JOURNAL OF APPROXIMATION THEORY 2, 136–151 (1969)

Duality in Nonlinear Approximation*

WERNER KRABS

Institut für Angewandte Mathematik der Universität, Hamburg, Germany

1. INTRODUCTION

We are mainly concerned with the problem of characterizing the distance $d_x(W)$ between a fixed point x and a nonempty subset W of a real normed linear space X. We intend to develop a purely geometric concept by which $d_x(W)$ can be estimated from below and even obtained as the maximum of certain lower bounds.

In Section 2 we start with the consideration of families $\mathscr{R}$ of half spaces R in X such that

$$W \subseteq \bigcup_{R \in \mathscr{R}} R \qquad \text{and} \qquad x \notin \bigcup_{R \in \mathscr{R}} R.$$

By $\mathscr{H}$ we denote the family of corresponding hyperplanes H. It is geometrically evident that the infimum of all the distances $d_x(H)$ from x to $H \in \mathscr{H}$ is a lower bound for $d_x(W)$. This is proved as Lemma 2.2.

The main result of Section 2 is a duality theorem which states that $d_x(W)$ is the maximum of all such infima. This generalizes the well-known fact that, if W is convex, $d_x(W)$ is the maximum of all the distances $d_x(H)$ where H is a hyperplane separating x and W (Theorem 2.5).

In Section 3 we introduce supporting systems and strong supporting systems for W. The latter play the major role since they serve as an important tool in the characterization of projection points $\hat{w} \in W$, i.e., points $\hat{w}$ such that $\|\hat{w} - x\| = d_x(W)$.

A strong supporting system for W is a family $\mathscr{R}$ of half spaces R such that

$$W \subseteq \bigcup_{R \in \mathscr{R}} R \qquad \text{and} \qquad S = \bigcap_{H \in \mathscr{H}} \{H \cap W\}$$

is nonempty, where $\mathscr{H}$ is the family of corresponding hyperplanes. The elements of S are called supporting points. For instance, if W is convex, each supporting hyperplane defines a strong supporting system which consists of only one half space.

In Theorem 3.3 we obtain a well-known sufficient condition for a point $\hat{w} \in W$ to be a projection point. If W is convex then a restricted form of this

* This research was supported by the Air Force Office of Scientific Research under grant AF-AFOSR-937-67.

condition is also necessary. The formulation we give makes use of supporting systems and thus yields a purely geometric viewpoint.

We conclude Section 3 with the consideration of the following situation which generalizes several special cases: Let Y be another real normed linear space and A a nonempty open subset of Y. We consider a Fréchet-differentiable mapping $F: A \to X$ and put $W = F(A)$. We have investigated this case in [11] and we give a short review of the results at the end of Section 3.

At the beginning of Section 4 we present an algebraic version of Lemma 2.2. Then we consider the special case of the Chebychev approximation problem where X is the vector space $C(M)$ of all continuous, real valued functions on a compact Hausdorff space M with the maximum norm. The concept of H-sets in M, due to Collatz [2], can be formulated in terms of strong supporting systems. Furthermore, a result of Collatz concerning lower bounds for $d_x(W)$ and a similar one of Meinardus and Schwedt [12] turn out to be special cases of Lemma 2.2. The case $W = F(A)$ where A is a nonempty open subset of the real euclidean n-space and $F: A \to C(M)$ is a Fréchet-differentiable mapping has been investigated in [10], so that we content ourselves with a short review of the results. Finally, we treat the case of discrete L_p-approximation and give a simple method to verify the assumptions of Lemma 2.2 for the generalized rational approximation problem.

2. A Duality Theorem

We consider a normed linear space X over the reals and denote the norm by $\|\cdot\|$. Let X^* be the dual space of X, that is, the set of all continuous linear functionals L mapping X into the reals. X^* becomes a real Banach space if we define the norm by

$$\|L\| = \sup_{\|x\|=1} |L(x)|, \qquad L \in X^*.$$

By S^* we denote the unit sphere of X^*, that is, the set of all $L \in X^*$ such that $\|L\| = 1$. In the following, a half space R of X is always defined by a pair (L, α) with $L \in S^*$ and α real, so that

$$R = \{y \in X : L(y) \geqq \alpha\}. \tag{2.1}$$

We call

$$H = \{h \in X : L(h) = \alpha\}$$

the corresponding hyperplane.

Lemma 2.1 [3]. *For a given $L \in S^*$, let the half space R be defined by* (2.1) *and let H be the corresponding hyperplane. Then for each $x \notin R$, the distance*

$$d_x(H) = \inf_{h \in H} \|h - x\|$$

from x to H is given by

$$d_x(H) = L(h - x) = \alpha - L(x)$$

for all $h \in H$.

Proof. $x \notin R$ implies $x \notin H$. Since H is closed we have $d_x(H) > 0$. By assumption the closed ball

$$K = \{y \in X : \|y - x\| \leqq d_x(H)\}$$

is contained in the half space

$$\{y \in X : L(y) \leqq \alpha\}.$$

This implies

$$L(h) = \alpha \geqq \sup_{y \in K} L(y) = d_x(H) + L(x)$$

or

$$\alpha - L(x) = L(h - x) \geqq d_x(H) \qquad \text{for all } h \in H.$$

On the other hand we have for each $h \in H$

$$\alpha - L(x) = L(h - x) \leqq \|h - x\|,$$

and therefore

$$L(h - x) = \alpha - L(x) \leqq d_x(H).$$

This completes the proof.

LEMMA 2.2. *Let* $\mathscr{R}$ *be a family of half spaces and* $\mathscr{H}$ *the family of corresponding hyperplanes. For a nonempty subset W of X and for an* $x \in X$ *we assume*

$$W \subseteq \bigcup_{R \in \mathscr{R}} R \tag{2.2}$$

and

$$x \notin \bigcup_{R \in \mathscr{R}} R. \tag{2.3}$$

Then we have

$$\inf_{H \in \mathscr{H}} d_x(H) \leqq d_x(W) = \inf_{w \in W} \|w - x\|. \tag{2.4}$$

Proof. Every $R \in \mathscr{R}$ is given by (2.1) for some $L \in S^*$. Let $\mathscr{L}$ be the collection of all these L. (2.2) then implies that for each $w \in W$ there exists $H \in \mathscr{H}$ and $L \in \mathscr{L}$ such that

$$L(w) \geqq L(h) \qquad \text{for all } h \in H. \tag{2.5}$$

(2.3) implies, by Lemma 2.1, that for each $R \in \mathscr{R}$

$$d_x(H) = L(h - x) \qquad \text{for all } h \in H,$$

where H is the corresponding hyperplane and L the corresponding element of $\mathscr{L}$.

Suppose (2.4) is false. Then by the definition of $d_x(W)$ there is a $\hat{w} \in W$ such that

$$d_x(H) > \|\hat{w} - x\| \qquad \text{for all } H \in \mathscr{H}.$$

Hence for each $H \in \mathscr{H}$ and the corresponding $L \in \mathscr{L}$ we have

$$L(h - \hat{w}) = L(h - x) - L(\hat{w} - x) \geqq L(h - x) - \|\hat{w} - x\|$$
$$= d_x(H) - \|\hat{w} - x\| > 0 \qquad \text{for all } h \in H.$$

This contradicts (2.5). Therefore (2.4) must be true.

Now we assume W to be a nonempty convex subset of X and $x \in X$ to be a point not belonging to the closure of W. According to a well-known separation theorem [9] there is a half space R such that $W \subseteq R$ and $x \notin R$. By Lemma 2.2 we therefore have $d_x(H) \leqq d_x(W)$ where H is the corresponding hyperplane.

LEMMA 2.3. *In addition to the assumptions of Lemma* 2.2 *we require* $W \subseteq X$ *to be convex and*

$$d = \inf_{H \in \mathscr{H}} d_x(H) > 0.$$

Then there is a half space $\hat{R}$ *such that* $W \subseteq \hat{R}$, $x \notin \hat{R}$ *and*

$$d \leqq d_x(\hat{H}) \leqq d_x(W), \tag{2.6}$$

where $\hat{H}$ *is the corresponding hyperplane.*

Proof. By the above separation theorem [9] the closed ball

$$K_d = \{y \in X : \|y - x\| \leqq d\}$$

and W can be separated by a hyperplane

$$\hat{H} = \{h \in X : \hat{L}(h) = \hat{\alpha}\},$$

where $\hat{L} \in S^*$ and $\hat{\alpha}$ is a real scalar; i.e.,

$$W \subseteq \hat{R} = \{y \in X : \hat{L}(y) \geqq \hat{\alpha}\} \qquad \text{and} \qquad K_d \subseteq \{y \in X : \hat{L}(y) \leqq \hat{\alpha}\};$$

in particular, $x \notin \hat{R}$. This implies

$$\inf_{w \in W} \hat{L}(w) \geqq \hat{\alpha} \geqq \sup_{y \in K_d} \hat{L}(y) = d + \hat{L}(x),$$

and applying Lemma 2.1 we conclude

$$d_x(W) \geqq \inf_{w \in W} \hat{L}(w - x) \geqq \hat{\alpha} - L(x) = d_x(\hat{H}) \geqq d$$

which completes the proof.

LEMMA 2.4. *Let* $x \in X$ *and a nonempty subset* W *of* X *be given such that*

$$d_x(W) = \inf_{w \in W} \|w - x\| > 0.$$

Then there exists a family $\mathscr{R}$ of half spaces such that (2.2) *and* (2.3) *hold and*

$$d_x(H) = d_x(W) \tag{2.7}$$

for all H of the family $\mathscr{H}$ of corresponding hyperplanes.

Proof. Put $K = \{y \in X : \|y - x\| \leqq d_x(W)\}$. If we choose an arbitrary $w \in W$ and define

$$z = (1 - \lambda) x + \lambda w \tag{2.8}$$

where

$$\lambda = \frac{d_x(W)}{\|w - x\|},$$

then

$$\|z - x\| = \lambda \|w - x\| = d_x(W).$$

Furthermore, there exists a hyperplane supporting K at z [9], which is given by

$$H_z = \{h \in X : L_z(h) = L_z(z)\}$$

where $L_z \in S^*$. With no loss of generality we may assume

$$L_z(x) < L_z(z) = L_z(h) \qquad \text{for all } h \in H_z. \tag{2.9}$$

On the other hand, we have

$$L_z(w - z) = \frac{1 - \lambda}{\lambda} L_z(z - x) \geqq 0. \tag{2.10}$$

If we define

$$R_z = \{y \in X : L_z(y) \geqq L_z(z)\}$$

and denote by $\mathscr{R}$ the family of all such half spaces R_z where z is defined by (2.8) and w varies over W, then (2.2) and (2.3) are an immediate consequence of (2.9) and (2.10). Furthermore, we have

$$L_z(z) \geqq \sup_{y \in K} L_z(y) = d_x(W) + L_z(x),$$

whence by Lemma 2.1

$$d_x(H_z) = L_z(z - x) \geqq d_x(W).$$

On the other hand

$$d_x(H_z) = L_z(z - x) \leqq \|z - x\| = d_x(W).$$

This completes the proof.

Lemma 2.2 and 2.3 yield the following

DUALITY THEOREM. *If for $x \in X$ and a nonempty subset W of X we have $d_x(W) > 0$, then*

$$d_x(W) = \max_{\mathscr{R}} \inf_{H \in \mathscr{H}} d_x(H),$$

where the maximum is taken over all families $\mathscr{R}$ of half spaces satisfying (2.2) *and* (2.3), *and $\mathscr{H}$ is the family of corresponding hyperplanes.*

From now on we assume W to be convex. If for some $x \in X$ we have $d_x(W) > 0$, then by the Duality Theorem there is a family $\mathscr{R}$ of half spaces satisfying (2.2) and (2.3) such that

$$d_x(W) = \inf_{H \in \mathscr{H}} d_x(H) > 0,$$

where $\mathscr{H}$ is the corresponding family of hyperplanes. However, by Lemma 2.3 there exists a half space $\hat{R}$ such that $W \subseteq \hat{R}$, $x \notin \hat{R}$, and

$$d_x(\hat{H}) = \inf_{H \in \mathscr{H}} d_x(H)$$

where $\hat{H}$ is the corresponding hyperplane.

Using this result and Lemma 2.2 we get

THEOREM 2.5. *Let W be a nonempty convex subset of X and assume $d_x(W) > 0$ for some $x \in X$. Then*

$$d_x(W) = \max_R d_x(H)$$

where the maximum is taken over all the half spaces R such that $W \subseteq R$, $x \notin R$, and H is the corresponding hyperplane.

This result is well known (compare, for instance, [3], [5], [7], where equivalent results are obtained) and can now be considered as a special case of the above Duality Theorem.

3. Supporting Systems and a Sufficient Condition for Projection Points

As in Section 2, we start with a real normed linear space X and consider a family $\mathscr{R}$ of half spaces R defined by (2.1), where $L \in \mathscr{L}$, and $\mathscr{L}$ is the corresponding set of linear forms in S^*.

DEFINITION. $\mathscr{R}$ is called a supporting system for a nonempty subset W of X if (2.2) holds and if for all $H \in \mathscr{H}$ we have

$$H \cap W \neq \varnothing, \tag{3.1}$$

where $\mathscr{H}$ is the corresponding family of hyperplanes and $\varnothing$ denotes the empty set.

If $\mathscr{R}$ is a supporting system for the nonempty subset W of X and $x \in X$ is such that (2.3) holds then Lemma 2.2 yields a lower bound for the distance

$$d_x(W) = \inf_{w \in W} \|w - x\|$$

between x and W.

LEMMA 3.1. *If for $x \in X$ and for a nonempty subset W of X we have $d_x(W) > 0$, then there is a supporting system $\mathscr{R}$ for W such that* (2.3) *is satisfied.*

Proof. If in the proof of Lemma 2.4 we substitute the family $\mathscr{R} = \{R_z\}$ by $\hat{\mathscr{R}} = \{\hat{R}_z\}$, where

$$\hat{R}_z = \{y \in X : L_z(y) \geqq L_z(w)\}$$

and z is given by (2.8), then it is easy to verify that $\hat{\mathscr{R}}$ satisfies (2.2), (2.3), and (3.1).

DEFINITION. A supporting system $\mathscr{R}$ for the nonempty subset W of X is called a strong supporting system if

$$S = \bigcap_{H \in \mathscr{H}} \{H \cap W\} \neq \varnothing, \tag{3.2}$$

where $\mathscr{H}$ is the corresponding family of hyperplanes.

The elements of S are called supporting points.

Now let $\mathscr{R}$ be a strong supporting system for the nonempty subset W of X and let $\hat{w} \in W$ be an arbitrary, but fixed, supporting point. Then each $R \in \mathscr{R}$ is of the form

$$R = R_L = \{y \in X : L(y) \geqq L(\hat{w})\} \tag{3.3}$$

where $L \in \mathscr{L}$.

The condition (2.2) is therefore equivalent to the following: For each $w \in W$ there is an $L \in \mathscr{L}$ such that

$$L(w) \geqq L(\hat{w}). \tag{3.4}$$

The condition (2.3) is equivalent to

$$L(x) < L(\hat{w}) \tag{3.5}$$

for all $L \in \mathscr{L}$. By Lemmas 2.1 and 2.2 the inequalities (3.4) and (3.5) imply

$$\inf_{L \in \mathscr{L}} L(\hat{w} - x) \leqq d_x(W). \tag{3.6}$$

If we, furthermore, assume that $\mathscr{L}$ is a nonempty weakly* closed subset of S^*, hence weakly* compact, then (3.4) is equivalent to

$$\min_{L \in \mathscr{L}} L(\hat{w} - w) \leqq 0 \tag{3.7}$$

for all $w \in W$, and in (3.6) "inf" can be replaced by "min."

LEMMA 3.2. *Let $\mathscr{R}$ be a family of half spaces R_L defined by* (3.3), *where $L \in \mathscr{L}$, and $\mathscr{L}$ is a nonempty weakly* closed subset of S^*. Let $E(\mathscr{L})$ be the (nonempty* [9]) *set of extreme points of $\mathscr{L}$, and $\tilde{\mathscr{R}}$ the family of all $R_L \in \mathscr{R}$ such that the corresponding L is an element of $E(\mathscr{L})$. If $\mathscr{R}$ is a strong supporting system for a nonempty subset W of X then the same is true for $\tilde{\mathscr{R}}$.*

Proof (as in [8]). Let $\hat{w} \in W$ be an arbitrary supporting point. For each $w \in W$ we define a linear functional g_w mapping $\mathscr{L}$ into the reals by $g_w(L) = L(w - \hat{w})$ where $L \in \mathscr{L}$. As g_w is weakly* continuous, $g_w(\mathscr{L})$ is a compact subset of the reals and, therefore, has an extreme point $r \geqq 0$, since $\mathscr{R}$ is a strong supporting system. It is well known [9] that r is the image of an extreme point $L_e \in E(\mathscr{L})$. Hence, for each $w \in W$, there exists an $L_e \in E(\mathscr{L})$ such that $L_e(w - \hat{w}) \geqq 0$, which completes the proof.

Let $x \in X$ be a fixed point, and W a nonempty subset of X. $\hat{w} \in W$ is called a projection point of x in W if

$$\|\hat{w} - x\| = d_x(W).$$

In the following, we assume $d_x(W) > 0$. For each $w \in W$ we define a set:

$$E_w = \{L \in S^* : L(w - x) = \|w - x\|\}.$$

By the Hahn–Banach Theorem, E_w is nonempty, and, furthermore, E_w is obviously a weakly* closed (hence weakly* compact) convex subset of S^*.

THEOREM 3.3. *For some $\hat{w} \in W$, let $\hat{\mathscr{R}}$ be the family of half spaces R_L defined by* (3.3), *where $L \in \mathscr{L}$, and $\mathscr{L}$ is a nonempty subset of $E_{\hat{w}}$. If $\hat{\mathscr{R}}$ is a strong supporting system for W, then $\hat{w}$ is a projection point of x in W.*

Proof. By assumption, (3.4) and (3.5) are satisfied, the latter because of

$$0 < d_x(W) \leqq \|\hat{w} - x\| = L(\hat{w} - x) \qquad \text{for all } L \in \mathscr{L}.$$

Hence (3.6) holds, implying

$$\|\hat{w} - x\| = \inf_{L \in \mathscr{L}} L(\hat{w} - x) \leqq d_x(W).$$

If $\mathscr{L}$ is a nonempty weakly* closed subset of E_w, then $\mathscr{L}$ is also weakly* closed in S^*, and therefore the assumption of Theorem 3.3 is equivalent to (3.7). By Lemma 3.2, the assumption of Theorem 3.3 remains true if we replace $\mathscr{L}$ by the set $E(\mathscr{L})$ of its extreme points.

For applications it is important to know whether, in this case, $E(\mathscr{L})$ is contained in the set $E(K^*)$ of extreme points of the unit ball K^* of X^*, because in various special cases, $E(K^*)$ has a rather simple structure.

The statement of Theorem 3.3 is not new. In [*13*] Nikolskiĭ considers the case where $\mathscr{L}$ is the intersection of $E_{\hat{w}}$ and a so-called fundamental system Γ, which is a weakly* closed subset of K^* such that for each $y \neq 0$ in X there exists an $L \in \Gamma$ with $L(y) = \|y\|$. Examples of fundamental systems are the unit sphere S^* and the closure of $E(K^*)$.

Nikolskiĭ proves that $W \subseteq \bigcup_{L \in E(\mathscr{L})} R_L$ is a sufficient condition for $\hat{w} \in W$ to be a projection point of x in W, and a necessary condition in the case where W is convex.

In [*6*] Garkavi obtains the same result as Nikolskiĭ, with the only differences that instead of $E(\mathscr{L})$, the set $E(E_{\hat{w}}) = E_{\hat{w}} \cap E(K^*)$ of extreme points of $E_{\hat{w}}$ is considered and that X is a Banach space. Recently, Deutsch and Maserick [*3*] reproved this result for a normed linear space X.

In [*7*], Havinson gives the same characterization for projection points in convex sets as Garkavi. Furthermore, he obtains the following criterion which is a simple consequence of Theorem 2.5: If W is a nonempty convex subset of X, and $x \in X$ is such that $d_x(W) > 0$, then $\hat{w} \in W$ is a projection point of x in W if and only if there exists an element $L \in S^*$ such that $L \in E_{\hat{w}}$ and $L(\hat{w}) \leq L(w)$ for all $w \in W$.

However, as Deutsch and Maserick point out in [*3*], this L cannot, in general, be chosen to be an element of $E(K^*)$.

Brosowski considers the case $\mathscr{L} = E_{\hat{w}} \cap \Gamma$, where Γ is a fundamental system, and studies the question: For what nonempty subsets W of X other than convex subsets is the condition $W \subseteq \bigcup_{L \in E(\mathscr{L})} R_L$ necessary for $\hat{w}$ to be a projection point of x in W? He states that the condition is necessary for so-called Γ-regular subsets of X. For details we refer to [*1*], where the results are given without proofs. These are to appear in a forthcoming paper.

In [*11*] we have investigated the following situation which occurs in various special cases: Let Y be a real normed linear space, A a nonempty open subset of Y, and $F: A \to X$ a mapping such that for each $a \in A$ the Fréchet derivative F_a' exists. For W we take the image $F(A)$, and we consider an element $x \in X$ such that

$$d_x(W) = \inf_{a \in A} \|F(a) - x\| > 0.$$

We then obtain the following necessary condition for a projection point; we assume that for every fixed $h \in Y$ the mapping $a \to F_a'(h)$, $a \in A$, is continuous. If $F(\hat{a})$, $\hat{a} \in A$, is a projection point of x in W, then for each $h \in Y$ there exists an $L \in E_{F(\hat{a})}$ such that

$$L(F_{\hat{a}}'(h)) \leqslant 0. \tag{3.8}$$

This result has also been given by Henze in [*8*], however, without the above continuity assumption on the mappings $a \to F_a'(h)$. But this is indispensable.

If we define for each $a \in A$ the linear manifold

$$T_a = \{F(a) - F_a'(h) : h \in Y\}$$

and consider the family $\mathscr{R}$ of all half spaces

$$R_L = \{y \in X : L(y) \geqslant L(F(\hat{a}))\}, \qquad L \in E_{F(\hat{a})},$$

then condition (3.8) is equivalent to $\mathscr{R}$ being a strong supporting system for $T_{\hat{a}}$, with $F(\hat{a})$ as supporting point.

Furthermore, Henze shows in [*8*] that condition (3.8) remains true if we take the set $E(E_{F(\hat{a})})$ of extreme points of $E_{F(\hat{a})}$ instead of $E_{F(\hat{a})}$. This is also an immediate consequence of Lemma 3.2.

In order to prove that condition (3.8) with $E(E_{F(\hat{a})})$ instead of $E_{F(\hat{a})}$ is sufficient for $F(\hat{a})$ to be a projection point of x in $F(A)$, we assumed in [*11*] that F has the following property: For each pair $(a,b) \in A \times A$, there is a positive continuous functional $\phi_{a,b} : E(K^*) \to \Re$ and an element $h = h(a,b) \in Y$ such that

$$L(F(a) - F(b)) = \phi_{a,b}(L) \cdot L(F_a'(h)). \tag{3.9}$$

In the case of the Chebychev approximation problem (compare Section 4) this property is essentially equivalent to the asymptotic convexity of F introduced by Meinardus and Schwedt in [*12*]. This was shown in [*10*]. We therefore call F asymptotically convex if it has the property (3.9).

Under the assumptions that F is asymptotically convex and that for each fixed $h \in Y$ the mapping $a \to F_a'(h)$, $a \in A$, is continuous, we have shown in [*11*] that for $F(\hat{a})$, $\hat{a} \in A$, to be a projection point of x in $W = F(A)$, the following condition is necessary: For each $a \in A$ there is an $L \in E(E_{F(\hat{a})})$ such that

$$L(F(a)) \geqslant L(F(\hat{a})). \tag{3.10}$$

Finally, we give a somewhat negative result which is also contained in [*11*]. We assume X to be flat convex [*9*], that is, at each point of the unit sphere of X there is exactly one supporting hyperplane of the unit ball. Examples of flat convex normed linear spaces are Hilbert spaces and L_p-spaces with $1 < p < \infty$.

If, furthermore, $F : A \to X$ is asymptotically convex, if for each fixed $h \in Y$ the mapping $a \to F_a'(h)$, $a \in A$, is continuous, and if for each $x \in X$ there exists a projection point in $W = F(A)$, then W is a linear manifold.

4. Special Cases and Examples

We start with an algebraic version of Lemma 2.2. Let X be a real normed linear space, X^* its dual, and $\mathscr{L}$ a nonempty subset of the unit sphere S^* of X^*. To each $L \in \mathscr{L}$ we assign a real number α_L. Then we have the following:

Let W be a nonempty subset of X, and let x be an arbitrary point of X. If for each $w \in W$ there is an $L \in \mathscr{L}$ such that

$$L(w) \geqslant \alpha_L, \tag{4.1}$$

then

$$\inf_{L \in \mathscr{L}} \{\alpha_L - L(x)\} \leqslant d_x(W). \tag{4.2}$$

If $L(x) \geqslant \alpha_L$ for some $L \in \mathscr{L}$, the assertion (4.2) is trivial. We therefore assume

$$L(x) < \alpha_L \qquad \text{for all } L \in \mathscr{L}. \tag{4.3}$$

We put for each $L \in \mathscr{L}$

$$R_L = \{y \in X : L(y) \geqslant \alpha_L\} \tag{4.4}$$

and define

$$\mathscr{R} = \{R_L : L \in \mathscr{L}\}.$$

Then (4.1) and (4.3) are equivalent to (2.2) and (2.3), thus implying (4.2) by Lemmas 2.1 and 2.2.

Now we consider special cases.

(a) Uniform approximation: Let X be the vector space $C(M)$ of real valued continuous functions defined on a compact Hausdorff space M. The norm in $X = C(M)$ will be the maximum norm

$$\|g\| = \max_{P \in M} |g(P)|, \qquad g \in C(M).$$

Let W be a nonempty subset of $C(M)$ and let f be a point of $C(M)$ not belonging to the closure of W.

The problem of finding projection points $\hat{w} \in W$ of f in W is the well-known nonlinear Chebychev approximation problem.

For $X = C(M)$, the set of extreme points $E(K^*)$ of the unit ball K^* of X^*, is given by

$$E(K^*) = \{\epsilon_P \delta_P : P \in M, \epsilon_P = +1 \text{ or } -1\},$$

where δ_P is the point measure in P, i.e.,

$$\delta_P(g) = g(P) \qquad \text{for all } g \in C(M).$$

Let D be a nonempty subset of M. To each $P \in D$ we assign a number $\epsilon_P \in \{-1, +1\}$ and define $\mathscr{L}$ by

$$\mathscr{L} = \{\epsilon_P \delta_P : P \in D\}. \tag{4.5}$$

A family $\mathscr{R}$ of half spaces R_L of the form (4.4) with $L \in \mathscr{L}$ is then given by

$$\mathscr{R} = \{R_P : P \in D\} \tag{4.6}$$

where

$$R_P = \{g \in C(M) : \epsilon_P g(P) \geqslant \alpha_P\} \tag{4.7}$$

and α_P is a real scalar assigned to P.

Collatz [2] calls D an H-set if D is the disjoint union of two nonempty sets D_1 and D_2 such that for no pair $w, \hat{w} \in W$ it is true that

$$w(P) - \hat{w}(P) \begin{cases} <0 & \text{for all } P \in D_1 \\ >0 & \text{for all } P \in D_2. \end{cases} \tag{4.8}$$

Assume D to be an H-set and consider an arbitrary but fixed $\hat{w} \in W$. Then (4.8) implies that for each $w \in W$ there is a $P \in D$ with

$$\epsilon_P(w(p) - \hat{w}(P)) \geqslant 0$$

where

$$\epsilon_P = \begin{cases} +1 & \text{for } P \in D_1 \\ -1 & \text{for } P \in D_2. \end{cases}$$

Defining $\alpha_P = \epsilon_P \hat{w}(P)$, $P \in D$, and $\mathscr{R}$ by (4.6), (4.7), we get the result that $\mathscr{R}$ is a strong supporting system for w with $\hat{w}$ as supporting point.

Besides (4.8), we assume that for some fixed $\hat{w} \in W$, we have

$$\epsilon_P(\hat{w}(P) - f(P)) > 0 \qquad \text{for all } P \in D.$$

Then it follows that

$$\inf_{P \in D} |\hat{w}(P) - f(P)| \leqslant d_f(W).$$

This is the contents of Theorem 2 of [2] and a special case of the fact that (3.4) and (3.5) imply (3.6). In [2], Collatz gives various examples of H-sets and develops a method by which H-sets can be systematically constructed for certain subsets W of $C(M)$.

Let M be a finite closed interval $[a,b]$. Then, for example, in the case of rational or exponential approximation, W has the following property: There is a number r such that no difference $w - \hat{w}$ of functions $w, \hat{w} \in W$ has more than r zeros in $[a,b]$. In this case, obviously, each set of $r+2$ points $P_i \in [a,b]$, with $a \leqslant P_1 < \ldots < P_{r+2} \leqslant b$, is an H-set.

Now we consider the following situation: Let Y be a normed linear space and A a nonempty subset of Y. Let $F: A \to C(M)$ be a given map, and put $W = F(A)$. We require D to be a nonempty closed subset of M and assume that for some $\hat{a} \in A$ the following two conditions are satisfied:

$$\min_{P \in D} (F(\hat{a},P) - f(P))(F(\hat{a},P) - F(a,P)) \leqslant 0$$

for all $a \in A$, and

$$|F(\hat{a},P) - f(P)| > 0 \qquad \text{for all } P \in D.$$

We define $\epsilon_P = \operatorname{sgn}(F(\hat{a},P) - f(P))$, where $P \in D$, and $\mathscr{S}$ by (4.5). Then (3.4) and (3.5) are satisfied, implying (3.6) with "min" instead of "inf", since

$\mathscr{L}$ is a weakly* closed subset of K^* and hence weakly* compact. (3.6), in turn, is equivalent to

$$\min_{P \in D} |F(\hat{a}, P) - f(P)| \leqslant d_f(W).$$

This is exactly Theorem 1 of [*12*], for the case of real valued functions.

In [*10*] we considered the case where Y is the real euclidean space $\mathfrak{R}^n$ and A is a nonempty open subset of Y. If F is Fréchet-differentiable, then for each $P \in M$ and $a = (a_1, \ldots, a_n) \in A$, there exist the partial derivatives

$$\frac{\partial F}{\partial a_j}(a, P), \qquad j = 1, \ldots, n,$$

and we have

$$F_a'(h) = \sum_{j=1}^{n} h_j \frac{\partial F}{\partial a_j}(a),$$

where $h = (h_1, \ldots, h_n) \in Y = \mathfrak{R}^n$.

We have shown in [*10*] that for each fixed $h \in Y$, the mapping $a \to F_a'(h)$, $a \in A$, is continuous if and only if the partial derivatives depend continuously on $(a, P) \in A \times M$.

Condition (3.9) immediately leads to the following property of F (compare condition (5) in [*10*]): For each pair $(a, b) \in A \times A$, there is a positive function $\phi(a, b) \in C(M)$ and an element $h = h(a, b) \in \mathfrak{R}^n$ such that

$$F(a) - F(b) = \phi(a, b) \sum_{j=1}^{n} h_j \frac{\partial F}{\partial a_j}(a).$$

As to the relationship with the asymptotic convexity of F, introduced by Meinardus and Schwedt in [*12*], and the discussion of further special cases, we refer to [*10*].

(b) Discrete L_p-approximation: Let X be the m-dimensional space $\mathfrak{R}^m$ with norms

$$\|y\|_p = \left(\sum_{i=1}^{m} |y_i|^p \right)^{1/p}$$

for $1 \leqslant p < \infty$, and

$$\|y\|_\infty = \max_{i=1, \ldots, m} |y_i|,$$

where $y = (y_1, \ldots, y_m) \in \mathfrak{R}^m$.

X^* can be identified with $\mathfrak{R}^m$, and each element $L \in X^*$ is given by

$$L(y) = \langle l, y \rangle = \sum_{i=1}^{m} l_i y_i, \tag{4.9}$$

where $l = (l_1, \ldots, l_m) \in \mathfrak{R}^m$ is uniquely defined by L. The norm of X^* is given by

$$\|L\| = \sup_{\|y\|_p=1} |L(y)| = \|l\|_q,$$

where $(1/p) + (1/q) = 1$ for $1 < p < \infty$, $q = 1$ if $p = \infty$, and $q = \infty$ if $p = 1$.

In order to apply Lemma 2.2, we consider a nonempty subset $W \subseteq \mathfrak{R}^m$, an element $x \in \mathfrak{R}^m$, and a nonempty subset $\mathscr{L}$ of

$$S_q = \{l \in \mathfrak{R}^m : \|l\|_q = 1\}.$$

To each $l \in \mathscr{L}$ we assign a real number α_l, and define $L \in X^*$ by (4.9). Then $L \in S^*$, and (4.1) is equivalent to the following statement: For each $w \in W$, there is an $l \in \mathscr{L}$ such that

$$\langle l, w \rangle \geqslant \alpha_l. \tag{4.10}$$

By (4.2), we then have

$$\inf_{l \in \mathscr{L}} \{\alpha_l - \langle l, x \rangle\} \leqslant d_x{}^p(W) = \inf_{w \in W} \|w - x\|_p.$$

A very simple way of realizing (4.10) is the following: Let $\mathscr{L}$ consist of m vectors of the form $l^i = \epsilon_i e^i$, where $e^i = (e_1{}^i, \ldots, e_m{}^i)$, $e_j{}^i = \delta_{ij}$, and $\epsilon_i = +1$ or -1. Then, obviously, $\|l^i\|_q = 1$ for every q, $1 \leqslant q \leqslant \infty$. Putting $\alpha_i = \alpha_{l^i}$, we can express condition (4.10) by

$$\min_{i=1,\ldots,m} \{\alpha_i - \epsilon_i w_i\} \leqslant 0 \tag{4.11}$$

for all $w = (w_1, \ldots, w_m) \in W$.

Finally, we demonstrate in the case of rational approximation, how (4.11) can be realized:

Let U and V be subspaces of $X = \mathfrak{R}^m$, spanned by $u^0, \ldots, u^r$ and $v^0, \ldots, v^s$, respectively, where $r + s + 2 \leqslant m$. We assume

$$V^+ = \{v \in V : v^i > 0,\ i = 1, \ldots, m\}$$

to be nonempty, and put

$$W = \left\{ \frac{u}{v} : u \in U,\ v \in V^+ \right\}.$$

(4.11) is equivalent to the following statement: There is no vector $(a_0, \ldots, a_r, b_0, \ldots, b_s) \in \mathfrak{R}^{r+s+2}$ such that

$$\left.\begin{aligned} &\sum_{k=0}^{s} \alpha_i v_i{}^k b_k - \sum_{j=0}^{r} \epsilon_i u_i{}^j a_j > 0 \\ &\sum_{k=0}^{s} v_i{}^k b_k > 0 \end{aligned}\right\} \text{for } i = 1, \ldots, m.$$

By Theorem 2.9 of [*4*], this is equivalent to the existence of a vector $(y_1,\ldots, y_m, p_1,\ldots, p_m) \in \mathfrak{R}^{2m}$ such that

$$\left.\begin{array}{ll} y_i \geqslant 0 \quad \text{and} \quad p_i \geqslant 0, & i = 1,\ldots,m, \\ \sum_{i=1}^{m} v_i^k \alpha_i y_i + \sum_{i=1}^{m} v_i^k p_i = 0, & k = 0,\ldots,s, \\ \sum_{i=1}^{m} u_i^j \epsilon_i y_i \qquad\qquad = 0, & j = 0,\ldots,r. \end{array}\right\} \tag{4.12}$$

We put $c_i = -\epsilon_i y_i$ and $\lambda_i = \alpha_i - \epsilon_i x_i$, for $i = 1,\ldots,m$. Then (4.11) is equivalent to the existence of vectors $(c_1,\ldots,c_m)$ and $(p_1,\ldots, p_m)$ of $\mathfrak{R}^m$ such that

$$\left.\begin{array}{ll} p_i \geqslant 0 & \text{for } i = 1,\ldots,m, \\ \sum_{i=1}^{m} u_i^j c_i = 0, & j = 0,\ldots,r, \\ \sum_{i=1}^{m} v_i^k x_i c_i = \sum_{i=1}^{m} v_i^k(\lambda_i |c_i| + p_i), & k = 0,\ldots,s. \end{array}\right\} \tag{4.13}$$

Since V^+ is assumed to be nonempty, it can easily be shown that not all c_i can vanish. If (4.13) is satisfied, we have by the definition of the λ_i's,

$$\min_{i=1,\ldots,m} \lambda_i \leqslant d_x^p(W).$$

Under the natural assumption that the matrix

$$\begin{pmatrix} u_i^j \\ v_i^k x_i \end{pmatrix}$$

has the rank $r + s + 2$, it is easy to satisfy (4.13). One merely has to choose $z_i \geqslant 0$, $i = 1,\ldots,m$, and compute a nontrivial solution $(c_1,\ldots,c_m)$ of

$$\sum_{i=1}^{m} u_i^j c_i = 0, \qquad j = 0,\ldots,r,$$

$$\sum_{i=1}^{m} v_i^k x_i c_i = \sum_{i=1}^{m} v_i^k z_i, \qquad k = 0,\ldots,s,$$

which is always possible if not all the z_i vanish. We define

$$\lambda_i = \begin{cases} \dfrac{z_i}{|c_i|} & \text{if } c_i \neq 0, \\ \max\limits_{c_i \neq 0} \dfrac{z_i}{|c_i|} & \text{if } c_i = 0, \end{cases}$$

and

$$p_i = \begin{cases} 0 & \text{if } c_i \neq 0, \\ z_i & \text{if } c_i = 0. \end{cases}$$

Then (4.13) is satisfied, and we have

$$\min_{c_i \neq o} \frac{z_i}{|c_i|} \leqslant d_x^{\,p}(W).$$

REFERENCES

1. B. BROSOWSKI, Einige Bemerkungen zum verallgemeinerten Kolmogoroffschen Kriterium. Max-Planck-Institut für Physik und Astrophysik. MPI-PEA/Astro 7/68.

2. L. COLLATZ, Inclusion theorems for the minimal distance in rational Tschebyscheff approximation with several variables. In: "Approximation of Functions," (H. L. Garabedian, Ed.). Elsevier, Amsterdam-London-New York, 1965, 43–56.

3. F. R. DEUTSCH AND P. H. MASERICK, Applications of the Hahn–Banach theorem in approximation theory. *SIAM Review* **9** (1967), 516–530.

4. D. GALE, "Linear Economic Models." McGraw-Hill Book Company, New York-Toronto-London, 1960.

5. A. L. GARKAVI, Duality theorems for approximation by elements of convex sets (Russian). *Uspehi Mat. Nauk* **16** (1961), 141–145.

6. A. L. GARKAVI, On a criterion for an element of best approximation (Russian). *Sibirsk Mat. Z.* **5** (1964), 472–476.

7. S. J. HAVINSON, Approximation by elements of convex sets. *Soviet Math. Dokl.* **8** (1967), 98–101.

8. D. HENZE, Über die Menge der Minimallösungen bei linearen und nichtlinearen Approximationsproblemen. Schriften des Rheinisch-Westfälischen Instituts für instrumentelle Mathematik an der Universität Bonn (Herausgeber: E. Peschl, H. Unger), Serie A, Nr. 17, Köln-Opladen: Westdeutscher Verlag.

9. G. KÖTHE, "Topologische Lineare Räume I." Springer, Berlin-Göttingen-Heidelberg, 1960.

10. W. KRABS, Über differenzierbare asymptotisch konvexe Funktionenfamilien bei der nicht-linearen gleichmäßigen Approximation. *Arch. Rat. Mech. Anal.* **27** (1967), 275–288.

11. W. KRABS, Asymptotische Konvexität bei der Approximation in normierten Vektorräumen. *Math. Z.* **108** (1969), 368–376.

12. G. MEINARDUS AND D. SCHWEDT, Nichtlineare Approximationen. *Arch. Rat. Mech. Anal.* **17** (1964), 297–326.

13. V. N. NIKOLSKIĬ, Best approximation by elements of convex sets in normed linear spaces (Russian). *Kalinin. Gos. Ped. Inst. Ucen. Pap.* **29** (1963), 85–119.