Appl. Set-Valued Anal. Optim. 3 (2021), No. 3, pp. 325-340
Available online at http://asvao.biemdas.com
https://doi.org/10.23952/asvao.3.2021.3.07

CHARACTERIZING THE UPPER SET RELATION BY GENERAL FUNCTIONALS

ELISABETH KÖBIS, MARKUS A. KÖBIS*

Institute of Mathematics, Norwegian University of Science and Technology, Trondheim, Norway

Abstract. In this paper, we investigate very general scalarization concepts for solving set-valued optimization problems where the pre-order is induced by the upper set-less relation as introduced by Kuroiwa. The scalarization functionals that we consider in the analysis are not given explicitly, but are rather general functions that satisfy certain properties, such as monotonicity, separation, transitivity, translation invariance, and the transfer of inclusion. Underlined by several examples, we show how certain combinations of such properties can be used to characterize the set relation by simple inequalities. We further propose a derivative-free descent method based on our theoretical findings.
Keywords. Descent methods; Monotonicity; Set optimization; Scalarization; Set Relations.

1. Introduction

Set optimization is a modern dynamic field that subsumes scalar and vector optimization, and therefore provides an important extension in optimization theory. Due to a large number of applications, such as duality principles in vector optimization, gap functions for vector variational inequalities, inverse problems for partial differential equations and variational inequalities, fuzzy optimization, image processing, optimal control problems with differential inclusions, viability theory, medical image registration or in mathematical economics, set optimization has recently expanded as a distinct branch of applied mathematics. As a result, set optimization became a bridge between different areas in optimization. We refer to Khan, Tammer and Zălinescu [1] for an exhaustive introduction to set optimization and its applications.

In this paper, we study quite general scalarization approaches for set optimization problems where the image set is equipped with the so-called upper set less relation. We focus on the properties that such functionals must have to give rise for formulating set inclusions by means of inequality statements.

The rest of this manuscript is organized as follows. In Section 2, we give some preliminary definitions and introduce the notations and main ordering concepts used throughout our analysis. The main results will be presented in Section 3. We formulate how abstract properties of separation functions can be used to characterize set relations. These characterizations will

*Corresponding author.
E-mail addresses: elisabeth.kobis@ntnu.no (E. Köbis), markus.koebis@gmail.com (M.A. Köbis).
Received May 30, 2021; Accepted August 12, 2021.

further be used for finding minimal elements of a family of sets (using the so-called set approach) in Section 4 before they can be applied within a descent method which we will outline in Section 5 before giving an outlook on potential extensions.

2. Preliminaries

By $\mathscr{P}(Y)$, we denote the power set of the space Y without the empty set. At this point, Y is just assumed to be a real linear vector space, an assumption to be refined at various places throughout this manuscript. The dual space of Y is denoted by Y^*, and $0 \in Y$ ($0 \in Y^*$) denotes the null element in Y (Y^* resp.).

For two elements A, B of $\mathscr{P}(Y)$, we denote the sum of sets by

$$A + B := \{a + b \mid a \in A,\ b \in B\}.$$

We say that a set $A \in \mathscr{P}(Y)$ is *additively self-inclusive* if it fulfills

$$A + A \subseteq A.$$

The set $C \subseteq Y$ is a *cone* if, for all $c \in C$ and $\lambda \geq 0$, $\lambda c \in C$ holds true. The cone C is *convex* if it is additively self-inclusive. The dual cone $\{\ell \in Y^* : \forall y \in C : \ell(y) \geq 0\}$ is denoted by C^*. We say that a set C is *proper* (or *nontrivial*) if $C \notin \{\emptyset, \{0\}, Y\}$. The cone C is called *pointed* if $C \cap (-C) = \{0\}$ holds.

An important part of set optimization includes comparing sets by means of set relations, which are binary relations among sets. There is a variety of set relations based on convex cones known in the literature (for an overview, see [1, Chapter 2.6.2]), and several authors have discussed which set relations are appropriate for certain applications (compare [2]). In this paper, we study the following set relation, which was originally introduced by Kuroiwa [3, 4] in case that D in the following definition is a convex cone.

Definition 2.1 (Upper Set Less Relation (see [3, 4, 5]))**.** Let Y be a real linear space and $D \subset Y$. The **upper set less relation** $\preceq^u_D$ is defined for two nonempty sets $A, B \in Y$ by

$$A \preceq^u_D B :\Longleftrightarrow A \subseteq B - D.$$

Remark 2.1. To interpret the notion of upper set less relation in actual applications, it is useful to impose further assumptions on the set D. It is, for example, clear that there is no use of this concept in set optimization if D is empty or a singleton. On the other hand, if D is a convex cone, the relation $\preceq^u_D$ is one extension of the partial ordering

$$y_1 \leq_D y_2 \quad \Leftrightarrow \quad y_1 \in y_2 - D$$

used in vector optimization. For the purpose of generality in this article, we will only specify certain additional properties of the set D whenever they are necessary for the characterization by separation functionals without strictly particularizing to specific cases. For more detailed discussions on the interpretation of $\preceq^u_D$, we refer to the literature [1, 4].

In this paper, we show how to characterize the set inclusion given in the relation $\preceq^u_D$ by means of very general functionals. In the literature, first approaches to represent a relation like $A \preceq^u_D B$ have been proposed and investigated by means of particular functionals, such as a Gerstewitz-type nonlinear scalarization functional [6], oriented distance functional [7], or linear functionals [8]. It is our goal in this work to point out the properties that general functionals must fulfill in order to be useful for a representation of $\preceq^u_D$.

3. Representation of the Upper Set Relation by Means of General Functionals

In the following assumption we collect several properties of functionals that will be used for representing the set relation $\preceq^u_D$.

Assumption 3.1. Let Y be a real linear space, $\{\emptyset,\{0\},Y\} \not\ni D \subseteq Y$, and let a functional $z\colon Y \to \mathbb{R}\cup\{\pm\infty\} =: \overline{\mathbb{R}}$ be given. Assume that z fulfills the properties:

(M) for all $a,b \in Y$, $a \in b - D$ implies $z(a) \le z(b)$ (Monotonicity);

(S) $y \in -D \iff z(y) \le 0$ (Separation Property);

(T1) for all $a,b,c \in Y$, $a \in b-D$ and $b \in c-D$ imply $z(a) \le z(c)$ (Transitivity);

(I) there is a $k \in Y$ such that for any $\varepsilon > 0$ and all $a \in Y$

$$z(a-\varepsilon k) = z(a) - \varepsilon \quad \text{(Translation Invariance)}.$$

Alternatively, one can alter Assumption 3.1(T1) to

(T2) for all $a,b \in Y$, $a \in b-D$ is equivalent to $z(a) \le z(b)$. (Transfer of Inclusion)

Remark 3.1. Note that we are dealing with extended real-valued functionals. In particular, this may result in the values being infinite. In this case, we will use the convention that $\infty \le \infty$ holds true.

Remark 3.2. Note that, in the literature, Assumption 3.1(I) is often formulated for general $\varepsilon \in \mathbb{R}$. It is easy to see that requiring $\varepsilon > 0$ suffices to show the validity for all ε. In pursuit of weak conditions, we formulate (I) only for positive scaling factors.

Remark 3.3. It is straightforward to see that Assumption 3.1(M) implies Assumption 3.1(T1) and vice versa if the involved set D contains the zero element $0 \in Y$. Likewise, it is obvious that Assumption 3.1(T2) is just Assumption 3.1(M) with the implication being replaced by an equivalence. In Figure 1, we depict how the different requirements interconnect with each other, potentially under appropriate additional assumptions.[1] It is apparent how Assumptions 3.1(T2) and (I) are the most stringent requirement while (S) is arguably the weakest, only providing a binary classification and no scalarization in the actual sense. The translation invariance property (I) is not directly implied by or a generalization of any of the other concepts. In contrast to the others, it actually involves some scaling of the values of z by means of the chosen vector $k \in Y$. Note, in this respect, that the translation invariance does not explicitly contain information regarding the set $D \subseteq Y$ but instead only the existence of such a vector k.

Before we continue with the characterization, we will illustrate the above assumption by the study of several well-known functionals from the literature. For a recent survey that reviews the relations thereof, we refer to [9].

Example 3.1. Let Y be a real linear space and $M \subseteq Y$. For the simplest case, consider the indicator function

$$\psi_M(y) = \begin{cases} 1, & y \in M, \\ 0, & y \notin M. \end{cases}$$

[1] We only depict direct implications from one of the assumptions to one another. The picture becomes significantly more complex if implications like ((I) $\wedge$ (S)) $\Rightarrow$ (T1) for additively self-inclusive sets D or similar ones are added, cf. [10, Proposition 4.1.1].

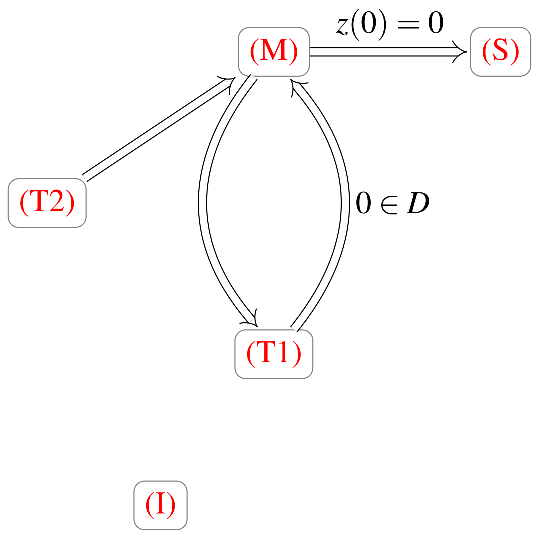


FIGURE 1. Interrelations between the assumptions on separation functionals in Assumption 3.1

Then, for $y \in Y$, $z^{\text{ind}}(y) := 1 - \psi_{-D}(y)$ fulfills Assumption 3.1(S). If D is additively self-inclusive, then $z^{\text{ind}}(y)$ also satisfies Assumption 3.1(M): To see this, let

$$a \in b - D, \tag{3.1}$$

but suppose that $z^{\text{ind}}(a) > z^{\text{ind}}(b)$. Because z^{ind} can only attain two values, we have $z^{\text{ind}}(a) = 1$ and $z^{\text{ind}}(b) = 0$. This corresponds to $a \notin -D$ and $b \in -D$. From (3.1), we have $a \in b - D \subseteq -D - D \subseteq -D$, a contradiction. A similar argument shows that, in this case, Assumption (T1) is fulfilled. It is obvious that this functional does not satisfy Assumption 3.1(I), because the functional z^{ind} only attains one of two values. Also Assumption (T2) is not satisfied in general. As a counterexample, take the additively self-inclusive set $D = \mathbb{R}^2_+ := \{(y_1, y_2) \in \mathbb{R}^2 : y_1, y_2 \geq 0\}$ and $a := (1, -1)$, $b := (-1, 1)$. Then $z^{\text{ind}}(a) = z^{\text{ind}}(b) = 1$ but $a \notin b - D$.

Example 3.2. In a similar manner as in the example above, let Y be a real linear space, $D \subseteq Y$, and consider the characteristic function

$$\chi_M(y) = \begin{cases} 0, & y \in M, \\ +\infty, & y \notin M. \end{cases}$$

Then, for $y \in Y$, $z^{\text{char}}(y) := \chi_{-D}(y)$ fulfills Assumption 3.1(S). If D is additively self-inclusive, then $z^{\text{char}}(y)$ also satisfies Assumptions 3.1(M) and (T1), which can be proven in a similar manner as above. Also, Assumptions (I) and (T1) are not fulfilled for this separation concept.

From now on, we assume that the set D is closed.

Example 3.3 (Distance Function)**.** The indicator and characteristic function both only give a binary characterization of the inclusion without providing any additional information on how much the argument deviates from the set under consideration. A, in that respect more evolved, way to quantify this deviation can be achieved by considering the distance

$$d_M(y) := \inf_{m \in M} \|y - m\|, \quad y \in Y, M \subseteq Y,$$

where Y is now assumed to be a normed vector space with norm $\|\cdot\|$. A straightforward separation functional based on this norm is given by

$$z^{\text{dist}}(y) := d_{-D}(y)\,.$$

Because trivially $z^{\text{dist}}(y) = 0$ for all $y \in -D$ and we assume D to be closed, this function fulfills Assumption (S). Likewise, it is obvious that (I) does not hold in general: If $\varepsilon > 0$ is chosen small enough and $a \in Y$ lies in the interior of $-D$ (assuming here that this interior is not empty), the function value z^{dist} will be exactly zero as $z^{\text{dist}}(a)$. To analyze the validity of Assumption (M), it is worth noting that the triangle inequality of the norm $\|\cdot\|$ transfers to a triangle inequality of $d_{-D}(\cdot)$ if D is additively self-inclusive. If one assumes this and $a \in b - D$, Assumption (M) can be deduced as follows: Since $a \in b - D$, there is an element $\tilde{d} \in D$ such that $a = b - \tilde{d}$. One furthermore obtains

$$z^{\text{dist}}(a) = d_{-D}(a) = d_{-D}(b + (-\tilde{d})) \overset{\text{triangl. ineq.}}{\leq} d_{-D}(b) + \underbrace{d_{-D}(-\tilde{d})}_{=0} = z^{\text{dist}}(b)\,.$$

A similar argument shows that Assumption (T1) holds for additively self-inclusive D. On the other hand, Assumption (T2) cannot be guaranteed even for the case of D being a convex cone in Y. This is again due to the fact that the values of z^{dist} inside $-D$ are all the same.

Example 3.4 (Quadratic Distance Function)**.** As a scalarization method, the functional $z^{\text{dist}}(y)$ has two downsides: Firstly, it does not provide any distance information for $y \in -D$ but is instead constant there (As seen above, this rules out the use of (T2) and/or (I).) and secondly, it does not include any deviation information within its first-order derivatives. To remedy the latter downside, one can adapt z^{dist} to

$$z^{\text{d-sq.}}(y) := \left(z^{\text{dist}}(y)\right)^2, \quad y \in Y.$$

From this adaptation, one can expect some numerical benefits as the quadratic terms increase smoothness and provide more ‘distance information’ on the level of derivatives, which is the main ingredient in most (sub-) gradient-based methods. The properties from z^{dist} in the sense of Assumption 3.1 remain valid for the quadratic distance but, apart from potential numerical gains, $z^{\text{d-sq.}}$ has no further advantages.

Example 3.5 (Oriented Distance Function)**.** To approach the first downside of z^{dist}, one can alter the separation function in the following sense: Let M be a closed nonempty subset of a real and again normed linear space $(Y, \|\cdot\|)$. The function $\triangle_M \colon Y \to \overline{\mathbb{R}}$ defined by

$$\triangle_M(y) := d_M(y) - d_{Y\setminus M}(y) = \begin{cases} d_M(y) & \text{if } y \in Y \setminus M \\ -d_{Y\setminus M}(y) & \text{if } y \in M \end{cases}, \quad \forall y \in Y \tag{3.2}$$

is called *oriented distance function* (or Hiriart-Urruty functional), where $d_M(y) := \inf\limits_{m \in M} \|y - m\|$ is the distance function from $y \in Y$ to the set M, see the above definition of z^{dist}. In this definition, the conventions

$$\triangle_Y(y) = -\infty, \qquad \triangle_\emptyset(y) = +\infty$$

are used. Hiriart-Urruty [7] introduced the notion of oriented distance functions to study optimality conditions of nonsmooth optimization problems from the geometric point of view. Later, Zaffaroni [11] discussed several properties of the oriented distance function.

In contrast to a simple deviation measure like in z^{dist}, this separation concept includes additional information for the case that the argument is contained in the set M, leading to the following candidate for a separation functional of the above type:

$$z^{\text{H-U}}(y) := \triangle_{-D}(y).$$

It holds that $z^{\text{H-U}}(y) \le 0$ if and only if $y \in -D$, showing that Property (S) in Assumption 3.1 is satisfied by this separation concept.

Furthermore, it can be shown [12] that $z^{\text{H-U}}$ is $(-D)$-monotone if D is additively self-inclusive, such that Properties (M) and in turn (T1) of Assumption 3.1 hold. Note that, at least from a numerical viewpoint, the norm $\|\cdot\|$ in the definition of $\triangle_M$ can be replaced by any equivalent norm $\|\cdot\|_*$ providing some flexibility in the definition of this family of separations. It is exactly this freedom of choice that allows for the construction of norms such that Property (I) is in general not fulfilled for the oriented distance function $z^{\text{H-U}}$.

Example 3.6 (Nonlinear Scalarizing Functional of Gerstewitz-Type)**.** Let Y be a real linear topological space. We assume that $D \subseteq Y$ is a nonempty closed proper set satisfying the inclusion

$$D + [0, +\infty) \cdot k \subseteq D \tag{3.3}$$

for some $k \in Y \setminus \{0\}$. If the relation (3.3) is fulfilled, the functional $z^{\text{G}} = z^k \colon Y \to \overline{\mathbb{R}}$ defined by

$$z^k(y) := \inf\{t \in \mathbb{R} \mid y \in tk - D\} \tag{3.4}$$

is well-defined provided the convention $\inf \emptyset = +\infty$. The functional z^{G} was used to obtain a separation theorem for not necessarily convex sets, see Gerstewitz [13], Gerstewitz and Iwanow [14] and also Gerth and Weidner [15]. Additionally, numerous applications of z^{G} are known in the literature, for instance, coherent risk measures in financial mathematics (see [16]) and uncertain optimization (in particular, in robustness theory, compare [17]). Many properties of z^{G} can be found in [14, 15, 18, 19] and the recent monograph [10]. It is interesting to notice that the construction of z^G was mentioned by Krasnoselskiĭ [20] (see Rubinov [21]) in the context of operator theory. It can be shown that z^{G} is monotone if and only if D is additively self-inclusive. As defined above, this is the case if D is a convex cone, for example. Thus, Property (M) in Assumption 3.1 is satisfied if $D + D \subseteq D$. Moreover, we have for all $y \in Y$ and for all $r \in \mathbb{R}$ that $z^{\text{G}}(y) \le r$ if and only if $y \in rk - D$. Thus, Property (S) in Assumption 3.1 is fulfilled as well, see also [10, Proposition 4.1.1]. Moreover, it can be shown [18] that Assumption 3.1(I) is satisfied by the functional z^{G} if the vector k in (I) is chosen as the vector in the definition of z^{G}. Even more, it can be shown, see also [10, Proposition 4.1.1], that *any* functional $z \colon Y \to \overline{\mathbb{R}}$ defined on a linear space Y fulfilling (I) must be of the form (3.4) for a particularly chosen set D.

The functional z^{G} additionally has the benefit that, despite the somewhat complex definition in (3.4), its computation is simple and computationally efficient for many practically relevant cases, see [22] for an example.

For the following two examples, D is assumed to be a proper closed convex cone in Y.

Example 3.7 (Linear Functional)**.** The historically speaking oldest way to apply separation techniques in optimization theory is by means on linear functionals $\ell \in D^*$, where $D^* := \{\ell \in Y^* : \ell(d) \ge 0 \;\forall d \in D\}$. Such a linear functional can directly be used for separation/scalarization as

$$z^{\text{lin}}(y) = \ell(y)$$

but, in general, does not even fulfill the plain property (S). For numerical considerations, however, linear functionals have played and continue to play a pinnacle role in numerical and analytical optimization.

It is a simple application of the linearity of z^{lin} if one wants to show Properties (M) and (I) the first one of which implying (T1). Property (T2), however, cannot be guaranteed in general.

Example 3.8 (Support Functional)**.** A way to narrow down the separation property of linear separation functionals is to use the maximum or supremum value of several linear functionals to quantify the separation. Practically, the set from which these functionals are drawn is usually constrained to not include zero or, what is numerically relevant, to all that have norm one. This motivates the following functional:

$$z^{\text{supp}}(y) = \sigma_{D^*}(y) := \sup_{\ell \in D^*, \|\ell\|=1} \ell(y)$$

Compared to linear functionals, the introduction of the separation functional z^{supp} adds Property (S) while the generality of the norm used in the definition of z^{supp} entails, like for the Hiriart-Urruty functional, that Property (I) does no longer hold in gereral.

Remark 3.4. In the recent work [9], Bouza et al. provide a full characterization (by means of inclusions) of the functions $z^{\text{H}-\text{U}}$ and z^{G} for a Banach space setting. The analysis also includes separation functions of Drummond-Svaiter type [23], which form a generalization of z^{supp} and a new concept in which the construction principles of support functions and oriented distance functions are combined.

To summarize the above examples, we collected the properties of the previously mentioned separation functionals in Table 1. It is apparent how z^{G} in general combines most properties of

		(S)	(M)	(T2)	(T1)	(I)
general D	indicator $z^{\text{ind}} = 1 - \psi_{-D}$	$+$	$*$	$-$	$*$	$-$
	characteristic $z^{\text{char}} = \chi_{-D}$	$+$	$*$	$-$	$*$	$-$
D closed	distance $z^{\text{dist}} = d_{-D}$	$+$	$*$	$-$	$*$	$-$
	squared distance $z^{\text{d}-\text{sq}} = (d_{-D})^2$	$+$	$*$	$-$	$*$	$-$
	Hiriart-Urruty $z^{\text{H}-\text{U}} = \triangle_{-D}$	$+$	$*$	$-$	$*$	$-$
	nonlinear scalarization $z^{\text{G}} = z^k$	$+$	$*$	$-$	$*$	$+$
D cone	linear $z^{\text{lin}} = \ell \in D^*$	$-$	$+$	$-$	$+$	$+$
	support functional $z^{\text{supp}} = \sigma_{D^*}$	$+$	$+$	$-$	$+$	$-$

$+$ ... fulfilled
$-$ ... not fulfilled
$*$ ... fulfilled if D is additively self-inclusive

TABLE 1. Interrelation of the functionals presented throughout Examples 3.1 to 3.8

Assumption 3.1, making it a good candidate for scalarization in set optimization applications. Given the equivalence of Property (I) and the definition of z^{G}, see [10, Proposition 4.1.1], any scalarization result relying on this property will in some way require the involvement of z^{G} as defined in (3.4). Another noticeable observation is that, unless further assumptions are imposed, none of the so far introduced separation concepts fulfill Property (T2) and only the simplest, the linear separation, does not fulfill the basic separation property (S).

In a next step, we will show how the introduced properties can be used for determining algebraic characterizations of the upper set less relation. Similar results can be found in [8] and [24] for the special cases of D being a cone, in [25] for more general sets D and in [26] for the case of general linear spaces without assumptions on an underlying topology.

We start with Property (M), where the following implication can be deduced by an element-wise consideration of the inclusion $A \subseteq B - D$.

Theorem 3.1. *Let Assumption 3.1(M), be satisfied and consider $A, B \in \mathscr{P}(Y)$. Then it holds*

$$A \preceq_D^u B \Leftrightarrow A \subseteq B - D \implies \sup_{a \in A} z(a) \leq \sup_{b \in B} z(b). \tag{3.5}$$

The following corollary is a simple consequence of the observations in Figure 1.

Corollary 3.1. *Let Assumption 3.1(T1) be satisfied and $0 \in D$. Then (3.5) holds for $A, B \in \mathscr{P}(Y)$.*

Using the transfer of inclusion property (T2), an equivalent implication follows immediately, but also the inverse can be shown:

Theorem 3.2. *Let Assumption 3.1(T2), be satisfied and consider $A, B \in \mathscr{P}(Y)$. Then it holds*

$$A \preceq_D^u B \Leftrightarrow A \subseteq B - D \implies \sup_{a \in A} z(a) \leq \sup_{b \in B} z(b).$$

If, in addition, $\sup_{b \in B} z(b)$ is attained, then

$$A \preceq_D^u B \Leftrightarrow A \subseteq B - D \impliedby \sup_{a \in A} z(a) \leq \sup_{b \in B} z(b).$$

Proof. As outlined, the first implication is immediate. To verify the second implication, let $\sup_{a \in A} z(a) \leq \sup_{b \in B} z(b)$ and assume that there is some $a \in A$ with $a \notin B - D$, then $z(a) > z(b)$ for all $b \in B$. Because $\sup_{b \in B} z(b)$ is attained, we get that $\sup_{a \in A} z(a) > \sup_{b \in B} z(b)$, in contradiction to the assumption. □

Example 3.9. Let Y be a real linear space, $D \subseteq Y$ be a halfspace, and $A, B \subseteq Y$ be nonempty. We consider for $y \in Y$ the functional $z(y) := \ell(y)$, with $\ell \in D^*$ (that is, the separation defining the halfspace), and assume that $\sup_{b \in B} z(b)$ is attained. Then z fulfills Assumption 3.1(T2), and we have

$$A \preceq_D^u B \Leftrightarrow A \subseteq B - D \iff \sup_{a \in A} z(a) \leq \sup_{b \in B} z(b). \tag{3.6}$$

In order to show a corresponding result and omitting the attainment property, we can require a strong inequality, as proposed in the following theorem.

Theorem 3.3. *Let Assumption 3.1(T2), be satisfied and consider $A, B \in \mathscr{P}(Y)$. Then*

$$A \preceq_D^u B \Leftrightarrow A \subseteq B - D \impliedby \sup_{a \in A} z(a) < \sup_{b \in B} z(b).$$

Proof. Let $\sup_{a \in A} z(a) < \sup_{b \in B} z(b)$ and assume that there is some $a \in A$ with $a \notin B - D$, then $z(a) > z(b)$ for all $b \in B$. We get that $\sup_{a \in A} z(a) \geq \sup_{b \in B} z(b)$, in contradiction to the assumption. □

Remark 3.5. An equivalence similar to (3.6) is proven in [8] and [24] for convex sets A,B without requiring (T2). The main difference in this analysis is that a *family* of scalarizations is considered in these works which then allows to obtain global results by taking further infima and suprema over these families. Since it is our goal here to base the separation on single expressions, we cannot (directly) apply these techniques.

The above examples showed that (T2) is an exceptionally hard requirement on the separation function. It gives, however, a global scalarization result in terms of the equivalence (3.6) which allows to fully transform a set-valued optimization problem to a (family of) scalar one(s). When working with 'weaker' separation functionals, the set relation '$\preceq_D^u$' can still be transformed into an inequality. The involved scalarizations are, however, no longer separated such that a plain application of 'out-of-the-box' optimization routines is no longer possible.

Theorem 3.4. *Let Assumption 3.1(S), be satisfied and consider $A,B \in \mathscr{P}(Y)$. Then it holds*

$$A \subseteq B - D \implies \sup_{a\in A} \inf_{b\in B} z(a-b) \leq 0.$$

If, additionally, for all $a \in A$, $\inf_{b\in B} z(a-b)$ is attained, then

$$A \subseteq B - D \impliedby \sup_{a\in A} \inf_{b\in B} z(a-b) \leq 0.$$

Proof. (1) Assume that for all $a \in A$, there exists some $b \in B$ such that $a-b \in -D$. This immediately yields $\sup_{a\in A} \inf_{b\in B} z(a-b) \leq 0$.

(2) Let $\sup_{a\in A} \inf_{b\in B} z(a-b) \leq 0$. If, by contradiction, there is some $a \in A$ with $a \notin B-D$, then $z(a-b) > 0$ for all $b \in B$. Because for all $a \in A$, $\inf_{b\in B} z(a-b)$ is attained, we get that $\sup_{a\in A} \inf_{b\in B} z(a-b) > 0$, a contradiction.

□

In the second part of Theorem 3.4, we need the assumption that $\inf_{b\in B} z(a-b)$ is attained for all $a \in A$. Sufficient conditions for such an attainment property, i.e., assertions concerning the existence of solutions of the corresponding optimization problems (extremal principles) are given in the literature. The well-known Theorem of Weierstrass says that a lower semi-continuous function on a nonempty weakly compact set in a reflexive Banach space has a minimum. An extension of the Theorem of Weierstrass is given by Zeidler [27, Proposition 9.13]: A proper lower semi-continuous and quasi-convex function on a nonempty closed bounded convex subset of a reflexive Banach space has a minimum.

In case that we cannot rely on continuity assumptions of the function z, we propose the following theorem which is able to omit any attainment property and is loosely based on [26, Theorem 3.6].

Theorem 3.5. *Let Assumption 3.1,(S) and (I) be satisfied with $k \in -D$ and consider $A,B \in \mathscr{P}(Y)$. Assume further that D is a convex cone. Then*

$$A \subseteq B - D \iff \sup_{a\in A} \inf_{b\in B} z(a-b) \leq 0.$$

Proof. The first implication holds, as has been already shown. Now, let $\sup_{a\in A} \inf_{b\in B} z(a-b) \leq 0$. This means that for all $\varepsilon > 0$ and for all $a \in A$, $\inf_{b\in B} z(a-b) < \varepsilon$. This means that

$$\forall \varepsilon > 0,\ \forall a \in A,\ \exists b \in B: \quad z(a-b) - \varepsilon = z(a-b-\varepsilon k) < 0.$$

This yields

$$A \subseteq B + \varepsilon k - D \overset{\substack{k\in -D\\ D \text{ cone}}}{\subseteq} B - D - D \overset{D \text{ add. self-incl.}}{\subseteq} B - D,$$

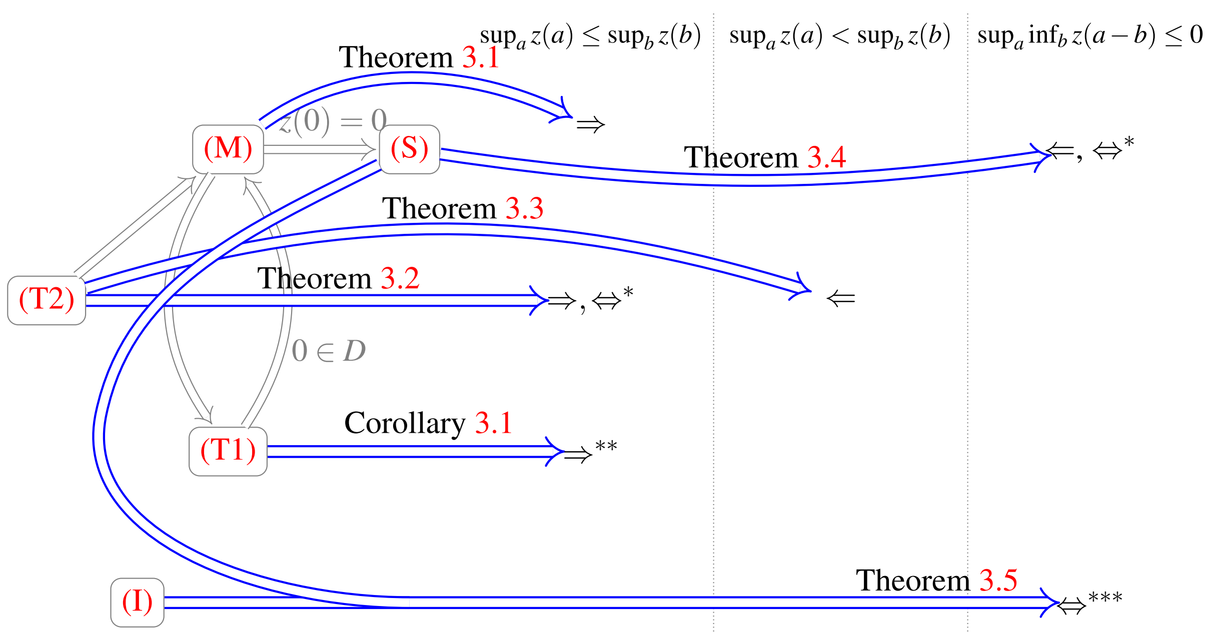
and the proof is complete. □

Short summary. We have in general two types of characterizations for the fulfilled set relation between $A \in \mathscr{P}(Y)$ and $B \in \mathscr{P}(Y)$:

(1) $\sup\limits_{a\in A} z(a) \leq \sup\limits_{b\in B} z(b)$ like in (3.6) gives the strongest and most practical result. If this is fulfilled, set optimization problems can be solved directly using algorithms for scalar problems, cf. Theorem 4.2 below.

(2) $\sup\limits_{a\in A} \inf\limits_{b\in B} z(a-b) \leq 0$ gives the means for a descent method based on local comparisons, cf. Section 5 below.

In Figure 2 we summarize the findings of the preceding theorems and corollaries graphically. The implication/equivalence arrows in the columns correspond to the main claims of the theorems; such that, for example, the '$\Leftrightarrow^{***}$' in the bottom right corner can be read as: If (I) and (S) are fulfilled, with $k \in -D$ and D is a convex cone, then Theorem 3.5 says that $A \preceq^u_D B$ is equivalent to $\sup_a \inf_b z(a-b) \leq 0$.



* ... attainment property is fulfilled.

** ... $0 \in D$ holds.

*** ... D is a convex cone, $k \in -D$.

FIGURE 2. Assumptions on separation functions revisited: Implications on characterizations of $\preceq^u_D$

4. Finding Minimal Elements of a Family of Sets by Means of General Functionals

In this section, we will shortly review notions of minimality as they are used in set optimization using the so-called set approach. For the purpose of an easier reference to the literature, we will only work in $\mathscr{P}(Y)$ throughout this section and use a generic set relation $\preceq$. Later on, the sets A, $\bar{A}$ etc. will be replaced by the image sets of a set-valued mapping f and $\preceq$ specified to the upper set less relation.

Definition 4.1. Let $\mathscr{A}$ be a family of elements of $\mathscr{P}(Y)$, $D \subseteq Y$ and $\preceq$ be a binary relation on $\mathscr{P}(Y)$. We call $\bar{A} \in \mathscr{A}$ a *minimal element* of $\mathscr{A}$ w.r.t. $\preceq$ if

$$A \preceq \bar{A},\, A \in \mathscr{A} \quad \Longrightarrow \quad \bar{A} \preceq A\,.$$

The set of all minimal elements of $\mathscr{A}$ w.r.t. $\preceq$ will be denoted by $\mathscr{A}_{\preceq}$.

Definition 4.2. Let $z\colon Y \to \overline{\mathbb{R}}$ be a functional, which shall later be equipped with some of the properties from Assumption 3.1. Let $\mathscr{A}$ be a family of nonempty elements of the power set of Y and $D \subseteq Y$. We call $\bar{A} \in \mathscr{A}$ a *strict minimal solution* of the problem

$$\min_{A \in \mathscr{A}} \sup_{a \in A} z(a) \tag{4.1}$$

if there does not exist an $A \in \mathscr{A} \setminus \{\bar{A}\}$ such that $\sup_{a \in A} z(a) \leq \sup_{\bar{a} \in \bar{A}} z(\bar{a})$. An element $\bar{A} \in \mathscr{A}$ is called a *minimal solution* of the problem (4.1) if there does not exist an $A \in \mathscr{A}$ such that $\sup_{a \in A} z(a) < \sup_{\bar{a} \in \bar{A}} z(\bar{a})$. The set of all (strict) minimal solutions of the problem (4.1) will be denoted by $\mathscr{A}_{\sup}$ ($\mathscr{A}^s_{\sup}$, respectively).

Remark 4.1. Note that $\mathscr{A}^s_{\sup} \subseteq \mathscr{A}_{\sup}$.

Theorem 4.1. *Let Assumption 3.1,(M), be fulfilled, $\mathscr{A}$ be a family of elements of $\mathscr{P}(Y)$, $D \subseteq Y$ and $\preceq$ be $\preceq^u_D$ the upper set relation on $\mathscr{P}(Y)$. Then we have $\mathscr{A}_{\sup} \subseteq \mathscr{A}_{\preceq^u_D}$.*

Proof. Let $\bar{A} \in \mathscr{A}_{\sup}$, but suppose by contradiction that $\bar{A} \notin \mathscr{A}_{\preceq^u_D}$. Then there exists some $A \in \mathscr{A}$ such that $A \preceq^u_D \bar{A}$ and $\bar{A} \not\preceq^u_D A$. By Theorem 3.1, this implies $\sup_{a \in A} < \sup_{\bar{a} \in \bar{A}} z(\bar{a})$, in contradiction to $\bar{A} \in \mathscr{A}_{\sup}$. □

The reverse inclusion ($\mathscr{A}_{\sup} \supseteq \mathscr{A}_{\preceq}$) is possible under stricter assumptions on the functional z:

Theorem 4.2. *Let Assumption 3.1,(T2), be fulfilled, $\mathscr{A}$ be a family of elements of $\mathscr{P}(Y)$, $D \subseteq Y$ and $\preceq^u_D$ be the upper set relation on $\mathscr{P}(Y)$. Then we have $\mathscr{A}_{\sup} = \mathscr{A}_{\preceq^u_D}$.*

Proof. The first inclusion is shown analogously to the proof of Theorem 4.1. Conversely, let $\bar{A} \in \mathscr{A}_{\preceq^u_D}$ and assume that $\bar{A} \notin \mathscr{A}_{\sup}$. Then there exists some $A \in \mathscr{A}$ such that $\sup_{a \in A} z(a) < \sup_{\bar{a} \in \bar{A}} z(\bar{a})$. By Theorem 3.3, this implies $A \preceq^u_D \bar{A}$, which in turn implies by the minimality of $\bar{A}$ (in terms of Definition 4.1) that $\bar{A} \preceq^u_D A$, leading to $\sup_{\bar{a} \in \bar{A}} z(\bar{a}) \leq \sup_{a \in A} z(a)$, a contradiction. □

Theorem 4.2 shows that looking for minimal elements in terms of Definition 4.1 with $\preceq = \preceq^u_D$ reduces to solving for solutions of the scalarized problem (4.1), i.e., the problem $\min_{A \in \mathscr{A}} \sup_{a \in A} z(a)$, assuming that Assumption 3.1,(T2) holds.

It is possible to show a similar assertion as in Theorem 4.2 for strict minimal solutions of (4.1). For this, we need the following assumption.

Assumption 4.1. Let $z\colon Y \to \overline{\mathbb{R}}$ be given. For all $A \in \mathscr{A}$, suppose that $\sup_{a\in A} z(a)$ is attained.

Definition 4.3. Let a set relation $\preceq$ and a family of sets $\mathscr{A}$ be given. We say that $\mathscr{A}$ is an *antisymmetric* family of sets if for all $A, \bar{A} \in \mathscr{A}$ we have that

$$\left(A \preceq \bar{A}\right) \wedge \left(\bar{A} \preceq A\right) \Longrightarrow \bar{A} = A$$

Note that the set relation $\preceq$ itself does not need to be antisymmetric itself for $\mathscr{A}$ to be an antisymmetric family of sets.

Theorem 4.3. *Let Assumption 3.1,(T2), and Assumption 4.1 be fulfilled. Furthermore, let $\mathscr{A}$ be an antisymmetric family of elements of $\mathscr{P}(Y)$, $D \subseteq Y$. Then we have $\mathscr{A}^s_{\sup} = \mathscr{A}_{\preceq^u_D}$.*

Proof. The first inclusion follows from Theorem 4.2, as $\mathscr{A}^s_{\sup} \subseteq \mathscr{A}_{\sup}$. To show the other inclusion, let $\bar{A} \in \mathscr{A}_{\preceq^u_D}$ and assume that $\bar{A} \notin \mathscr{A}^s_{\sup}$. Then there exists some $A \in \mathscr{A}$, $A \neq \bar{A}$ such that $\sup_{a\in A} z(a) \leq \sup_{\bar{a}\in \bar{A}} z(\bar{a})$. By Theorem 3.2, this implies $A \preceq^u_D \bar{A}$, which in turn implies by the minimality of $\bar{A}$ (in terms of Definition 4.1) that $\bar{A} \preceq^u_D A$. Because of the antisymmetry of $\preceq^u_D$, this leads to $\bar{A} = A$, a contradiction. □

Theorem 4.3 above shows that under suitable assumptions, it is sufficient to solve for strict minimal solutions of problem (4.1) in order to obtain all minimal elements of $\mathscr{A}$.

Algorithms based on Theorems 4.1, 4.2 and 4.3 are straightforward and can be derived in various ways. As the main goal is to solve (4.1), one can derive numerical procedures using, for example, trust-region methods, Quasi-Newton methods, or other well-established methods for solving an unconstrained problem. Such methods shall thus not be considered further here. Instead, we will focus on derivative-free descent methods in Section 5, which will be based on Theorems 3.4 and 3.5.

The theorem below can be easily proven by means of Theorem 3.4.

Theorem 4.4. *Let Assumption 3.1,(S), and Assumption 4.1 be fulfilled. Let $\mathscr{A}$ be a family of elements of $\mathscr{P}(Y)$, $D \subseteq Y$. Then $\bar{A} \in \mathscr{A}$ is a minimal element if and only if*

$$\nexists A \in \mathscr{A} : \sup_{a\in A} \inf_{\bar{a}\in\bar{A}} z(a-\bar{a}) \leq 0 \wedge \sup_{\bar{a}\in\bar{A}} \inf_{a\in A} z(\bar{a}-a) > 0.$$

Theorem 3.5 directly leads to the following result.

Theorem 4.5. *Let Assumption 3.1,(S) and (I) with $k \in -D$, and assume that $D \subseteq Y$ is a convex cone. Let $\mathscr{A}$ be a family of elements of $\mathscr{P}(Y)$. Then $\bar{A} \in \mathscr{A}$ is a minimal element if and only if*

$$\nexists A \in \mathscr{A} : \sup_{a\in A} \inf_{\bar{a}\in\bar{A}} z(a-\bar{a}) \leq 0 \wedge \sup_{\bar{a}\in\bar{A}} \inf_{a\in A} z(\bar{a}-a) > 0.$$

Remark 4.2. Hernández and Rodríguez-Marín [28] introduce an extension of the nonlinear scalarization functional z^k as given in Example 3.6 which uses pairs of sets as input in the first place. Adapted to the upper set less relation, this extension reads

$$Z^{D,k}\colon \mathscr{P} \times \mathscr{P} \to \overline{\mathbb{R}} \quad Z^{D,k}(A,B) := \sup_{a\in A} \inf\{t \in \mathbb{R} : a \in t\cdot k + B - D\}.$$

Under appropriate assumptions, it is in this respect possible to show the characterization

$$A \preceq^u_D B \quad \Leftrightarrow \quad Z^{D,k}(A,B) \leq 0,$$

which could be used just like the above characterization $\sup_{a\in A}\inf_{b\in B} z(a-b)\leq 0$ for the numerical procedures below as a quantified version of the set relation, see also [29, 30] for this approach.

5. A Descent Method

In the literature, there already exist some algorithms for solving set-valued optimization problems based on descent methods. For example, Jahn [31] proposes a descent method that generates approximations of minimal elements of set-valued optimization problems under convexity assumptions on the considered sets. In [31], the set less relation is characterized by means of linear functionals. More recently, in [24], the authors propose a similar descent method for obtaining approximations of minimal elements of set-valued optimization problems. In [24], several set relations are characterized by the nonlinear scalarizing functional z^{G}, where D is assumed to be a proper convex cone. Since the nonlinear functional z^{G} is used in [24], no convexity assumptions on the considered outcome sets $F(x)$ are needed. Note that the approaches in [31, 24] all rely on set relations where the involved domination structure is given by cones, whereas in this section, we consider arbitrary nonemtpy sets $D\subset Y$.

Here we are trying to solve the set-valued problem

$$\min_{x\in\mathbb{R}^n} F(x)\,, \tag{P}$$

where $F\colon \mathbb{R}^n \rightrightarrows Y$ is a set-valued mapping.

Definition 5.1 (Minimal Solutions of Problem (P))**.** We say that $\bar{x}\in S$ is a **minimal solution** of (P) w.r.t. $\preceq$ if $F(\bar{x})$ is a minimal element of the family of sets $F(x)$, $x\in\mathbb{R}^n$ w.r.t. $\preceq$.

The results in Section 3 give rise to develop an algorithmic procedure to compute approximations of minimal solutions of (P). A natural way of constructing an algorithm for solving problem (P) is an iterative pattern search where in each iteration the minimal function value is determined to specify the locally best search direction. For this reason we refer to Algorithm 5.1 below as a *descent method*, cf. [31].

The following algorithm calculates an approximation of a minimal solutions of the set-valued problem (P) w.r.t. $\preceq^u_D$, where D is assumed to be a convex cone, such that $\preceq^u_D$ is a preorder. The main tool is a general functional z which fulfills either

- Assumption 3.1,(S) and for all $x\in\mathbb{R}^n$ and $y\in F(x)$, $\inf_{b\in F(x)} z(y-b)$ is attained; or
- Assumption 3.1,(S) and (I) with $k\in -D$, and D is a convex cone (or additively self-inclusive respectively).

Algorithm 5.1. (A descent method for finding an approximation of a minimal solution of the set-valued problem (P))

Input: $F\colon\mathbb{R}^n\rightrightarrows\mathbb{R}^m$, starting point $x^0\in\mathbb{R}^n$, functional z fulfilling one of the above assumptions, maximal number i_{max} of iterations, number of search directions n_s, maximal number j_{max} of iterations for the determination of the step size, initial step size h_0 and minimum step size $h_{\min}$

% initialization
$i := 0$, $h := h_0$
choose n_s points $\tilde{x}^1, \tilde{x}^2, \ldots, \tilde{x}^{n_s}$ on the unit sphere around $0_{\mathbb{R}^n}$

% iteration loop
while $i \leq i_{max}$ **do**
check $F(x^i + h\tilde{x}^j) \preceq F(x^i)$ for every $j \in \{1, \dots, n_s\}$ by evaluating the term $\sup_{a\in A} \inf_{b\in B} z(a-b)$ for $A = F(x^i + h\tilde{x}^j)$ and $B = F(x^i)$. Choose the index $n_0 := j$ with the smallest function value $\sup_{a\in A} \inf_{b\in B} z(a-b)$.
if $\sup_{a\in F(x^i+h\tilde{x}^j)} \inf_{b\in F(x^i)} z(a-b) \leq 0$, **then**
$x^{i+1} := x^i + h\tilde{x}^{n_0}$ % new iteration point
$j := 1$
while $\sup_{a\in F(x^i+(j+1)h\tilde{x}^{n_0})} \inf_{b\in F(x^i+jh\tilde{x}^{n_0})} z(a-b) \leq 0$ and $j \leq j_{max}$ **do**
$j := j+1$
$x^{i+1} := x^{i+1} + h\tilde{x}^{n_0}$ % new iteration point
end while
else
$h := h/2$
if $h \leq h_{\min}$, **then**
STOP $x := x^i$
end if
end if
$i := i+1$
end while

Output: An approximation x of a minimal solution of the set-valued problem (P) w.r.t. $\preceq^u_D$.

Remark 5.1. An important ingredient of Algorithm 5.1 is the choice of the index n_0, which is the important point where the *scalarization/descent* idea is actively used. Note that taking such a minimum is not reasonable for 'simple' separation functionals such as z^{ind}.

For one given starting point x^0, Algorithm 5.1 approximates one minimal solution of problem (P). To find more than one approximation of minimal solutions, one needs to vary the input parameters, such as choosing a different starting point $x^0 \in \mathbb{R}^n$, or choosing another functional z which should fulfill the required attainment property.

Algorithms like this have been proposed and applied in [31] and [24] where also numerical examples for some small-scale benchmarks can be found. Note, however, that the algorithms in these references are usually tailor-made and studied for a particular choice of the separation functional whereas within this contribution, we point out how the construction of a more general-purpose solution procedure can be based on abstract properties of z.

6. Summary and Outlook

In this paper, we proposed characterizations of the upper set less relation by means of inequalities containing separation functionals. The main contribution lies in not focusing on well-established examples of such separation concepts but instead build upon general properties of such. We showed that simple separation properties lead to a local characterization that can be exploited by a local grid search algorithm whereas stronger properties lead to a global scalarization result. A practitioner is in that light faced by a trade-off between easy-to-compute scalarizations that can then only be used locally or complex ones allowing for powerful established algorithms from scalar optimization. Out findings have been underlined by several well-studied separation concepts from the literature.

Further research shall include the following aspects:

(1) further properties of D and the image sets like closedness and/or even compactness may allow for further algorithmic improvement.
(2) The same is true for further properties of z like Lipschitz-continuity, (quasi-)/(sub-) differentiability such that derivative-information can improve the descent-finding in Algorithm 5.1.
(3) Finally, to fully make use of the obtained results, in particular the '$\Rightarrow$' implications in the first column of Figure 2, the algorithm can be assisted by a pre-sorting routine. Similar ideas have been used for discrete set optimization problems, see [32].

Acknowledgments
The work of M. Köbis was carried out during the tenure of an ERCIM 'Alain Bensoussan' Fellowship of the author at the Norwegian University of Science and Technology. The authors are grateful to the reviewer for providing helpful suggestions that led to an improvement of the first version of the manuscript.

REFERENCES

[1] A. A. Khan, Chr. Tammer, C. Zălinescu, Set-valued Optimization, Springer Berlin Heidelberg, 2015.
[2] J. Ide, E. Köbis, D. Kuroiwa, A. Schöbel, C. Tammer, The relationship between multi-objective robustness concepts and set-valued optimization, Fixed Point Theory Appl. 2014 (2014), 83.
[3] D. Kuroiwa, The natural criteria in set-valued optimization, Sūrikaisekikenkyūsho Kōkyūroku, Research on nonlinear analysis and convex analysis, pp. 85-90, 1997.
[4] D. Kuroiwa, Some duality theorems of set-valued optimization with natural criteria, In T. Tanaka, (ed.) Proceedings of the International Conference on Nonlinear Analysis and Convex Analysis, pp. 221-228. World Scientific, 1999.
[5] E. Köbis, D. Kuroiwa, C. Tammer, Generalized set order relations and their numerical treatment, Appl. Anal. Optim. 1 (2017), 45-65.
[6] C. Gerstewitz (Tammer), Nichtkonvexe Dualitität in der Vektoroptimierung, Wissenschaftliche Zeitschrift Technische Hochschule Leuna-Merseburg, 25 (1983), 357-364.
[7] J. B. Hiriart-Urruty, Tangent cones, generalized gradients and mathematical programming in banach spaces, Math. Oper. Res. 4 (1979), 79-97.
[8] J. Jahn, Vectorization in set optimization, J. Optim. Theory Appl. 167 (2015), 783–795.
[9] G. Bouza, E. Quintana, C. Tammer, A unified characterization of nonlinear scalarizing functionals in optimization, Vietnam J. Math. 47 (2019), 683-713.
[10] C. Tammer, P. Weidner, Scalarization and Separation by Translation Invariant Functions, Springer International Publishing, 2020.
[11] A. Zaffaroni, Degrees of efficiency and degrees of minimality, SIAM J. Control Optim. 42 (2003), 1071-1086.
[12] L. Tang, Y. Gao, Some properties of nonconvex oriented distance function and applications to vector optimization problems, J. Indust. Mang. Optim. 17 (2021), 485-500.
[13] C. Gerstewitz (Tammer), Beiträge zur Dualitätstheorie der nichtlinearen Vektoroptimierung, PhD thesis, Technische Hochschule Leuna-Merseburg, 1984.
[14] C. Gerth (Tammer), E. Iwanow, Dualität für nichtkonvexe Vektoroptimierungsprobleme, Wissenschaftliche Zeitschrift Technische Hochschule Ilmenau, 2 (1985), 61-81.
[15] C. Gerth (Tammer), P. Weidner, Nonconvex separation theorems and some applications in vector optimization, J. Optim. Theory Appl. 67 (1990), 297-320.

[16] F. Heyde, Coherent risk measures and vector optimization, In K.-H. Küfer, (ed.) Multicriteria decision making and fuzzy systems theory, methods and applications, chapter 1, pp. 3-12. Shaker, Aachen, 2006.
[17] K. Klamroth, E. Köbis, A. Schöbel, C. Tammer, A unified approach for different concepts of robustness and stochastic programming via non-linear scalarizing functionals, Optimization 62 (2013), 649-671.
[18] A. Göpfert, H. Riahi, C. Tammer, C. Zălinescu, Variational Methods in Partially Ordered Spaces, Springer-Verlag, 2003.
[19] P. Weidner, Ein Trennungskonzept und seine Anwendung auf Vektoroptimierungsverfahren, Dissertation B, Martin Luther University Halle–Wittenberg, 1990.
[20] M. A. Krasnoselskiĭ, Positive solutions of operator equations, P. Noordhoff Ltd., Groningen, 1964. (Translated from Russian by Richard E. Flaherty.)
[21] A. M. Rubinov, Sublinear operators and their applications Russian Math. Surveys 32 (1977), 115-175.
[22] C. Tammer, K. Winkler, A new scalarization approach and applications in multicriteria d. c. optimization, J. Nonlinear Convex Anal. 4 (2003), 365-380.
[23] L. M. Graña Drummond, B. F. Svaiter, A steepest descent method for vector optimization, J. Comput. Appl. Math. 175 (2005), 395-414.
[24] E. Köbis and M. A. Köbis, Treatment of set order relations by means of a nonlinear scalarization functional: a full characterization, Optimization, 65 (2016), 1805-1827.
[25] E. Köbis, M. A. Köbis, J.-C. Yao, Generalized upper set less order relation by means of a nonlinear scalarizing functional, J. Nonlinear Convex Anal. 17 (2016), 725-734.
[26] N. Hebestreit, E. Köbis, Representation of set relations in real linear spaces, J. Nonlinear Convex Anal. 19 (2018), 287-296.
[27] E. Zeidler, Nonlinear Functional Analysis and its Applications I: Fixed-Point Theorems, Springer New York, 1986.
[28] E. Hernández, L. Rodríguez-Marín, Nonconvex scalarization in set optimization with set-valued maps, J. Math. Anal. Appl. 325 (2007), 1-18.
[29] C. Gutiérrez, B. Jiménez, E. Miglierina, E. Molho, Scalarization in set optimization with solid and nonsolid ordering cones, J. Global optim. 61 (2014), 525-552.
[30] C. Gutiérrez, B. Jiménez, V. Novo, Nonlinear scalarizations of set optimization problems with set orderings, In: Set Optimization and Applications - The State of the Art, pp. 43-63. Springer Berlin Heidelberg, 2015.
[31] J. Jahn, A derivative-free descent method in set optimization, Comput. Optim. Appl. 60 (2015), 393-411.
[32] C. Günther, N. Popovici, New algorithms for discrete vector optimization based on the Graef-Younes method and cone-monotone sorting functions, Optimization, 67 (2018), 975-1003.