Discussiones Mathematicae
Differential Inclusions, Control and Optimization 22 (2002) 33–66

APPROXIMATION OF SET-VALUED FUNCTIONS BY SINGLE-VALUED ONE

Ivan Ginchev

Technical University of Varna
BG-9010 Varna, Bulgaria

e-mail: ginchev@ms3.tu-varna.acad.bg

and

Armin Hoffmann

Technical University of Ilmenau
D-98684 Ilmenau, PF 100565, Germany

e-mail: armin.hoffmann@tu-ilmenau.de

Abstract

Let $\Sigma : M \to 2^Y \setminus \{\emptyset\}$ be a set-valued function defined on a Hausdorff compact topological space M and taking values in the normed space $(Y, \|\cdot\|)$. We deal with the problem of finding the best Chebyshev type approximation of the set-valued function Σ by a single-valued function g from a given closed convex set $V \subset C(M, Y)$. In an abstract setting this problem is posed as the extremal problem $\sup_{t\in M} \ \rho(g(t),\ \Sigma(t)) \to \inf,\ g \in V$. Here ρ is a functional whose values $\rho(q, S)$ can be interpreted as some distance from the point q to the set $S \subset Y$. In the paper, we are confined to two natural distance functionals $\rho = H$ and $\rho = D$. $H(q, S)$ is the Hausdorff distance (the excess) from the point q to the set $\operatorname{cl} S$, and $D(q, S)$ is referred to as the oriented distance from the point q to set $\operatorname{cl}\operatorname{conv} S$. We prove that both these problems are convex optimization problems. While distinguishing between the so called regular and irregular case problems, in particular the case $V = C(M, Y)$ is studied to show that the solutions in the irregular case are obtained as continuous selections of certain set-valued maps. In the general case, optimality conditions in terms of directional derivatives are obtained of both primal and dual type.

Keywords: Chebyshev approximation, set-valued functions, convex optimization.

2000 Mathematics Subject Classification: 41A50, 41A65, 49J53, 90C25.

1. Introduction

In the paper, M is a Hausdorff compact topological space, Y is a normed space with its dual Banach space Y^*, $C(M,Y)$ is the space of the continuous functions from M to Y supplied with the sup-norm, $V \subset C(M,Y)$ is a nonempty closed and convex subset of $C(M,Y)$ and $\Sigma : M \to 2^Y$ is a set-valued function from M into the subsets of Y with $\operatorname{dom}\Sigma := \{t \in M \,|\Sigma(t) \neq \emptyset\}$ and $\operatorname{range}\Sigma = \bigcup_{t\in M}\Sigma(t)$.

We consider two **Chebyshev type approximation problems**

$$\textbf{(P}^H\textbf{)} \qquad \varphi^H(g) := \sup_{t\in M} H(g(t), \Sigma(t)) \to \inf, \quad g \in V,$$

$$\textbf{(P}^D\textbf{)} \qquad \varphi^D(g) := \sup_{t\in M} D(g(t), \Sigma(t)) \to \inf, \quad g \in V,$$

where the functionals $H, D : Y \times 2^Y \to \overline{\mathbb{R}} := \mathbb{R} \cup \{-\infty\} \cup \{+\infty\}$ are defined for $q \in Y$ and $S \subset Y$ as follows

$$H(q,S) := \sup_{\|y^*\|=1} \left(\sigma_S(y^*) - \langle y^*, q\rangle\right), \tag{1}$$

$$D(q,S) := \sup_{\|y^*\|=1} \left(\iota_S(y^*) - \langle y^*, q\rangle\right). \tag{2}$$

Here y^* varies in the unit sphere ∂B^* of the dual space Y^* of Y. The support functionals of S

$$\sigma_S(y^*) := \sup_{s\in S} \langle y^*, s\rangle \tag{3}$$

and

$$\iota_S(y^*) := \inf_{s\in S} \langle y^*, s\rangle \tag{4}$$

are functions of the variable $y^* \in Y^*$ and the set $S \subset Y$ being in y^* positively homogeneous, subadditive and positively homogeneous, superadditive, respectively. We put $\sigma_S(y^*) = -\infty$ and $\iota_S(y^*) = +\infty$ if $S = \emptyset$. We summarize well-known properties of σ, ι, H, D and add some further suitable properties. We give a geometric interpretation in case $S = \operatorname{cl}\operatorname{conv} S$, whence it is clear that $H(q,S)$ and $D(q,S)$ can be referred to as the **Hausdorff distance** (the excess) and the **oriented distance** from q to S, respectively. The values of $H(q,S)$ are nonnegative, while $D(q,S)$ can take both nonnegative and negative values. While the Hausdorff distance $H(q,S)$ appears often in the literature, e.g. [14] or [10], concerning $D(q,S)$ mostly the case $D(q,S) \geq 0$

is studied in relation to the metric projection. Some prerequisites for the case of an arbitrary $D(q, S)$ are given in [12] and the one-dimensional case $Y = \mathbb{R}$ appeared in [9].

Practical aspects of the problems under consideration are the following. If we deal with Problem (P^H) we may consider $\Sigma(t)$ as a set of possible empirical data at $t \in M$ (if $M \subset \mathbb{R}$ we may interpret t as time), the set-valuedness occurs as a result of noise or errors in measurement. If V is the set of the pattern functions, then the solution to Problem (P^H) could be referred to as the "*worst case*" best approximation. In contrast to Problem (P^H) the method of the least squares, usually used in practice to approximate empirical data, only averages but does not particularize the worst data for each given t. Some related problems concerning best approximation of a given function or empirical data can be found in [17], [11] and [16]. Problem (P^D) occurs in the continuous selection theory, where for some reasons the best in some sense continuous selection of the given set-valued function Σ must be chosen. Such continuous selections appear for instance in the viability theory, see [1].

Similarities in the definitions and properties of the functionals H and D and of Problems (P^H) and (P^D) are the reason not to consider the two problems separately.

In the present paper, we generalize [9], where the particular case $Y = \mathbb{R}$ is considered. The paper is organized as follows. In Section 2, we summarize essential properties of the functionals σ, ι, H and D and also of the functionals φ^H and φ^D and state that Problems (P^H) and (P^D) are convex which allows in particular directional derivatives to describe their solutions. In Section 3, a concept of regularity is developed. In particular, we handle the special irregular case $V = C(M, Y)$, which demonstrates that in the irregular case the solutions are obtained as continuous selections of certain level set maps. In Section 4, we calculate exemplarily the directional derivative of φ^H and φ^D and establish optimality criteria of primal and dual type for both problems. As tools standard arguments from variational analysis, see e.g. [20], [16], are used. We point out that in the irregular case the optimality criteria collapse to conditions for single points of M.

2. Support and distance functionals

The functionals H and D defined by formulas (1) and (2), respectively, use σ_S and ι_S. The function σ_S is well-known as the **support function** of S.

The counter part ι_S is not explicitly exposed in the literature. However, see [18, p. 188], the difference $\sigma_S(y^*) - \iota_S(y^*)$ is known as the width of the body S with respect to the direction $y^* \in Y^*$. The representation

$$\iota_S(y^*) = -\sigma_S(-y^*) \tag{5}$$

directly follows from their definitions in (4) and (3). Hence it is convenient to obtain properties of ι_S from adequate properties of σ_S. We denote by B and B^* the **open unit balls** in Y, Y^*, respectively. The r-**ball** of S is defined by $S_r = S + rB$ for each $r \geq 0$, cl S is the **topological closure** of S and $\operatorname{conv} S$ is the **convex hull** of S.

Many of results for H and D are similar, since they are derived from corresponding results of $\sigma_S(y^*)$ and $\iota_S(y^*)$. To shorten our notation we apply here and in the sequel the following conventions. We write ρ instead of H or D if a property is valid for both H and D or if the proof in both cases is similar or if indexes of variables should denote the case H and D. The similarity of properties and of proofs is sometimes a matter of interchanging σ and ι. For this reason, we agree to use γ in statements if properties are valid for σ and ι. In proofs for properties of H and D we use γ instead of σ if H is considered and instead of ι if D is considered. We use further the convenient notations $S - q := S + \{-q\}$, $\gamma_q := \gamma_{\{q\}}$, $\rho(q_1, q_2) := \rho(q_1, \{q_2\})$.

Although H and D have similarities in their definitions and properties, we get sometimes strong differences in their proofs. Here the case $\rho = D$ often demands, because of its "max min" – type definition and of the use of inscribed balls, stronger assumptions and more complex proofs. So we cannot hold everywhere the full generality and cannot always prove similarity of the statements concerning H and D.

In our considerations we mainly study the rich case where S is bounded. It is equivalent to the finiteness of the support functional γ_S on Y^*, we say $\operatorname{dom} \gamma_S = Y^*$. If S is *unbounded in direction* $y^* \in Y^*$ for $\gamma = \sigma[-y^* \in Y^*$ for $\gamma = \iota]$ that means $\sup_{s \in S} \langle y^*, s \rangle$ $[= -\inf_{s \in S} \langle -y^*, s \rangle] = +\infty$ or if $S = \emptyset$ then γ_S is infinite. We shortly mention now some well-known properties of the calculus of $\gamma_S(\cdot)$ with respect to y^* and to the index set S, see e.g. [13, Chapter V]. The proofs given there can be translated to Banach spaces without any problem. Since σ_S and ι_S are Fenchel-conjugate functions we have immediately its lower and upper semi-continuity on Y^*, respectively, whenever $\operatorname{dom} \gamma_S \neq Y^*$. The boundedness of S yields the Lipschitz continuity of γ_S on Y^*. From $\sigma_S(y^*) \leq \|y^*\| \sup_{s \in S} \|s\|$ we get the Lipschitz constant $H(0, S)$. The mapping $S \longmapsto \gamma_S(y^*)$ is additive,

whenever at most one of the summands is infinite, is positively homogeneous and isotone ($\gamma_S = \sigma_S$) or antitone ($\gamma_S = \iota_S$). It is easy to see that $\gamma_S(y^*) = \gamma_{\mathrm{cl}\,S}(y^*) = \gamma_{\mathrm{conv}\,S}(y^*)$. Standard examples are support functions of a half space $S = \mathcal{H}_-(n^*) := \{y \in Y \,|\langle n^*, y\rangle \le 0\}$ with $\sigma_S(y^*) = 0$ for $y^* = \lambda n^*,\ \lambda \ge 0$ and $\sigma_S(y^*) = +\infty$ else, of the unit ball with $\sigma_B(y^*) = \|y^*\|$, $\iota_B(y^*) = -\|y^*\|$ and of a point q with $\sigma_{\{q\}}(y^*) = \iota_{\{q\}}(y^*) = \langle y^*, q\rangle$. We add here the support function of an r-ball of the set S.

Proposition 2.1. $\sigma_{S_r}(y^*) = \sigma_S(y^*) + r\,\|y^*\|$, $\iota_{S_r}(y^*) = \iota_S(y^*) - r\,\|y^*\|$.

Proof. $\sigma_{S_r}(y^*) = \sigma_{S+rB}(y^*) = \sigma_S(y^*) + r\,\sigma_B(y^*) = \sigma_S(y^*) + r\,\|y^*\|$. Further use (5). ■

Now we mention some properties of the functionals H and D. The proofs follow directly and straightforwardly from the associated properties of the support functions γ_S. The functional H introduces the Hausdorff distance and both H and the underlying support functions σ_S are widely studied [5], [10], [13], [17], which is not the case for D and ι_S respectively. We see by definition that both $H(q_1, S)$ and $D(q_1, S)$ reduce to the distance $\|q_1 - q_2\|$ when S is the singleton $\{q_2\}$. Therefore both $H(q, S)$ and $D(q, S)$ generalize the usual distance between points to a point-to-set distance. The distances ρ are translation invariant, $\rho(q+y, S+y) = \rho(q, S)$, isotone ($\rho = H$) / antitone ($\rho = D$) w.r.t. S, positively homogeneous in the sense $\rho(\lambda q, \lambda S) = \lambda\,\rho(q, S)$ and subadditive according to $\rho(q, S_1 + S_2) \le \rho(q_1, S_1) + \rho(q_2, S_2)$ for each partition $q = q_1 + q_2$, whenever one of the summands is finite. Further it is true $\rho(q, S) = \rho(q, \mathrm{cl}\ S) = \rho(q, \mathrm{conv}\,S) = \rho(q, \mathrm{cl}\,\mathrm{conv}\,S)$. For r-balls of S we have

Proposition 2.2. *$H(q, S_r) = H(q, S) + r$ and $D(q, S_r) = D(q, S) - r$ for $r \ge 0$.*

Proof. $\rho(q, S_r) = \sup_{\|y^*\|=1} (\gamma_{S-q}(y^*) + \gamma_{rB}(y^*))$. ■

Now we deal with the *geometric interpretations* of $H(q, S)$ and $D(q, S)$ which are derived from the properties of the support functions. The *smallest circumscribed ball* of $S \ne \emptyset$ with a given center q is the ball $q + r^H(q, S)\ \mathrm{cl}\,B$ where the radius $r^H(q, S)$ is defined by $r^H(q, S) := \inf\{R \in \mathbb{R}_+ \cup \{+\infty\} \mid S \subset q + R\,\mathrm{cl}\,B\}$. We do not exclude the case $r^H(q, S) = +\infty$ where $Y = q + (+\infty)\,\mathrm{cl}\,B$. Further $r^H(S) := \inf_{q\in Y} r^H(q, S)$ is said to be the *Chebyshev radius* of S and each $q \in \{q \in Y \mid S \subset q + r^H\ \mathrm{cl}\ B\}$ is called a *Chebyshev*

center of S [6]. The Chebyshev center of a bounded set in a Hilbert space is uniquely determined and belongs to its convex hull [7]. For the existence of a Chebyshev center for each bounded set in a given normed space we need the reflexivity of the Banach space Y, see [6]. The *largest inscribed ball* of S with a given center $q \in S$ is the ball $q + r^D(q, S)\, B$ where the radius $r^D(q, S)$ is defined by $r^D(q, S) := \sup\{r \in \mathbb{R}_+ \cup \{+\infty\} \mid q + r\,B \subset S\}$. We also use $r^D(S) := \sup_{q \in Y} r^D(q, S)$ the radius of the largest inscribed ball which exists in reflexive Banach spaces, see e.g. [2]. $r^H(q, S) = \sup_{s \in S} \|q - s\|$ and $r^D(q, S) = \inf_{s \in Y \setminus S} \|q - s\|$ can be derived straightforwardly from their definitions. The *oriented distance* $d(q, S)$ from the point q to the boundary ∂S of the set S is defined by

$$d(q, S) = \begin{cases} -r^D(q, S) & \text{if} \quad q \in S, \\ r^D(q, Y \setminus S) & \text{if} \quad q \notin S. \end{cases}$$

and can be expressed by the formula

$$d(q, S) = \inf_{s \in S} \|q - s\| - \inf_{s \in Y \setminus S} \|q - s\|. \tag{6}$$

Note that for $q \in S$ the first term turns into zero and for $q \notin S$ this happens to the second term. The following lemma is a simple consequence of the calculus of support functions and the definition of the distance D.

Lemma 2.3. *Let $y_0^* \in \partial B^*$ q, $q_0 \in Y$ and consider the affine (closed !) half space $S = \mathcal{H}_-(y_0^*) + q_0$. Then $D(q, S) = -\sigma_{S-q}(y_0^*) = \langle y_0^*, q - q_0 \rangle$.*

Proposition 2.4. i) $H(q, S) = \sup_{s \in S} \|q - s\| = r^H(q, S)$.
ii) $D(q, S) = d(q, \operatorname{cl}\operatorname{conv} S)$ $[= d(q, \operatorname{conv} S)$ *for* $\dim Y < \infty]$.

Proof. i) is obvious by commutation of suprema. ii) Let $S = \operatorname{cl}\operatorname{conv} S$ (in finite dimensions we do not need the closure of S for using the separation argument by working with the relative interior) and consider three cases, $q \in \operatorname{int} S$, $q \in \partial S$ and $q \notin S$.

Case 1. Let $q \in \operatorname{int} S \neq \emptyset$, i.e. $d(q, S) < 0$. Consider arbitrary $y^* \in \partial B^*$ and $r > 0$ with $r < -d(q, S)$. Because of $q + rB \subset S$ and the monotony of the support function we get $r = \sup_{s \in B} \langle y^*, rB \rangle = \sigma_{rB}(y^*) = \sigma_{(q+rB)-q}(y^*) \leq \sigma_{S-q}(y^*)$. Taking the infimum over $y^* \in \partial B^*$ we find out that

$$r \leq \inf_{y^* \in \partial B^*} \sigma_{S-q}(y_0^*) = -D(q,S). \tag{7}$$

The arbitrariness of r yields $-d(q,S) \leq -D(q,S)$. Assume now the strict inequality

$$d(q,S) - D(q,S) > \varepsilon > 0. \tag{8}$$

For the ball $\hat{B} := q + (-d(q,S) + \varepsilon/2)\, B$ and arbitrary $y^* \in \partial B^*$ we obtain with (7) and (8) that $-d(q,S) + \varepsilon/2 = \sigma_{\hat{B}-q}(y^*) < -D(q,S) - \varepsilon/2 \leq \sigma_{S-q}(y^*) - \varepsilon/2 < \sigma_{S-q}(y^*)$ which implies the contradiction $\hat{B} \subset S$ since $-d\,(q,S)$ is the largest radius of an inscribed ball of S around q.

Case 2. $q \in \partial S$ ($\operatorname{int} S$ may be empty). Since S is closed and convex there is by separation theorem a supporting closed hyperplane $\mathcal{H}$ where q belongs to its support. Hence $D(q,S) = D\,(q,\mathcal{H}) = 0$. But $d(q,S) = 0$ by definition.

Case 3. $q \notin S$ (i.e. $d\,(q,S) > 0$). By well-known separation arguments there is some $y_0^* \in \partial B^*$ such that $\mathcal{H}_S\,(y_0^*) := \{y \in Y \mid \langle y_0^*, y\rangle \leq \sigma_S(y_0^*)\}$ is a supporting half space of S, $S \subset \mathcal{H}_S(y_0^*)$ and $(q + d(q,S)) \cap \mathcal{H}_S\,(y_0^*) = \emptyset$. The first equality follows from Singer [21]. Observe for the second one that the desired result is already proven for a closed half space S by using $\operatorname{cl}(Y \setminus S)$ instead of S in Case 1 and 2) $d(q,S) = d\,(q, \mathcal{H}_S\,(y_0^*)) = D\,(q, \mathcal{H}_S\,(y_0^*)) = -\sigma_{S-q}\,(y_0^*) = -\sigma_{\mathcal{H}_S(y_0^*)-q}\,(y_0^*) \leq D\,(q,S) = -\inf_{y^* \in \partial B^*} \sigma_{S-q}(y^*)$. If we assume strict inequality, then there is some $y_1^* \in \partial B^*$, $y_1^* \neq y_0^*$ such that $-\sigma_{S-q}\,(y_0^*) < -\sigma_{S-q}\,(y_1^*)$, $\mathcal{H}_S(y_1^*)$ separates q and S and is a supporting half space of S. This yields the contradiction $d\,(q,S) \geq d\,(q, \mathcal{H}_S\,(y_1^*)) = D\,(q, \mathcal{H}_S\,(y_1^*)) = -\sigma_{\mathcal{H}_S(y_1^*)-q}\,(y_1^*) = -\sigma_{S-q}\,(y_1^*) > -\sigma_{S-q}\,(y_0^*) = d\,(q,S)$. Hence, we conclude $d\,(q,S) = D\,(q,S)$. ■

We learned from the last proposition that $H(q,S)$ gives the radius of the smallest circumscribed for S ball with center q and that $D(q,S)$ gives the oriented distance from q to the boundary of $\operatorname{cl}\operatorname{conv} S$. To derive the Kolmogorov optimality conditions we will come back to this geometric property of the distances. The equalities $\rho(q,S) = \rho(q, \operatorname{cl} S) = \rho(q, \operatorname{conv} S) = \rho\,(q, \operatorname{cl}\operatorname{conv} S)$ and the monotony show that these two distances coincide for all sets $\hat{S}$ with $S \subset \hat{S} \subset \operatorname{cl}\operatorname{conv} S$. For closed convex sets S the distance D

is reduced to $D(q,S) = d(q,S)$, while already the following simple examples show that in general this equality does not hold.

Example 2.1. Let $Y = \mathbb{R}$ and $S = \mathbb{Q}$ be the set of rational numbers. Then $-\infty \equiv D(q,S) = d\,(q, \operatorname{conv} S) < d(q,S) \equiv 0$.

Example 2.2. Let Y be an infinite dimensional Banach space and u^* a discontinuous linear functional on Y. Then its kernel $N\,(u^*)$ is a linear non closed subspace of Y which is dense in Y. Put $S = (\operatorname{cl} B) \cap N\,(u^*)$. Obviously, S is convex. We get in this case $d(q,S) = d\,(q, \operatorname{conv} S) = 0 > D\,(q,S) = d\,(q, \operatorname{cl} S) = d\,(q, \operatorname{cl}\operatorname{conv} S)$ for each $q \in B$.

Since, in contrast to $D(q,S)$, $d(q,S)$ is the oriented distance directly to the boundary of S and not to that of $\operatorname{cl}\operatorname{conv} S$, one may wish to implement d instead of D in a problem like the one considered in this paper. This was the authors' initial intention. However as far as nonconvex sets are under consideration, computational difficulties arise immediately and most of the useful properties of D are not true or are only partially true. For instance consider the property $D(q,S_r) = D(q,S) - r$, $r \geq 0$. The following example shows that if $d(q,S) < r$ such an equality might not be true. The validity can be proven provided that $d(q,S) \geq r > 0$.

Example 2.3. Let Y be an Euclidean space and R be a positive number. Put $Q = R\,B$, $a \in Y \setminus \{0\}$ and let $0 < \mu < R$. Denote by O the center of Q. Let $I = Q \cap \{y \in Y \mid y = \lambda a, \lambda > 0\}$. We consider the set $S = Q \setminus (I + \mu B)$. Then for $r > \mu$ it is easy to calculate that $d(O,S) = \mu$ and $d(O,S_r) = \sqrt{R^2 - \mu^2} + \sqrt{r^2 - \mu^2}$. Obviously $d(O,S_r) \neq d(O,S) - r$.

Proposition 2.5. *If $d(q,S) \geq r > 0$ then $d(q,S_r) = d(q,S) - r$.*

Proof. We prove first that $d(q,S) \geq d(q,S_r) + r$. Note that in this case $r \leq \|q - s\|$. Then $d(q,S_r) = \inf_{s \in S_r} \|q - s\| \leq \inf_{s \in S} \|q - (s + r(q-s)\,/\,\|q-s\|)\| = \inf_{s \in S} \|(q-s)(1 - r\,/\,\|q-s\|)\| = \inf_{s \in S} \|q-s\| - r = d(q,S) - r$, which proves $d(q,S) \geq d(q,S_r) + r$. The opposite inequality follows from $d(q,S_r) = \inf_{s \in S,\ u \in B} \|q - (s + ru)\| \geq \inf_{s \in S,\ u \in B} (\|q - s\| - r\,\|u\|) = d(q,S) - r$. ∎

Now, we collect some results of ρ w.r.t. the variable q for a fixed set S. The *domain* $\operatorname{dom} \rho(\cdot, S)$ be the set of all $q \in Y$ such that $\rho(q,S)$ is finite. The set S is said to be *proper* if it is neither $S = \emptyset$ nor $\operatorname{cl}\operatorname{conv} S = Y$. Some of the

results w.r.t. q concerning H can be found e.g. in [15, Propositions X.14, p. 562 and X.15, p. 565] and [4, Proposition 2.4.1], nevertheless we formulate also these statements for D, since the accepted notations make it possible to state simultaneously their validity both for H and D. The proofs follow straightforwardly from the geometric interpretation and from the properties of γ_S and are omitted here. The interesting case $\operatorname{dom}\rho(\cdot,S)\neq\emptyset$ can easily be described since either $\operatorname{dom}\rho(\cdot,S)=Y$ or $\operatorname{dom}\rho(\cdot,S)=\emptyset$. The first case is given if and only if S is nonempty and bounded for $\rho=H$ and if and only if S is proper for $\rho=D$. If $\operatorname{dom}\rho(\cdot,S)=Y$ we get that the distances are nonexpansive in q on Y, i.e. $|\rho(q_1,S)-\rho(q_2,S)|\leq\|q_1-q_2\|$. The convexity of $\rho(\cdot,S)$ follows from the homogeneity and subadditivity of ρ w.r.t. S. Indeed, $\rho(\lambda_1q_1+\lambda_2q_2,S)=\rho(\lambda_1q_1+\lambda_2q_2,\operatorname{conv}S)=\rho(\lambda_1q_1+\lambda_2q_2,\lambda_1\operatorname{conv}S+\lambda_2\operatorname{conv}S)\leq\rho(\lambda_1q_1,\lambda_1\operatorname{conv}S)+\rho(\lambda_2q_2,\lambda_2\operatorname{conv}S)=\lambda_1\rho(q_1,\operatorname{conv}S)+\lambda_2\rho(q_2,\operatorname{conv}S)=\lambda_1\rho(q_1,S)+\lambda_2\rho(q_2,S)$. Obviously, we have $D(q,S)\leq H(q,S)$ if S is nonempty. The coercivity $\rho(q,S)\geq\|q\|-\rho(0,S)$ follows immediately from $\rho(q,S)\geq\sup_{\|y^*\|=1}\langle y^*,q\rangle+\inf_{\|y^*\|=1}\gamma_S(y^*)$. Additionally, we can prove some representation formulas and stability results.

Proposition 2.6. i) *If S is bounded, then the level set $\Phi^\rho(r)=\{q\in Y\,|\rho(q,S)\leq r\}$ is convex, closed, bounded and increasing in r. Further the representation holds*

$$\Phi^\rho(r)=\begin{cases}\bigcap_{s\in S}(s+r\operatorname{cl}B) & \text{if}\quad \rho=H,\\ Y\setminus\bigcup_{s\in Y\setminus\operatorname{cl}S}(s-rB) & \text{if}\quad \rho=D, r<0, S \text{ closed and convex.}\end{cases}$$

ii) $\rho(q,S)=\sup_{b\in B}\rho(q+\varepsilon b,S)-\varepsilon$ *and*

$$\inf_{b\in B}\rho(q+\varepsilon b,S)\leq\rho(q,S)\leq\inf_{b\in B}\rho(q+\varepsilon b,S)+\varepsilon. \tag{9}$$

for each $\varepsilon>0$.

Proof. i) If the level set is non-empty, then the monotony is obvious and the closedness and convexity follow from the continuity and convexity of $\rho(\cdot,S)$. If $S\subset RB$ for some $R>0$ then the level set is contained in any case in $(r+R)\operatorname{cl}B$.

Let $H(q,S)\leq r$. Then $\|q-s\|\leq r$ for all $s\in S$ and hence $q\in s+r\operatorname{cl}B$ for all $s\in S$. But the last means that vice versa $\|q-s\|\leq r$ for all $s\in S$ which implies again $H(q,S)\leq r$.

Let $D(q,S)\leq -r$, and S be closed, convex with $\operatorname{int}S\neq\emptyset$ and $r\geq 0$. Then $\|q-s\|\geq r$ for all $s\in Y\setminus\operatorname{cl}S$ and hence $q\notin s+rB$ for all $s\in Y\setminus\operatorname{cl}S$.

This means $q \in \bigcap_{s\in Y\setminus \operatorname{cl} S} Y\setminus (s+rB) = Y\setminus \bigcup_{s\in Y\setminus \operatorname{cl} S} (s+rB)$. Again we can invert the conclusion.

ii) $\sup_{b\in B} \rho(q+\varepsilon\, b, S) = \sup_{b\in B} \rho(q, S-\varepsilon b) = \sup_{b\in B} \rho(q, S+\varepsilon b) = \sup_{b\in B} \sup_{y^*\in\partial B^*} \gamma_{S-q+\varepsilon b}(y^*) = \sup_{b\in B} \sup_{y^*\in\partial B^*} (\gamma_{S-q}(y^*) + \varepsilon \langle y^*, b\rangle) = \sup_{y^*\in\partial B^*} [\gamma_{S-q}(y^*) + \varepsilon \sup_{b\in B} \langle y^*, b\rangle] = \sup_{y^*\in\partial B^*} [\gamma_{S-q}(y^*) + \varepsilon \|y^*\|] = \rho(q,S) + \varepsilon$, $\inf_{b\in B} \rho(q+\varepsilon\, b, S) = \inf_{b\in B} \sup_{y^*\in\partial B^*} (\gamma_{S-q}(y^*) + \varepsilon \langle y^*, b\rangle) \le \sup_{y^*\in\partial B^*} \gamma_{S-q}(y^*) + \varepsilon \inf_{b\in B} \sup_{y^*\in\partial B^*} \langle y^*, b\rangle = \rho(q,S) + \varepsilon \inf_{b\in B} \|b\| = \rho(q,S)$.

Otherwise, interchanging inf-sup we get $\inf_{b\in B} \rho(q+\varepsilon b, S) = \inf_{b\in B} \sup_{y^*\in\partial B^*}(\gamma_{S-q}(y^*) + \varepsilon \langle y^*, b\rangle) \ge \sup_{y^*\in\partial B^*} \inf_{b\in B}(\gamma_{S-q}(y^*) + \varepsilon \langle y^*, b\rangle) = \sup_{y^*\in\partial B^*} (\gamma_{S-q}(y^*) - \varepsilon \|y^*\|) = \rho(q,S) - \varepsilon$. ■

Remark 2.1. Minimax-theorems do not work in ii) since ∂B^* is not convex. Indeed, $0 = \inf_{b\in B} \sup_{y^*\in\partial B^*} \langle y^*, b\rangle \gneqq \sup_{y^*\in\partial B^*} \inf_{b\in B} \langle y^*, b\rangle = -1$.

Since $\varphi^\rho(g) := \sup_{t\in M} \rho(g(t), \Sigma(t))$, $\rho = H, D$, the most properties of ρ can be translated into similar properties of φ^ρ. We need the following assumptions with respect to the effective domain **(D)** and the image **(I)** of Σ for the basic properties of the objective φ^ρ.

(D):	$\operatorname{dom}\Sigma = M$.
(I), case $\rho = H$:	$\Sigma(t)$ is uniformly bounded on M and $\operatorname{dom}\Sigma \neq \emptyset$.
(I), case $\rho = D$:	$\operatorname{dom}\Sigma = M$, $\forall_{t\in M}$ $\operatorname{cl}\operatorname{conv}\Sigma(t) \neq Y$ and $\exists_{r>0}\forall_{t\in M}$ $\operatorname{cl}\operatorname{conv}\Sigma(t) \cap r\,B \neq \emptyset$.

Proposition 2.7 (Properties of the objective)**.** *Let $\Sigma : M \to 2^Y$ be a given set-valued function. Then the following statements are true.*

i) *If* (D) *and* (I) *are valid then for each $g \in C(M,Y)$ all the values $\rho(g(t), \Sigma(t))$, $t \in M$, are finite.* (I) *implies the finiteness of φ^ρ on $C(M,Y)$. If* (I) *is violated, then φ^ρ is either $+\infty$ or $-\infty$ on $C(M,Y)$.*

ii) *If* (D) *and* (I), *then $|\varphi^\rho(g_1) - \varphi^\rho(g_1)| \le \|g_1 - g_2\|_{C(M,Y)}$.*

iii) *φ^ρ is convex on $C(M,Y)$.*

iv) *$\varphi^\rho(g) \ge \|g\|_{C(M,Y)} - \sup_{t\in M} \rho(0, \Sigma(t))$ and in particular φ^ρ is coercive if φ^ρ is finite, i.e. if* (I) *holds.*

v) (D) *implies $\varphi^D(g) \le \varphi^H(g)$.*

Proof. i) straightforward. ii) The expansivity of ρ yields the result. iii) The convexity of ρ and the subadditivity of suprema yields the statement.

iv) The coercivity of ρ implies the first result. If φ^ρ is finite, then we get $-\infty < \sup_{t\in M} \rho(0, \Sigma(t)) \leq r < \infty$ and consequently φ^ρ is coercive. v) Obvious. ■

In order to investigate the conditions implying finiteness we have considered in Proposition 2.7 i) an arbitrary set-valued function Σ. The finiteness of the value $\varphi^\rho(g)$ and *all the values* $\rho(g(t), \Sigma(t))$, $t \in M$, are necessary for useful results. Therefore we assume at least (D) and (I) in our further investigations. In the sequel, we define and investigate the semi-continuity properties of certain set-valued functions Φ^ρ named level set maps. In the next section, we use the level set maps in connection with the concept of regularity introduced there. As a prerequisite to the level set maps we define the functions $f^\rho : Y \times M \to \mathbb{R}$ by

$$f^\rho(q, t) = \rho(q, \Sigma(t)),$$

$q \in Y$, $t \in M, \rho = H, D$. The values of f are in $\mathbb{R}$ if our general assumptions (D) and (I) are satisfied. The next proposition establishes semi-continuity properties of f^ρ. We use there and further on the abbreviation u.s.c. for upper semi-continuous and l.s.c. for lower semi-continuous functions. Further, we use the standard definitions of upper semi-continuity (u.s.c.), Hausdorff upper semi-continuity (H.u.s.c.), lower semi-continuity (l.s.c.) and Hausdorff lower semi-continuity (H.l.s.c.) for set-valued mappings which can be found e.g. in [3, p. 26] and in [1, p. 40–46] (here the H.l.s.c. / H.u.s.c. is called ε-lower/upper semi-continuity). The standard definitions imply that Σ is l.s.c. and H.l.s.c. at each t_0 with $\Sigma(t_0) = \emptyset$ and that $\operatorname{dom}\Sigma$ of a l.s.c. and a H.l.s.c. map is open. Further, Σ is u.s.c. and H.u.s.c. at t_0 with $\Sigma(t_0) = \emptyset$ if $\Sigma(t) = \emptyset$ in some neighborhood of t_0. Hence, if Σ is u.s.c. or H.u.s.c. on M, then $\operatorname{dom}\Sigma$ is closed. The following consequences are well-known for set-valued maps:

i) *u.s.c.* $\Longrightarrow$ *H.u.s.c.*,

ii) *H.u.s.c.* and Σ compact valued $\Longrightarrow$ *u.s.c.*,

iii) *H.l.s.c.* $\Longrightarrow$ *l.s.c.*,

iv) *l.s.c.* and Σ compact valued $\Longrightarrow$ *H.l.s.c.*

Statement iv) yields that the most general concept seems to be the use of H.u.s.c. and l.s.c. set-valued maps. However, in the case of lower semi-continuity we need in our concept either compact values of the map Σ and lower semi-continuity or directly Hausdorff lower semi-continuity. Hence by

iv) the Hausdorff concept of semi-continuity is more suitable for us. The following proposition is a standard result of set valued analysis if we assume that Σ is u.s.c or l.s.c. and can be traced back to statements on marginal functionals (cf. [1, Chapter 1.2. Theorem 4 / Theorem 5], [3, Chapter 4])

Proposition 2.8. *Let the set-valued function* $\Sigma : M \to 2^Y$ *satisfy conditions* (D) *and* (I).

i) *If* Σ *is H.u.s.c. then* f^H *is u.s.c. and* f^D *is l.s.c. on* $Y \times M$.

ii) *If* Σ *is H.l.s.c. then* f^H *is l.s.c. and* f^D *is u.s.c. on* $Y \times M$.

Proof. By using Proposition 2.1 the proofs are straightforward. We give only one example where the compactness can be avoided by the use of Hausdorff semi-continuity. The remaining three proofs are very similar. Case i), f^D: Fix $(q_0, t_0) \in Y \times M$. Take $\varepsilon > 0$. Let U be a neighborhood of t_0 such that $\Sigma(t) \subset \Sigma(t_0) + \frac{1}{2}\varepsilon\, B$ for $t \in U$. Then for $(q,t) \in (q_0 + \frac{1}{2}\varepsilon\, B) \times U$ we have $f^D(q,t) = \sup_{\|y^*\|=1} \inf_{s\in\Sigma(t)} \langle y^*, s-q\rangle \geq \sup_{\|y^*\|=1} \inf_{s\in\Sigma(t_0)+\frac{1}{2}\varepsilon\, B} \langle y^*, s-q\rangle$ $\geq \sup_{\|y^*\|=1} \inf_{s_0\in\Sigma(t_0),\, b\in B} \big(\langle y^*, s_0 - q_0\rangle + \frac{1}{2}\,\varepsilon\, \langle y^*, b\,\rangle + \langle y^*, q_0 - q\rangle\big) \geq \sup_{\|y^*\|=1} \inf_{s_0\in\Sigma(t_0)} \langle y^*, s_0 - q_0\rangle - \frac{1}{2}\varepsilon - \frac{1}{2}\varepsilon = f^D(q_0,t_0) - \varepsilon$. ■

As a simple consequence we get the following semi-continuity properties of the composed function $\psi^\rho : M \longrightarrow \mathbb{R}$ defined by $\psi^\rho(t) = \rho(g(t), \Sigma(t))$.

Corollary 2.9. *Let the set-valued function* $\Sigma : M \to 2^Y$ *satisfy conditions* (D), (I). *Then for each* $g \in C(M,Y)$ *the following assertions hold.*

i) *If* Σ *is H.u.s.c. then* ψ^H *is u.s.c. and* ψ^D *is l.s.c. on* M.

ii) *If* Σ *is H.l.s.c. then* ψ^H *is l.s.c. and* ψ^D *is u.s.c. on* M.

We define now the *level set maps* $\Phi^\rho, \Phi_0^\rho : M \times \mathbb{R} \to 2^Y$ by $\Phi^\rho(t,r) = \{q \in Y \mid \rho(q, \Sigma(t)) \leq r\}$ and $\Phi_0^\rho(t,r) = \{q \in Y \mid \rho(q, \Sigma(t)) < r\}$. We consider some *characteristic set* $I^\rho(r) = \{t \in M \mid \Phi_0^\rho(t,r) = \emptyset\}$, which is the set of such $t \in M$ where the *Slater condition* of the corresponding system $\rho(q, \Sigma(t)) \leq r$ is not satisfied.

Definition 2.1. The set-valued map $\Phi^\rho(\cdot, r)$ is said to be **stable** at $t_0 \in I^\rho(r)$ if for each $\varepsilon > 0$ there is a $\delta(\varepsilon, t_0) > 0$ and some neighborhood $U(\varepsilon)$ of t_0 such that $\Phi^\rho(t,r) + \varepsilon B \supset \Phi^\rho(t, r + \delta(\varepsilon, t_0))$ for all $t \in U(\varepsilon)$.

From now we replace assumption (I) by the sharper **condition (B)** that the range Σ *is bounded*, already known from (I) in the case $\rho = H$. In the next

proposition we formulate semi-continuity properties of the level set maps w.r.t. $t \in M$.

Proposition 2.10. *Assume that* $\Sigma : M \to 2^Y$ *satisfies* (D), (B), *that* $r \in \mathbb{R}$ *is fixed and* $\operatorname{dom} \Phi^\rho(\cdot, r) = M$. *Then the following statements are true.*

i) $\Phi^\rho(t,r)$ *is closed, convex for all* $t \in M$. *Further* $\Phi^\rho(t,r) \subset \operatorname{cl}\operatorname{conv}\Sigma(t) + r_0 \operatorname{cl} B$, $r_0 = \max(0,r)$, *whence* $\Phi^\rho(\cdot,r)$ *inherits from* Σ *each of the properties: bounded-valued, bounded range,* $\forall_{t\in M} \operatorname{cl}\operatorname{conv}\Sigma(t) \neq Y$, $\exists_{r>0}\forall_{t\in M} \operatorname{cl}\operatorname{conv}\Sigma(t) \cap r\,B \neq \emptyset$.

ii) *Let* $t \in M \setminus I^\rho(r)$. $\Phi^H(\cdot,r)$ *is l.s.c. at* t *if* Σ *is H.u.s.c. at* t, *similarly* $\Phi^D(\cdot,r)$ *is l.s.c. at* t *if* Σ *is H.l.s.c. at* t.

iii) $\Phi^\rho(\cdot,r)$ *is l.s.c. at* $t_0 \in I^\rho(r)$ *if additionally to the semi-continuity assumptions in ii) one of the following conditions is satisfied.*

   a) $Y = \mathbb{R}^n$, Σ *is H-continuous on some neighborhood* U *of* t_0 *and* $\Phi^\rho(t_0,r)$ *is single valued.*

   b) $\Phi^\rho(\cdot,r)$ *is stable at* t_0.

iv) *Let* $Y = \mathbb{R}^n$. *Then* $\Phi^H(\cdot,r)$ *is (H.)u.s.c. if* Σ *is H.l.s.c., similarly* $\Phi^D(\cdot,r)$ *is (H.)u.s.c. if* Σ *is H.u.s.c.*

Proof. The proof of this proposition needs decisively the following results of [3] which can be slightly extended to a compact Hausdorff space T for the set of parameters. Let us shortly formulate in our notations some simple corollaries of the statements given there working under our conditions. We omit the index ρ and the fixed variable r for Φ, Φ_0 and f, where $f(q,t) = \rho(q, \Sigma(t))$.

Corollary 2.11 [3, Theorem 3.1.5]**.** *Let* f *be u.s.c. on* $\Phi(t_0) \times \{t_0\}$ *and* $\Phi(t_0) \subset \operatorname{cl}\Phi_0(t_0)$. *Then* Φ *is l.s.c. at* t_0.

Corollary 2.12 [3, Theorem 3.1.6]**.** *Let* f *be u.s.c. on* $\Phi(t_0) \times \{t_0\}$, *let* $\Phi_0(t_0) \neq \emptyset$ *and let* $f(\cdot,t_0)$ *be convex on* Y. *Then* $\Phi(t_0) \subset \operatorname{cl}\Phi_0(t_0)$.

Corollary 2.13 [3, Theorem 3.2.2]**.** *Assume that* $Y = \mathbb{R}^n$, U *is some neighborhood of* t_0 *and let* f *be continuous on* $Y \times U$. *If* Φ *is bounded valued on* U *and single valued at* t_0, *then* Φ *is l.s.c. at* t_0.

Corollary 2.14 [3, Theorem 3.1.2]**.** *If* f *is l.s.c. on* $\Phi(t_0) \times \{t_0\}$ *and there is some compact set* K *such that* $\Phi(t) \subset K$ *for all* t *in some neighborhood* U *of* t_0, *then* Φ *is H.u.s.c. at* t_0.

Proof of the Proposition 2.10. We omit again the index ρ for Φ, Φ_0 and $I(r)$.

i) The value $\Phi(t,r) = \{q \in Y \mid \rho(q, \Sigma(t)) \leq r\}$ is closed and convex since $\rho(\cdot, \Sigma(t))$ is globally Lipschitzean and convex. Assume $q \notin \operatorname{cl}\operatorname{conv} \Sigma(t) + r_0 \operatorname{cl} B$. Then $0 < D(q, (\operatorname{cl}\operatorname{conv} \Sigma(t))_{r_0}) = D(q, \Sigma(t)) - r_0$ and $H(q, (\operatorname{cl}\operatorname{conv} \Sigma(t))_{r_0}) = H(q, \Sigma(t)) + r_0 > r_0$. In both cases $\rho(q, \Sigma(t)) > r_0 = \max(0, r) \geq r$ which implies $q \notin \Phi(t,r)$, that is $\Phi^r(t,r) \subset \operatorname{cl}\operatorname{conv} \Sigma(t) + r_0 \operatorname{cl} B$, $r_0 = \max(0,r)$.

ii) Proposition 2.8 and the convexity of $\rho(\cdot, S)$ imply that $f(q,t) := \rho(q, \Sigma(t))$ is convex in q on Y and u.s.c. in (q,t) on $Y \times M$. Corollary 2.12 and Corollary 2.11 can be applied on $M \setminus I(r)$ to get the statement.

iii) a) Let $t_0 \in I(r)$ be given and let Σ be H-continuous on some neighborhood U of t_0. The continuity of f on U follows from Proposition 2.8. $\Phi(t,r)$ is bounded valued for all $t \in U$ and single valued at t_0. Applying Corollary 2.13 we get that $\Phi(\cdot, r)$ is l.s.c. at t_0.

b) Let $q \in \Phi(t_0, r)$ and take the basis of neighborhoods $\{\lambda B + q\}_{\lambda > 0}$ of q. We have to prove that for each $\lambda > 0$ there is a neighborhood $U(\lambda)$ of t_0 such that $\Phi(t,r) \cap (\lambda B + q) \neq \emptyset$ for all $t \in U(\lambda)$.
First we have the monotony for all $\delta > 0$

$$\Phi(t_0, r) \subset \Phi(t_0, r + \delta). \tag{10}$$

$\Phi(\cdot, r + \delta)$ is l.s.c. at t_0 i.e. for arbitrary $q \in \Phi(t_0, r + \delta)$ and all $\lambda > 0$ there is some $U(q, \lambda, \delta)$ of t_0 such that for all $t \in U(q, \lambda, \delta)$

$$\Phi(t, r + \delta) \cap (\lambda B + q) \neq \emptyset. \tag{11}$$

From the stability we have for each $\tau > 0$ some neighborhood $U_0(\tau)$ of t_0 and some $\delta = \delta(\tau) > 0$ such that

$$\Phi(t,r) + \tau B \supset \Phi(t, r + \delta(\tau)) \tag{12}$$

for all $t \in U(\tau)$.

From (10) we have $\Phi(t_0, r + \delta) \cap (\lambda B + q) \neq \emptyset$ for all $\lambda > 0$. Using (11) we get for all $t \in U(q, \lambda, \delta)$ and all $\delta, \lambda > 0$ that $\Phi(t, r + \delta) \cap (\lambda B + q) \neq \emptyset$. The upper semi-continuity (12) yields then for $\delta = \delta(\lambda/2)$ and all $\lambda > 0$ and all $t \in U(\lambda) := U(q, \lambda/2, \delta(\lambda/2)) \cap U_0(\lambda/2)$ the desired property $\Phi(t,r) \cap (\lambda B + q) \neq \emptyset$.

iv) range $\Phi(\cdot, r)$ is bounded by i) for each fixed r and $K = \operatorname{cl}$ range Φ is a compact set containing the values of $\Phi(\cdot, r)$. The lower semi-continuity of the function f with $f(q,t) = \rho(q, \Sigma(t))$ is given from the assumptions

for Σ by Proposition 2.8. The (Hausdorff) upper semi-continuity of $\Phi(\cdot, r)$ follows now from Corollary 2.14. ■

Remark 2.2. Upper semi-continuity of level set maps w.r.t. r is closely related to our stability assumption. We demand some what stronger additional uniformity. It is well-known by a theorem of Dolecki see e.g. [3, Theorem 3.3.1] that $\Phi^{\rho}(t, \cdot)$ is u.s.c. at each r where $\Phi_0^{\rho}(t, r) \neq \emptyset$, i.e. where the Slater condition is satisfied for the constraint $\rho(q, \Sigma(t)) \leq r$.

It is open whether a) implies b). Obviously there are elementary examples that b) can be satisfied for Σ where $\Phi^{\rho}(t_0, r)$ is not single valued which may be the case if the normed space is not rotund ($\rho = H$) or if the boundary of $\Sigma(t) \subset \mathbb{R}^2$ contains some line segment ($\rho = D$). We now shortly discuss the stability assumption in iii) b) for the cases $\rho = H$ and $\rho = D$ in a few examples and statements to illustrate our opinion that this condition may be satisfied in lots of problems. Without loss of generality we assume that the values of Σ are bounded, closed and convex. We investigate the value $S = \Sigma(t_0)$ where $t_0 \in I^{\rho}(r) \neq \emptyset$ and $\operatorname{dom} \Phi^{\rho} = M$.

Example 2.4. i) We consider the case $\rho = H$ using the Euclidean norm in $Y = \mathbb{R}^2$, suppress the upper index H and assume that $\Phi(t_0, r) = \{q\}$. By the representation formula Proposition 2.6 ii), we get $\Phi(t_0, r + \delta) = \bigcap_{m \in \partial S \cap (q + r \operatorname{cl} B)} (m + (r + \delta) \operatorname{cl} B)$. The worst case is given whenever the uniquely existing smallest circumscribed ball with midpoint q in S has only two common points with the boundary of S. Then $\Phi(t_0, r + \delta)$ is a convex set of which boundary is a biarc with smallest width $W = 2\delta$ and diameter $\Delta = 2\varepsilon$. Obviously in this case $r^2 + \varepsilon^2 = (r + \delta)^2$ and $\delta \leq \varepsilon$ since $\sin\alpha + \cos\alpha \geq 1$ for all $\alpha \in [0, \pi/2]$. Hence $\Phi(t_0, r + \delta) \subset q + \varepsilon B$ and usually $(r > 0, \varepsilon \to 0)$ from $r \gg \varepsilon$ we get by expansion $\delta(\varepsilon) = -r + \sqrt{r^2 + \varepsilon^2} = \frac{\varepsilon^2}{2r} + O\left(\frac{\varepsilon^4}{r^3}\right)$.

ii) This result remains valid if Y is an infinite dimensional Hilbert space.

iii) If we consider in $Y = \mathbb{R}^2$ the usual l_p norms $1 \leq p \leq \infty$, then we obtain by similar calculations $\delta(\varepsilon) = -r + \sqrt[p]{r^p + \varepsilon^p} = \frac{\varepsilon^p}{pr^{p-1}} + O\left(\frac{\varepsilon^{2p}}{r^{2p-1}}\right)$ for $1 \leq p < \infty$ and $\delta(\varepsilon) = \varepsilon$ for $p = 1, \infty$. The case $p = \infty$ is not covered for $p \to \infty$ since $\delta(\varepsilon)$ tends to zero if p tends to infinity. The considerations only depend on the type of the norm and do not depend on the structure of S. If we take any point t of some neighborhood U of t_0, then the same considerations can be done with a smallest circumscribed ball of $\Sigma(t)$ with

Chebyshev center $q(t)$ and its radius $r(t)$. We obtain $\Phi(t, r(t) + \delta(\varepsilon, t)) \subset q(t) + \varepsilon B$ with $\delta(\varepsilon, t)$ as above where in $\delta(\varepsilon)$ the parameter r is replaced by $r(t)$. For our examples we get immediately $\delta(\varepsilon, t) \geq \delta(\varepsilon) = \delta(\varepsilon, t_0)$ since $r(t) \leq r = r(t_0)$ for all $t \in M$. The next statement shows that the continuity of the radii $r(t)$ at t_0 implies the above formulated stability condition b). The lower semi-continuity of the radii follows from the Hausdorff lower semi-continuity of Σ (see below Proposition 3.3 i)).

Proposition 2.15. *Let $\varepsilon > 0$ be arbitrary, assume $\delta(\varepsilon, t) \geq \delta(\varepsilon, t_0)$ for all t of some neighborhood $U_0(\varepsilon)$ of $t_0 \in I^H(r)$. If $\Phi^H(t, r(t) + \delta(\varepsilon, t)) \subset \Phi^H(t, r(t)) + \varepsilon B$ for all $t \in U_0(\varepsilon)$ and if $r(\cdot)$ is l.s.c. at t_0, then $\Phi^H(\cdot, r)$ is stable at t_0.*

Proof. We have for each $\varepsilon, \alpha > 0$ and some neighborhood $U(\alpha)$ of t_0 for all $t \in U(\alpha)$ the inequality $r(t_0) - \alpha \leq r(t)$. Using the monotony property of the level sets we obtain for each $t \in U(\alpha) \cap U_0(\varepsilon)$ the chain $\Phi(t, r(t_0)) + \varepsilon B \supset \Phi(t, r(t)) + \varepsilon B \supset \Phi(t, r(t) + \delta(\varepsilon, t)) \supset \Phi(t, r(t) + \delta(\varepsilon, t_0)) \supset \Phi(t, r(t_0) - \alpha + \delta(\varepsilon, t_0))$. If we choose $\alpha = \delta(\varepsilon, t_0)/2$, then we get with $\delta(\varepsilon) = \delta(\varepsilon, t_0)/2$ and $U(\varepsilon) = U(\delta(\varepsilon, t_0)/2) \cap U_0(\varepsilon)$ the relation $\Phi(t, r(t_0)) + \varepsilon B \supset \Phi(t, r(t_0) + \delta(\varepsilon))$ for all $t \in U(\varepsilon)$. ■

Example 2.5. Now we consider the case $\rho = D$ only for $Y = \mathbb{R}^2$ and the Euclidean norm which is already much more delicate. Interesting is the case whenever $\operatorname{int} S \neq \emptyset$, i.e. $r = -r^D < 0$. We assume that the largest inscribed ball B_0 is uniquely defined and q is its midpoint. Then $\Phi(t_0, r + \delta) = S \setminus \bigcup_{m \in \partial S} \left(m + \left(r^D - \delta\right) \operatorname{cl} B\right)$, or equivalently, $\Phi(t_0, r + \delta)$ is the set of all midpoints p such that the ball $p + \left(r^D - \delta\right) \operatorname{cl} B$ is a subset of S. The curvature of the extreme ball is equal to or larger than the curvature κ of ∂S on $\partial S \cap B_0$. Consider at such a boundary point as illustration what happens with the normal – it may be the ordinate – at this point and the balls B_0, $B_\delta := q + \left(r^D - \delta\right) B$ and the ball B_α with radius $\alpha = 1/\kappa$ having its midpoint on this normal and being tangential to ∂S. Neglecting effects of higher order, ε is the most possible shift of B_δ in B_α in the direction of abscissa. Elementary calculations give a second order approximation of $\delta(\varepsilon)$

$$\delta(\varepsilon) = -\left(\alpha - r^D\right) + \sqrt{\varepsilon^2 + \left(\alpha - r^D\right)^2} = \frac{\varepsilon^2}{2\left(\alpha - r^D\right)} + O\left(\frac{\varepsilon^4}{\left(\alpha - r^D\right)^3}\right).$$

If B_0 has some boundary curve of S in common, then we get $\delta(\varepsilon) = \varepsilon$. If α tends to infinity, then $\delta(\varepsilon)$ tends to zero. This is probably the case where

difficulties can be expected. However, if some part of ∂S is a line ($\kappa \equiv 0$) then it has no influence on $\delta(\varepsilon)$ since it is determined by the other points of $\partial S \cap B_0$. Since r^D is larger for all points being not in $I^D(r)$ we obtain again some $\delta(\varepsilon, t) \geq \delta(\varepsilon, t_0) = \delta(\varepsilon)$ till errors of maybe higher order. We must bear in mind that α depends on t. We do not continue these more sophisticated considerations since they require lots of differential geometric results already in finite dimensions with the Euclidean norm. We do not know a stability statement for $\rho = D$ similar to Proposition 2.15 for $\rho = H$.

The above example offers some hope that the stability property is also valid in both the cases of ρ for a large class of mappings Σ. However, this must be seriously investigated and it is not our intention in this paper.

For deriving necessary and sufficient conditions of optimality we have to calculate the directional derivative of the functionals φ^ρ. In the case φ^H we have the "max max max" structure which can be dealt with in a well-known manner. The handling of φ^D is because of its "max min max" structure more troublesome. In classical books and papers it is not dealt with, see e.g. in [17] and also in [5, Kap. I.5, Kap. I.6.2], [20] where investigations are only done in spaces with scalar product which yields considerable simplifications. We could not find other relevant references in order to shorten some of the following considerations. We only need a special treatment for $x_0 \in \operatorname{int} S$. The case $x_0 \notin \operatorname{int} S$ can be dealt with the reduction result from Proposition 2.2. We look for equivalent conditions ensuring that the directional derivative defined by $D'(x_0, S; h) = \lim_{t \to +0} (D(x_0 + th, S) - D(x_0, S)) / t$ is negative which we need for the formulation of the Kolmogorov – criterion of (P^D) for $x_0 \in \operatorname{int} S$ and $S = \operatorname{cl} \operatorname{conv} S$. We use the largest inscribed ball of S with the given midpoint x_0. Hence the tangent cones $T(s)$ of S at active points s, that means the points s of the intersection of the boundary of S with an inscribed ball, and its (negative) polar cone $N^*(s) = T(s)^*$, the normal cones of S at s, play further on an essential role.

The following considerations are also possible for the distance H. Here the set S is the "inscribed" set of the smallest outer ball of S. Thus, the tangent cone of the outer ball and its normal cone at active points has to be used here. However, the method of the proof essentially uses the compactness of the unit sphere which immediately implies the finite dimensionality of Y. This can be avoided for (P^H) since we are able to use stability results of [16].

Let $x_0 \in \operatorname{int} S$. Then we call $F(x_0) := \{s \in S | \|s - x_0\| = \inf_{y \in Y \setminus S} \|y - x_0\| =: r_0\}$ the **active index set** of S with respect to the largest in-

scribed ball with a given midpoint x_0, the **tangential cone** $T_S(s) :=$ cl cone $(S-s)$ and the **normal cone** (negative polar cone) $N_S^*(s) := T_S(s)^* = \{y^* \in Y^* \,|\langle y^*, y\rangle \le 0 \;\forall y \in T_S(s)\}$ of S at $s \in F(x_0)$. We omit the index S and write simply $T(s)$ and $N^*(s)$ whenever the reference to S is evident. Under the assumption $x \in \operatorname{int} S$ with a proper closed convex set S and finite dimensional space Y we have $D(x,S) = -\inf_{s\in S^c}\|x-s\| = -\min_{s\in S^c}\|x-s\|$.

Lemma 2.16. *Let Y be a finite dimensional normed space. Assume that $S \subset Y$ is closed, convex and bounded and that $x_0 \in \operatorname{int} S$. Then*

$$h \in \bigcap_{s\in F(x_0)} \operatorname{int} T(s) \tag{13}$$

if and only if there exist numbers $r > r_0$ and $t_0 > 0$, such that for all $0 < t < t_0$

$$x_0 + th + (r_0 + t(r-r_0))\, B \subset S. \tag{14}$$

Proof. *Sufficiency:* With the abbreviation $B_t = x_0 + th + (r_0 + t(r-r_0))\, B$ the support functions of the balls B_t and B_0 satisfy the equality $\sigma_{B_t}(y^*) = \sigma_{B_0}(y^*) + t\,(\langle y^*, h\rangle + (r-r_0))$, $y^* \in \bar{B}^*$. The inclusion $B_0 \subset S$ is equivalent to $\sigma_{B_0}(y^*) - \sigma_S(y^*) \le 0$ for all $y^* \in \bar{B}^*$. Similarly, to prove that $B_t \subset S$ we must show that $\sigma_{B_t}(y^*) - \sigma_S(y^*) \le 0$ for all $y^* \in \bar{B}^*$. We prove first that for each $y_0^* \in \bar{B}^*$ there exist a norm-neighborhood $U(y_0^*)$ of y_0^* and numbers $\varepsilon(y_0^*) > 0$, $r(y_0^*) > r_0$, such that $\sigma_{B_t}(y^*) - \sigma_S(y^*) \le 0$, for all $y^* \in U(y_0^*)$, $0 < t < \varepsilon(y_0^*)$, $r_0 \le r \le r(y_0^*)$. We have to consider two cases.

1) $y_0^* \in N^*(s)$, $s \in F(x_0)$. Since $h \in \operatorname{int} T(s)$, we have $\langle y_0^*, h\rangle < -\delta(y_0^*) < 0$ for some positive $\delta(y_0^*)$. Therefore $\langle y^*, h\rangle < -\delta(y_0^*) < 0$ for all y^* of some neighborhood $U(y_0^*)$ of y_0^*. Put $r(y_0^*) = r_0 + \delta(y_0^*)$. Then for $t > 0$ and $r_0 \le r \le r(y_0^*)$ we have $\sigma_{B_t}(y^*) - \sigma_S(y^*) = \sigma_{B_0}(y^*) - \sigma_S(y^*) + t\,(\langle y^*, h\rangle + (r-r_0)) \le t\,(\langle y^*, h\rangle + (r-r_0)) \le t\,(-\delta(y_0^*) + (r-r_0)) \le t\,(-\delta(y_0^*) + r(y_0^*) - r_0) = 0$.

2) $y_0^* \notin \bigcup_{s_0\in F(x_0)} N^*(s_0)$. In such a case $\sigma_{B_0}(y_0^*) - \sigma_S(y_0^*) < -\Delta(y_0^*) < 0$ for some positive number $\Delta(y_0^*)$. From the continuity of the support function there exists a norm-neighborhood $U(y_0^*)$ of y_0^* with radius smaller than one such that $\sigma_{B_0}(y^*) - \sigma_S(y^*) < -\Delta(y_0^*) < 0$ for all $y^* \in U(y_0^*)$. Choose $r(y_0^*) = r_0 + \|h\|$ and $\varepsilon(y_0^*) = \Delta(y_0^*)/(3\|h\|)$. Then for $r < r(y_0^*)$ and $t < \varepsilon(y_0^*)$ it follows $\sigma_{B_t}(y^*) - \sigma_S(y^*) < -\Delta(y^*) + t\,(\langle y^*, h\rangle + (r-r_0)) \le$

$-\Delta(y^*)+t\left(\langle y^*,h\rangle+\|h\|\right)\leq 0$ for all $y^*\in U(y_0^*)$. Because of the compactness of $\bar{B}^*$ there exists a finite number of points y_i^*, $i=1,2,\ldots,n$, and related above norm-neighborhoods $U(y_i^*)$ such that $\bar{B}^*\subset U^*(y_1^*)\cup\ldots\cup U^*(y_n^*)$. Then the assertion holds for $r_0<r<\min\left(r(y_1^*),\ldots,r(y_1^*)\right)$ and $t_0=\min\left(\varepsilon(y_1^*),\ldots,\varepsilon(y_1^*)\right)$.

Necessity: We have to show

If $h\notin\bigcap_{s\in F(x_0)}\operatorname{int}T(s_0)$, $\|h\|\neq 0$, then for arbitrary $r>r_0$ and $t>0$ $B_t\cap(S)^c\neq\emptyset$ holds.

Please note that for all $y^*\in\bigcup_{s\in F(x_0)}N^*(s)$ the equality $\sigma_{\bar{B}_0}(y^*)=\sigma_S(y^*)$ is valid. If $h\notin\bigcap_{s\in F(x_0)}\operatorname{int}T(s), h\neq 0$ then there exist $s\in F(x_0)$ and $y^*\in N^*(s)$ such that $\langle y^*,h\rangle\geq 0$. We obtain $\sigma_{B_t}(y^*)-\sigma_S(y^*)=\sigma_{B_0}(y^*)-\sigma_S(y^*)+t\left(\langle y^*,h\rangle+(r-r_0)\right)>0$ for all $t>0$ and all $r>r_0$ i.e. there are points of B_t which do not belong to S. ■

Proposition 2.17. *Let the assumptions of Lemma* 2.16 *be satisfied. Then* (13) *is equivalent to*

$$D'\left(x_0,S,h\right)<0 \tag{15}$$

and to the assertion that $\langle n^*\left(s\right),h\rangle<0$ *for all* $s\in F\left(x_0\right)$ *and* $n^*\left(s\right)\in N^*\left(s\right)$.

Proof. 1) If $h\in\bigcap_{s\in F(x_0)}\operatorname{int}T(s_0)$, then there exist numbers $r>r_0$ and $t_0>0$, such that (14) is valid for all $0<t<t_0$. Hence we obtain for each $t_0>0$ with $\psi\left(t\right):=\left(D\left(x_0+th,S\right)-D\left(x_0,S\right)\right)\ /\ t$ the estimations $\psi\left(t\right)\leq-\left(r_0+t(r-r_0)\right)-(-r_0))\ /\ t=r_0-r<0$. Going to the limit for $t\to+0$ we get (15).

2) Now assume (15) and define $r_0-\tilde{r}:=D'\left(x_0,S,h\right)<0$. Since ψ is increasing in t we get $\psi\left(t\right)t=t\left(r_0-\tilde{r}\right)+o\left(t\right)$ with $o\left(t\right)\geq 0$ and $\lim_{t\to+0}o\left(t\right)/t=0$. Choose $t_0>0$ such that $o\left(t\right)\leq\delta\,t$ for all $t\in(0,t_0)$ and such that $r_0-\tilde{r}+\delta<0$. Then we obtain $D\left(x_0+th,S\right)\leq-\left(r_0+t(\tilde{r}-\delta-r_0)\right))$ for all $t\in(0,t_0)$ and thus with $r=\tilde{r}-\delta$ the relation (14) is satisfied.

3) It is well-known that $h\in\operatorname{int}T\left(s\right)$ is equivalent to $\langle n^*,h\rangle<0$ for all $n^*\in N^*\left(s\right)$. ■

Corollary 2.18. *Assume that Y is finite dimensional. Then the ball $B(x_0, r_0) \subset S$ is a largest inscribed ball for S if and only if $\bigcap_{s_0 \in F(x_0)} \operatorname{int} T(s_0) = \emptyset$. Consequently, this is a necessary and sufficient condition in order that $D'(x_0, S, h) \geq 0$ for all $h \in Y$.*

3. Existence of solutions and regularity

First, we want to make some statements about the existence of solutions to the convex optimization problems

$$\textbf{(P}^\rho\textbf{)} \qquad \varphi^\rho(g) \equiv \sup_{t \in M} \rho(g(t), \Sigma(t)) \to \inf, \ g \in V,$$

for $\rho = H, D$. Because of the continuity of φ^ρ we can formulate the following standard statement which covers a lot of applications.

Proposition 3.1. *If $\Sigma : M \to 2^Y$ satisfies condition* (D), (B) *and $V \subset C(M, Y)$ is closed, convex and $V \cap \lambda \operatorname{cl} B$ is nonempty and weakly compact for sufficiently large $\lambda > 0$, then Problem* (P^ρ) *has an optimal solution.*

Proof. φ^ρ is weakly l.s.c. and convex. Observing its coercivity, V can be replaced by the weakly compact set $V \cap \lambda \operatorname{cl} B$ for suitable $\lambda > 0$. Hence the existence follows from the Weierstraß Theorem. ■

The scope of this paper is to describe the behavior of possible solutions of Problem (P^ρ) and to show that on the one hand, the notion of triviality of an approximation problem – we call it here better regularity – must be extended and on the other hand, the standard optimality conditions of Kolmogorov can be established in a well-known manner. These optimality conditions directly follow from the sufficient and necessary optimality condition that the directional derivative of φ^ρ is nonnegative on the cone of admissible directions of V at the solution g_0. To compute the directional derivative of φ^ρ the solutions of the lower level problems of (P^ρ) play an essential role as so-called active index sets. Here we have two approaches, the first one by use of the definition of ρ and the other by use of its geometric interpretation. In both cases we need for the solvability of the associated lower level problems additional assumptions which are

$$\textbf{(H)} \qquad \Sigma \text{ is } \left\{ \begin{array}{ll} \text{H.u.s.c.} & \text{if } \rho = H, \\ \text{H.l.s.c.} & \text{if } \rho = D, \end{array} \right\}$$

$$\textbf{(C)} \qquad \forall_{t \in M}\, \Sigma(t) \text{ is relatively compact.}$$

We assume at the beginning that (D), (B), (H) are satisfied and assume also that Problem (P^ρ) has an optimal solution $\hat{g}$ and start with the first approach. Proposition 2.9 yields that the set

$$T := \left\{ \hat{t} \in M \mid \varphi(\hat{g}) = \sup_{t \in M} \rho(\hat{g}(t), \Sigma(t)) \right\}$$

of the active indices is nonempty, compact and contains the accumulation points of each net $\{t_\alpha\}$ in M with $\lim_\alpha \rho(\hat{g}(t_\alpha), \Sigma(t_\alpha)) = \varphi(\hat{g})$. Taking into account $\rho(\hat{g}(t), \Sigma(t)) = \rho(\hat{g}(t), \operatorname{cl}\operatorname{conv}\Sigma(t))$ we deduce that Problem (P^ρ) is equivalent to

$$\varphi(g) \equiv \sup_{t \in M} \rho(g(t), \operatorname{cl}\operatorname{conv}\, \Sigma(t)) \to \inf,\ g \in V\,. \tag{16}$$

Fixing $\hat{t} \in T$ and keeping in mind the support function structure of ρ we have two further lower level problems. It can be shown by standard separation arguments that under (C) these problems have optimal solutions and the corresponding active index set is compact. However, if we use the geometric interpretation of ρ, then we have only one lower level problem. Since the complications by the derivation of directional derivatives increase with the number of lower level problems we prefer in our further considerations the geometric interpretation where $\Sigma\,(t)$ has to be replaced by $\operatorname{cl}\operatorname{conv}\Sigma\,(t)$ which is without consequences for the original problems. We need for the attainment of the supremum in $\varphi(\hat{g}) = H(\hat{g}(\hat{t}), \Sigma(\hat{t})) = \sup_{s \in \operatorname{cl}\operatorname{conv}\Sigma(\hat{t})} \left\|\hat{g}(\hat{t}) - s\right\|$ and of the related infimum in $\varphi(\hat{g}) = D(\hat{g}(\hat{t}), \Sigma(\hat{t})) = \inf_{s \in \operatorname{cl}\operatorname{conv}\Sigma(\hat{t})} \left\|\hat{g}(\hat{t}) - s\right\| - \inf_{s \in Y \setminus \operatorname{cl}\operatorname{conv}\Sigma(\hat{t})} \left\|\hat{g}(\hat{t}) - s\right\|$ the compactness of $\operatorname{cl}\operatorname{conv}\Sigma(\hat{t})$, i.e. condition (C) just as in the first approach. The replacement of $\Sigma\,(t)$ by $\operatorname{cl}\operatorname{conv}\Sigma\,(t)$ inherits the semi-continuity properties.

Lemma 3.2. *If M is a topological space, Y a normed space and $\Sigma : M \to 2^Y$ a given set-valued function, which is either H.l.s.c. or H.u.s.c., then the functions $t \in M \to \operatorname{cl}\Sigma(t)$ and $t \in M \to \operatorname{conv}\Sigma(t)$ possess the same property.*

Proof. The proof is straightforward by use of $\operatorname{cl}\Sigma(t_0) \subset \Sigma(t_0) + \varepsilon B$ and $\operatorname{conv}(\Sigma(t) + \varepsilon B) \subset \operatorname{conv}\Sigma(t) + \varepsilon B$ ■

We study Problem (P^ρ) assuming that the set-valued function Σ satisfies conditions (D), (B) and (H). Considering the classical case of the approximation theory of functions one can notice the general assumption that the

function f to be approximated by some closed set of functions V does not belong to V. If $f \in V$, then one could say that some case of irregularity is satisfied. A similar situation can be described for the approximation of a set-valued function by a single-valued one. To define the irregularity, inherited from the classical case, we introduce some estimation for the optimal value $\hat{\varphi} = \inf\{\varphi(g) \mid g \in V\}$ which is independent of the constraint set V.

Let W and S be nonempty subsets of Y. We define a *non symmetric distance* between W and S by $\hat{\rho}(W,S) = \inf_{q\in W} \rho(q,S)$ and for given $t \in M$, $V \subset C(M,Y)$ the *section* of V at t by $V(t) = \{g(t) \mid g \in V\}$.

Obviously, $\hat{\rho}(\cdot,S)$ is monotonically decreasing in the sense that $W_1 \subset W_2$ implies $\hat{\rho}(W_1,S) \geq \hat{\rho}(W_2,S)$ and the following estimation holds for all $\tau \in M$

$$\hat{\varphi} = \inf_{g\in V} \sup_{t\in M} \rho(g(t),\Sigma(t)) \geq \inf_{g\in V} \rho(g(\tau),\Sigma(\tau)) = \hat{\rho}(V(\tau),\Sigma(\tau)). \tag{17}$$

Taking the supremum over $\tau \in M$ we find

$$\hat{\varphi} = \inf_{g\in V} \varphi(g) \geq \sup_{t\in M} \hat{\rho}(V(t),\Sigma(t)) \tag{18}$$

and finally by monotonicity of $\hat{\rho}(\cdot,S)$ we obtain

$$\hat{\varphi} = \inf_{g\in V} \varphi(g) \geq \sup_{t\in M} \hat{\rho}(Y,\Sigma(t)) =: \bar{\varphi}. \tag{19}$$

Definition 3.1. We say that (P^ρ) is a **regular case problem** if in (19) a strict inequality occurs and it is an **irregular case problem** if there is an equality in (19). $\bar{\varphi} := \sup_{t\in M} \hat{\rho}(Y,\Sigma(t))$ is called the **irregular case level**.

To obtain the irregular case level $\bar{\varphi}$ in (19) we have first to find the optimal value of the lower level optimization problem

$$\rho(q,\Sigma(t)) = \rho(q,\operatorname{cl}\operatorname{conv}\Sigma(t)) \to \inf,\ q \in Y. \tag{20}$$

for all $t \in M$. We denote, *in the case* $\rho = H$, with $r^H(t)$ the optimal value (infimum of the radii of all circumscribed balls – an optimal solution $q(t)$ is called the *Chebyshev center* of $\Sigma(t)$) and, *in the case* $\rho = D$, with $r^D(t)$ the *absolute* value of the optimal value (supremum of the radii of all inscribed balls of $\operatorname{cl}\operatorname{conv}\Sigma(t)$) of Problem (20).

The next proposition concerns the solutions to both problems. Note that we do not predispose condition (C).

Proposition 3.3. *Let* Σ *satisfy conditions* (D), (B), (H).

i) *The function* $r^H(\cdot)$ *is u.s.c. and* $r^D(\cdot)$ *is l.s.c. on* M. *Hence* $r^H_{\max} = \max_{t\in M} r^H(t)$ *and* $r^D_{\min} = \min_{t\in M} r^D(t)$ *are attained on compact subsets of* M. *If* Σ *is Hausdorff-continuous at* t_0, *then both radii are continuous at* t_0.

ii) *If* Y *is a reflexive Banach space, then Problem* (20) *has an optimal solution* $q^\rho(t)$.
Case $\rho = H$: *The optimal solution belongs to* $\operatorname{cl}\operatorname{conv}\Sigma(t)$ *whenever* Y *is a Hilbert space or* $\dim Y \le 2$.
Case $\rho = D$: *A (generally non unique) optimal solution* $q^D(t)$ *belongs to* $\operatorname{cl}\operatorname{conv}\Sigma(t)$. *If* $\operatorname{int}\operatorname{cl}\operatorname{conv}\Sigma(t) = \emptyset$, *then each* $q \in \operatorname{cl}\operatorname{conv}\Sigma(t)$ *is a solution and* $r^D(t) = 0$.

iii) $\bar\varphi = r^H_{\max}$ *in case* $\rho = H$ *and* $\bar\varphi = -r^D_{\min}$ *in case* $\rho = D$ *are the irregular case levels.*

iv) *If* Y *is a reflexive Banach space, then the sets* $\Phi^H(t, r^H_{\max})$ *and* $\Phi^D(t, -r^D_{\min})$ *are nonempty for all* $t \in M$.

Proof. i) Let $t_0 \in M$ be fixed.
For $\rho = H$ we have for each $\tau > 0$ some $q^H_\tau(t_0)$ and radius $r^H_\tau(t_0) > 0$ such that $\Sigma(t_0) \subset q^H_\tau(t_0) + r^H_\tau(t_0)\ \operatorname{cl} B$. Further, $\lim_{\tau\to 0} r^H_\tau(t_0) = r^H(t_0)$ holds.

By the Hausdorff upper semi-continuity of Σ for each $\varepsilon > 0$ there exists a neighborhood U_ε of t_0 such that for $t \in U_\varepsilon$ it holds $\Sigma(t) \subset \Sigma(t_0) + \varepsilon\ \operatorname{cl} B \subset q^H_\tau(t_0) + (r^H_\tau(t_0) + \varepsilon)\ \operatorname{cl} B$. Since $r^H(t)$ is the infimum of the radii of all circumscribed balls for $\Sigma(t)$ we obtain $r^H(t) \le r^H_\tau(t_0) + \varepsilon$. For $\tau \to 0$ we get the upper semi-continuity of r^H.

From the additional Hausdorff lower semi-continuity of Σ we find analogously for arbitrary $\varepsilon > 0$, $t \in U_\varepsilon$ and each $\tau > 0$ the inclusions $\Sigma(t_0) \subset \Sigma(t) + \varepsilon\ \operatorname{cl} B \subset q^H_\tau(t) + \big(r^H_\tau(t) + \varepsilon\big)\ \operatorname{cl} B$, i.e. $r^H(t_0) \le \lim_{\tau\to 0} r^H_\tau(t) + \varepsilon = r^H(t) + \varepsilon$ and hence l.s.c at t_0.

The proof for $\rho = D$ is very similar.

ii) Use e.g. our geometric interpretation and take Belobrov [2, Paragraf 1, Statement 2] for q^H and Garkavi [8, Theorem 4] for q^D, see also [21, p. 288] for more information and [6, Theorem 6] or [8, Theorem 3] for uniqueness.

iii) Actually this point is proven in i). The formulation is rather emphatic to fix the formula for the irregular case level.

iv) The assertion is obvious in view of the definition $\Phi^r(t, r) = \{q \in Y \mid \rho(q, \Sigma(t) \le r\}$. ■

Proposition 3.4. *Assume that Σ satisfies* (D), (B), (H). *If Σ is single-valued function, then $\bar{\varphi} = 0$. If additionally $V \subset C(M, Y)$ is closed, then* (P^ρ) *is an irregular case problem if and only if $\Sigma \in V$.*

Proof. (H) implies continuity for single-valued mappings which immediately gives the above statements. ■

In case of a set-valued function Σ the situation is more complicated. Already the case $V = C(M, Y)$, trivial for single-valued Σ, is of principal importance and non-trivial for set-valued Σ as the next theorem shows.

Theorem 3.5. *Consider Problem (P^ρ) with $V = C(M, Y)$. Assume that the set-valued function Σ satisfies conditions* (D), (B), (H) *and let $\bar{\varphi}$ be the irregular case level. Then the following statements are true.*

i) *The set $L(\varepsilon)$ of all elements g of $C(M, Y)$ with $\bar{\varphi} \leq \varphi(g) \leq \bar{\varphi} + \varepsilon$ is equal to the set $S(\varepsilon)$ of all continuous selections of the map $\Phi^\rho(\cdot, \bar{\varphi} + \varepsilon)$ for all $\varepsilon \geq 0$. If $\varepsilon > 0$, then $L(\varepsilon) \neq \emptyset$.*

ii) *Problem (P^ρ) is an irregular case problem.*
The set $T := \{t \in M \,|\hat{\rho}(Y, \Sigma(t)) = \bar{\varphi}\}$ of active points is non empty.

iii) *$L(0) \neq \emptyset$, i.e. Problem (P^ρ) has a solution, if Y is a reflexive Banach space and if for each $t_0 \in T$ one of the following conditions is true.*

   a) *$Y = \mathbb{R}^n$, Σ is Hausdorff-continuous on some neighborhood U of t_0 and $\Phi(t_0, \bar{\varphi})$ is single valued.*

   b) *$\Phi^\rho(\cdot, \bar{\varphi})$ is stable at t_0.*

Proof. i) Take $\varepsilon > 0$ and $t \in M$ arbitrarily and define $\rho(t) := \hat{\rho}(Y, \Sigma(t))$ then $\Phi^\rho(t, \rho(t) + \varepsilon) = \{q \in Y \,|\rho(q, \Sigma(t)) \leq \rho(t) + \varepsilon\} \neq \emptyset$ by definition of $\rho(t)$. Since $\rho(t) \leq \bar{\varphi}$ for all $t \in M$ we get also $\Phi^\rho(t, \bar{\varphi} + \varepsilon) \neq \emptyset$ and above that $\Phi_0^\rho(t, \bar{\varphi} + \varepsilon) \neq \emptyset$. Hence, by Proposition 2.10 ii) $\Phi^\rho(\cdot, \bar{\varphi} + \varepsilon)$ is l.s.c. which implies by Michael's Selection Theorem that $\Phi^\rho(\cdot, \bar{\varphi} + \varepsilon)$ admits a continuous selection g_ε. It follows $\bar{\varphi} = \sup_{t \in M} \inf_{q \in Y} \rho(q, \Sigma(t)) \leq \sup_{t \in M} \rho(g_\varepsilon(t), \Sigma(t)) = \varphi(g_\varepsilon) \leq \bar{\varphi} + \varepsilon$. and therefore $\emptyset \neq S(\varepsilon) \subset L(\varepsilon)$. Now let be $g \in C(M, Y)$ and $\bar{\varphi} \leq \varphi(g) \leq \bar{\varphi} + \varepsilon$. Then $\bar{\varphi} \leq \rho(g(t), \Sigma(t)) \leq \bar{\varphi} + \varepsilon$ for all $t \in M$. This implies $g \in S(\varepsilon)$.

ii) The first statement follows from $\lim_{\varepsilon \to 0} \varphi(g_\varepsilon) = \bar{\varphi}$ and the second from Proposition 3.3 i).

iii) Using a) and b) it follows from Proposition 3.3 iv) that $\Phi^\rho(t, \bar{\varphi}) \neq \emptyset$ for all $t \in M$ and finally from Proposition 2.10 that $\Phi^\rho(\cdot, \bar{\varphi})$ is l.s.c. on M.

With the same arguments as in i) we get the existence of a solution to Problem (P^ρ) for $V = C(M, Y)$ and $L(0) = S(0) \neq \emptyset$. ■

Remark 3.1. We use Michael's Selection Theorem to prove the existence of a continuous selection. There is a large number of generalizations which use weaker conditions then the lower semi-continuity of $\Phi^\rho(\cdot, r)$. For level set maps we can find in [19] several statements for continuous selections. However, all statements use similar conditions as our Slater's condition on M which means $I^\rho(r) = \emptyset$. The troublesome case $I^\rho(r) \neq \emptyset$ seems to be not considered, as far as we noticed.

Dealing with Problem (P^ρ) we have to determine whether it is a regular or irregular case problem and in the case of irregular case problem we can find its solutions like in Theorem 3.5 as continuous selections of the level set map $\Phi^\rho(\cdot, \bar{\varphi})$, which belong to V. Then the optimal value $\hat{\varphi} = \bar{\varphi}$ according to Proposition 3.3 is equal to $r^H_{\max}$ in case $\rho = H$ and $-r^D_{\min}$ in case $\rho = D$. When distinguishing between regular and irregular case according to inequality (18) the following comparison argument may be useful. Let $V_1 \subset V_2$ be two subsets of $C(M, Y)$ and consider the two problems (P^ρ) with $V = V_1$ and $V = V_2$. Call these problems respectively $P(V_1)$ and $P(V_2)$. For the optimal values of these problems we have obviously $\hat{\varphi}_{V_1} = \inf_{g \in V_1} \varphi(g) \geq \inf_{g \in V_2} \varphi(g) = \hat{\varphi}_{V_2} \geq \bar{\varphi}$. Therefore if $P(V_1)$ is an irregular case problem, so is $P(V_2)$. Conversely, if $P(V_2)$ is a regular case problem, so is $P(V_1)$.

4. Kolmogorov optimality conditions

In the whole section, we assume conditions (B), (D), (H), (C), that $V \subset C(M, Y)$ is a closed convex set and that in case $\rho = D$ the space Y is finite dimensional. Further, we denote a solution of (P^ρ) by g_0. Because of our geometrical approach we assume that $\Sigma(t) = \operatorname{cl}\operatorname{conv}\Sigma(t)$. In other cases, $\Sigma(t)$ must be replaced in the following formulas by $\operatorname{cl}\operatorname{conv}\Sigma(t)$.

We exemplarily demonstrate in this section that the derivation of optimality conditions of the Kolmogorov-type is straightforward as done in the classical theory of Chebyshev approximation, see e.g. [16], [17], [11], [21] or in the simple case $Y = \mathbb{R}$, considered in [9]. Therefore we do not grasp at full generality and make suitable simplifications, especially for deriving the dual statements. If the problem is not regular, then the Kolmogorov conditions are of a less complicated structure. We omit the index ρ as long as a specification is not necessary or is evident. The following well-known statement

is true *independent of the above regularity* of (P^H) or (P^D). In the case of $\inf_{g\in V}\varphi(g)=\bar{\varphi}$ the selection argument gives an additional possibility of the characterization of optimal solutions.

Theorem 4.1 ([16, p. 66])**.** *$g_0\in V$ is a solution of (P^H) or (P^D) if and only if the one-sided directional derivative of the objective is nonnegative on admissible directions, i.e. $\varphi'_+(g_0;g-g_0)\geq 0$ for each $g\in V$.*

Remark 4.1. Let S be convex and bounded. From our continuity and convexity results it follows that the sequences of difference quotients for $\varphi, D(\cdot,S)$, $H(\cdot,S)$ and $\|\cdot\|$ are non increasing ([16, Theorem 1, p. 47]).

First we study the case $\rho=H$. The irregular, classical case $\varphi(g_0)=0$, i.e. $\Sigma(t)=\{g_0(t)\}$ for all $t\in M$, is excluded and will not be considered. Let $E(g):=\{t\in M\mid H(g(t),\Sigma(t))=\varphi(g)\}$ and $F(g,t):=\{s\in\Sigma(t)\mid\|g(t)-s\|=\varphi(g)\}, t\in E(g)$ be the **sets of active indices** for an element $g\in V$. We use the notations $\Omega:=\{(t,s)\in M\times Y\ |s\in\Sigma(t)\}$, $\omega=(t,s),\Omega_0:=\{(t,s)\in\Omega\ |t\in E(g_0),s\in F(g_0,t)\}=\{\omega\in\Omega\,|\varphi(g_0)=\|g_0(t)-s\|\}$.

Proposition 4.2 ([16, p. 58–62])**.** *If $\{f_n\}$ is a non increasing sequence of u.s.c. functions of the compact metric space Y into the extended reals $\overline{\mathbb{R}}:=\mathbb{R}\cup\{-\infty,\infty\}$ converging pointwise to the function f, then f is u.s.c. and $\lim_{n\to\infty}\max_{y\in Y}f_n(y)=\max_{y\in Y}f(y)$.*

Remark 4.2. It is well-known for arbitrary q and h of the normed space Y for $q\neq 0$ that

$$\lim_{n\to\infty}n\left(\|q+h/n\|-\|q\|\right)=\max_{q^*}\{\langle q^*,h\rangle\,|\langle q^*,q\rangle=\|q\|,\|q^*\|=1\}\tag{21}$$

and for $q=0$ that $\lim_{n\to\infty}n\left(\|q+h/n\|-\|q\|\right)=\|h\|$ where $q^*\in Y^*$ is uniquely determined whenever Y is a reflexive Banach space. If Y is a Hilbert space with the scalar product $\bullet$ and $q\neq 0$ then

$$\lim_{n\to\infty}n\left(\|q+h/n\|-\|q\|\right)=(q\bullet h)/\|q\|.\tag{22}$$

Proposition 4.3. *If M is compact and the set-valued mapping Σ from M to the Banach space Y is convex compact valued and u.s.c. on M, then we get $\varphi'_+(g_0,h)=\max_{(t,s)\in\Omega_0}\lim_{n\to\infty}n(\|g_0(t)-s+h(t)/n\|-\|g_0(t)-s\|)$ and in case of a Hilbert space $\varphi'_+(g_0,h)=\max_{(t,s)\in\Omega_0}(g_0(t)-s)\bullet h(t)/\|g_0(t)-s\|$ for each direction $h\in C(M,Y)$.*

Proof. Using $f_n(\omega) = n(\|g_0(t) - s + h(t)/n\| - \varphi(g_0))$ we obtain $\varphi'_+(g_0, h) = \lim_{n\to\infty} \max_{\omega\in\Omega} f_n(\omega)$. The set-valued mapping $t \longmapsto \Omega(t) := \{t\} \times \Sigma(t)$ is u.s.c. on M with compact values in $M \times Y$. Hence (see e.g. [1, Proposition 3 in Chapter 1.1]) $\Omega = \Omega(M)$ is compact. It remains to show that f_n is non increasing. Let $\omega = (t, s)$ be fixed. Since the sequence of difference quotients of norms is non increasing we get $n\|g_0(t) - s + h(t)/n\| - (n+1)\|g_0(t) - s + h(t)/(n+1)\| \geq -\|g_0(t) - s\| \geq -\varphi(g_0)$. Adding $-n\varphi(g_0)$ to the inequality we obtain $f_n(\omega) \geq f_{n+1}(\omega)$. Proposition 4.2 yields now $\varphi'_+(g_0, h) = \max_{\omega\in\Omega} \lim_{n\to\infty} f_n(\omega)$. It is easy to see that $\lim_{n\to\infty} f_n(\omega) = -\infty$ whenever $(t, s) \notin \Omega_0$. To finish the proof apply (21) and (22). ■

The negation of the formulation that the descent is nonnegative for all directions $g - g_0$, $g \in V$ immediately yields the following **Kolmogorov**-type **primal** optimality condition for problem (P^H).

Theorem 4.4. *Assume M is compact and the point-to-set mapping Σ from M to the Banach space Y is convex compact valued and u.s.c. on M and assume $\varphi(g_0) > 0$. g_0 is a solution of (P^H) if and only if there is no $g \in V$ such that for each (t, s) of the active index set Ω_0 the inequality $\max\{\langle q^*, g(t) - g_0(t)\rangle \mid \langle q^*, g_0(t) - s\rangle = \|g_0(t) - s\|, \|q^*\| = 1\} < 0$ or in the case of a Hilbert space Y with the scalar product $\bullet$ the inequality $(g_0(t) - s) \bullet (g(t) - g_0(t)) < 0$ is satisfied. Obviously Ω_0 is compact as a closed subset of the compact set Ω.*

Example 4.1 [9, Example 1]**.** Let $0 \leq a < b$, $M = [-1, 1]$ be given. Consider (P^H) for $V = \{g \in C(M, \mathbb{R}) \mid g(t) = \alpha + \beta t,\ \alpha,\ \beta \in \mathbb{R}\}$, $\Sigma(t) = \{0\}$ if $-1 \leq t \leq 0$ and $\Sigma(t) = [at, bt]$ if $0 < t \leq 1$. We get the irregular case level $r_{\max} = \bar{\varphi}^H = (b-a)/2$. It follows $\Phi^H(t) = [-(b-a)/2, (b-a)/2]$ if $-1 \leq t \leq 0$ and $\Phi^H(t) = [bt - (b-a)/2, at + (b-a)/2]$ if $0 < t \leq 1$. Both mappings Σ and Φ^H are continuous. Φ^H admits at least one selection belonging to V if and only if $b \geq 2a$. The set of all such selections $S(0)$ is given by $S(0) = \{g_0 \in C(M, \mathbb{R}) \mid g_0(t) = k(t-1) + (a+b)/2,\ a \leq k \leq b/2\}$. Now we illustrate the primal Kolmogorov conditions for this example in the irregular case $b \geq 2a$. We find $E(g_0) = [0, 1]$ if $k = a$, $E(g_0) = \{1\}$ if $a < k < b/2$ and $E(g_0) = \{-1, 1\}$ if $k = b/2$. Further, we get $F(g_0, t) = \{0\}$ if $t = -1, F(g_0, t) = \{at\}$ if $t \in [0, 1)$ and $F(g_0, t) = \{a, b\}$ if $t = 1$. For each $k \in [a, b/2]$ we find $\Omega_0 \supset \{(1, a), (1, b)\}$. The two resulting conditions $\pm(b-a)(\alpha + \beta - (a+b)/2) < 0$ cannot be simultaneously satisfied for any

$\alpha, \beta \in \mathbb{R}$. Hence, each $g_0 \in V_0$ satisfies the Kolmogorov optimality condition stated above.

In the irregular case the consideration of the active points of Ω_0 defined by some $t \in M$, where $r_{\max}$ is attained, seems to be enough to show optimality. The following statement assures this aspect.

Proposition 4.5. *Let S be a compact, convex subset of $Y = \mathbb{R}^n$. r is the optimal value and y_0 is a solution to the problem* (cf. (20)) $H(y, S) \to \min$, *subject to $y \in S$ if and only if there is no $y \in S$ such that the inequality* $\max\{\langle q^*, y - y_0\rangle \,|\langle q^*, y_0 - s\rangle = \|y_0 - s\|, \|q^*\| = 1\} < 0$ *or in the case of a Hilbert space the inequality $(y_0 - s) \bullet (y - y_0) < 0$ is satisfied for all $s \in S_H(y_0) := \{y \in S \,|\|y_0 - y\| = r\}$.*

Proof. Repeat the proof of the Proposition 4.3 with $M = \{t\}, y = g(t)$, $S = \Sigma(t)$. ■

Now we study the case $\rho = D$. Distinguishing the sign of $\varphi(g)$ we use the similar active index sets $E(g) := \{t \in M \mid D(g(t), \Sigma(t)) = \varphi(g)\}$, in case $\varphi(g) > 0$: $F(g,t) := \{s \in \Sigma(t) \mid \|g(t) - s\| = \varphi(g)\}$, in case $\varphi(g) = 0$: $F(g,t) := \{g(t)\}$ and in case $\varphi(g) < 0$: $F(g,t) := \{s \in \mathrm{cl}(Y \setminus \Sigma(t)) \mid -\|g(t) - s\| = \varphi(g)\}$ for an element $g \in V$, $t \in E(g)$ and put again $\Omega_0 = \{(t,s) \in M \times Y \,|t \in E(g), s \in F(g,t)\}$. The first step with respect to the max-part is analogous to the calculation for (P^H). The further discussion for the distance function D is much more complicated as in the case of the distance function H.

Lemma 4.6. *Let $\tau \in E(g_0)$ be arbitrarily given. Then $\varphi'_+(g_o, h) = \max\limits_{t \in E(g_0)} D'_+(g_o(t), \Sigma(t); h(t))$.*

Proof. Take $f_n(t) := n\left(D\left(g_o(t) + \frac{1}{n}h(t), \Sigma(t)\right) - \varphi(g_0)\right)$, use the decreasing behavior of the sequence $\left(n\left(D\left(q + \frac{1}{n}p, S\right) - D(q, S)\right)\right)_{n \in \mathbb{N}}$ and $\varphi(g_0) \geq D(g_o(t), \Sigma(t))$ for all $t \in M$ and proceed as in problem (P^H). ■

Corollary 4.7. *g_0 is a solution of (P^D) if and only if there is no direction $h \in V - g_0$ such that $D'_+(g_o(t), \Sigma(t); h(t)) < 0$ for all $t \in E(g_0)$.*

Theorem 4.8. *g_0 is a solution of (P^D) with $\varphi(g_0) < 0$ if and only if there is no direction $h \in V - g_0$ such that*

$$\langle n^*, h(t)\rangle < 0 \; \forall n^* \in N^*_{\Sigma(t)}(s) \tag{23}$$

for all $(t,s) \in \Omega_0$. *In a case of a Hilbert space* (23) *is equivalent to*

$$(s - g_0(t)) \bullet h(t) < 0 \tag{24}$$

for all $(t,s) \in \Omega_0$.

Proof. $\varphi(g_0) < 0$ implies $g_0(t) \in \mathrm{int}\Sigma(t)$ for all $t \in E(g_0)$. Let $t \in E(g_0)$ be fixed. Choose $r_0 = D(g_0(t), \Sigma(t))$, $S = \Sigma(t)$, $x_0 = g_0(t)$, $T(s) = T(t,s)$ and $N^*(s) = N^*(t,s)$. Then Proposition 2.17 implies immediately (23). If Y is a Hilbert space, then the tangent space of the ball at $s \in \Sigma(t)$ with the radius $\|s - g_0(t)\| = -\varphi(g_0)$ is a closed half space and is contained in $T_{\Sigma(t)}(s)$. Hence $T_{\Sigma(t)}(s)$ is a closed half space and $N^*_{\Sigma(t)}(s) = \{n^* \,|n^* = \lambda(s - g_0(t)), \lambda \geq 0\}$. ■

Remark 4.3. If Y has a smooth norm, then again $N^*_{\Sigma(t)}(s)$ is a closed half space and it suffices to use the $N^*_{\Sigma(t)}(s)$ generating element n^* uniquely defined by $\langle n^*, s - g_0(t)\rangle = \|s - g_0(t)\|$, $\|n^*\|_{Y^*} = 1$.

Now we trace back the case $\varphi(g_0) \geq 0$ to the previous one $\varphi(g_0) < 0$ by using Proposition 2.2 vii) and well-known separation arguments.

Theorem 4.9. g_0 *is a solution of* (P^D) *with* $\varphi(g_0) \geq 0$ *if and only if there is no direction* $h \in V - g_0$ *such that*

$$\langle n^*, h(t)\rangle < 0 \;\forall n^* \in N^*_{\Sigma(t)}(s) \cap \left(N^*_{\varphi(g_0)\,\mathrm{cl}\,B}(g_0(t) - s)\right) \tag{25}$$

for all $(t,s) \in \Omega_0$, *where in case* $\varphi(g_0) = 0$ *the cone* $N^*_{\{0\}}(0)$ *is the entire space* Y^*. *In a case of a Hilbert space and* $\varphi(g_0) > 0$ *the relation* (25) *is equivalent to* $(g_0(t) - s(t)) \bullet h(t) < 0$ *for all* $(t, s(t)) \in \Omega_0$ *where* $s(t)$ *is the unique element of* $F(g_0, t)$.

Proof. We treat first the case $\varphi(g_0) > 0$. Now let $g_0(t) \notin \mathrm{cl}\Sigma(t)$. Then we consider $r > 0$ such that $g_0(t) \in \mathrm{int}\,\Sigma(t)_r$. Because of $D(q, \Sigma(t)_r) = D(q, \Sigma(t)) - r$ we obtain $D'_+(g_o(t),\ \Sigma(t);\ h(t)) = D'_+(g_o(t),\ \Sigma(t)_r;\ h(t))$ for each $r \geq 0$. Hence $D'_+(g_o(t),\ \Sigma(t);\ h(t)) < 0$ is equivalent to $D'_+(g_o(t), \Sigma(t)_r; h(t)) < 0$ for all $r \geq 0$. We can apply the Proposition 4.8 to $g_o(t), \Sigma(t)_r$ for suitable $r > 0$ where the optimal value is $\varphi(g_0) - r < 0$. The optimality condition gives instead of (23) $\langle n^*, h(t)\rangle < 0 \;\forall n^* \in N^*_r(t,s)$ for all $(t,s) \in \Omega_r$, where $F_r(g_0,t) = \left\{s \in \overline{(\Sigma(t)_r)^c} \,|\|s - g_0(t)\| = r - \varphi(g_0)\right\}$,

$\Omega_r = \{(t,s) \in E(g_0) \times Y \,|s \in F_r\,(g_0,t)\}$ and $N_r^*(t,s) = [\text{cone}\ (\Sigma(t)_r - s)]^*$. For fixed $t \in E\,(g_0)$, the set $F\,(g_0,t)$ is convex and closed (this is wrong for $\varphi\,(g_0) < 0$). By usual separation arguments it can be shown (straightforwardly but rather technically) that for each $s \in F_r(g_0,t)$, $n^* \in N_r^*(t,s) \cap \partial B^*$ the hyperplane $\mathcal{H}$ defined by $\langle n^*, x-s\rangle = 0$ is a supporting hyperplane of $\Sigma(t)_r$ and $\mathcal{H} + r\frac{(g_0(t)-s)}{\|g_0(t)-s\|}$ is a separating hyperplane of $g\,(t_0) + \varphi(g_0)B$ and of S where all of them having the same support $\text{cl}\,(g_0\,(t) + \varphi\,(g_0)\,B) \cap S = F\,(g_0,t)$. Further, each separating hyperplane of $(g_0(t) + \varphi(g_0)B)$ and S is with suitable parallel shift a supporting hyperplane of $\Sigma(t)_r$ with support in $F_r(g_0,t)$ and normal $n^* \in N_r^*(t,s) \cap \partial B^*$. Thus we have $\bigcup_{s \in F_r(t,g_0)} N_r^*(t,s) = N^*\,(t,\tilde{s}) \cap \left(-N^*_{g_0(t)+\varphi(g_0)B}\,(\tilde{s})\right)$. On the right hand side we do not need to take the union over all $\tilde{s} \in (g_0(t) + \varphi(g_0)B) \cap S = F(g_0,t)$ since the common convex support is contained entirely in each separating hyperplane. In a Hilbert space we have the unique determined normal $n^* = (g_0(t) - s)\ /\ \|g_0(t) - s\|$.

The case $\varphi_0 = 0$ works similarly by using the closed ball $\varphi\,(g_0)\,\text{cl}\,B = \{0\}$. Again the normal cone at $g_0(t)$ with respect to $\{g_0(t)\}$ is the entire space Y^*. ■

Remark 4.4. If Y^* is a strictly convex normed space, then n^* is uniquely determined and we get similarly as above but with a different sign $\langle n^*, g_0(t) - s\rangle = \|g_0(t) - s\| = \varphi(g_0); \|n^*\|_{Y^*} = 1$.

Example 4.2. Consider $\Sigma(t) = [-2,2] \times [-1,1] + t[-1,1]^2, t \in [0,1]$ and $\mathbb{R}^2$ with the usual Euclidean norm. Then for (P^D) we get $r_{\min} = 1$ and $\Phi^D(t) = [-1,1] \times \{0\} + t[-1,1]^2$. If we have $V = \{(f,g) \in C([0,1],\mathbb{R}^2) | f(t) = a + bt, g(t) = c + dt, a,b,c,d \in \mathbb{R}\}$ then $S(0) = \{(f_0,g_0) \in C([0,1],\mathbb{R}^2) | f_0(t) = a + bt, g_0(t) = d\,t, a \in [-1,1],\ a + b \in [-2,2], d \in [-1,1]\}$ and $\varphi^D\,((f_0,g_0)) = -1 = -r_{\min}$ for all $(f_0,g_0) \in S\,(0)$. Again we have got the irregular case. Obviously, the set Ω_0 contains $\{(0,a,1),(0,a,-1)\}$. Since V is a subspace we have to show there is no $(h,k) \in V$ such that $(f_0(t) - s_1)h(t) + (g_0 t) - s_2)k(t) > 0$ for all $(t,s) \in \Omega_0$. We already find for the subset $\{(0,a,1),(0,a,-1)\} \subset \Omega_0$ at $t = 0$ the system $(a-a)h - k > 0$ and $(a-a)h + k > 0$ which cannot be satisfied for any $(h,k) \in \mathbb{R}^2$.

We see again in the irregular case (! $\varphi(g_0) > 0$ is ever regular), that already the system of inequalities cannot be satisfied for any admissible direction h, whenever $t \in E(g_0)$ and the largest inscribed ball of $\Sigma(t)$ has

the radius $r_{\min}$. The following corollary illustrates this fact. For simpler notations we assume that Y is a Hilbert space.

Corollary 4.10. *Let S be a compact, convex subset of the Hilbert space $Y = \mathbb{R}^n$. Then r is the optimal value and $y_0 \in \operatorname{int} S$ is the solution to the problem $D(y, S) \to \min$, subject to $y \in S$ if and only if there is no $h \in Y$ such that $(y_0 - s) \bullet h > 0$ for all $s \in S(y_0, r) := \{y \in S \,|\, \|y_0 - y\| = r\}$.*

Now let us say something about the *dual Kolmogorov optimality conditions.* For simplicity we assume that V is an n-dimensional linear subspace of $C(M, Y)$ generated by the linearly independent elements $\{g_1, g_2, \ldots, g_n\} \subset C(M, Y)$ and that Y is a Hilbert space. We follow now the standard construction in the Chebyshev-approximation (cf e.g. [16], [17], [9]).

Let a be the continuous mapping from $M \times Y$ to R^n defined by $a(t, s) := ((g_0(t) - s) \bullet g_1(t), (g_0(t) - s) \bullet g_2(t), \ldots, (g_0(t) - s) \bullet g_n(t))$, $t \in M$, $s \in Y$. Then the assertion in Theorem 4.4 is equivalent to $\left\{\lambda \in \mathbb{R}^n \,\middle|\, \forall l \in L : l^T \lambda < 0\right\} = \emptyset$ where $L = a(\Omega_0)$ is compact as continuous image of a compact set. Hence, this is equivalent to $0 \in \operatorname{conv}(L)$. By Caratheodory's Theorem we get the existence of vectors $l^1, \ldots, l^r \in L$ and numbers $\rho_1, \ldots, \rho_r \geq 0$ with $\sum_{i=1}^r \rho_i = 1, r \leq n + 1$, such that $0 = \sum_{i=1}^r \rho_i l^i$. This implies the following result.

Theorem 4.11. *g_0 is a solution to the problem (P^H) if and only if there are $i = 1, 2, \ldots, r \leq n+1$ points $(t_i, s_i) \in \Omega_0$ such that $0 = \sum_{i=1}^r \rho_i (g_0(t_i) - s_i) \bullet g(t_i)$ for all $g \in V$. For all $i = 1, 2, 3, \ldots, r$ we have $\|g_0(t_i) - s_i\| = \varphi(g_0)$, which is the optimal value. Further, $s_i - g_0(t_i)$ and $g_0(t_i) - s_i$ belong to the normal cone of $\Sigma(t_i)$ if $g_0(t_i)$ is an element of $\Sigma(t_i)$ and $\Sigma(t_i)^c$, respectively.*

Example 4.3 [9, Example 1]**.** Now we want to illustrate the dual Kolmogorov conditions for our Example 4.1. We have to find $\rho_i > 0$, $(t_i, s_i) \in \Omega_0, i \leq 3$, such that $\Sigma_i \rho_i = 1$, $0 = \sum_{i=1}^r \rho_i (g_0(t_i) - s_i)$ and $0 = \sum_{i=1}^r \rho_i (g_0(t_i) - s_i) t_i$. It suffices to use the active points $\{(1, a), (1, b)\} \subset \Omega_0$. The resulting equality $0 = \rho_1((a + b)/2 - a) + \rho_2((a + b)/2 - b) = \rho_1(b - a)/2 + \rho_2(a - b)/2$ is solved by $\rho_1 = \rho_2 = 1/2$.

Again we have used the fact that in the irregular case the active points of Ω_0 which are defined by some $t \in M$, where $r_{\max}$ is attained, are enough to show optimality. The following statement assures this aspect. We recall that in case of a Hilbert space the Chebyshev center belongs to $\operatorname{cl} \operatorname{conv} S$.

Proposition 4.12. *Let S be a compact, convex subset of Hilbert space $Y = \mathbb{R}^n$. r is the optimal value and $y_0 \in S$ is the solution to the problem*

$H(y,S) \to \min$, *subject to* $y \in S$, *if and only if there are* $\rho_i > 0$, $s_i \in S(y_0, r)$, $i = 1, 2, \ldots, p \le n+1$ *such that* $\sum_{i=1}^{p} \rho_i = 1$ *and* $\sum_{i=1}^{p} \rho_i (y_0 - s_i) \bullet y = 0$ *for all* $y \in Y$.

In the case $\rho = D$ we get with the same constructions a similar statement regarding the three parts of the primal Kolmogorov conditions.

Theorem 4.13. (*case* $\varphi(g_0) \neq 0$) $g_0 \in V$ *is a solution of* (P^D) *if and only if there are* $i = 1, 2, \ldots, r \le n+1$ *points* $(t_i, s_i) \in \Omega_0$ *and* $\rho_i > 0$, *such that* $\sum_{i=1}^{r} \rho_i = 1$ *and* $0 = \sum_{i=1}^{r} \rho_i (g_0(t_i) - s_i) \bullet g(t_i)$ *for all* $g \in V$.

Theorem 4.14. (*case* $\varphi(g_0) = 0$) $g_0 \in V$ *is a solution of* (P^D) *if and only if there are* $i = 1, 2, \ldots, r \le n+1$ (t_i, n_i) *with elements* n_i *of the polar cone of* $\Sigma(t_i)$ *at* $g_0(t_i)$ *and* $\rho_i > 0$, *such that* $\sum_{i=1}^{r} \rho_i = 1$ *and* $0 = \sum_{i=1}^{r} \rho_i n_i \bullet g(t_i)$ *for all* $g \in V$.

In the irregular case $\varphi(g_0) = -r_{\min} < 0$ it suffices to apply the next Proposition.

Proposition 4.15. *Let* S *be a compact, convex subset of the Hilbert space* $Y = \mathbb{R}^n$. *Then* r *is the optimal value and* $y_0 \in \operatorname{int} S$ *is the solution to the problem* $D(y,S) \to \min$, *subject to* $y \in S$ *if and only if there are* $\rho_i > 0$, $s_i \in S(y_0, r)$, $i = 1, 2, \ldots, p \le n+1$ *such that* $\sum_{i=1}^{p} \rho_i = 1$ *and* $\sum_{i=1}^{p} \rho_i (y_0 - s_i) \bullet y = 0$ *for all* $y \in Y$.

Example 4.4. *The above example yields with the subset* $\{(0, a, 1), (0, a, -1)\} \subset \Omega_0$ $(t_1 = t_2 = 0)$ *and the multipliers* $\rho_1 = \rho_2 = 0.5$ *the identity* $\rho_1 k + \rho_1(-k) = 0$ *for arbitrary* $(h, k) \in \mathbb{R}^2$.

References

[1] J.P. Aubin and A. Cellina, Differential Inclusions, Springer, Berlin, Heidelberg, New York, Tokyo 1984.

[2] P.K. Belobrov, *K voprosu o chebyshevskom sentre mnozhestva*, Izvestija vysshich uchebnych zavedenij **38** (1) (1964), 3–9. (in Russian)

[3] B. Bank, J. Guddat, D. Klatte, K. Kummer, K. Tammer, Non-Linear Parametric Optimization, Akademie Verlag Berlin 1982.

[4] F.H. Clarke, Optimization and Nonsmooth Analysis, Canadian Mathematical Society Series of Monographs and Advanced Texts, A Wiley-Interscience Publication, John Wiley & Sons, New York 1983.

[5] V.F. Demjanov and A.M. Rubinov, Constructive Nonsmooth Analysis, Verlag Peter Lang, Frankfurt am Main 1995.

[6] A.L. Garkavi, *On the best net and the best cut of a set in a normed space*, Izv. Akad. Nauk SSSR, Ser. Mat. **26** (1962), 87–106. (in Russian)

[7] A.L. Garkavi, *On the Chebyshev center and the convex hull of a set*, Usp. Mat. Nauk **19** (6) (1964), 139–145, (120). (in Russian)

[8] A.L. Garkavi, *Minimax balayge theorem and an inscribed ball problem*, Matematicheskie Zametki **30** (1) (1981), 109–121. (in Russian)

[9] I. Ginchev and A. Hoffmann, *On the best approximation of set-valued functions*, in: P. Gritzmann, R. Horst, E. Sachs, R. Tichatschke (eds.), Recent Advances in Optimization (Proc. of the 8th French-German Conference on Optimization, Trier, July 21–26, 1996, Lect. Notes Econ. Math. Syst. 452, Springer, Berlin Heidelberg 1997, 61–74.

[10] P.M. Gruber, *The space of convex bodies*, in: P.M. Gruber, J.M. Wills (eds.), Handbook of Convex Geometry, Volume A, North-Holland, Amsterdam 1993, 301–318.

[11] R. Hettich and P. Zencke, Numerische Methoden der Approximation und Semi-Infiniten Optimierung, B.G. Teubner, Leipzig 1982.

[12] A. Hoffmann, *The distance to the intersection of two convex sets expressed by the distances to each of them*, Math. Nachr. **157** (1992), 81–98.

[13] J.B. Hiriart Urruty and C. Lemaréchal, Convex Analysis and Minimization Algorithms I, Springer Verlag, Berlin 1993.

[14] R.B. Holmes, Geometric Functional Analysis and its Applications, Graduate Texts in Mathematics 24, Springer, New York-Heidelberg-Berlin 1975.

[15] R. Horst and H. Tuy, Global Optimization, Deterministic Approaches, Springer, Berlin etc. 1990.

[16] P. Kosmol, Optimierung und Approximation, Walter de Gruyter, Berlin 1991.

[17] P.J. Laurent, Approximation et Optimisation, Enseignement des Sciences 13, Hermann, Paris 1972.

[18] K. Leichtweiss, Konvexe Mengen, Deutscher Verlag der Wissenschaften, Berlin 1980.

[19] L.E. Rybiński, Continuous Selections and Variational Systems, Monografie 61, Institute of Mathematics, Higher College of Engineering, Zielona Góra, Poland 1992, 100 pp, ISSN 0239-7390.

[20] R.T. Rockafellar and R.J.B. Wets, Variational Analysis, Grundlehren der mathematischen Wissenschaften 317, Springer, Berlin 1998.

[21] I. Singer, Best Approximation in Normed Linear Spaces by Elements of Linear Subspaces, Grundlehren der mathematischen Wissenschaften 171, Springer, Berlin 1970.

Received 25 January 2002