoconvex functions. *Acta Sci.Math. (Szeged)*, 57:569–586, 1993.

[21] J. Van Tiel. *Convex Analysis an Introductory Text*. Wiley, New York, 1984.

[22] U.Passy and E.Z.Prisman. Cionjugacy in quasi-convex programming. *Mathematical Programming*, 30:121–146, 1984.

[23] W.E.Diewert. Alternative characterisations of six kind of quasiconcavity in the nondifferentiable case with applications to nonsmooth programming. In S.Schaible and W.T.Ziemba, editors, *Generalized Concavity in Optimization and Economics*, New York, 1981. Academic Press.

[24] Y.Nesterov and A.Nemirovski. *Interior Point Polynomial Algorithms in Convex Programming*. SIAM Studies in Applied Mathematics. Siam, Philadelphia, 1994.

Econometric Institute,Erasmus University,Rotterdam