



# Characterizations of the Set Less Order Relation in Nonconvex Set Optimization

Johannes Jahn<sup>1</sup>

Received: 2 March 2021 / Accepted: 2 December 2021 / Published online: 30 January 2022  
© The Author(s) 2022

## Abstract

For nonconvex set optimization problems based on the set less order relation, this paper presents characterizations of optimal sets and gives necessary conditions for set inequalities and non-optimal sets using directional derivatives. For specific order cones, the directional derivatives of known functionals describing the negative cones are also given.

**Keywords** Set less order relation · Set optimization · Necessary conditions

**Mathematics Subject Classification** 90C48 · 90C47 · 90C29

## 1 Introduction

Set optimization problems, i.e. optimization problems with set-valued objective map, are difficult to treat because inequalities of nonconvex sets have a complex structure and they are not simple to characterize. For convex sets, the characterization of set inequalities is known (for basic results see [16]) and can be simply done using certain linear functionals. But for nonconvex sets one needs more elaborate approaches, which are sometimes named scalarization (e.g. see [13] for early results and also [5,9,10,12,21,22] together with [20] for related results). These scalarization approaches work with appropriate sup inf problems.

The present paper starts with the known fact that a set inequality can be characterized by an inequality in  $\mathbb{R}$  where these sup inf problems play an important role. We restrict ourselves to the well-known set less order relation (introduced by Young [27] and named by Chiriaev/Walster [3]) so that the presented results also subsume the weaker

---

Communicated by Boris S. Mordukhovich.

---

✉ Johannes Jahn  
johannes.jahn@fau.de

<sup>1</sup> Department Mathematik, Universität Erlangen-Nürnberg, Erlangen, Germany

l-type less and u-type less order relations (see Kuroiwa [23] for first results). The so-called minmax less order relation (introduced in [19]) is only investigated for a special result.

The results of this paper consist of conditions under which the inequalities in  $\mathbb{R}$  are well-defined, characterizations of nonoptimal sets, necessary optimality conditions for the afore-mentioned sup inf problems, necessary conditions for a set inequality and a necessary condition for the non-optimality of a set. The necessary conditions of the main results work with the directional derivative of a functional being used for the description of the negative order cone in the considered real linear space. For different real linear spaces, the directional derivative is calculated for various standard functionals of this type.

This paper is organized as follows: Background material and first results are summarized in Sect. 2. Investigations of non-optimal sets and optimality conditions in nonconvex set optimization are given in Sect. 3. The last section contains the main results like necessary conditions for set inequalities.

## 2 Basic Results

Throughout this paper we use the following standard assumption.

**Assumption 2.1** Let  $Y$  be a real linear space, let  $C \subset Y$  be a convex cone, and let a functional  $\psi : Y \rightarrow \mathbb{R}$  be given with

$$\psi(y) \leq 0 \iff y \in -C. \quad (1)$$

Such a functional  $\psi$  (compare [7, Remark 3.3]), which characterizes the cone  $-C$ , is not uniquely defined. With  $\psi$  the functional  $\alpha\psi$  also has the required properties for every  $\alpha > 0$ .

For convenience, known examples of these functionals are now recalled for different spaces.

**Example 2.1** ([7, Example 3.4] for (a)–(c))

- (a) Let the real linear space  $Y := \mathbb{R}^m$  (with  $m \in \mathbb{N}$ ) be given. For the polyhedral cone

$$C := \{y \in \mathbb{R}^m \mid a_i^T y \leq 0 \text{ for all } i \in \{1, \dots, k\}\}$$

with  $k \in \mathbb{N}$  and nonzero vectors  $a_1, \dots, a_k \in \mathbb{R}^m$  the functional  $\psi : \mathbb{R}^m \rightarrow \mathbb{R}$  with

$$\psi(y) = \max_{i \in \{1, \dots, k\}} \{-a_i^T y\} \text{ for all } y \in \mathbb{R}^m$$

fulfills the equivalence (1). In the special case  $C := \mathbb{R}_+^m$  we get the functional  $\psi$  given by

$$\psi(y) = \max\{y_1, \dots, y_m\} \text{ for all } y = (y_1, \dots, y_m) \in \mathbb{R}^m.$$

The functional  $\psi : \mathbb{R}^m \rightarrow \mathbb{R}$  with

$$\psi(y) = \|(y_1, \dots, y_{m-1})\|_2 + y_m \text{ for all } y \in \mathbb{R}^m$$

(where  $m \geq 2$  and  $\|\cdot\|_2$  denotes the Euclidean norm in  $\mathbb{R}^{m-1}$ ) is associated to the *Lorentz cone*

$$C := \{y \in \mathbb{R}^m \mid \|(y_1, \dots, y_{m-1})\|_2 \leq y_m\}.$$

- (b) We now consider the real linear space  $Y := \mathcal{S}^n$  (with  $n \in \mathbb{N}$ ) of all real symmetric  $(n, n)$  matrices. Then the functional  $\psi : \mathcal{S}^n \rightarrow \mathbb{R}$  with

$$\psi(M) = \max\{\lambda_1, \dots, \lambda_n\} \text{ for all } M \in \mathcal{S}^n$$

(here  $\lambda_1, \dots, \lambda_n$  denote the  $n$  eigenvalues of the matrix  $M$ ) is associated to the well-known *Löwner cone*

$$C := \{M \in \mathcal{S}^n \mid M \text{ positive semidefinite}\} =: \mathcal{S}_+^n.$$

And the functional  $\psi : \mathcal{S}^n \rightarrow \mathbb{R}$  with

$$\psi(M) = \max_{\substack{x \in \mathbb{R}_+^n \\ \|x\|_2=1}} x^T M x \text{ for all } M \in \mathcal{S}^n$$

is associated to the *copositive cone*

$$C := \{M \in \mathcal{S}^n \mid x^T M x \geq 0 \text{ for all } x \in \mathbb{R}_+^n\}$$

(here  $\|\cdot\|_2$  denotes the Euclidean norm in  $\mathbb{R}^n$ ).

- (c) In the infinite dimensional real linear space  $Y := \mathcal{C}[a, b]$  of real-valued continuous functionals on  $[a, b]$  with  $-\infty < a < b < \infty$  the functional  $\psi : \mathcal{C}[a, b] \rightarrow \mathbb{R}$  given by

$$\psi(y) = \max_{t \in [a, b]} y(t) \text{ for all } y \in \mathcal{C}[a, b]$$

is associated to the natural ordering cone

$$C := \{y \in \mathcal{C}[a, b] \mid y(t) \geq 0 \text{ for all } t \in [a, b]\}.$$

- (d) Let  $(Y, \|\cdot\|)$  be a real normed space and let some continuous linear functional  $\ell \in Y^*$  be arbitrarily chosen. Then the functional  $\psi : Y \rightarrow \mathbb{R}$  with

$$\psi(y) = \ell(y) + \|y\| \text{ for all } y \in Y$$

is associated to the so-called *Bishop-Phelps cone*

$$C(\ell) := \{y \in Y \mid \ell(y) \geq \|y\|\}$$

(compare [11,15] for Bishop-Phelps cones).

There are various order relations, which can be used for the comparison of sets in the real linear space  $Y$  (e.g. compare [19]). In this paper we restrict ourselves to the well-known set less order relation introduced by Young [27] in 1931.

**Definition 2.1** Let Assumption 2.1 be satisfied. For nonempty subsets  $A, B \subset Y$  the set less order relation  $\preceq_s$  is defined by

$$A \preceq_s B \iff B \subset A + C \text{ and } A \subset B - C.$$

Next we characterize the set inclusions in Definition 2.1 by certain inequalities.

**Proposition 2.1** Let Assumption 2.1 be satisfied, and let the nonempty sets  $A, B \subset Y$  be arbitrarily given. Then we assert

$$B \subset A + C \begin{cases} \implies \sup_{b \in B} \inf_{a \in A} \psi(a - b) \leq 0 \\ \iff \sup_{b \in B} \min_{a \in A} \psi(a - b) \leq 0, \text{ if this min term exists} \end{cases}$$

and

$$A \subset B - C \begin{cases} \implies \sup_{a \in A} \inf_{b \in B} \psi(a - b) \leq 0 \\ \iff \sup_{a \in A} \min_{b \in B} \psi(a - b) \leq 0, \text{ if this min term exists.} \end{cases}$$

**Proof** For arbitrary nonempty sets  $A, B \subset Y$  we have

$$\begin{aligned} B \subset A + C &\iff \forall b \in B \exists a \in A : \underbrace{b \in \{a\} + C}_{\iff a - b \in -C \iff \psi(a - b) \leq 0} \\ &\begin{cases} \implies \sup_{b \in B} \inf_{a \in A} \psi(a - b) \leq 0 \\ \iff \sup_{b \in B} \min_{a \in A} \psi(a - b) \leq 0, \text{ if this min term exists.} \end{cases} \end{aligned}$$

In analogy, one can prove the second part of the assertion.  $\square$

The advantage of this simple result is that one can check the validity of the aforementioned set inclusions using appropriate optimization problems. Such a rewriting of these set inclusions as inequalities has been already given by Hernández and Rodríguez-Marín [13, Thm. 3.10,(iii)] in 2007 using an extension of the Tammer (formerly Gerstewitz) scalarization approach [8]. Later such a rewriting of set inclusions as inequalities was also given in [22, Thm. 3.3 and 3.8] for scalarizing functionals introduced by Tammer. Proposition 2.1 is parameter free and it subsumes these scalarization approaches.

Next, we investigate under which conditions the sup min problems in Proposition 2.1 are solvable.

**Proposition 2.2** *Let Assumption 2.1 be satisfied and, in addition, let  $(Y, \|\cdot\|)$  be a real normed space. Let  $A \subset Y$  be a nonempty weakly compact set, let  $B \subset Y$  be a nonempty set, let  $\psi$  be bounded on  $A - B$ , and let  $\psi(\cdot - b)$  be weakly semicontinuous for every  $b \in B$ . Then the problem  $\sup_{b \in B} \min_{a \in A} \psi(a - b)$  is solvable.*

**Proof** First, we show the solvability of the problem  $\min_{a \in A} \psi(a - b)$  for an arbitrary  $b \in B$ . Since  $\psi(\cdot - b)$  is weakly semicontinuous for every  $b \in B$  and the set  $A$  is weakly compact, there is at least one minimal solution  $a_b \in A$  with  $\psi(a_b - b) = \min_{a \in A} \psi(a - b)$  (e.g., compare [18, Thm. 2.3]). Now we consider the problem  $\sup_{b \in B} \psi(a_b - b)$ . Because the functional  $\psi$  is assumed to be bounded on the set  $A - B$ , it is evident that  $\sup_{b \in B} \psi(a_b - b) < \infty$ , which has to be shown.  $\square$

**Remark 2.1** In a reflexive Banach space  $(Y, \|\cdot\|)$  the min subproblem in Proposition 2.2 is also solvable, if the set  $A$  is nonempty convex closed and bounded and the functional  $\psi(\cdot - b)$  is continuous and quasiconvex for all  $b \in B$  (compare [18, Thm. 2.12]).

For specific scalarizing functionals such existence results are also remarked in [22, Remark 1].

**Example 2.2** Consider the function  $\psi : \mathbb{R}^m \rightarrow \mathbb{R}$  (with  $m \in \mathbb{N}$ ) with

$$\psi(y) := \max\{y_1, \dots, y_m\} \text{ for all } y = (y_1, \dots, y_m) \in \mathbb{R}^m$$

(compare Example 2.1,(a)). The function  $\psi$  is continuous and so, it is also weakly lower semicontinuous because weak and strong convergence coincide in  $\mathbb{R}^m$ . If  $A, B \subset \mathbb{R}^m$  are nonempty sets where  $A$  is weakly compact and  $B$  is norm bounded, then the set difference  $A - B$  is norm bounded. This implies that the function  $\psi$  is bounded on the set  $A - B$ . In this case the function  $\psi$  satisfies the assumptions of Proposition 2.2.

The following proposition is a direct consequence of the Propositions 2.1 and 2.2.

**Proposition 2.3** *Let Assumption 2.1 be satisfied and, in addition, let  $(Y, \|\cdot\|)$  be a real normed space. Let  $A, B \subset Y$  be nonempty weakly compact sets, let  $\psi$  be bounded on  $A - B$ , and let the functionals  $\psi(a - \cdot)$  and  $\psi(\cdot - b)$  be weakly lower semicontinuous for every  $a \in A$  and  $b \in B$ , respectively. Then*

$$\max \left\{ \sup_{b \in B} \min_{a \in A} \psi(a - b), \sup_{a \in A} \min_{b \in B} \psi(a - b) \right\} \leq 0 \implies A \preceq_s B.$$

This proposition gives a sufficient condition for the set less order relation, which may be helpful in practice (see also [22, Corollary 3.11] for specific scalarizing functionals).

### 3 Optimality

We now turn our attention to problems of set optimization. Based on the set less order relation in a real linear space  $Y$ , we consider a family  $\mathcal{F}$  of nonempty subsets of  $Y$  and we investigate optimal sets of  $\mathcal{F}$ .

**Definition 3.1** Let Assumption 2.1 be satisfied, and let  $\mathcal{F}$  be a family of nonempty subsets of  $Y$ . A set  $\bar{A} \in \mathcal{F}$  is called an *optimal* set of  $\mathcal{F}$ , iff

$$A \preceq_s \bar{A}, A \in \mathcal{F} \implies \bar{A} \preceq_s A.$$

Under Assumption 2.1 recall for any nonempty subset  $A$  of  $Y$  that the set

$$\min A := \{a \in A \mid (\{a\} - C) \cap A \subset \{a\} + C\}$$

denotes the set of all *minimal elements* of  $A$ , and the set

$$\max A := \{a \in A \mid (\{a\} + C) \cap A \subset \{a\} - C\}$$

is called the set of *maximal elements* of  $A$ . Of course, if the convex cone  $C$  is pointed (i.e.  $C \cap (-C) = \{0_Y\}$ ), then  $\min A = \{a \in A \mid (\{a\} - C) \cap A = \{a\}\}$  and  $\max A = \{a \in A \mid (\{a\} + C) \cap A = \{a\}\}$ .

With the following proposition we investigate the question: What does it mean, if for some set  $A \in \mathcal{F}$  the two inequalities  $A \preceq_s \bar{A}$  and  $\bar{A} \preceq_s A$  given in Definition 3.1 hold?

**Proposition 3.1** Let Assumption 2.1 be satisfied and, in addition, let the convex cone  $C$  be pointed. Let  $\mathcal{F}$  be a family of nonempty subsets of  $Y$ , for which the set of minimal elements and the set of maximal elements are nonempty. For every  $A \in \mathcal{F}$  let the set equalities

$$A + C = (\min A) + C \text{ and } A - C = (\max A) - C \quad (2)$$

be satisfied. For some  $\bar{A} \in \mathcal{F}$ , we then have

$$\begin{aligned} \bar{A} \text{ optimal} &\iff \text{for every } A \in \mathcal{F} \text{ with } A \preceq_s \bar{A}: \\ &\quad \min A = \min \bar{A} \text{ and } \max A = \max \bar{A}. \end{aligned}$$

**Proof** For some  $\bar{A} \in \mathcal{F}$  and an arbitrary  $A \in \mathcal{F}$  it holds by definition

$$A \preceq_s \bar{A} \text{ and } \bar{A} \preceq_s A \iff \bar{A} \subset A + C, A \subset \bar{A} - C, A \subset \bar{A} + C, \bar{A} \subset A - C.$$

By [17, Lemma 2.4] this is equivalent to the inclusions

$$\bar{A} + C \subset A + C, A - C \subset \bar{A} - C, A + C \subset \bar{A} + C, \bar{A} - C \subset A - C$$

again being equivalent to

$$\bar{A} + C = A + C, \bar{A} - C = A - C.$$

By the equalities (2) this can be written as

$$(\min \bar{A}) + C = (\min A) + C, (\max \bar{A}) - C = (\max A) - C. \quad (3)$$

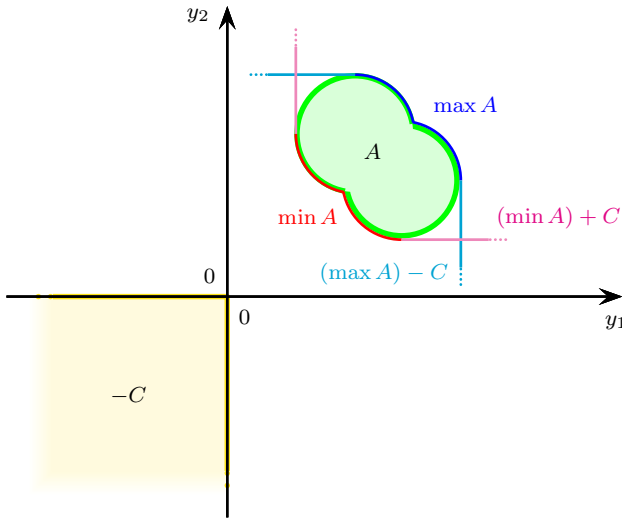


Fig. 1 Illustration of the set equalities (2)

The first equality implies  $\min \bar{A} \subset (\min A) + C$ , which means

$$\forall \bar{a} \in \min \bar{A} \exists a \in \min A, c \in C : \bar{a} = a + c.$$

Since the first equality in (3) also implies  $\min A \subset (\min \bar{A}) + C$ , there are  $\hat{a} \in \min \bar{A}$  and  $\hat{c} \in C$  with  $a = \hat{a} + \hat{c}$ . Hence, we get  $\bar{a} = a + c = \hat{a} + \hat{c} + c$ . Because of  $\bar{a}, \hat{a} \in \min \bar{A}$  we obtain  $\hat{c} + c = 0_Y$ , i.e.  $c = -\hat{c}$ . Since  $C$  is pointed, we conclude  $c \in C \cap (-C) = \{0_Y\}$ . Consequently, we have  $\bar{a} = a$  and thereby  $\min \bar{A} \subset \min A$ . By renaming we also get  $\min A \subset \min \bar{A}$ , and so we have  $\min A = \min \bar{A}$ . The equality  $\max A = \max \bar{A}$  can be proven in analogy. The assertion then follows with the definition of optimality.  $\square$

Figure 1 illustrates that the set equalities (2) may also be satisfied for nonconvex sets. For  $Y := \mathbb{R}^2, C := \mathbb{R}_+^2$  and  $A$  given in Fig. 1 we have  $A + C = (\min A) + C$  and  $A - C = (\max A) - C$ .

But the set inequalities (2) do not hold in general; for instance, if we choose  $Y := \mathbb{R}^2, C := \mathbb{R}_+^2$  and  $A := \{(y_1, y_2) \in \mathbb{R}^2 \mid y_1^2 + y_2^2 < 1\} \cup \{(-1, 0)\}$  we have  $\min A = \{(-1, 0)\}$  and we get  $(\min A) + C = \{(y_1, y_2) \in \mathbb{R}^2 \mid y_1 \geq -1, y_2 \geq 0\} \neq A + C$ , i.e. the first set inequality in (2) is not satisfied.

**Remark 3.1** Under the assumptions of Proposition 3.1 we have

$$\begin{aligned} \bar{A} \text{ not optimal} &\iff \exists A \in \mathcal{F} \text{ with } A \preceq_s \bar{A} : \\ &\quad \min A \neq \min \bar{A} \text{ or } \max A \neq \max \bar{A}. \end{aligned}$$

For instance, if one works with a descent method for the calculation of an optimal set, one can use this result in order to decide whether a set  $\bar{A}$  is not optimal. Let  $\bar{A} \in \mathcal{F}$

be given and let some set  $A \in \mathcal{F}$  with  $A \preceq_s \bar{A}$  be calculated. If one can check that  $\min A \neq \min \bar{A}$  or  $\max A \neq \max \bar{A}$ , one knows that the set  $\bar{A}$  cannot be optimal.

The assumption in Proposition 3.1 that the set equalities (2) are fulfilled for all  $A \in \mathcal{F}$  can be avoided, if one works with a more strict order relation introduced in [19, Def. 3.5].

**Definition 3.2** Let Assumption 2.1 be satisfied.

- (a) For subsets  $A, B \subset Y$ , for which the set of minimal elements  $\min(\cdot)$  and the set of maximal elements  $\max(\cdot)$  are nonempty, the *minmax less* order relation  $\preceq_m$  is defined by

$$A \preceq_m B \iff \min A \preceq_s \min B \text{ and } \max A \preceq_s \max B.$$

- (b) Let  $\mathcal{F}$  be a family of subsets of  $Y$ , for which the set of minimal elements and the set of maximal elements are nonempty. A set  $\bar{A} \in \mathcal{F}$  is called a *minmax optimal* set of  $\mathcal{F}$ , iff

$$A \preceq_m \bar{A}, A \in \mathcal{F} \implies \bar{A} \preceq_m A.$$

**Proposition 3.2** Let Assumption 2.1 be satisfied and, in addition, let the convex cone  $C$  be pointed. Let  $\mathcal{F}$  be a family of nonempty subsets of  $Y$ , for which the set of minimal elements and the set of maximal elements are nonempty. For some  $\bar{A} \in \mathcal{F}$ , we then have

$$\begin{aligned} \bar{A} \text{ minmax optimal} &\iff \text{for every } A \in \mathcal{F} \text{ with } A \preceq_m \bar{A} : \\ &\min A = \min \bar{A} \text{ and } \max A = \max \bar{A}. \end{aligned}$$

**Proof** In analogy to the proof of Proposition 3.1 we obtain for some  $\bar{A} \in \mathcal{F}$  and an arbitrary  $A \in \mathcal{F}$

$$\begin{aligned} &A \preceq_m \bar{A} \text{ and } \bar{A} \preceq_m A \\ &\iff \min A \preceq_s \min \bar{A}, \max A \preceq_s \max \bar{A}, \min \bar{A} \preceq_s \min A, \max \bar{A} \preceq_s \max A \\ &\iff (\min \bar{A}) + C \subset (\min A) + C, (\min A) - C \subset (\min \bar{A}) - C, \\ &\quad (\max \bar{A}) + C \subset (\max A) + C, (\max A) - C \subset (\max \bar{A}) - C, \\ &\quad (\min A) + C \subset (\min \bar{A}) + C, (\min \bar{A}) - C \subset (\min A) - C, \\ &\quad (\max A) + C \subset (\max \bar{A}) + C, (\max \bar{A}) - C \subset (\max A) - C \\ &\iff (\min A) + C = (\min \bar{A}) + C, (\min A) - C = (\min \bar{A}) - C, \\ &\quad (\max A) + C = (\max \bar{A}) + C, (\max A) - C = (\max \bar{A}) - C \\ &\iff \min A = \min \bar{A}, \max A = \max \bar{A}. \end{aligned}$$

This leads to the assertion. □

The proof of this proposition follows the lines in [17, Lemma 2.8,(a)].



**Remark 3.2** Notice that Proposition 3.2 does not require the assumption (2) because  $(\min A) + C = (\min(\min A)) + C$  and the other set equalities follow similarly.

### 4 Necessary Conditions for Set Inequalities

The set less order relation is defined by certain set inclusions; by Proposition 2.1 these set inclusions are characterized by appropriate inequalities. And now we investigate necessary conditions for these inequalities.

Under Assumption 2.1 we consider two arbitrary nonempty subsets  $A, B \subset Y$ . For an arbitrarily chosen  $\bar{b} \in B$  we use the abbreviation

$$\hat{A}_{\bar{b}} := \left\{ \hat{a} \in A \mid \psi(\hat{a} - \bar{b}) = \min_{a \in A} \psi(a - \bar{b}) \right\}.$$

$\hat{A}_{\bar{b}}$  denotes the set of all minimal solutions of the optimization problem  $\min_{a \in A} \psi(a - \bar{b})$ .

For the first result we need a technical lemma.

**Lemma 4.1** *Let Assumption 2.1 be satisfied, and let  $h \in Y$  be arbitrarily chosen. For nonempty subsets  $A, B \subset Y$  let  $\bar{b} \in B$  and  $\hat{a} \in \hat{A}_{\bar{b}}$  be arbitrarily given. Let  $\hat{A}_{\bar{b}} \neq A$ . Suppose that*

$$\begin{aligned} \exists \bar{\lambda} > 0, \alpha \in \mathbb{R} : \psi(a - \bar{b} + \lambda h) > \alpha > \psi(\hat{a} - \bar{b}) \text{ for all } \lambda \in [0, \bar{\lambda}], a \in A \setminus \hat{A}_{\bar{b}} \\ \text{and} \\ \psi(\hat{a} - \bar{b} + \cdot h) : [0, \bar{\lambda}] \rightarrow \mathbb{R} \text{ is continuous at } \lambda = 0. \end{aligned} \tag{4}$$

Then there exists some  $\tilde{\lambda} > 0$  so that

$$\min_{a \in A} \psi(a - \bar{b} + \lambda h) = \min_{a \in \hat{A}_{\bar{b}}} \psi(a - \bar{b} + \lambda h) \text{ for all } \lambda \in [0, \tilde{\lambda}]$$

provided that these min terms exist.

**Proof** Let  $h \in Y, \bar{b} \in B$  and  $\hat{a} \in \hat{A}_{\bar{b}}$  be arbitrarily chosen. By the definition of the set  $\hat{A}_{\bar{b}}$  we obtain  $\psi(\hat{a} - \bar{b}) = \min_{a \in A} \psi(a - \bar{b})$ . By the assumptions there is a sufficiently small  $\varepsilon > 0$  and some  $\tilde{\lambda} \in [0, \bar{\lambda}]$  with

$$\psi(a - \bar{b} + \lambda h) > \alpha > \alpha - \varepsilon \geq \psi(\hat{a} - \bar{b} + \lambda h) \text{ for all } \lambda \in [0, \tilde{\lambda}] \text{ and all } a \in A \setminus \hat{A}_{\bar{b}}.$$

Consequently, we obtain

$$\inf_{a \in A \setminus \hat{A}_{\bar{b}}} \psi(a - \bar{b} + \lambda h) \geq \alpha > \psi(\hat{a} - \bar{b} + \lambda h) \geq \min_{a \in \hat{A}_{\bar{b}}} \psi(a - \bar{b} + \lambda h) \text{ for all } \lambda \in [0, \tilde{\lambda}].$$

These inequalities imply

$$\min_{a \in A} \psi(a - \bar{b} + \lambda h) = \min_{a \in \hat{A}_{\bar{b}}} \psi(a - \bar{b} + \lambda h) \text{ for all } \lambda \in [0, \tilde{\lambda}],$$

which has to be shown. □

Under the condition (4) Lemma 4.1 states that the minimal value  $\min_{a \in \hat{A}_{\bar{b}}} [0] \psi(a - \bar{b} + \lambda h)$  remains unchanged for sufficiently small  $\lambda \geq 0$ , if the set  $\hat{A}_{\bar{b}}$  is replaced by its superset  $A$ . The continuity requirement in the second part of the condition (4) is fulfilled for many functionals  $\psi$  used in practice (compare Example 2.1). The inequalities in the first part of the condition (4) are decisive for Lemma 4.1. By the definition of the set  $\hat{A}_{\bar{b}}$  it obviously holds

$$\psi(a - \bar{b}) > \psi(\hat{a} - \bar{b}) \text{ for all } a \in A \setminus \hat{A}_{\bar{b}}.$$

But in (4) we replace the argument  $a - \bar{b}$  by  $a - \bar{b} + \lambda h$ , i.e. we consider elements with respect to the direction of  $h$ , and we require a stronger inequality. Besides the functional  $\psi$  the special choice of the set  $A$  plays a central role for this condition.

For simplicity we use the following notation for the next result. Under Assumption 2.1 we consider nonempty subsets  $A, B \in Y$  and define the functional  $\varphi : B \rightarrow \mathbb{R}$  with

$$\varphi(b) = \min_{a \in A} \psi(a - b) \text{ for all } b \in B$$

provided that the min term exists. Next, we investigate the question under which conditions the functional  $\varphi$  has a directional derivative.

**Theorem 4.1** *Let Assumption 2.1 be satisfied, and let  $A, B \subset Y$  be arbitrarily chosen. Let  $\bar{b} \in B$  be a solution of the optimization problem  $\max_{b \in B} \min_{a \in A} \psi(a - b)$ , and let  $\hat{A}_{\bar{b}} \neq A$ . Let  $\varphi$  be directionally differentiable at  $\bar{b}$  in every direction  $\bar{b} - b$  with arbitrary  $b \in B$ . For an arbitrary  $b \in B$  and an arbitrary  $\hat{a} \in \hat{A}_{\bar{b}}$  suppose that*

$$\begin{aligned} \exists \bar{\lambda} > 0, \alpha \in \mathbb{R} : \psi(a - \bar{b} + \lambda(\bar{b} - b)) > \alpha > \psi(\hat{a} - \bar{b}) \text{ for all } \lambda \in [0, \bar{\lambda}], a \in A \setminus \hat{A}_{\bar{b}} \\ \text{and} \\ \psi(\hat{a} - \bar{b} + \cdot(\bar{b} - b)) : [0, \bar{\lambda}] \rightarrow \mathbb{R} \text{ is continuous at } \lambda = 0. \end{aligned}$$

In addition, suppose that for all  $b \in B$

$$\begin{aligned} & \lim_{\lambda \rightarrow 0_+} \min_{a \in \hat{A}_{\bar{b}} \frac{1}{\lambda}} \left( \psi(a - \bar{b} + \lambda(\bar{b} - b)) - \psi(a - \bar{b}) \right) \\ & = \min_{a \in \hat{A}_{\bar{b}}} \lim_{\lambda \rightarrow 0_+} \frac{1}{\lambda} \left( \psi(a - \bar{b} + \lambda(\bar{b} - b)) - \psi(a - \bar{b}) \right). \end{aligned}$$

Then it follows

$$\min_{a \in \hat{A}_{\bar{b}}} \psi'(a - \bar{b})(\bar{b} - b) \leq 0 \text{ for all } b \in B$$

provided that the arising min terms exist ( $\psi'(a - \bar{b})(\bar{b} - b)$  denotes the directional derivative of  $\psi$  at  $a - \bar{b}$  in the direction  $\bar{b} - b$ ).

**Proof** First, we determine the directional derivative of  $\varphi$  at  $\bar{b}$  in every direction  $b - \bar{b}$  with arbitrary  $b \in B$ . With Lemma 4.1 and the assumptions of this theorem, we then get for all  $b \in B$

$$\begin{aligned} \varphi'(\bar{b})(b - \bar{b}) &= \lim_{\lambda \rightarrow 0^+} \frac{1}{\lambda} \left( \varphi(\bar{b} + \lambda(b - \bar{b})) - \varphi(\bar{b}) \right) \\ &= \lim_{\lambda \rightarrow 0^+} \frac{1}{\lambda} \left( \min_{a \in A} \psi(a - \bar{b} + \lambda(\bar{b} - b)) - \underbrace{\min_{a \in A} \psi(a - \bar{b})}_{= \min_{a \in \hat{A}_{\bar{b}}} \psi(a - \bar{b})} \right) \\ &= \lim_{\lambda \rightarrow 0^+} \frac{1}{\lambda} \left( \min_{a \in \hat{A}_{\bar{b}}} \psi(a - \bar{b} + \lambda(\bar{b} - b)) - \min_{a \in \hat{A}_{\bar{b}}} \underbrace{\psi(a - \bar{b})}_{= \text{const. } \forall a \in \hat{A}_{\bar{b}}} \right) \\ &= \lim_{\lambda \rightarrow 0^+} \frac{1}{\lambda} \min_{a \in \hat{A}_{\bar{b}}} \left( \psi(a - \bar{b} + \lambda(\bar{b} - b)) - \psi(a - \bar{b}) \right) \\ &= \lim_{\lambda \rightarrow 0^+} \min_{a \in \hat{A}_{\bar{b}}} \frac{1}{\lambda} \left( \psi(a - \bar{b} + \lambda(\bar{b} - b)) - \psi(a - \bar{b}) \right) \\ &= \min_{a \in \hat{A}_{\bar{b}}} \lim_{\lambda \rightarrow 0^+} \frac{1}{\lambda} \left( \psi(a - \bar{b} + \lambda(\bar{b} - b)) - \psi(a - \bar{b}) \right) \\ &= \min_{a \in \hat{A}_{\bar{b}}} \psi'(a - \bar{b})(\bar{b} - b). \end{aligned}$$

Since  $\bar{b} \in B$  is a maximal solution of the problem  $\max_{b \in B} \varphi(b) = -\min_{b \in B} -\varphi(b)$  and  $\varphi$  has a directional derivative at  $\bar{b}$  in every direction  $b - \bar{b}$  with arbitrary  $b \in B$ , by [18, Thm. 3.8,(a)] we obtain

$$0 \leq -\varphi'(\bar{b})(b - \bar{b}) = -\min_{a \in \hat{A}_{\bar{b}}} \psi'(a - \bar{b})(\bar{b} - b) \text{ for all } b \in B,$$

which leads to the assertion. □

Early investigations on directional derivatives of sup inf functions can be found in [4] (compare also [2]).

With this necessary condition for certain min max problems, we then obtain necessary conditions for set inequalities as well.

**Corollary 4.1** *Let Assumption 2.1 be satisfied, and let  $A, B \subset Y$  be arbitrarily chosen. Let  $\bar{a} \in A$  be a solution of the optimization problem  $\max_{a \in A} \min_{b \in B} \psi(a - b)$  and let  $\bar{b} \in B$  be a solution of the optimization problem  $\max_{b \in B} \min_{a \in A} \psi(a - b)$ . Let  $\hat{A}_{\bar{b}} \neq A$  and  $\hat{B}_{\bar{a}} \neq B$ . Let  $\min_{b \in B} \psi(\cdot - b)$  be directionally differentiable at  $\bar{a}$  in every direction  $a - \bar{a}$  with arbitrary  $a \in A$ , and let  $\min_{a \in A} \psi(a - \cdot)$  be directionally differentiable at  $\bar{b}$  in every direction  $\bar{b} - b$  with arbitrary  $b \in B$ . For an arbitrary  $a \in A$  and an arbitrary  $\hat{b} \in \hat{B}_{\bar{a}}$  suppose that*

$$\exists \bar{\lambda}_1 > 0, \alpha_1 \in \mathbb{R} : \psi(\bar{a} - b + \lambda(a - \bar{a})) > \alpha_1 > \psi(\bar{a} - \hat{b})$$

for all  $\lambda \in [0, \bar{\lambda}_1], b \in B \setminus \hat{B}_{\bar{a}}$

and

$$\psi(\bar{a} - \hat{b} + \cdot(a - \bar{a})) : [0, \bar{\lambda}_1] \rightarrow \mathbb{R} \text{ is continuous at } \lambda = 0.$$

Moreover, for an arbitrary  $b \in B$  and an arbitrary  $\hat{a} \in \hat{A}_{\bar{b}}$  suppose that

$$\exists \bar{\lambda}_2 > 0, \alpha_2 \in \mathbb{R} : \psi(a - \bar{b} + \lambda(\bar{b} - b)) > \alpha_2 > \psi(\hat{a} - \bar{b})$$

for all  $\lambda \in [0, \bar{\lambda}_2], a \in A \setminus \hat{A}_{\bar{b}}$

and

$$\psi(\hat{a} - \bar{b} + \cdot(\bar{b} - b)) : [0, \bar{\lambda}_2] \rightarrow \mathbb{R} \text{ is continuous at } \lambda = 0.$$

In addition, suppose that for all  $a \in A$

$$\begin{aligned} & \lim_{\lambda \rightarrow 0^+} \min_{b \in \hat{B}_{\bar{a}}} \frac{1}{\lambda} \left( \psi(\bar{a} - b + \lambda(a - \bar{a})) - \psi(\bar{a} - b) \right) \\ & = \min_{b \in \hat{B}_{\bar{a}}} \lim_{\lambda \rightarrow 0^+} \frac{1}{\lambda} \left( \psi(\bar{a} - b + \lambda(a - \bar{a})) - \psi(\bar{a} - b) \right) \end{aligned}$$

and for all  $b \in B$

$$\begin{aligned} & \lim_{\lambda \rightarrow 0^+} \min_{a \in \hat{A}_{\bar{b}}} \frac{1}{\lambda} \left( \psi(a - \bar{b} + \lambda(\bar{b} - b)) - \psi(a - \bar{b}) \right) \\ & = \min_{a \in \hat{A}_{\bar{b}}} \lim_{\lambda \rightarrow 0^+} \frac{1}{\lambda} \left( \psi(a - \bar{b} + \lambda(\bar{b} - b)) - \psi(a - \bar{b}) \right). \end{aligned}$$

If the inequality  $A \preceq_S B$  is satisfied, then we have

$$\min_{a \in \hat{A}_{\bar{b}}} \psi'(a - \bar{b})(\bar{b} - b) \leq 0 \text{ for all } b \in B \tag{5}$$

$$\min_{b \in \hat{B}_{\bar{a}}} \psi'(\bar{a} - b)(a - \bar{a}) \leq 0 \text{ for all } a \in A \tag{6}$$

provided that the arising min terms exist.

**Proof** Let the set inequality  $A \preceq_s B$  hold for arbitrary  $\emptyset \neq A, B \subset Y$ . By Proposition 2.1 and the assumptions, we then have

$$0 \geq \sup_{b \in B} \inf_{a \in A} \psi(a - b) = \max_{b \in B} \min_{a \in A} \psi(a - b) = \min_{a \in A} \psi(a - \bar{b})$$

and

$$0 \geq \sup_{a \in A} \inf_{b \in B} \psi(a - b) = \max_{a \in A} \min_{b \in B} \psi(a - b) = \min_{b \in B} \psi(\bar{a} - b).$$

Hence, the inequalities (5) and (6) follow from Theorem 4.1. □

**Remark 4.1** The inequalities (5) and (6) are necessary conditions for the set inequality  $A \preceq_s B$  but in general, they are not sufficient. Even if these inequalities imply  $\max_{b \in B} \min_{a \in A} \psi(a - b) = \min_{a \in A} \psi(a - \bar{b})$  and  $\max_{a \in A} \min_{b \in B} \psi(a - b) = \min_{b \in B} \psi(\bar{a} - b)$ , one has to assume that  $\psi(\bar{a} - \bar{b}) \leq 0$  because in this case we have

$$\max_{b \in B} \min_{a \in A} \psi(a - b) = \min_{a \in A} \psi(a - \bar{b}) \leq \psi(\bar{a} - \bar{b}) \leq 0$$

and

$$\max_{a \in A} \min_{b \in B} \psi(a - b) = \min_{b \in B} \psi(\bar{a} - b) \leq \psi(\bar{a} - \bar{b}) \leq 0.$$

An application of Proposition 2.1 then gives  $A \preceq_s B$ .

By Corollary 4.2 we now present a necessary condition for sets to be non-optimal.

**Corollary 4.2** *Let Assumption 2.1 be satisfied, and let the convex cone  $C$  be pointed. Let  $\mathcal{F}$  be a family of nonempty subsets of the real linear space  $Y$ , for which the set of minimal elements and the set of maximal elements are nonempty and the set equalities (2) are satisfied (for every  $A \in \mathcal{F}$ ). In addition, let some set  $B \in \mathcal{F}$  and every set  $A' \in \mathcal{F}$  with  $\min A' \neq \min B$  or  $\max A' \neq \max B$  satisfy the assumptions of Corollary 4.1. If the set  $B$  is not an optimal set of  $\mathcal{F}$ , then there exists some set  $A \in \mathcal{F}$  so that the inequalities (5) and (6) hold and  $\min A \neq \min B$  or  $\max A \neq \max B$ .*

**Proof** The assertion is a simple consequence of Remark 3.1 and Corollary 4.1. □

The assumptions of Corollaries 4.1 and 4.2 concerning the considered sets and the functional  $\psi$  are very strong. For the two sets it seems to be helpful, if they consist of finitely many elements. But for the functional  $\psi$  we need the directional derivative, which depends on the order cone  $C$ . In the following proposition the directional derivative of  $\psi$  is investigated for various special real linear spaces and order cones.

**Proposition 4.1** *Let Assumption 2.1 be satisfied.*

(a) [18, Exercise 8.3]

Let  $Y := \mathbb{R}^m$  and  $C := \mathbb{R}_+^m$  be chosen. Then the functional  $\psi : \mathbb{R}^m \rightarrow \mathbb{R}$  with

$$\psi(y) = \max\{y_1, \dots, y_m\} \text{ for all } y = (y_1, \dots, y_m) \in \mathbb{R}^m$$

has the directional derivative at an arbitrary  $\bar{y} \in \mathbb{R}^m$  given by

$$\psi'(\bar{y})(h) = \max_{i \in I(\bar{y})} \{h_i\} \text{ for all } h \in \mathbb{R}^m$$

with

$$I(\bar{y}) := \{i \in \{1, \dots, m\} \mid \bar{y}_i = \psi(\bar{y})\}.$$

(b) Let the real linear space  $Y := C[0, 1]$  of continuous real-valued functionals on  $[0, 1]$  and the natural order cone  $C := \{y \in C[0, 1] \mid y(t) \geq 0 \text{ for all } t \in [0, 1]\}$  be chosen. Then the functional  $\psi : C[0, 1] \rightarrow \mathbb{R}$  with

$$\psi(y) = \max_{t \in [0, 1]} y(t) \text{ for all } y \in C[0, 1]$$

has the directional derivative at an arbitrary  $\bar{y} \in C[0, 1]$  given by

$$\psi'(\bar{y})(h) = \max_{t \in M(\bar{y})} h(t) \text{ for all } h \in C[0, 1]$$

with

$$M(\bar{y}) := \{t \in [0, 1] \mid \bar{y}(t) = \psi(\bar{y})\}.$$

(c) [14, page 275] Let the real Hilbert space  $Y := S^n$  of real symmetric  $(n, n)$  matrices (with  $n \in \mathbb{N}$ ) and the Löwner cone  $C := S_+^n := \{A \in S^n \mid A \text{ positive semidefinite}\}$  be chosen. Then the functional  $\psi : S^n \rightarrow \mathbb{R}$  with

$$\psi(A) = \max_{\|x\|_2=1} x^T A x \left( = \text{maximal eigenvalue of } A \right) \text{ for all } A \in S^n$$

( $\|\cdot\|_2$  denotes the Euclidean norm in  $\mathbb{R}^n$ ) has the directional derivative at an arbitrary  $\bar{A} \in S^n$  given by

$$\psi'(\bar{A})(H) = \max_{x \in X(\bar{A})} x^T H x \text{ for all } H \in S^n$$

with

$$X(\bar{A}) := \{x \in \mathbb{R}^n \mid \|x\|_2 = 1 \text{ and } \bar{A}x = \psi(\bar{A})x\}$$

( $X(\bar{A})$  denotes the set of all normed eigenvalues associated to the maximal eigenvalue of  $\bar{A}$ ).

(d) Let the real Hilbert space  $Y := S^n$  of real symmetric  $(n, n)$  matrices (with  $n \in \mathbb{N}$ ) and the copositive cone

$$C := \{M \in S^n \mid x^T Mx \geq 0 \text{ for all } x \in \mathbb{R}_+^n\}$$

be chosen. Then the functional  $\psi : S^n \rightarrow \mathbb{R}$  with

$$\psi(M) = \max_{\substack{x \in \mathbb{R}_+^n \\ \|x\|_2=1}} x^T Mx \text{ for all } M \in S^n$$

( $\|\cdot\|_2$  denotes the Euclidean norm in  $\mathbb{R}^n$ ) has the directional derivative at an arbitrary  $\bar{A} \in S^n$  given by

$$\psi'(\bar{A})(H) = \max_{x \in X(\bar{A})} x^T Hx \text{ for all } H \in S^n$$

with

$$X(\bar{A}) := \{x \in \mathbb{R}_+^n \mid \|x\|_2 = 1 \text{ and } \psi(\bar{A}) = x^T \bar{A}x\}.$$

(e) Let a real normed space  $(Y, \|\cdot\|)$  and the Bishop-Phelps cone

$$C(\ell) := \{y \in Y \mid \ell(y) \geq \|y\|\}$$

for an arbitrary continuous linear functional  $\ell \in Y^*$  be chosen. Then the functional  $\psi : Y \rightarrow \mathbb{R}$  with

$$\psi(y) = \ell(y) + \|y\| \text{ for all } y \in Y$$

has the directional derivative at an arbitrary  $\bar{y} \in Y$  given by

$$\psi'(\bar{y})(h) = \ell(h) + \max_{\tilde{\ell} \in \partial \|\bar{y}\|} \tilde{\ell}(h) \text{ for all } h \in Y$$

with the subdifferential

$$\partial \|\bar{y}\| = \begin{cases} \{\tilde{\ell} \in Y^* \mid \tilde{\ell}(\bar{y}) = \|\bar{y}\| \text{ and } \|\tilde{\ell}\|_{Y^*} = 1\} & \text{if } \bar{y} \neq 0_Y \\ \{\tilde{\ell} \in Y^* \mid \|\tilde{\ell}\|_{Y^*} \leq 1\} & \text{if } \bar{y} = 0_Y \end{cases}$$

of the norm  $\|\cdot\|$  at  $\bar{y}$ .

**Proof** (b) For arbitrarily chosen  $\bar{y}, h \in C[0, 1]$  we obtain

$$\psi'(\bar{y})(h) = \lim_{\lambda \rightarrow 0^+} \frac{1}{\lambda} (\psi(\bar{y} + \lambda h) - \psi(\bar{y}))$$

$$\begin{aligned}
&= \lim_{\lambda \rightarrow 0^+} \frac{1}{\lambda} \left( \underbrace{\max_{t \in [0,1]} \bar{y}(t) + \lambda h(t)}_{\geq \max_{t \in M(\bar{y})} \bar{y}(t) + \lambda h(t)} - \underbrace{\max_{t \in [0,1]} \bar{y}(t)}_{= \max_{t \in M(\bar{y})} \bar{y}(t)} \right) \\
&\geq \lim_{\lambda \rightarrow 0^+} \frac{1}{\lambda} \left( \max_{t \in M(\bar{y})} (\bar{y}(t) + \lambda h(t)) - \underbrace{\max_{t \in M(\bar{y})} \bar{y}(t)}_{=\text{const.}} \right) \\
&= \lim_{\lambda \rightarrow 0^+} \frac{1}{\lambda} \max_{t \in M(\bar{y})} \bar{y}(t) + \lambda h(t) - \bar{y}(t) \\
&= \max_{t \in M(\bar{y})} h(t). \tag{7}
\end{aligned}$$

Next, we prove the converse inequality. For every  $\lambda > 0$  there is some  $t_\lambda \in [0, 1]$  with

$$\bar{y}(t_\lambda) + \lambda h(t_\lambda) = \max_{t \in [0,1]} \bar{y}(t) + \lambda h(t). \tag{8}$$

For some  $\tilde{t} \in M(\bar{y})$  we obtain for all  $\lambda > 0$

$$\max_{t \in [0,1]} (\bar{y}(t) + \lambda h(t)) - \max_{t \in [0,1]} \bar{y}(t) \geq \bar{y}(\tilde{t}) + \lambda h(\tilde{t}) - \bar{y}(\tilde{t}) = \lambda h(\tilde{t})$$

and

$$\begin{aligned}
\max_{t \in [0,1]} (\bar{y}(t) + \lambda h(t)) - \max_{t \in [0,1]} \bar{y}(t) &\leq \max_{t \in [0,1]} \bar{y}(t) + \lambda \max_{t \in [0,1]} h(t) - \max_{t \in [0,1]} \bar{y}(t) \\
&= \lambda \max_{t \in [0,1]} h(t).
\end{aligned}$$

So, we conclude for some  $\alpha > 0$

$$\left| \max_{t \in [0,1]} (\bar{y}(t) + \lambda h(t)) - \max_{t \in [0,1]} \bar{y}(t) \right| \leq \lambda \alpha \text{ for all } \lambda > 0,$$

and with the equality (8) we get

$$\lim_{\lambda \rightarrow 0^+} \bar{y}(t_\lambda) + \lambda h(t_\lambda) = \lim_{\lambda \rightarrow 0^+} \max_{t \in [0,1]} \bar{y}(t) + \lambda h(t) = \max_{t \in [0,1]} \bar{y}(t).$$

Then there is a sequence  $(\lambda_k)_{k \in \mathbb{N}}$  of positive real numbers converging to 0 with  $\lim_{k \rightarrow \infty} t_{\lambda_k} =: \hat{t} \in [0, 1]$ , and it is evident that  $\hat{t} \in M(\bar{y})$ . Hence, we conclude

$$\begin{aligned}
\psi'(\bar{y})(h) &= \lim_{k \rightarrow \infty} \frac{1}{\lambda_k} \left( \underbrace{\max_{t \in [0,1]} (\bar{y}(t) + \lambda_k h(t))}_{= \bar{y}(t_{\lambda_k}) + \lambda_k h(t_{\lambda_k})} - \underbrace{\max_{t \in [0,1]} \bar{y}(t)}_{\geq \bar{y}(t_{\lambda_k})} \right) \\
&\leq \lim_{k \rightarrow \infty} \frac{1}{\lambda_k} \left( \bar{y}(t_{\lambda_k}) + \lambda_k h(t_{\lambda_k}) - \bar{y}(t_{\lambda_k}) \right) \\
&= \lim_{k \rightarrow \infty} h(t_{\lambda_k})
\end{aligned}$$



$$\begin{aligned}
 &= h(\hat{t}) \\
 &\leq \max_{t \in M(\bar{y})} h(t).
 \end{aligned}
 \tag{9}$$

The inequalities (7) and (9) lead to the assertion.

(d) For arbitrarily chosen  $\bar{A}, H \in \mathcal{S}^n$  we have

$$\begin{aligned}
 \psi'(\bar{A})(H) &= \lim_{\lambda \rightarrow 0_+} \frac{1}{\lambda} \left( \psi(\bar{A} + \lambda H) - \psi(\bar{A}) \right) \\
 &= \lim_{\lambda \rightarrow 0_+} \frac{1}{\lambda} \left( \underbrace{\max_{\substack{x \in \mathbb{R}_+^n \\ \|x\|_2=1}} x^T (\bar{A} + \lambda H)x}_{\geq \max_{x \in X(\bar{A})} x^T (\bar{A} + \lambda H)x} - \underbrace{\max_{\substack{x \in \mathbb{R}_+^n \\ \|x\|_2=1}} x^T \bar{A}x}_{= \max_{x \in X(\bar{A})} x^T \bar{A}x} \right) \\
 &\geq \lim_{\lambda \rightarrow 0_+} \frac{1}{\lambda} \max_{x \in X(\bar{A})} x^T \bar{A}x + \lambda x^T Hx - x^T \bar{A}x \\
 &= \max_{x \in X(\bar{A})} x^T Hx.
 \end{aligned}
 \tag{10}$$

For the proof of the converse inequality choose an arbitrary  $\lambda > 0$ . Then there exists some  $x_\lambda \in \mathbb{R}_+^n$  with  $\|x_\lambda\|_2 = 1$  so that

$$x_\lambda^T (\bar{A} + \lambda H)x_\lambda = \max_{\substack{x \in \mathbb{R}_+^n \\ \|x\|_2=1}} x^T (\bar{A} + \lambda H)x.$$

For an arbitrary  $\tilde{x} \in X(\bar{A})$  we obtain

$$\begin{aligned}
 \max_{\substack{x \in \mathbb{R}_+^n \\ \|x\|_2=1}} x^T (\bar{A} + \lambda H)x - \max_{\substack{x \in \mathbb{R}_+^n \\ \|x\|_2=1}} x^T \bar{A}x &\geq \tilde{x}^T (\bar{A} + \lambda H)\tilde{x} - \tilde{x}^T \bar{A}\tilde{x} \\
 &= \lambda \tilde{x}^T H\tilde{x} \text{ for all } \lambda > 0
 \end{aligned}$$

and with

$$\begin{aligned}
 &\max_{\substack{x \in \mathbb{R}_+^n \\ \|x\|_2=1}} x^T (\bar{A} + \lambda H)x - \max_{\substack{x \in \mathbb{R}_+^n \\ \|x\|_2=1}} x^T \bar{A}x \\
 &\leq \max_{\substack{x \in \mathbb{R}_+^n \\ \|x\|_2=1}} x^T \bar{A}x + \lambda \max_{\substack{x \in \mathbb{R}_+^n \\ \|x\|_2=1}} x^T Hx - \max_{\substack{x \in \mathbb{R}_+^n \\ \|x\|_2=1}} x^T \bar{A}x \\
 &= \lambda \max_{\substack{x \in \mathbb{R}_+^n \\ \|x\|_2=1}} x^T Hx \text{ for all } \lambda > 0
 \end{aligned}$$

we conclude for some  $\alpha \geq 0$

$$\left| \max_{\substack{x \in \mathbb{R}_+^n \\ \|x\|_2=1}} x^T (\bar{A} + \lambda H)x - \max_{\substack{x \in \mathbb{R}_+^n \\ \|x\|_2=1}} x^T \bar{A}x \right| \leq \lambda \alpha \text{ for all } \lambda > 0.$$

Hence, we have

$$\lim_{\lambda \rightarrow 0_+} x_\lambda^T (\bar{A} + \lambda H)x_\lambda = \max_{\substack{x \in \mathbb{R}_+^n \\ \|x\|_2=1}} x^T \bar{A}x.$$

So, there is a sequence  $(\lambda_k)_{k \in \mathbb{N}}$  of positive real numbers converging to 0 with  $\lim_{k \rightarrow \infty} x_{\lambda_k} =: \hat{x} \in X(\bar{A})$ . And we conclude

$$\begin{aligned} \psi'(\bar{A})(H) &= \lim_{k \rightarrow \infty} \frac{1}{\lambda_k} \left( \underbrace{\max_{\substack{x \in \mathbb{R}_+^n \\ \|x\|_2=1}} x^T (\bar{A} + \lambda_k H)x}_{=x_{\lambda_k}^T (\bar{A} + \lambda_k H)x_{\lambda_k}} - \underbrace{\max_{\substack{x \in \mathbb{R}_+^n \\ \|x\|_2=1}} x^T \bar{A}x}_{\geq x_{\lambda_k}^T \bar{A}x_{\lambda_k}} \right) \\ &\leq \lim_{k \rightarrow \infty} \frac{1}{\lambda_k} \left( x_{\lambda_k}^T \bar{A}x_{\lambda_k} + \lambda_k x_{\lambda_k}^T Hx_{\lambda_k} - x_{\lambda_k}^T \bar{A}x_{\lambda_k} \right) \\ &= \lim_{k \rightarrow \infty} x_{\lambda_k}^T Hx_{\lambda_k} \\ &= \hat{x}^T H\hat{x} \\ &\leq \max_{x \in X(\bar{A})} x^T Hx. \end{aligned} \tag{11}$$

The inequalities (10) and (11) imply the assertion.

- (e) Since the norm  $\|\cdot\|$  is continuous and convex, its directional derivative is given by [18, Theorem 3.28]. Then for arbitrarily chosen  $\bar{y}, h \in Y$  we obtain

$$\psi'(\bar{y})(h) = \ell(h) + \max_{\tilde{\ell} \in \partial \|\bar{y}\|} \tilde{\ell}(h)$$

where the subdifferential  $\partial \|\bar{y}\|$  is calculated in [18, Example 3.24,(b)].

□

The proof of part (b) follows the lines of the proof of the directional derivative of the maximum norm (compare [18, Exercise 3.2]). Investigations of the directional derivative of the maximal eigenvalue of a symmetric matrix have been already given in [25] (compare also [1,6,24,26,28]).

**Remark 4.2** If in Proposition 4.1,(c) the maximal eigenvalue of a matrix  $\bar{A} \in \mathcal{S}^n$  is simple, then the associated normed eigenvector  $\bar{x}$  is unique and the set  $X(\bar{A})$  consists only of this vector. In this special case the directional derivative is simply given as

$$\psi'(\bar{A})(H) = \bar{x}^T H\bar{x} \text{ for all } H \in \mathcal{S}^n.$$

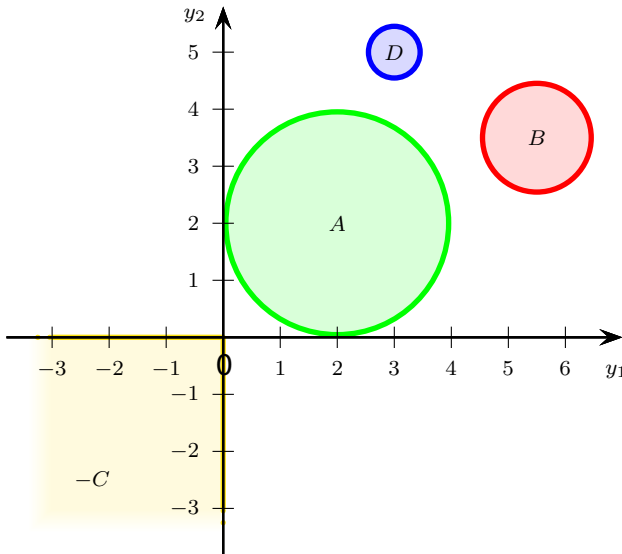


Fig. 2 Illustration of the closed balls  $A$ ,  $B$  and  $D$  defined in Example 4.1

If the maximal eigenvalue of  $\bar{A}$  is a double eigenvalue, then the set  $X(\bar{A})$  is a circle in a two dimensional subspace of  $\mathbb{R}^n$  spanned by the associated eigenvectors. Let  $\bar{x}^1$  and  $\bar{x}^2$  be orthogonal normed eigenvectors associated to the maximal eigenvalue of  $\bar{A}$ . Then this subspace is spanned by  $\bar{x}^1$  and  $\bar{x}^2$ , and we have

$$X(\bar{A}) = \{\alpha\bar{x}^1 + \beta\bar{x}^2 \mid \alpha, \beta \in \mathbb{R} \text{ and } \|\alpha\bar{x}^1 + \beta\bar{x}^2\|_2 = 1\}.$$

In this special case we obtain the directional derivative

$$\psi'(\bar{A})(H) = \max_{\substack{\alpha, \beta \in \mathbb{R} \\ \|\alpha\bar{x}^1 + \beta\bar{x}^2\|_2 = 1}} (\alpha\bar{x}^1 + \beta\bar{x}^2)^T H (\alpha\bar{x}^1 + \beta\bar{x}^2) \text{ for all } H \in S^n.$$

This maximization problem with only two real variables is very simple to solve.

**Example 4.1** We now apply the necessary conditions of Corollary 4.1 and Proposition 2.1 to the very simple case  $Y := \mathbb{R}^2$  and  $C := \mathbb{R}_+^2$ , and we choose the functional  $\psi : \mathbb{R}^2 \rightarrow \mathbb{R}$  with

$$\psi(y) = \max\{y_1, y_2\} \text{ for all } y = (y_1, y_2) \in \mathbb{R}^2$$

(compare Example 2.1.(a)). For convenience only we consider the closed balls  $A := \mathcal{B}(2, 2)$ ,  $B := \mathcal{B}(5.5, 3.5, 1)$  and  $D := \mathcal{B}(3, 5, 0.5)$  illustrated in Fig. 2. Then we investigate two cases.

- (a) It is geometrically obvious from the definition of the set less order relation that  $A \preccurlyeq_s B$ . We investigate the necessary conditions (5) and (6) for

this order relation. The optimization problem  $\max_{a \in A} \min_{b \in B} \psi(a - b)$  equals  $\max_{a \in A} \min_{b \in B} \max\{a_1 - b_1, a_2 - b_2\}$ , and  $\bar{a} := (2, 4) \in A$  is a maximal solution of this problem. Moreover, we obtain  $\bar{b} := (5.5, 2.5) \in B$  as maximal solution of the problem  $\max_{b \in B} \min_{a \in A} \max\{a_1 - b_1, a_2 - b_2\}$ . By Proposition 4.1,(a) we can write for all  $b \in B$

$$\min_{a \in \hat{A}_{\bar{b}}} \psi'(a - \bar{b})(\bar{b} - b) = \min_{a \in \hat{A}_{\bar{b}}} \max_{i \in I(a - \bar{b})} \{\bar{b}_i - b_i\} \tag{12}$$

with

$$\begin{aligned} \hat{A}_{\bar{b}} &= \left\{ \hat{a} \in A \mid \psi(\hat{a} - \bar{b}) = \min_{a \in A} \psi(a - \bar{b}) \right\} \\ &= \left\{ \hat{a} \in A \mid \max\{\hat{a}_1 - 5.5, \hat{a}_2 - 2.5\} = \min_{a \in A} \max\{a_1 - 5.5, a_2 - 2.5\} \right\} \\ &= \{(2, 0)\} \end{aligned}$$

and

$$\begin{aligned} I(a - \bar{b}) &= \{i \in \{1, 2\} \mid a_i - \bar{b}_i = \psi(a - \bar{b})\} \\ &= \{i \in \{1, 2\} \mid a_i - \bar{b}_i = \max\{a_1 - 5.5, a_2 - 2.5\}\}. \end{aligned}$$

For  $a := (2, 0)$  it follows

$$I(a - \bar{b}) = I((-3.5, -2.5)) = \{2\},$$

and then we get with (12)

$$\min_{a \in \hat{A}_{\bar{b}}} \psi'(a - \bar{b})(\bar{b} - b) = 2.5 - b_2 \leq 0 \text{ for all } b \in B.$$

So, the necessary condition (5) is shown. With  $\hat{B}_{\bar{a}} = \{(5.5, 4.5)\}$  and  $I(\bar{a} - b) = \{2\}$  for  $b := (5.5, 4.5)$  we get

$$\min_{b \in \hat{B}_{\bar{a}}} \psi'(\bar{a} - b)(a - \bar{a}) = \min_{b \in \hat{B}_{\bar{a}}} \max_{i \in I(\bar{a} - b)} \{a_i - 2, a_2 - 4\} = a_2 - 4 \leq 0 \forall a \in A,$$

and the necessary condition (6) is satisfied as well.

(b) It is evident from Fig. 2 that  $A \not\prec_s D$ . For these two sets we obtain  $\bar{a} := (4, 2)$  as maximal solution of the problem

$$\max_{a \in A} \min_{d \in D} \max\{a_1 - d_1, a_2 - d_2\} = \min_{d \in D} \max\{4 - d_1, 2 - d_2\} = 0.5 > 0.$$

By Proposition 2.1, we then conclude  $A \not\subset D - C$ .

## 5 Conclusions

The investigation of set inequalities generally leads to nontrivial necessary or sufficient conditions. Starting from known sup inf problems characterizations are given for set inequalities and optimal and non-optimal sets. The decisive key for the main results is the use of directional derivatives. For various standard cones the directional derivative of the functionals describing the negative cone is recalled or calculated. It seems to be that the presented technique of proof can also be applied to additional functionals used in practice. A simple example shows the usefulness of this theory but it also demonstrates that the decision, whether a set inequality holds or not, is a difficult task.

**Acknowledgements** The author thanks two anonymous referees for useful comments that helped improve this paper.

**Funding** Open Access funding enabled and organized by Projekt DEAL.

## Declaration

**Data sharing** Data sharing not applicable to this article as no datasets were generated or analysed during the current study.

**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

## References

1. Andrew, A.L., Chu, K.-W.E., Lancaster, P.: Derivatives of eigenvalues and eigenvectors of matrix functions. *SIAM J. Matrix Anal. Appl.* **14**, 903–926 (1993)
2. Borisenko, O.F., Minchenko, L.I.: Directional derivative of the maximum function. *Kibernet. Sistem. Anal.* **2**, 171–173 (1992)
3. Chiriaev, A., Walster, G.W.: Interval Arithmetic Specification, Technical Report (1998)
4. Correa, R., Seeger, A.: Directional derivative of a minimax function. *Nonlinear Anal.* **9**, 13–22 (1985)
5. Dhingra, M., Lalitha, C.S.: Approximate solutions and scalarization in set-valued optimization. *Optimization* **66**, 1793–1805 (2017)
6. Ding, C., Sun, D., Toh, K.-C.: An introduction to a class of matrix cone programming. *Math. Program., Ser. A* **144**, 141–179 (2014)
7. Eichfelder, G., Jahn, J.: Optimality conditions in discrete-continuous nonlinear optimization. *Minimax Theory Appl.* **6**, 127–144 (2021)
8. Gerstewitz (Tammer), C.: Nichtkonvexe Dualität in der Vektoroptimierung. *Wissenschaftliche Zeitschrift der TH Leuna-Merseburg* **25**, 357–364 (1983)
9. Gutiérrez, C., Jiménez, B., Miglierina, E., Molho, E.: Scalarization in set optimization with solid and nonsolid ordering cones. *J. Global Optim.* **61**, 525–552 (2015)
10. Gutiérrez, C., Jiménez, B., Novo, V.: Nonlinear scalarizations of set optimization problems with set orderings. In: Hamel, A.H., Heyde, F., Löhne, A., Rudloff, B., Schrage, C. (eds.): *Set Optimization and Applications - The State of the Art*, p. 43–63. Springer, Berlin (2015)

11. Ha, T.X.D., Jahn, J.: Properties of Bishop-Phelps cones. *J. Nonlinear Convex Anal.* **18**, 415–429 (2017)
12. Han, Y.: Nonlinear scalarizing functions in set optimization problems. *Optimization* **68**, 1685–1718 (2019)
13. Hernández, E., Rodríguez-Marín, L.: Nonconvex scalarization in set optimization with set-valued maps. *J. Math. Anal. Appl.* **325**, 1–18 (2007)
14. Hiriart-Urruty, J.-B., Lemaréchal, C.: *Convex Analysis and Minimization Algorithms I - Fundamentals*. Springer, Berlin (1993)
15. Jahn, J.: Bishop-Phelps cones in optimization. *Int. J. Optim. Theory Methods Appl.* **1**, 123–139 (2009)
16. Jahn, J.: Vectorization in set optimization. *J. Optim. Theory Appl.* **167**, 783–795 (2015)
17. Jahn, J.: A derivative-free rooted tree method in nonconvex set optimization. *Pure Appl. Funct. Anal.* **3**, 603–623 (2018)
18. Jahn, J.: *Introduction to the Theory of Nonlinear Optimization*. Springer, Cham (2020)
19. Jahn, J., Ha, T.X.D.: New order relations in set optimization. *J. Optim. Theory Appl.* **148**, 209–236 (2011)
20. Karaman, E., Güvenç, I.A., Soyertem, M., Tozkan, D., Küçük, M., Küçük, Y.: A vectorization for nonconvex set-valued optimization. *Turkish J. Math.* **42**, 1815–1832 (2018)
21. Khushboo, Lalitha, C.S.: Scalarizations for a set optimization problem using generalized oriented distance function. *Positivity* **23**, 1195–1213 (2019)
22. Köbis, E., Köbis, M.A.: Treatment of set order relations by means of a nonlinear scalarization functional: a full characterization. *Optimization* **65**, 1805–1827 (2016)
23. Kuroiwa, D.: *Natural Criteria of Set-Valued Optimization*. Shimane University, Japan, Manuscript (1998)
24. Lancaster, P.: On eigenvalues of matrices dependent on a parameter. *Numer. Math.* **6**, 377–387 (1964)
25. Magnus, J.R.: On differentiating eigenvalues and eigenvectors. *Economet. Theor.* **1**, 179–191 (1985)
26. Shapiro, A.: First and second order analysis of nonlinear semidefinite programs. *Math. Program.* **77**, 301–320 (1997)
27. Young, R.C.: The algebra of many-valued quantities. *Math. Ann.* **104**, 260–290 (1931)
28. Zhang, L., Zhang, N., Xiao, X.: The second order directional derivative of symmetric matrix-valued functions. [www.optimization-online.org/DB\\_FILE/2011/04/3010.pdf](http://www.optimization-online.org/DB_FILE/2011/04/3010.pdf) (2011)

**Publisher's Note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.