THE THEORY OF BEST APPROXIMATION AND FUNCTIONAL ANALYSIS

IVAN SINGER

THE THEORY OF BEST APPROXIMATION AND FUNCTIONAL ANALYSIS

IVAN SINGER
Institute of Mathematics
Bucharest

SOCIETY for INDUSTRIAL and APPLIED MATHEMATICS
PHILADELPHIA, PENNSYLVANIA 1910[illegible]

Printed for the Society for Industrial and Applied Mathematics by
J. W. Arrowsmith Ltd., Bristol 3, England

Preface

In this monograph we present some results and problems in the modern theory of best approximation, i.e., in which the methods of functional analysis are applied in a consequent manner. This modern theory constitutes both a unified foundation for the classical theory of best approximation (which treats the problems with the methods of the theory of functions) and a powerful tool for obtaining new results. Within the general framework of normed linear spaces the problem of best approximation amounts to the minimization of a distance, which permits us to use geometric intuition (but rigorous analytic proofs), and the connections of the phenomena become clearer and the arguments simpler than those of the classical theory of best approximation in the various particular concrete function spaces. We hope that this has been proved convincingly enough in the monograph [168] (which was the first of this kind in the literature) and in the lecture notes [175], and will be proved again in the present monograph (see, for example, § 1, the remarks made after Theorem 1.8, or § 3, the remark to Theorem 3.5).

Naturally, the interaction between the theory of best approximation and functional analysis works also in the other direction, for example, problems of best approximation in normed linear spaces have led to the discovery of the theorem on extremal extension of extremal functionals, of the concrete representations for the extremal points of the unit cell in certain conjugate spaces, of new results on semi-continuity of set-valued mappings, etc. However, we shall not consider this side of their interaction in the present monograph.

Since June 1966, when the Romanian version of the monograph [168] went to print, the theory of best approximation in normed linear spaces has developed rapidly. However, except for the expository paper, covering the period up to 1967, of A. L. Garkavi [67], the only attempt for a comprehensive survey material was made in [175]. The latter constitutes also the basis for the present monograph, which, though self-contained, may be regarded as an up-to-date complement to [168] and [175] ; we have essentially conserved the structure of [175], but improved the parts which overlap with [175] and added much new material, part of which appeared after [175] went to print (the bibliography of [175] contained only 99 items).

We note also the appearance, since 1970, of the lecture notes of A. L. Brown [34], P. D. Morris [132] and R. B. Holmes [82] and of the book of P. J. Laurent [118], which treat some topics in the theory of best approximation in normed linear spaces and optimization in locally convex spaces [118], [82] ; let us also mention the lecture notes of F. Deutsch (*Theory of Approximation in Normed Linear Spaces*, Fall 1972), which we received when the present work was completed. We take the liberty of suggesting that it would be more appropriate to replace "Approximation" by "Best Approximation" in the titles of all these works with the exception of [82]; indeed, since these works (except for part of [118]) are actually concerned with the theory of best approximation, which is only a part of approximation theory, their present titles might mislead the reader to expect more topics to be covered. Let us note that the use of the word "best" in the titles of [168], [67], [175], [82] and of the present monograph, which serves also for the delimitation of their scope (namely, to emphasize that some other important parts of approximation theory are not considered), is not a lack of modesty, but a universally accepted classical term for approximation problems related to nearest points; this is the term which is used also in the texts of the abovementioned works of Brown, Morris, Laurent and Deutsch.

In order to be able to enter more deeply into the problems, without increasing the size of the present monograph, we have restricted ourselves to consider only five basic topics in the theory of best approximation, shown in the titles of the five sections. Thus, various problems of best approximation treated in [168] (for example, the distance functionals, metric projections onto sequences of subspaces, deviations of sets from linear subspaces, finite-dimensional diameters and secants, Chebyshev centers, etc.) and not treated in [168] (for example, methods of computation of elements of best approximation, applications to extremal problems of the theory of analytic functions, etc.) have been deliberately omitted. On the other hand, [168] contained only a few isolated results on metric projections (for example, we did not include in [168] some of the results of [166], [122], [32]), since their theory was only beginning to develop at that time, but in [175] and in the present monograph the largest section is devoted to metric projections (see § 4). Also, [168] contained only a short appendix on best approximation by elements of nonlinear sets, since this topic was beyond the scope of [168], but in the present monograph these problems, although treated briefly, occupy the second largest section (see § 5).

Furthermore, in order to limit the size of the present monograph, we have included here only a few short proofs (some of them being very simple and some, which are less elementary, being particularly interesting), but for all results we give references. Also, we do not present all known applications of the general theory in concrete spaces, but only some examples of some more important ones. Finally, we tried to reduce the overlapping with [168] to the very minimum necessary for self-containedness and thus we often refer to [168] for complementary results and bibliography. In the bibliography of the present monograph we wanted to emphasize those works which have appeared after [168] went to print and some earlier papers which have been omitted from [168]; therefore, sometimes we give here some results (and their authors' names) only with reference to the bibliography of [168]. In exchange for these limitations, we have tried to give the up-to-date stand and literature of the problems treated herein.

We hope that the present monograph will be useful for a large circle of readers, including those who are not specialists in these problems and those specialists who work in the theory of best approximation with the classical function-theoretic methods or using the methods of functional analysis. Also, for specialists in functional analysis, in particular, in the geometry of normed linear spaces, this monograph may offer a new field of applications.

The reader is assumed to know some elements of functional analysis and integration theory, but we recall (giving also a reference to a treatise), whenever necessary, the notions and results which we use.

We acknowledge with pleasure that we benefited from attending the seminar lectures of Dr. G. Godini (at the Institute of Mathematics of the Academy, Bucharest, in 1971–1973) and from our visits at the universities of Bonn (1971, 1972) and Grenoble (1972), where we had useful discussions with Dr. W. Pollul and Professors J. Blatter and P.-J. Laurent on some problems treated in this monograph. Finally, we wish to express our thanks to Professor R. S. Varga for the invitation to write this monograph and to present it at the National Science Foundation Regional Conference on Theory of Best Approximation and Functional Analysis at Kent State University, June 11–June 15.

IVAN SINGER

Bucharest

June 1, 1973

Contents

The Theory of Best Approximation and Functional Analysis

Ivan Singer

1. Characterizations of elements of best approximation.

(a) Throughout the sequel, without any special mention, we shall denote by ρ the distance in a metric space E and, in particular, if E is a normed linear space, ρ will denote the distance in E induced by the norm, i.e.,

$$\rho(x, y) = \|x - y\|, \qquad x, y \in E. \tag{1.1}$$

DEFINITION 1.1. Let E be a metric space, G a set in E and $x \in E$. An element $g_0 \in G$ is called an *element of best approximation* of x (by the elements of the set G) if we have

$$\rho(x, g_0) = \inf_{g \in G} \rho(x, g), \tag{1.2}$$

i.e., if g_0 is "nearest" to x among the elements of G; we shall denote by $\mathscr{P}_G(x)$ the set of all such elements g_0, i.e.,

$$\mathscr{P}_G(x) = \{g_0 \in G \mid \rho(x, g_0) = \inf_{g \in G} \rho(x, g)\}. \tag{1.3}$$

It is natural to consider first the problem of characterization of elements of best approximation, i.e., the problem of finding necessary and sufficient conditions in order that $g_0 \in \mathscr{P}_G(x)$, since these results will be applied to solve the other problems on best approximation (for example, those of existence and uniqueness of elements of best approximation, etc.). Also, the characterization theorems in concrete spaces (see, for example, the "alternation theorem" 1.13 below) are convenient tools for verifying whether or not a given g_0 satisfies $g_0 \in \mathscr{P}_G(x)$, since they are easier to use than (1.2).

Since we have obviously

$$\mathscr{P}_G(x) = \begin{cases} x & \text{for } x \in G, \\ \varnothing & \text{for } x \in \overline{G} \setminus G, \end{cases} \tag{1.4}$$

it will be sufficient to characterize the element of best approximation of the elements $x \in E \setminus \overline{G}$. In order to exclude the case when such elements do not exist, in the sequel we shall assume, without any special mention, that $\overline{G} \neq E$.

Unless otherwise stated, the field of scalars for all (general or concrete) normed linear spaces considered in the sequel can be either the field of complex numbers or the field of real numbers.

(b) The first main theorem of characterization of elements of best approximation by elements of linear subspaces in normed linear spaces is the following (see [168, p. 18]):

THEOREM 1.1. *Let E be a normed linear space, G a linear subspace of E, $x \in E \setminus \bar{G}$ and $g_0 \in G$. We have $g_0 \in \mathcal{P}_G(x)$ if and only if there exists an $f \in E^*$ such that*

$$\|f\| = 1, \tag{1.5}$$

$$f(g) = 0, \qquad g \in G, \tag{1.6}$$

$$f(x - g_0) = \|x - g_0\|. \tag{1.7}$$

We recall that E^* denotes the conjugate space of E, i.e., the space of all continuous linear functionals on E, endowed with the usual vector operations and with the norm $\|f\| = \sup_{\substack{x \in E \\ \|x\| \leqq 1}} |f(x)|$.

To prove Theorem 1.1, assume that $g_0 \in \mathcal{P}_G(x)$. Then, since $x \in E \setminus \bar{G}$, we have $\rho(x, G) = \|x - g_0\| > 0$ and hence, by a corollary of the Hahn–Banach theorem (see, for example, [55, p. 64, Lemma 12]), there exists an $f_0 \in E^*$ such that $\|f_0\| = 1/\|x - g_0\|$, $f_0(g) = 0$ $(g \in G)$, and $f_0(x) = 1$. Then the functional $f = \|x - g_0\| f_0 \in E^*$ satisfies (1.5)–(1.7). Conversely, if there is an $f \in E^*$ satisfying (1.5)–(1.7), then for any $g \in G$ we have

$$\|x - g_0\| = |f(x - g_0)| = |f(x - g)| \leqq \|f\| \, \|x - g\| = \|x - g\|$$

and hence $g_0 \in \mathcal{P}_G(x)$, which completes the proof.

It is easy to see that Theorem 1.1 admits the following *geometrical interpretation*: *We have $g_0 \in \mathcal{P}_G(x)$ if and only if there exists a closed hyperplane G' in E* (i.e., a closed linear subspace G' such that $\dim E/G' = 1$) *containing G, which supports the cell* $S(x, \|x - g_0\|) = \{y \in E \mid \|y - g_0\| \leqq \|x - g_0\|\}$ (i.e., $\rho(G', S(x, \|x - g_0\|)) = 0$ and $G' \cap \operatorname{Int} S(x, \|x - g_0\|) = \emptyset$).

Any functional $f \in E^*$ satisfying (1.5) and (1.7) is called a "maximal functional" of the element $x - g_0$ (because we have $\|x - g_0\| = \sup_{\substack{h \in E^* \\ \|h\| = 1}} |h(x - g_0)|$). The usefulness of Theorem 1.1 for applications in various concrete normed linear spaces is due to the fact that for these spaces the general form of maximal functionals of the elements of the space is well known and simple (see, for example, [154], [198]).

We recall that an element x of a closed convex set A in a topological linear space L is called an *extremal point* of A if the relations $y, z \in A$, $0 < \lambda < 1$, $x = \lambda y + (1 - \lambda)z$ imply $y = z = x$. The second main theorem of characterization of elements of best approximation by elements of linear subspaces in normed linear spaces is the following (see [168, p. 62]).

THEOREM 1.2. *Let E be a normed linear space, G a linear subspace of E, $x \in E \setminus \bar{G}$ and $g_0 \in G$. We have $g_0 \in \mathcal{P}_G(x)$ if and only if for every $g \in G$ there exists an extremal point f^g of the unit cell $S_{E^*} = \{f \in E^* \mid \|f\| \leqq 1\}$ such that*

$$\operatorname{Re} f^g(g - g_0) \leqq 0, \tag{1.8}$$

$$f^g(x - g_0) = \|x - g_0\|. \tag{1.9}$$

For a geometrical interpretation of Theorem 1.2 see [168, p. 75]. Theorem 1.2 is also convenient for applications in the usual concrete spaces because for these spaces the general form of the extremal points of S_{E^*} is well known and simple.

It is easy to see that the sufficiency part of Theorem 1.2 remains valid for an arbitrary set G in E. Indeed, if the condition is satisfied, then for every $g \in G$ we have

$$\begin{aligned}\|x - g_0\| &= \operatorname{Re} f^g(x - g_0) \leqq \operatorname{Re} f^g(x - g) \leqq |f^g(x - g)| \\ &\leqq \|f^g\|\,\|x - g\| = \|x - g\|,\end{aligned}$$

whence $g_0 \in \mathscr{P}_G(x)$, which proves the assertion. The problem of characterizing those sets $G \subset E$ for which the condition in Theorem 1.2 is also necessary, is important in nonlinear approximation (see § 5).

Since best approximation amounts, by definition, to the minimization of the convex functional $\varphi = \varphi_x$, on the linear subspace G of a normed linear space E, where

$$\varphi(y) = \|x - y\|, \qquad y \in E, \tag{1.10}$$

there naturally arises the problem of obtaining characterizations of the elements of best approximation $g_0 \in \mathscr{P}_G(x)$ with the aid of differential calculus. The main difficulty is that in general the norm in E is not necessarily Gâteaux differentiable at each nonzero point of E. Nevertheless, it is known (see, for example, [55, p. 445, Lemma 1]) that the limits

$$\tau(x, y) = \lim_{t \to 0^+} \frac{\|x + ty\| - \|x\|}{t}, \qquad x, y \in E, \tag{1.11}$$

always exist and one can use them to give the following characterizations of elements of best approximation (see [168, pp. 88–90]).

THEOREM 1.3. *Let E be a normed linear space, G a linear subspace of E, $x \in E \setminus \overline{G}$ and $g_0 \in G$. We have $g_0 \in \mathscr{P}_G(x)$ if and only if*

$$\tau(x - g_0, g) \geqq 0, \qquad g \in G. \tag{1.12}$$

If the norm in E is Gâteaux differentiable at $x - g_0$, this condition is equivalent to the following:

$$\tau(x - g_0, g) = 0, \qquad g \in G. \tag{1.13}$$

Let us observe that Theorem 1.1 can be also expressed as follows: $g_0 \in \mathscr{P}_G(x)$ if and only if $0 \in \{f|_G \in G^* \mid f \in S_{E^*}, f(x - g_0) = \|x - g_0\|\}$. We have also the following characterization theorem [171].

THEOREM 1.4. *Let E be a normed linear space, G a linear subspace of E, $x \in E \setminus \overline{G}$ and $g_0 \in G$. We have $g_0 \in \mathscr{P}_G(x)$ if and only if 0 belongs to the $\sigma(G^*, G)$-closure of the* *convex hull of the set*

$$\Lambda = \{\overline{f(x - g_0)}f|_G \in G^* | f \in \mathscr{E}(S_{E^*}), |f(x - g_0)| = \|x - g_0\|\}. \tag{1.14}$$

Theorems 1.1–1.4 above can be also expressed, more concisely, as formulae for $\mathscr{P}_G(x)$, namely as

$$\mathscr{P}_G(x) = \{g_0 \in G | \|x - g_0\| = \max_{\substack{f \in G^\perp \\ \|f\| = 1}} |f(x - g_0)|\}, \tag{1.15}$$

$$\mathscr{P}_G(x) = \{g_0 \in G | \min_{\substack{f \in \mathscr{E}(S_{E^*}) \\ f(x - g_0) = \|x - g_0\|}} \operatorname{Re} f(g - g_0) \leqq 0 \, (g \in G)\}, \tag{1.16}$$

$$\mathscr{P}_G(x) = \{g_0 \in G | \tau(x - g_0, g) \geqq 0 \, (g \in G)\}, \tag{1.17}$$

$$\mathscr{P}_G(x) = \{g_0 \in G | 0 \in \widetilde{\text{co}}\, \Lambda\}, \tag{1.18}$$

respectively, where we use the notation

$$G^\perp = \{f \in E^* | f(g) = 0 \, (g \in G)\}, \tag{1.19}$$

$\mathscr{E}(S_{E^*})$ for the set of all extremal points of S_{E^*}, and $\widetilde{\text{co}}$ for the weak* closed convex hull, and where max, min stand for such sup, inf which are attained. Note also the obvious relations

$$\mathscr{P}_G(x) = G \cap S(x, \rho(x, G)) = G \cap \operatorname{Fr} S(x, \rho(x, G)), \tag{1.20}$$

which are useful for geometrical interpretations. Finally, let us give a characterization theorem (which we shall express concisely as formulae for $\mathscr{P}_\Gamma(f)$ and $\mathscr{P}_G(x)$), which will be useful in the study of existence and uniqueness problems (§§ 2–3) and of the set-valued metric projection operators (§ 4).

THEOREM 1.5. *Let E be a normed linear space.*

(i) *If Γ is a $\sigma(E^*, E)$-closed linear subspace of E^* and $f \in E^* \setminus \Gamma$, then*

$$\mathscr{P}_\Gamma(f) = f - \{h \in E^* | h|_{\Gamma_\perp} = f|_{\Gamma_\perp}, \|h\| = \|f|_{\Gamma_\perp}\|\}, \tag{1.21}$$

where

$$\Gamma_\perp = \{x \in E | \gamma(x) = 0 \, (\gamma \in \Gamma)\}. \tag{1.22}$$

(ii) *If G is a closed linear subspace of E and $x \in E \setminus G$, then*

$$\mathscr{P}_G(x) = x - \{y \in E | \kappa(y)|_{G^\perp} = \kappa(x)|_{G^\perp}, \|y\| = \|\kappa(x)|_{G^\perp}\|\}, \tag{1.23}$$

where κ denotes the canonical isometrical embedding of E into E^{**}, that is,

$$(\kappa(x))(f) = f(x), \qquad x \in E, f \in E^*. \tag{1.24}$$

Part (i) has been given in [174, formula (25)′]. Here is a slightly more direct proof, using an argument of R. R. Phelps [145] : If $\gamma_0 \in \mathscr{P}_\Gamma(f)$, then for $h = f - \gamma_0$, the element $f - h$ is in the right-hand side of (1.21), since by the Hahn-Banach theorem

$$\|f|_{\Gamma_\perp}\| = \min_{\gamma \in (\Gamma_\perp)^\perp} \|f - \gamma\| = \min_{\gamma \in \Gamma} \|f - \gamma\|$$

(we have $(\Gamma_\perp)^\perp = \Gamma$, since Γ is $\sigma(E^*, E)$-closed). Conversely, if $\gamma_0 = f - h$ is in the right-hand side of (1.21), then $\gamma_0 = f - h \in (\Gamma_\perp)^\perp = \Gamma$ and, by the preceding formula for $\|f|_{\Gamma_\perp}\|$, we have $\gamma_0 \in \mathscr{P}_\Gamma(f)$. To see part (ii), let $g_0 \in \mathscr{P}_G(x)$. Then for $y = x - g_0$, the element $x - y$ is in the right-hand side of (1.23), by virtue of (1.15). Conversely, if $g_0 = x - y$ is in the right-hand side of (1.23), then, since G is closed, we have $g_0 = x - y \in (G^\perp)_\perp = \bar{G} = G$ and hence, again by (1.15), $g_0 \in \mathscr{P}_G(x)$, which proves part (ii). Note that Theorem 1.5 implies that

$$\kappa(\mathscr{P}_G(x)) \subset \kappa(x) - \{\Phi \in E^{**} | \quad \Phi|_{G^\perp} = \kappa(x)|_{G^\perp}, \|\Phi\| = \|\kappa(x)|_{G^\perp}\|\} = \mathscr{P}_{G^{\perp\perp}}(\kappa(x)). \tag{1.25}$$

For a characterization of elements of best approximation in terms of fixed points of a set-valued mapping, see § 5, Proposition 5.1.

(c) Let us give now examples of applications of some of the above theorems in concrete spaces.

We shall use the word "compact" in the sense of N. Bourbaki, i.e., bicompact Hausdorff. For a compact space Q we shall denote by $C(Q)$, respectively $C_R(Q)$, the space of all complex or real, respectively of all real, continuous functions on Q, endowed with the usual vector operations and with the norm $\|x\| = \max_{q \in Q}|x(q)|$. Using the general form of maximal functionals of the elements of $C_R(Q)$, from Theorem 1.1 we obtain (see [168, p. 33])

THEOREM 1.6. *Let $E = C_R(Q)$ (Q compact), G a linear subspace of E, $x \in E \setminus \bar{G}$ and $g_0 \in G$. We have $g_0 \in \mathscr{P}_G(x)$ if and only if there exist two disjoint closed subsets $Y_{g_0}^+$, $Y_{g_0}^-$ of Q and a Radon measure μ on Q, such that*

$$|\mu|(Q) = 1, \tag{1.26}$$

$$\int_Q g(q)\, d\mu(q) = 0, \qquad g \in G, \tag{1.27}$$

$$\mu \geqq 0 \text{ on } Y_{g_0}^+, \quad \mu \leqq 0 \text{ on } Y_{g_0}^- \text{ and } Y_{g_0}^+ \cup Y_{g_0}^- \supset S(\mu), \tag{1.28}$$

$$x(q) - g_0(q) = \begin{cases} \|x - g_0\| & \text{for} \quad q \in Y_{g_0}^+, \\ -\|x - g_0\| & \text{for} \quad q \in Y_{g_0}^-, \end{cases} \tag{1.29}$$

where $S(\mu)$ denotes the carrier of the measure μ.

One can also give a characterization theorem in the spaces $E = C(Q)$ (see 168, p. 29]). Theorem 1.6, which appeared in [165], has constituted the first theorem of characterization in $E = C_R(Q)$ (even in $E = C_R([a, b])$) of elements of best approximation by elements of linear subspaces G *of arbitrary dimension*. From Theorem 1.6 one obtains easily, for example, the following result of E. W. Cheney and A. A. Goldstein (see [168, p. 44]) : *Let L be a real topological linear space, Q a compact subset of L, $E = C_R(Q)$, $G = L^*|_Q$, $x \in E \setminus G$ and $g_0 = \Phi_0|_Q \in G$, where $\Phi_0 \in L^*$. We have $g_0 \in \mathscr{P}_G(x)$ if and only if $0 \in$* co *B, where*

$$B = \{q \in Q | x(q) - g_0(q) = \|x - g_0\|\} \cup - \{q \in Q | x(q) - g_0(q) = -\|x - g_0\|\},$$

and where co *stands stands for "convex hull".*

For a positive measure space (T, ν) (we shall not specify the σ-field of subsets of T on which the measure ν is defined; this will cause no confusion) and for $1 \leqq p < \infty$ (respectively, $p = \infty$) we shall denote by $L^p(T, \nu)$ the space of all equivalence classes of functions with ν-integrable pth power (respectively of ν-measurable and ν-essentially bounded functions on T), endowed with the usual vector operations and with the norm $\|x\| = (\int_T |x(t)|^p \, d\nu(t))^{1/p}$ (respectively, $\|x\| = \operatorname{ess\,sup}_{t \in T} |x(t)|$); for simplicity, we use here the same notation for a function and for its equivalence class in L^p. Again the subscript R will mean, both here and for the spaces occurring in the sequel, that we restrict ourselves to real scalars. For a function x' on T we shall use the notation

$$Z(x') = \{t \in T \mid x'(t) = 0\}. \tag{1.30}$$

Using the general form of maximal functionals of the elements of $L^1(T, \nu)$, we obtain from Theorem 1.1 the following theorem (see [168, p. 46]), which was obtained initially by B. R. Kripke and T. J. Rivlin [110] with different (function-theoretic) methods and with the above functional analytic methods in [167].

THEOREM 1.7. *Let $E = L^1(T, \nu)$ (where (T, ν) is a positive measure space), G a linear subspace of E, $x \in E \setminus \bar{G}$ and $g_0 \in G$. We have $g_0 \in \mathscr{P}_G(x)$ if and only if*

$$\left| \int_{T \setminus Z(x - g_0)} g(t) \operatorname{sign} [x(t) - g_0(t)] \, d\nu(t) \right| \leqq \int_{Z(x - g_0)} |g(t)| \, d\nu(t), \qquad g \in G. \tag{1.31}$$

We recall that for a complex number $\alpha \neq 0$, by definition $\operatorname{sign} \alpha = e^{-i \arg \alpha} = \bar{\alpha}/|\alpha|$ and that $\operatorname{sign} 0 = 0$.

For $E = L^p(T, \nu)$ with $1 < p < \infty$ and for an abstract inner product space H we obtain from Theorem 1.1 the following well-known results (see [168, pp. 56–57]).

THEOREM 1.8. (i) *Let $E = L^p(T, \nu)$ (with (T, ν) a positive measure space and $1 < p < \infty$), G a linear subspace of E, $x \in E \setminus \bar{G}$ and $g_0 \in G$. We have $g_0 \in \mathscr{P}_G(x)$ if and only if*

$$\int_T g(t) |x(t) - g_0(t)|^{p-1} \operatorname{sign} [x(t) - g_0(t)] \, d\nu(t) = 0, \qquad g \in G. \tag{1.32}$$

(ii) *Let $E = H$ be an inner product space, G a linear subspace of E, $x \in E \setminus \bar{G}$ and $g_0 \in G$. We have $g_0 \in \mathscr{P}_G(x)$ if and only if*

$$(g, x - g_0) = 0, \qquad g \in G, \tag{1.33}$$

where (x, y) denotes the scalar product in $E = H$.

Some applications of Theorem 1.1 in other concrete spaces are given in [168].

The above results in concrete spaces illustrate the power of the methods of functional analysis in the theory of best approximation. Indeed, although Theorem 1.1 in general normed linear spaces is obtained by a simple application of a corollary of the Hahn-Banach theorem, it gives in various concrete spaces Theorems 1.6–1.8 and other results as particular cases, simply by using the general form of maximal functionals in these spaces. A direct proof of Theorems 1.6–1.8 would require different methods in each of the concrete spaces involved, apparently having no connection with each other; however, they all turn out to be particular cases of

Theorem 1.1, and this unified method of obtaining them is simpler and clearer than the separate proofs in each concrete space. This unified method is carried out in a consequent manner for the whole theory of best approximation (e.g., for problems of uniqueness, etc.), in the monograph [168]. In the sequel we shall only indicate some examples of applications in concrete spaces (rather than mention all known applications) of some of the general results on best approximation in arbitrary normed linear spaces which we shall give. Let us give now an application of Theorem 1.2.

It is well known (see, for example, [55, p. 441, Lemma 6]) that if $E = C(Q)$, a functional $f \in E^*$ is an extremal point of S_{E^*} if and only if there exist a $q \in Q$ and a scalar α with $|\alpha| = 1$ such that

$$f(y) = \alpha y(q), \qquad y \in E = C(Q). \tag{1.34}$$

From Theorem 1.2, using the general form (1.34) of the extremal points of S_{E^*} for $E = C(Q)$, one obtains immediately the following classical theorem of characterization of elements of best approximation, due to A. N. Kolmogorov (see [168, p. 69]).

THEOREM 1.9. *Let $E = C(Q)$ (Q compact), G a linear subspace of E, $x \in E \setminus \bar{G}$ and $g_0 \in G$. We have $g_0 \in \mathscr{P}_G(x)$ if and only if for every $g \in G$ there exists a $q = q^g \in Q$ such that*

$$\operatorname{Re} \overline{[x(q) - g_0(q)]} g(q) \geqq 0, \tag{1.35}$$

$$|x(q) - g_0(q)| = \|x - g_0\|. \tag{1.36}$$

For applications of Theorem 1.2 in other concrete spaces, see [168, pp. 75–87]. In $E = C(Q)$, using (1.34), one obtains from Theorem 1.4 a theorem of Y. Ikebe [94].

(d) Now we shall consider the important particular case when $\dim G = n < \infty$, i.e., when $G = [x_1, \cdots, x_n]$ = the (closed) linear subspace of E spanned by n linearly independent elements $x_1, \ldots, x_n$. Naturally, the preceding results are also applicable in this particular case; however, by using effectively the assumption $\dim G = n < \infty$, we can obtain additional information. For example, using the classical theorem of Minkowski that the elements of the unit cell

$$S_{E^*} = \{f \in E^* \mid \|f\| \leqq 1\}$$

in a finite-dimensional space E^* can be expressed as finite convex combinations of extremal points of S_{E^*} and that if G is an arbitrary linear subspace of a normed linear space E, then every extremal point of the unit cell S_{G^*} can be extended to an extremal point of S_{E^*} (see, for example, [168, p. 168]), from Theorem 1.1 we obtain (see [168, p. 170]):

THEOREM 1.10. *Let E be a normed linear space, $G = [x_1, \cdots, x_n]$ an n-dimensional linear subspace of E, $x \in E \setminus G$ and $g_0 \in G$. We have $g_0 \in \mathscr{P}_G(x)$ if and only if there exist h extremal points $f_1, \cdots, f_h$ of the unit cell S_{E^*}, where $1 \leqq h \leqq n + 1$ if the scalars are real and $1 \leqq h \leqq 2n + 1$ if the scalars are complex and h numbers*

$\lambda_1, \cdots, \lambda_h > 0$ *with* $\sum_{j=1}^{h} \lambda_j = 1$, *such that*

$$\sum_{j=1}^{h} \lambda_j f_j(g) = 0, \qquad g \in G, \tag{1.37}$$

$$\sum_{j=1}^{h} \lambda_j f_j(x - g_0) = \|x - g_0\|. \tag{1.38}$$

In other words, the additional information to Theorem 1.1 which we obtain for $\dim G = n < \infty$ is that for such subspaces G one can take the functional f of Theorem 1.1 to be a convex combination $f = \sum_{j=1}^{h} \lambda_j f_j$ of $h \leqq n + 1$ (respectively $h \leqq 2n + 1$) extremal points of the unit cell $S_{E^*} = \{f \in E^* \mid \|f\| \leqq 1\}$.

Using the mapping $\psi \to \{\psi(x_i)\}_1^n$ of G^* onto the n-space, from Theorem 1.4 one obtains [171]:

THEOREM 1.11. *Let E be a normed linear space, $G = [x_1, \cdots, x_n]$ an n-dimensional linear subspace of E, $x \in E \setminus G$ and $g_0 \in G$. We have $g_0 \in \mathscr{P}_G(x)$ if and only if 0 belongs to the convex hull of the following set in the n-dimensional Euclidean space*:

$$B_1 = \{\{\overline{f(x - g_0)} f(x_1), \cdots, \overline{f(x - g_0)} f(x_n)\} \mid f \in \mathscr{E}(S_{E^*}), |f(x - g_0)| = \|x - g_0\|\}. \tag{1.39}$$

(e) Let us give some applications in concrete spaces.

We recall that a system of n functions $x_1, \cdots, x_n \in C(Q)$ (Q compact) is called a *Chebyshev system* (on Q) if every linear combination $\sum_{i=1}^{n} \alpha_i x_i \not\equiv 0$ has at most $n - 1$ zeros on Q. We have the following characterization theorem of E. Ya. Remez (see [168, p. 182]).

THEOREM 1.12. *Let $G = [x_1, \cdots, x_n]$ be an n-dimensional linear subspace of $E = C_R(Q)$ such that $x_1, \cdots, x_n$ form a Chebyshev system and let $x \in E \setminus G$ and $g_0 \in G$. We have $g_0 \in \mathscr{P}_G(x)$ if and only if there exist $n + 1$ distinct points $q_1, \cdots, q_{n+1} \in Q$ such that*

$$x(q_j) - g_0(q_j) = \left(\operatorname{sign} \frac{\Delta_j}{\Delta}\right) \|x - g_0\|, \qquad j = 1, \cdots, n + 1, \tag{1.40}$$

where

$$\Delta = \begin{vmatrix} x_1(q_1) & \cdots & x_1(q_{n+1}) \\ \cdots & \cdots & \cdots \\ x_n(q_1) & \cdots & x_n(q_{n+1}) \\ x(q_1) & \cdots & x(q_{n+1}) \end{vmatrix} \neq 0, \tag{1.41}$$

and where Δ_j is the cofactor of the element $x(q_j)$ in Δ.

This theorem follows as a particular case both from Theorem 1.6 for $\dim G = n < \infty$ and from Theorem 1.10 for $E = C_R(Q)$, using the general form (1.34) of the extremal points of S_{E^*} (naturally, since the scalars are real, we have now $\alpha = \pm 1$ in (1.34)). In particular, for $Q = [a, b]$ one obtains the following classical "alternation theorem" of P. L. Chebyshev–S. N. Bernstein (see [168, p. 184]).

THEOREM 1.13. *Let* $G = [x_1, \cdots, x_n]$ *be an n-dimensional linear subspace of* $E = C_R([a, b])$ *such that* $x_1, \cdots, x_n$ *form a Chebyshev system and let* $x \in E \setminus G$ *and* $g_0 \in G$. *We have* $g_0 \in \mathscr{P}_G(x)$ *if and only if there exist* $n + 1$ *points* $q_1 < q_2 < \cdots < q_{n+1}$ *of* $[a, b]$, *at which the difference* $x(q) - g_0(q)$ *takes the value* $\|x - g_0\|$ *with alternating signs (i.e., with opposite signs at consecutive points* $q_j, q_{j+1}, j = 1, \cdots, n$).

In $E = C(Q)$, taking into account (1.34), one obtains from Theorem 1.11 a result of E. W. Cheney [38, p. 73]; the necessity part of this latter result was observed, essentially, by T. J. Rivlin and H. S. Shapiro (see [168, p. 181]).

Some other characterizations of elements of best approximation by finite-dimensional subspaces in various concrete spaces are given in [168, pp. 201–206].

(f) A natural generalization of the problem of characterization of elements of best approximation is the problem of *simultaneous characterization* of a set of elements of best approximation: given E, $G \subset E$ and x as above and a subset M of G, what are the necessary and sufficient conditions in order that *every* element $g_0 \in M$ be an element of best approximation of x by the elements of G? The answer is given (see [168, p. 23]) by

THEOREM 1.14. *Let E be a normed linear space, G a linear subspace of E,* $x \in E \setminus \bar{G}$ *and* $M \subset G$. *We have* $M \subset \mathscr{P}_G(x)$ *if and only if there exists an* $f \in E^*$ *satisfying* (1.5), (1.6) *and*

$$f(x - g_0) = \|x - g_0\|, \qquad g_0 \in M. \tag{1.42}$$

In other words, this says that one can find the functional $f \in E^*$ of Theorem 1.1 to be common for all $g_0 \in M$. Theorem 1.14 is an immediate consequence of Theorem 1.1 and the observation that $\|x - g_1\| = \|x - g_2\|$ for all pairs $g_1, g_2 \in M \subset \mathscr{P}_G(x)$; naturally, the converse is also true, since Theorem 1.1 is even a particular case of Theorem 1.14. We shall see in § 3 that Theorem 1.14 has applications in the study of the uniqueness of elements of best approximation.

(g) Finally, let us mention two other tools for the study of problems of best approximation in normed linear spaces, which will be useful in the sequel. We mention them here since they also yield the "characterization theorems" (1.43) and (1.45) below.

Firstly, the canonical mapping $\omega_G : E \to E/G$ is clearly related to best approximation, since in a normed linear space E formula (1.3) can also be written in the form

$$\mathscr{P}_G(x) = \{g_0 \in G \mid \|x - g_0\| = \|\omega_G(x)\|\}. \tag{1.43}$$

Another useful tool in the study of best approximation will be the set

$$\mathscr{P}_G^{-1}(0) = \{x \in E \mid 0 \in \mathscr{P}_G(x)\}. \tag{1.44}$$

Some authors call the set $\mathscr{P}_G^{-1}(0)$ the *metric complement* of G (because of § 2, Proposition 2.1, § 3, Proposition 3.1 and § 4, Proposition 4.1).

For this set we have, obviously,

$$\mathscr{P}_G(x) = (x + \mathscr{P}_G^{-1}(0)) \cap G. \tag{1.45}$$

Some other properties of the set $\mathscr{P}_G^{-1}(0)$ are collected in