UNIVERZITET U BEOGRADU

prirodno-matematički fakultet

ILIJA B. LAZAREVIĆ

PRIMENA KONUSA U TEORIJI OPTIMIZACIJE

DOKTORSKA DISERTACIJA

BEOGRAD
1975

In Mathematics, every formula
awaits its time: to be used so-
mewhere or to be - forgotten.

THE APPLICATION OF CONES IN OPTIMIZATION THEORY

-Doctoral thesis-

Ilija R. Lazarević
University of Belgrade

CONTENTS

P R E F A C E

The application of the cone method to the Optimization problems seems more efficient than others, as for example the dual, variation methods and so on. LUSTERNIK's result [1934], based on FRECHET differentiability, was for a long time the sole result of that kind. Later, on the basis of differentiability in the GATEAUX sense, the cone theory develepy to a greater extend. But the applications of the cone methcd was still limited by certain shortcomings. Recently, (for example HOFFMANN & KOLUMBAN [1974]) the cone theory entered the third stuge of development, which, it seems, has the greatest possibilities.

The present thesis is devcted to certain questions of cone applications. The largest part of this manuscript has been written during my study stay in Roumania (1973/74) in "Institutul de calcul" of Cluj. I like this opportunity to thank Professors ELENA POPOVICIU and TIBERIU POPOVICIU for many useful suggestions and help which they extended to me during my stay of the institut.

Since in literature there is quite a confusion about the names of many notions, in the first chapter some definitions and notions which will to used or which are enable to provide better connection of topics are given. In the second and third chapters, besides the basic cone theory and its application some results, as for example those related to tangential properties, which we consider as a contribution of the paper, are given. Some of these results were obtained in co-operation with Professor I.KOLUMBAN (Cluj), and presented at 5th Balkan Mathmatical Congress, Belgrade 1974, (see LAZAREVIĆ & KOLUMBAN [1974]).

As far as the application is concerned, in chapter IV, some where the cone application may be seen, are presented. Some of these results were communicated at International Colloquium of Constructive and Functional Analisis, Cluj 1973., as well as at the Colloquium of Functional Equations, Iasi 1973., (see LAZAREVIĆ [1974]).

Mrs. NADA OBRADOVIĆ was kind enough to read this manuscript from the point of linguistic.

CHAPTER I

P R E L I M I N A R I E S

1. NOTATIONS

We list below some of the abbreviations and symbols to be employed throughout the text. Although not all used right away it is convenient to have them collected together for an easy reference. Notation of less frequent usage will be introduced as the need arises.

We write:

$\mathbb{R}^n$ - for real Euclidean n-space; $\mathbb{R}^1 = \mathbb{R}$;

θ_X - for the zero vector in linear space X; 0 for the zero vector in $\mathbb{R}$;

$X^\#$ - for the algebraic dual of a linear space X ;

X^* - for the continuous dual of a topological linear space X;

span(A) - for the linear hull of a set A;

co(A) - for the convex hull of a set A;

int(A) or $\mathring{A}$ - for the interior of a set A;

adh(A) or $\bar{A}$ - for the closure of a set A;

C[X]- for the space of continuous scalar-valued functions on a compact (HAUSDORFF) space X as well as for the linear space of this functions; C[[a,b]] = C[a,b];

:= - for "equals by definition";

Y ⊂ X - for " Y is a subspace of X" ;

( ) - for a open interval;

[ ] - for a closed interval;

⟨ ⟩ - for the inner product in an Inner product space;

$K(A; x_0)$ - for the admissible cone of a set A;

$K[A; x_0]$ - for the adherent cone of a set A.

□ - for the end of the proof.

On some other definitions and notions see Subject index.

2. THE NOTIONS OF TOTALLY POSITIVE KERNELS

Now consider a real-valued function, frequently called k e r n e l, $K(s,t)$ of two real variables ranging over linearly ordered sets S and T respectively. Given such a kernel $K(s,t)$, $(s,t) \in S \times T$, we can form "minors" of arbitrary order p by selections of $s_1 < \ldots < s_p$ from S and of $t_1 < \ldots < t_p$ from T. We denote the resulting matrix by $\|K(s_i,t_j)\|_{i,j=1}^{p}$ and its determinant by

$$(2.1) \quad K\begin{pmatrix} s_1,\ldots,s_p \\ t_1,\ldots,t_p \end{pmatrix} := \begin{vmatrix} K(s_1,t_1) & K(s_1,t_2) & \ldots & K(s_1,t_p) \\ K(s_2,t_1) & K(s_2,t_2) & & K(s_2,t_p) \\ \vdots & & & \\ K(s_p,t_1) & K(s_p,t_2) & & K(s_p,t_p) \end{vmatrix} .$$

The kernel $K(s,t)$ is said to be t o t a l l y p o s i t i v e o f o r d e r r (abbreviated TP_r) if for each $p=1,\ldots,r$ and for all choices

$$(2.2) \qquad s_1 < \ldots < s_p \; ; \quad t_1 < \ldots < t_p \; , \quad (s_i,t_j) \in S \times T$$

we have the inequalities

$$(2.3) \qquad K\begin{pmatrix} s_1,\ldots,s_p \\ t_1,\ldots,t_p \end{pmatrix} \geqq 0 .$$

Except where there is an explicit statement to the contrary, these conditions will be assumed to hold throughout for the determinants (2.3): that is the rows and columns corresponding to s_i and t_j, respectively, are ordered by increasing values.

If the determinants in (2.3) are always positive for all choices possible of (2.2), then we say that $K(s,t)$ is s t r i c t l y t o t a l l y p o s i t i v e o f o r d e r r (abbreviated STP_r).

Most often S and T are either intervals of the real line or a countable set of discrete values along the real line, such as the set of all integers of the set of non-negative integers. When either S or T is a set of integers we may use the term s e q u e n c e rather than function. If both S and T are finite sets then K can be considered as a matrix in which case we speak of TP- m a t r i c e s, and STP- m a t r i c e s.

Observe that although $K(s,t)$ is a function on $S \times T$ the p^{th} - order determinant (2.1) can be viewed as a function whose domain of definition is a set of ordered p-tuples from the open simplex

$$(2.4)\qquad \begin{aligned} \Delta_p(S) &:= \{ s:=(s_1,\dots,s_p) \mid s_i \in S, \ s_1 < \dots < s_p \} \\ \Delta_p(T) &:= \{ t:=(t_1,\dots,t_p) \mid t_j \in T, \ t_1 < \dots < t_p \}. \end{aligned}$$

If m is the maximum number of orders which we consider in our problem and $K(s,t)$ totally positive (strictly totally positive) kernel of order p for every $p=1,\dots,m$, we say $K(s,t)$ is t o t a l l y p o s i t i v e (s t r i c t l y t o t a l l y p o s i t i v e)(abbreviated TP and STP respectively).

A more general concept than total positivity is that of sign regularity.

We say that $K(s,t)$ is s i g n- c o n s i s t e n t o f o r d e r r (SC_r) if all r^{th} - order minors of $K(s,t)$ have the same sign, i.e.,

$$(2.5)\qquad \varepsilon_r K\begin{pmatrix} s_1,\dots,s_r \\ t_1,\dots,t_r \end{pmatrix} \geqq 0$$

for all choices $(s_1,\dots,s_r) \in \Delta_r(S)$ and all choices $(t_1,\dots,t_r) \in \Delta_r(T)$, whenever ε_r equals to +1 or -1. If in (2.5) strict inequality holds, then we say that $K(s,t)$ is s t r i c t l y s i g n- c o n s i s t e n t o f o r d e r r (SSC_r). A kernel $K(s,t)$ is called s i g n - r e g u l a r o f o r d e r r (SR_r) if it is sign-consistent of every order from 1 through r, i.e., if there exists a sequence of constants $\varepsilon_1,\dots,\varepsilon_r$ each equal to +1 or -1, such that

$$(2.6) \qquad \varepsilon_p K\begin{pmatrix} s_1,\ldots,s_p \\ t_1,\ldots,t_p \end{pmatrix} \geq 0$$

for all choices $(s_1,\ldots,s_p) \in \Delta_p(S)$ and for all choices $(t_1,\ldots,t_p) \in \Delta_p(T)$ as well as $p=1,\ldots,r$. We say that $K(s,t)$ is s t i c t l y s i g n - r e g u l a r o f o r d e r r (SSR_r) if strict inequality prevails in every one of the above r inequalities.

We note that when S and T consist of the integers, for example $S=\{1,\ldots,m\}$ and $T=\{1,\ldots,n\}$, then the determinants of SC_r - and SSC_r -matrices remain defined only for orders at most $\min\{m,n\}$. Further if $S=\{1,\ldots,m\}$ and T is an interval of the real line, the determinants are meaningful for $p \leqq m$, and we shall continue to use them in this "mixed" case. These definitions cannot be extended when both S and T are discrete but only when either S or T is a real continuum. It is needless to insist any longer on strict inequality in the domain of (2.2).

Now we extend the domains (2.4) of the open simplex: Let $S^p := S \times \ldots \times S$ and $T^p := T \times \ldots \times T$ be the Cartesian products of p copies of S or T respectively. We consider the relative closure of $\Delta_p(S)$ and $\Delta_p(T)$ in S^p and in T^p in the following

$$(2.7) \qquad \begin{aligned} \overline{\Delta}_p(S) &:= \{\overline{s} := (s_1,\ldots,s_p) \mid s_i \in S,\ s_1 \leqq \ldots \leqq s_p\}, \\ \overline{\Delta}_p(T) &:= \{\overline{t} := (t_1,\ldots,t_p \mid t_j \in T,\ t_1 \leqq \ldots \leqq t_p\}. \end{aligned}$$

If, for example, $s_i = s_{i+1} = \ldots = s_{i+\ell-1}$, $(1 \leqq \ell \leqq p-1)$, we say that s_i has the m u l t i p l i c i t y o f o r d e r ℓ. Suppose that the determinant (2.1) contains several equal s's and/or t's. We shall assume that S and T are open sets and that $K(s,t)$ is at least as many times differentiable in each variable as the order p of the compound $K\begin{pmatrix} s_1,\ldots,s_p \\ t_1,\ldots,t_p \end{pmatrix}$, is necessary. Now, when a block of ℓ equal values of s's and/or t's occurs in (2.1), the corresponding rows and/or columns, are determined by successive derivatives; as in the following example:

$$K\begin{pmatrix} s_1, s_2, s_2, s_3, s_4 \\ t_1, t_2, t_2, t_2, t_3 \end{pmatrix} =$$

$$= \begin{vmatrix} K(s_1,t_1) & K(s_1,t_2) & \frac{\partial}{\partial t}K(s_1,t_2) & \frac{\partial^2}{\partial t^2}K(s_1,t_2) & K(s_1,t_3) \\ K(s_2,t_1) & K(s_2,t_2) & \frac{\partial}{\partial t}K(s_2,t_2) & \frac{\partial^2}{\partial t^2}K(s_2,t_2) & K(s_2,t_3) \\ \frac{\partial}{\partial s}K(s_2,t_1) & \frac{\partial}{\partial s}K(s_2,t_2) & \frac{\partial^2}{\partial s \partial t}K(s_2,t_2) & \frac{\partial^3}{\partial s \partial t^2}K(s_2,t_2) & \frac{\partial}{\partial s}K(s_2,t_3) \\ K(s_3,t_1) & K(s_3,t_2) & \frac{\partial}{\partial t}K(s_3,t_2) & \frac{\partial^2}{\partial t^2}K(s_3,t_2) & K(s_3,t_3) \\ K(s_4,t_1) & K(s_4,t_2) & \frac{\partial}{\partial t}K(s_4,t_2) & \frac{\partial^2}{\partial t^2}K(s_4,t_2) & K(s_4,t_3) \end{vmatrix}.$$

In order to emphasize that we have not extended the definition simply by continuity, for determinant in which s_i has the multiplicity of order ℓ_i, and/or t_j the multiplicity of order k_j, $i,j=1,\dots,r$, we denote

$$(2.8) \qquad K\begin{pmatrix} \overbrace{s_1,\dots,s_1}^{\ell_1},\dots,\overbrace{s_q,\dots,s_q}^{\ell_q} \\ \underbrace{t_1,\dots,t_1}_{k_1},\dots,\underbrace{t_p,\dots,t_p}_{k_p} \end{pmatrix}$$

where $\ell_1+\dots+\ell_q = r$, $k_1+\dots+k_p = r$. The most extreme instance is

$$K\begin{pmatrix} \overbrace{s,\dots,s}^{r} \\ t,\dots,t \end{pmatrix} = \det \left(\left\| \frac{\partial^{i+j}}{\partial s^i \partial t^j} K(s,t) \right\|_{i,j=0}^{r-1} \right).$$

But if the multiplicity of s's and/or t's is not stressed, then (2.8), in analogy with derived determinant, can be expressed by

$$K^*\begin{pmatrix} s_1,\dots,s_r \\ t_1,\dots,t_r \end{pmatrix}.$$

We say that $K(s,t)$ is e x t e n d e d s i g n - r e g u l a r o f o r d e r r a n d d e g r e e k i n t h e v a r i a b l e s ($ESR_r(s_{(k)})$ if there exists a sequence of constants $\varepsilon_1,\dots,\varepsilon_r$ each equal to +1 or −1, such that for each $p=1,\dots,r$ we have

$$(2.9) \qquad \varepsilon_p K^*\begin{pmatrix} s_1,\dots,s_p \\ t_1,\dots,t_p \end{pmatrix} > 0$$

for all choices $(t_1,\dots,t_p)\in\Delta_p(T)$ and for all choices $(s_1,\dots,s_p)\in\bar{\Delta}_p(S)$ with the restriction $\max\{\ell_\nu\}=k$. Particulary, if $r=k$, we say that $K(s,t)$ is e x t e n d e d s i g n - r e g u l a r o f o r d e r r i n t h e s v a r i a b l e $(ESR_r(s))$. In this case we have

$$\varepsilon_p K^*\binom{s,\dots,s}{t_1,\dots,t_p} > 0, \quad p=1,\dots,r$$

for any $s\in S$ and for all choices $(t_1,\dots,t_p)\in\Delta_p(T)$. If in (2.9) each ε_p is $+1$, i.e., for all choices of $(s_1,\dots,s_p)\in\bar{\Delta}_p(S)$ and all $(t_1,\dots,t_p)\in\Delta_p(T)$, we have

$$K^*\binom{s_1,\dots,s_p}{t_1,\dots,t_p} > 0, \quad p=1,\dots,r$$

we say that $K(s,t)$ is e x t e n d e d t o t a l l y p o s i t i v e o f o r d e r r w i t h r e s p e c t t o s v a r i a b l e $(ETP_r(s))$. In similar fashion we define the analogous notions with respect to variable t. If $K(s,t)$ is simultaneously extended sign-regular of order r in both variables, we symbolize it by ESR_r. Similarly if we have simultaneously $ETP_r(s)$ and $ETP_r(t)$, we symbolize them by ETP_r.

At the end we wish to mention the basic binary operation enabling us to construct a TP or SR kernel from two such kernels. Let K,L,M be BORELL-measurable functions of two variables satisfying

$$(2.10)\qquad M(\xi,\eta)=\int_T K(\xi,\zeta)\,L(\zeta,\eta)\,d\sigma(\zeta),\quad \xi\in S,\ \eta\in Z$$

where the integral is assumed to converge absolutely. Here ξ,ζ,η traverse S,T,Z respectively and each domain is a subset of the real line, $\sigma(\zeta)$ denotes a sigma-finite measure defined on T. When T consists of a discrete set the integral is of course interpreted as a sum. If (2.10) is viewed as a continuous version of a matrix product, then the expression of the CAUCHY-BINET formula that evaluates subdeterminants of M in terms of those of K and L becomes

(2.11) $K^*\binom{\xi_1,\ldots,\xi_m}{\eta_1,\ldots,\eta_m} =$

$$= \int_{\zeta_1<\ldots<\zeta_m} \cdots \int K\binom{\xi_1,\ldots,\xi_m}{\zeta_1,\ldots,\zeta_m} L^*\binom{\zeta_1,\ldots,\zeta_m}{\eta_1,\ldots,\eta_m} d\sigma(\zeta_1)\ldots d\sigma(\zeta_m)$$

where $(\xi_1,\ldots,\xi_m) \in \Delta_m(S)$ and $(\eta_1,\ldots,\eta_m) \in \Delta_m(Z)$. An obvious consequence of (2.11) is as follows: if $K(s,t)$ is TP_r on $S \times Z$ and $L(z,t)$ is TP_k on $Z \times T$, then $M(s,t)$ is $TP_{\min\{r,k\}}$ on $S \times T$.

The structure of Total positivity has been entirerly studied in KARLIN [1968]. In this work (p.13) the terminology of notions is motivated in an interesting manner.

3. CHEBYSHEV SYSTEMS ON FINITE INTERVAL

Let

(3.1) $$\phi_1,\ldots,\phi_n$$

be a continuous real-valued functions defined on a closed finite interval $[a,b]$ and consider the set $F := \operatorname{span}\{\phi_1,\ldots,\phi_n\}$ which formes all (generalized) polynomials of the form

$$f = \sum_{i=1}^{n} \alpha_i \phi_i$$

where α_i are real numbers. The element $f \in F$ is said to be non-trivial or non-zero if $\sum_{i=0}^{n} |\alpha_i| \neq 0$.

The sequence (3.1) is said to constitute a C h e b y s h e v s y s t e m (abbreviated T-system)(without reference to the interval $[a,b]$) if any non-zero polynomial does not vanish more than n-1 times on the $[a,b]$. This implies that the determinant

(3.2) $$F\binom{\phi_1,\ldots,\phi_n}{t_1,\ldots,t_n} := \begin{vmatrix} \phi_1(t_1) & \ldots & \phi_1(t_n) \\ \phi_2(t_1) & & \phi_2(t_n) \\ \vdots & & \\ \phi_n(t_1) & & \phi_n(t_n) \end{vmatrix}$$

for all choices $(t_1,\ldots,t_n) \in \Delta_n([a,b])$ never vanishes, andtherefore maintains a fixed sign. Precisely, for some choices $(t_1,\ldots,t_n) \in \Delta_n([a,b])$, by multiplying the rows in (3.2) by a factor +1 or -1, we may without restricting the generality of the results, specify the sign in (3.2) as positive. We note that the determinants of the form (3.2) always have the columns arranged so that the t_i appear in increasing order. Also the set (3.1) always has a fixed arrangement and thus we use the term "sequence" rather than "set" of functions. The two formulations of T-systems are equivalent. The first with reference to the number of zeros of generalized polynomials, the second in terms of certain determinantal inequalities have their relative merits.

The sequence $\{\phi_i\}_{i=1}^{n}$ will be referred to as a complete Chebyshev system (CT-system)*) if $\phi_1,\ldots,\phi_r$ is a T-system for each $r=1,\ldots,n$. Evidently, any T-system is always a system of linearly independent functions on considered interval. In general, the inverse is not true; as PEANO had remarked. For example, the system $\{\phi_1, \phi_2\}$ where $\phi_1(t)= t|t|$ and $\phi_2(t)= t^2$, is linearly independent on every interval $[a,b]$ which does not contain the origin, and it is not T-system on that interval.

The system $\{\phi_i\}_{i=1}^{n}$ of continuous functions on $[a,b]$ is called a weak Chebyshev system (on $[a,b]$) (WT-system) provided the functions $\phi_1,\ldots,\phi_n$ are linearly independent (on $[a,b]$) and the determinant (3.2) is nonnegative whenever $(t_1,\ldots,t_n) \in \Delta_n([a,b])$.

Put

$$F(s,t) := \phi_s(t)$$

where s is a discrete variable and t is a continuous variable. According to §2 we may symbolize the determinant (3.2) as

$$F\binom{1,\ldots,n}{t_1,\ldots,t_n}.$$

We say that a T-system is a special circumstance of a

*) This system is called by some authors a Haar system.

kernel function, more precisely SC_n kernel. If every subsystem of $\{\phi_i\}_{i=1}^n$ also constitutes a T-system (WT-system), i.e., all determinants

$$F\begin{pmatrix} i_1, i_2, \ldots, i_p \\ t_1, t_2, \ldots, t_p \end{pmatrix} = \begin{vmatrix} \phi_{i_1}(t_1) & \phi_{i_1}(t_2) & \ldots & \phi_{i_1}(t_p) \\ \phi_{i_2}(t_1) & \phi_{i_2}(t_2) & & \phi_{i_2}(t_p) \\ \vdots & & & \\ \phi_{i_p}(t_1) & \phi_{i_p}(t_2) & & \phi_{i_p}(t_p) \end{vmatrix}$$

$$1 \leqq i_1 < i_2 < \ldots < i_p \leqq p; \quad (x_1, \ldots, x_p) \in \Delta_p([a,b])$$

for $p=1,\ldots,n$ are positive (non-negative), is called a M a r k o f f system (W e a k M a r k o f f system)*). In this case the function

$$K(t,s) := \phi_s(t) \qquad s=1,\ldots,n; \quad t \in [a,b]$$

is manifestly STP_n respectively TP_n.

In a naturaly way as in §2, we may extend the notions of total and of regular positivity for T-systems. The system (1.12) will be called an e x t e n d e d C h e b y s h e v s y s t e m o f o r d e r r (ET_r-system) provided $\phi_i \in C^{r-1}[a,b]$, $i=1,\ldots,n$ and

$$F^*\begin{pmatrix} 1, \ldots, n \\ t_1, \ldots, t_n \end{pmatrix} > 0$$

for all $(t_1,\ldots,t_n) \in \bar{\Delta}_n([a,b])$ where the order of multiplicity of each t's is at most r. If $r=n$ it will be simply referred to as an ET-system. In the last case the determinant $F^*\begin{pmatrix} 1, \ldots, n \\ t, \ldots, t \end{pmatrix}$ is known as the W r o n s k i a n of functions $\phi_1, \ldots, \phi_n$ (in the point $t \in [a,b]$), i.e., $W(\phi_1, \ldots, \phi_n)(t)$. That is why T.POPOVICIU, in his work calls the determinants of the form $F\begin{pmatrix} \phi_1, \ldots, \phi_n \\ t_1, \ldots, t_n \end{pmatrix}$ p r e v r o n s k i a n s. If $\{\phi_i\}_{i=1}^r$ is ET-system for any $r=1,\ldots,n$ we say that (3.1) constituted an e x t e n d e d

---

*) This systems are frequently called a D e s c a r t e s s y s t e m s.

c o m p l e t e   C h e b y s h e v   s y s t e m (ECT-system).

If

(3.3) $$\phi_i(t) = t^{i-1} \ , \quad i=1,\dots,n$$

we have the classical problem of T-system. In this case the determinants $F\binom{1,\dots,n}{t_1,\dots,t_n}$ and $F^*\binom{1,\dots,n}{\underbrace{t_1,\dots,t_1}_{k_1},\dots,\underbrace{t_p,\dots,t_p}_{k_p}}$ reduce to the VANDERMOND determinant $V(t_1,\dots,t_n)$ and prevronskian $V(\underbrace{t_1,\dots,t_1}_{k_1},\dots,\underbrace{t_p,\dots,t_p}_{k_p})$ respectively.

It is known

(3.4) $$V(\underbrace{t_1,\dots,t_1}_{k_1},\dots,\underbrace{t_p,\dots t_p}_{k_p}) = \left(\prod_{i=1}^{n}(k_i-1)!!\right)\prod_{i<j}^{1,\dots,p}(t_j-t_i)^{k_ik_j} \ , \quad k_1+\dots+k_p=n$$

where $\alpha!! := 1!2!\cdots\alpha!$, $0!! := 1$. Specially $W(1,t,\dots,t^{n-1}) = (n-1)!!$. Another well known example of T-system is in the case

(3.5) $$\{\phi_i\}_{i=1}^{2n+1} = \{1,\cos t,\ \sin t,\dots,\cos nt,\ \sin nt\}.$$

Then we have (see for example GONCHAROV [1954])

(3.6) $$F^*\binom{1,\dots,2n+1}{\underbrace{t_1,\dots,t_1}_{k_1},\dots,\underbrace{t_p,\dots,t_p}_{k_p}} = 2^{-n}\left(\prod_{i=1}^{n}(k_i-1)!!\right)\prod_{i<j}^{1,\dots,p}\left(2\sin\frac{t_j-t_i}{2}\right)^{k_ik_j} ,$$

$$k_1+\dots+k_p = 2n+1.$$

From (3.4) we see that the system (3.3) formes ET_r-system of any order $r=1,\dots,n$ on every interval $[a,b]\subset R$. Similarly, from (3.6) it results that the system of trigonometric polynomials (3.5) forms ET_r-system of any order $r=1,\dots,n$ on every interval $[a,b]\subset R$, $b-a<2\pi$.

We remark that the set span$\{\phi_1,\dots,\phi_n\}$ will be later called C h e b y s h e v   s p a c e (of dimension n).

"The history and development of the theory of T-systems and their applications is of a longstanding" KARLIN & & STUDDEN [@ 1966, p.13].

In similar fashion we can define the Chebyshev systems on discrete or arbitrary open set. This notions are more known as interpolatory systems and hence we will mention some of these topics.

4. INTERPOLATORY SETS AND GENERALIZED CONVEXITY

Let n (≥ 1) be a given natural number and let Ω be a fixed set on real line, which contains at least n distinct points. Let us choose a subset $X \subseteq \Omega$ containing also at least n distinct points. A set F of functions defined on Ω is called i n t e r p o l a t o r y s e t o f o r d e r n o n X (abbreviated I_n-set) if for every n-tuple $y_1,\ldots,y_n$ of real numbers there is one and only one element in F which takes the values $y_1,\ldots,y_n$ at points $x_1,\ldots,x_n$ respectively*). An element $f \in F$ determined by the points $(x_i,y_i) \in X \times \mathbb{R}$, $i=1,\ldots,n$ will be denoted by

(4.1) $$L(F;\ x_1,\ldots,x_n;\ y_1,\ldots,y_n).$$

If the numbers y_i are equal to $g(x_i)$ $i=1,\ldots,n$, g being a prescribed function defined on a certain over-set of X, then instead of (4.1) we have written

(4.2) $$L(F;\ x_1,\ldots,x_n;\ g),$$

and we say that this function is a interpolation of g.

It is evident that for an I-set F, the condition of continuity of its elements (on X) is not required. So let F be a family of functions f_ξ defined on $X:=[0,2] \subset \Omega := \mathbb{R}$ with

$$f_\xi(x) = \begin{cases} x+\xi, & 0 \leq x \leq 1 \\ x+\xi+\alpha, & 1 < x \leq 2, \end{cases}$$

---

*) If the order n is not stressed, then I_n-set can be expressed by I-set. In the sequel this convention will be used in different definitions.

where α is a fixed, but arbitrary number. The set F forms an I_1-set on X but f_T are discontinuous at the point x=1, (ELENA POPOVICIU [1972, p.20]).

If an I_n-set F is formed from continuous functions (on X) then we say that F is of the type $I_n\{X\}$. For example the space $\mathbb{P}_n$ of algebraic polynomials, $\deg(\mathbb{P}_n) \leqq n$, is of $I_n\{[a,b]\}$ type, [a,b] being arbitrars, including the cases $a=-\infty$ and/or $b=+\infty$. Similarly, the space $\mathbb{T}_n$ of trigonometric polynomials is of $I_{2n+1}\{[a,a+2\pi)\}$ type.

If F is of $I_n\{[a,b]\}$ type, $n \geqq 1$ and [a,b] is a finite interval, then the difference $g := f_1 - f_2$ of two distinct elements $f_1, f_2 \in F$, can vanish at most at n-1 distinct points from [a,b].

In the definition of the I-set it is not required to mention explicitly that the set depends on certain parameters.

Let

$$\mathbb{D} := [a_1,b_1] \times \dots \times [a_n,b_n] \in \mathbb{R}^n, \quad -\infty \leqq a_i < b_i \leqq +\infty$$

$$A := (\alpha_1,\dots,\alpha_n) \in \mathbb{D}, \quad \alpha_i \in [a_i,b_i], \ i=1,\dots,n.$$

We say that a function

$$F(A;\ x) := f(\alpha_1,\dots,\alpha_n;\ x)$$

generalizes n-parameter family (on [a,b]) if the set $\{F(A;x)\}\ A \in \mathbb{D}$ is of $I_n\{[a,b]\}$ type with respect to variable x, and the functions f are continuous in variables $\alpha_1,\dots,\alpha_n$, x on the whole set of definitions, hence on $\mathbb{D} \times [a,b]$. From the above definition it results that the parameters $\alpha_1,\dots,\alpha_n$ are well and uniquely determined for every $(x_1,\dots,x_n) \in \Delta_n([a,b])$ and for every set of numbers $y_1,\dots,y_n$. Later we will see that n-parameter family can be observed as TP special type kernel.

Now, let F be an $I_n\{[a,b]\}$ type, $n \geqq 2$ and let us con-

---

*) In the sequel we will frequently denote a set and an element from this set by the same symbol.

sider a given choice $(x_1,\dots,x_k)\in\Delta_k([a,b])$, $1\le k\le n-1$, and given values $y_1,\dots,y_n$. The set

$$S_k(F;\begin{matrix}x_1,\dots,x_k\\ y_1,\dots,y_k\end{matrix})$$

of functions from F which takes the values $y_1,\dots,y_k$ at the points $x_1,\dots,x_k$ respectively, is called a s p i k e i n t e r p o l a t o r y o f o r d e r $n-k$ (abbreviated I_{n-k}-spike).

In ELENA POPOVICIU [1972, Lemma 2.1.3] it is proved that an I_{n-k}-spike on $[a,b]$ is of $I_{n-k}\{X\}$ type where $X:=$ $:=[a,x_1)\cup(x_1,x_2)\cup\dots\cup(x_{k-1},x_k)\cup(x_k,b]$ when the first and/or the last of the intervals is omitted if $x_1=a$ and/ /or $x_k=b$.

Now, let $\{F_i\}$ $i=m,m+1,\dots,m+\ell$, $\ell\ge 0$, be a sequence of interpolatory sets such that for every $i=m,\dots,m+\ell$ F_i is of $I_i\{X\}$ type and $F_i\subset F_{i+1}$. The sequence $\{F_i\}$ is called an i n t e r p o l a t o r y s e g m e n t. If $\ell=+\infty$ we have an i n t e r p o l a t o r y c h a i n.

In a similar way as in E.P kernel we may introduce a notion of the extended I-set. Thus for so called e x t e n d e d I-set o f o r d e r n a n d d e g r e e o f r e g u l a r i t y k (abbreviated $EI_{n,(k)}$-set) the element

$$f:=L(F;\underbrace{x_1,\dots,x_1}_{k_1},\dots,\underbrace{x_p,\dots,x_p}_{k_p};g)\in F$$

$$k_1+\dots+k_p=n,\qquad \max_i k_i=k$$

is well determined by equalities

$$f^{(j)}(x_i)=y_i^{(j)},\quad j=0,1,\dots,k_i-1;\quad i=1,\dots,p$$

where $y_i^{(j)}$ are arbitrary but prescribed numbers. The notion of $EI_{n,(1)}$-set is equivalent to the common notion of interpolatory. If $k=n$ we say that F is an e x t e n d e d c o m p l e t e r e g u l a r I-set o f o r d e r n ($ECRI_n$-set) We note here that we will frequently omit the words "extended" or "of order" whenever it will not lead to a confusion.

Let F be a linear interpolatory set on X, defined by

(4.3) $$F(\Lambda;\ x) := \operatorname{span}\{f_1,\dots,f_n\}$$

where

(4.4) $$\{f_i\}_{i=1}^{n}$$

is a given sequence of functions. Then from the definition of interpolation it follows

(4.5) $$F\binom{f_1,\dots,f_n}{x_1,\dots,x_n} = \begin{vmatrix} f_1(x_1) & \dots & f_1(x_n) \\ \vdots & & \\ f_n(x_1) & & f_n(x_n) \end{vmatrix} \neq 0$$

for every $(x_1,\dots,x_n) \in \Delta_n(X)$. In this case the sequence (4.4) is called an *interpolatory system of order n* (I_n-system), and F is called an *interpolatory space of dimension n*. If insead of (4.5) we consider

(4.6) $$F^{*}\binom{f_1,\dots,f_n}{x_1,\dots,x_n} \neq 0$$

where the order of multiplicity of each x's is at most k ($1 \leq k \leq n$), then we have an *extended interpolatory system of order n and of degree of regularity k* ($EI_{n,(k)}$-system).

If $X:=[a,b]$ and taking into account (4.5) we conclude that every T-system on X, constitutes also an I-system. Conversely, an I-system of the type $I_n\{[a,b]\}$ is a T-system on $[a,b]$. Thus, an I-system of the type $I_n\{X\}$, X being arbitrary, is called by some authors a T-system (in generalized sense). The concept of interpolation considered above is a particular case of a notion of interpolating in abstract space. Indeed, instead of the operator L which an element g transforms into $L(F;\ x_1,\dots,x_n;\ g)$, we may consider an arbitrary operator $A:E \to E$ defined on linear space E. Let S be a given subset from E. The set S is called *interpolatory with respect to A*, if we have

$$\forall g \in E \Rightarrow A(g) \in E, \quad \forall g \in S \Rightarrow A(g) = g.$$

Now, let F be an $I_n\{X\}$-set. If $\tilde{g}$ is the restriction of the given function g to the given n-typle $(x_1,\ldots,x_n) \in \Delta_n(X)$, then the function $L(F;\ x_1,\ldots,x_n;\ \tilde{g})$ extends the function $\tilde{g}$ on the whole set of definition of the functions from F, i.e., at least on X. For the function g, all restrictions to the n-tuples $(x_1,\ldots,x_n) \in \Delta_n(X)$ can be considered and hence all corresponding extensions $L(F;\ x_1,\ldots,x_n;\ g)$ for every $(x_1,\ldots,x_n) \in \Delta_n(X)$ and a given set F. We will consider the differences

$$(4.7) \qquad D[F; x_1,\ldots,x_{n+1}; g] := g(x_{n+1}) - L(F;\ x_1,\ldots,x_n; g)(x_{n+1})$$

$$(x_1,\ldots,x_{n+1}) \in \Delta_{n+1}(X).$$

The function g is called c o n v e x (abbreviated F-convex) or n o n c o n c a v e or p o l y n o m i a l or n o n c o n v e x or c o n c a v e on X with respect to F, if the inequalities

$$(4.8) \qquad D[F;\ x_1,\ldots,x_{n+1};\ g] >, \geqq, =, \leqq, < 0$$

holds respectively for every $(x_1,\ldots,x_{n+1}) \in \Delta_{n+1}(X)$. The functions which have one of the properties from (4.8) are called c o m p a r a b l e w i t h F. This definition, introduced in ELENA POPOVICIU [1955], is a generalization of the notion of s u p e r i o r o r d e r c o n v e x i t y with respect to a linear family (4.3) introduced in POPOVICIU [1933, 1936]. In the quoted papers author simultaneously considers two I-systems

$$F_n := \operatorname{span}\{f_1,\ldots,f_n\}$$

and

$$F_{n+1} := \operatorname{span}\{f_1,\ldots,f_n,f_{n+1}\}$$

where the function f_{n+1} is conveniently chosen. In this case the functional by (4.7) may be written in the form

$$(4.9) \qquad \begin{bmatrix} f_1,\ldots,f_{n+1} \\ x_1,\ldots,x_{n+1} \end{bmatrix}; g \Big] := \frac{D\begin{pmatrix} f_1,\ldots,f_n, g \\ x_1,\ldots,x_n,x_{n+1} \end{pmatrix}}{D\begin{pmatrix} f_1,\ldots, f_{n+1} \\ x_1,\ldots, x_{n+1} \end{pmatrix}}.$$

The POPOVICIU's idea on introduction of the supplementary function f_{n+1}, i.e., of consideration of the interpolatory segment $F_n \subset F_{n+1}$, enables to express the d i v i d e d d i f f e r e n c e (4.9) of the function g with respect to the system $\{f_i\}_{i=1}^{n}$ in a symmetric form. In other words, this difference is a symmetricaly function in the elements f_i and in the points x_i, $i=1,\ldots,n+1$. In the cases when we may assume the positive denominator in (4.9), instead of (4.8) we may write

$$(4.10) \qquad F\left(\begin{matrix} f_1,\ldots, f_n, g \\ x_1,\ldots, x_n, x_{n+1}\end{matrix}\right) > (\geqq, =, \leqq, <)\ 0.$$

In the above mentioned definition of F-convexity the function $g \in C[X]$ is compared with elements f from F, this comparison being made by means of the functional D defined by (4.7). Instead of this functional D, an arbitrary functional $\mathscr{D}$ defined by

$$\mathscr{D}\,[F;\ x_1,\ldots,x_{n+1};\ g] := \mathscr{L}((g - L(F;\ x_1,\ldots,x_n;g)(x_{n+1}))$$

may be considered. This is a manner to generalise the above described notion of convexity. A summary survey of convexity notions was given in the monographe ELENA POPOVICIU [1972].

CHAPTER II

THE METHOD OF CONES IN OPTIMAL PROBLEMS

5. ADMISSIBLE CONES AND ADHERENT CONES

A set K from a linear space X is called a *cone* (with the vertex in θ_X) if $\lambda K=K$ for every $\lambda > 0$, i.e., if from $h \in K$, $\lambda h \in K$ results for every $\lambda > 0$. If K is a cone with the vertex in θ_X then $x_0 + K := \{x + h \mid h \in K\}$ is a cone with the vertex in x_0. The cone K which does not contain any line, i.e., $-h \notin K$ results from $h \in K$, $h \neq \theta_X$, is called an *accute* cone. The intersection of any sets of cones, as well as any set of $\alpha K_1 + \beta K_2$-form, where K_1, K_2 are cones with the same vertex, $\alpha, \beta \in \mathbb{R}$, are also cones. In the topological linear space, the interior, the adherence and the envelope are also cones. If the set K is convex, the cone is called a *convex cone*. The convex cone has the property that $\alpha h_1 + \beta h_2 \in K$ results from $h_1, h_2 \in K$, $(\alpha > 0,\ \beta > 0)$.

Let X be a real topological linear space, Q a given subset of X and x_0 a given point of X.

DEFINITION 5.1. The set of all vectors h from X defined by

$$K(Q;\ x_0) := \{h \in X \mid \exists\, V(h),\ \exists\, \varepsilon > 0,\ \forall y \in V(h),\ \forall \eta \in (0, \varepsilon):\ x_0 + \eta y \in Q\}$$

is a cone in X and it is called the *admissible cone of Q with respect to the point* x_0 (abbreviated $K(Q;\ x_0)$). The elements $h \in K(Q;\ x_0)$ are called the vectors of the *admissible dep-*

l a c e m e n t (with respect to set Q at the point x_0)

DEFINITION 5.2. The set of all vectors h from X defined by

$$K[Q;\, x_0] := \{ h \in X \mid \forall \mathcal{V}(h),\ \forall \varepsilon > 0,\ \exists\ y_\varepsilon \in \mathcal{V}(h),\ \exists \eta_\varepsilon \in (0,\varepsilon):\ x_0 + \eta_\varepsilon y_\varepsilon \in Q \}$$

is a cone and it is called the a d h e r e n t c o n e o f t h e s e t Q w i t h r e s p e c t t o t h e p o i n t x_0 (abbreviated $K[Q;\, x_0]$). The elements $h \in K[Q:\, x_0]$ are called a d h e r e n t v e c t o r s (with respect to x_0).

In other words we can say that for every $h \in K(Q;\, x_0)$ there exists a neighbourhood $\mathcal{U}(\theta_X)$ such that

$$x_0 + \mathcal{U}(\theta_X) \cap \Big(\bigcup_{\lambda > 0} \lambda \mathcal{V}(h) \Big) \subseteq Q$$

and about $h \in K[Q;\, x_0]$ we can say that it has the property that for every $\mathcal{V}(h)$ and for every $\mathcal{U}(\theta_X)$ we have

$$Q \cap \Big[x_0 + \mathcal{U}(\theta_X) \cap \Big(\bigcup_{\lambda > 0} \lambda \mathcal{V}(h) \Big) \Big] \neq \emptyset.$$

The above definition was given first in BOULIGAND [1932].

We will give some properties of the cones $K(Q;\, x_0)$ and $K[Q;\, x_0]$, (see for example LAURENT [1972] or HOFFMANN & KOLUMBAN [1974]).

(a) The cones $K(Q;\, x_0)$ and $K[Q;\, x_0]$ have the vertices in the origin θ_X. $K(Q;\, x_0)$ is an open set while $K[Q;\, x_0]$ is a closed set.

We have

$$K(Q;\, x_0) = K(\mathring{Q};\, x_0),$$

$$K[Q;\, x_0] = K[\bar{Q};\, x_0] .$$

For every neighbourhood $\mathcal{V}(x_0)$ equalities

$$K(Q;\, x_0) = K(Q \cap \mathcal{V}(x_0);\, x_0),$$

$$K[Q;\, x_0] = K[Q \cap \mathcal{V}(x_0);\, x_0]$$

are valid.

(b) Let I be an arbitrary set of indices. The element $h \in X$, $h \neq \theta_X$, is contained in $K[Q;\, x_0]$ if for every seque-

nce $(x_i), i \in I$ of elements from Q, there exists a seque-
nce $(r_i), i \in I$ of non-negative numbers such that

$$\lim_{i \in I} r_i(x_i - x_o) = h.$$

(c) If $x_o \in \mathring{Q}$, then we have

$$K(Q;\ x_o) = K[Q;\ x_o] = X,$$

while if $x_o \in \complement_X \overline{Q}$, then

$$K(Q;\ x_o) = K[Q;\ x_o] = \emptyset$$

is valid.

(d) $\complement_X K(\complement_X Q;\ x_o) = K[Q;\ x_o]$,

$\complement_X K[\complement_X Q;\ x_o] = K(Q;\ x_o)$.

(e) $K(Q;\ x_o) \subseteq K[Q;\ x_o]$,

$Q_1 \subseteq Q_2 \Rightarrow K(Q_1;\ x_o) \subseteq K(Q_2;\ x_o)$

$\Rightarrow K[Q_1;\ x_o] \subseteq K[Q_2;\ x_o]$.

(f) If Q is a convex set of X then for $x_o \in \overline{Q}$ we have

$$K(Q;\ x_o) = \bigcup_{\lambda > 0} \lambda(\mathring{Q} - x_o),$$

$$K[Q;\ x_o] = \operatorname{adh}\left\{\bigcup_{\lambda > 0} \lambda(\overline{Q} - x_o)\right\}.$$

If Q is a convex and $\mathring{Q} \neq \emptyset$ then for $x_o \in \overline{Q}$ we have

$$K(Q;\ x_o) = \operatorname{int}\left\{K[Q;\ x_o]\right\},$$

$$K[Q;\ x_o] = \operatorname{adh}\left\{K(Q;\ x_o)\right\}.$$

The equalities take place also in the hypothesis that $K[Q;\ x_o]$ is convex.

(g) If $\{Q_i\}, i \in I$ is a family of sets from X, then we have

(*) $K(\bigcap_{i \in I} Q_i;\ x_o) \subseteq \bigcap_{i \in I} K(Q_i;\ x_o)$, (*) $K[\bigcup_{i \in I} Q_i;\ x_o] \supseteq \bigcup_{i \in I} K[Q_i;\ x_o]$,

(*) $K(\bigcup_{i \in I} Q_i;\ x_o) \supseteq \bigcup_{i \in I} K(Q_i;\ x_o)$, (**) $K[\bigcap_{i \in I} Q_i;\ x_o] \subseteq \bigcap_{i \in I} K[Q_i;\ x_o]$.

If the set of indices is finite then the first two inclusions become equalities.

REMARK 5.3. If the space X has a basis of neighbourhood of Θ_X, then we can consider $I = \mathbb{N}$.

REMARK 5.4. Below we will use the property (b) also as a definition of $K[Q; x_o]$.

REMARK 5.5. The latest two inclusions from (g) do not become equalities even when I is a finite set. Further, there are four more implications analog to those in (g), which are obtainable by combining symbols: $K, \cup, \cap$. They will be dealt with later.

6. DIFFERENTIATION WITH RESPECT TO THE CONE

Let X be a topological linear space, S a subset of X and let $F: S \to G$ be an operator which acts from S in a linear topological space G. We consider a subset Q of S and a given point $x_o \in Q$.

DEFINITION 6.1. (i) If for a fixed element $h \in K[Q; x_o]$, $h \neq \Theta_X$, for every sequence $(x_i), i \in I$ of points from Q and for every sequence $(r_i), i \in I$ of non-negative numbers with the properties

$$\lim_{i \in I} x_i = x_o, \quad \lim_{i \in I} r_i(x_i - x_o) = h,$$

the limit

$$F'_Q(x_o)(h) := \lim_{i \in I} r_i[F(x_i) - F(x_o)]$$

exists, then this limit is called the *directional derivative with respect to cone* $K[Q; x_o]$ of the operator F (in the direction h) at a point x_o, HOFFMANN & KOLUMBAN [1974];

(ii) If F has a directional derivative with respect to cone $K[Q; x_o]$ in all directions $h \in K[Q; x_o]$, and if the operator $F'_Q(x_o) : K[Q; x_o] \to G$ is continuous (in the sense of the topology of G), then we say that $F'_Q(x_o)$ is *the directional derivative of F with respect to Q at a point* x_o. In this case we say that F is *directionally differentiable with respect to set* Q. By definition we take $F'_Q(x_o)(\Theta_X) := \Theta_G$.

(iii) Let $\mathfrak{S} := \{s\}$ be a compact HAUSDORF space and let $f\colon D \times \mathfrak{S} \to \mathbb{R}$ be a given functional. For every $s \in \mathfrak{S}$ the mapping $f(\cdot, s)$ has the directional derivative with respect to Q at a point x_0 in direction $h \in K[Q; x_0]$, $h \neq \Theta_X$. We say that $f(\cdot, s)$ is $\mathfrak{S}$-u n i f o r m l y d i f f e r e n t i a b l e w i t h r e s p e c t t o Q a t a p o i n t x_0 i n d i r e c t i o n h, if for every sequence (x_i) $i \in I$, $x_i \in Q$, $x_i \to x_0$, and for every sequence (r_i) $i \in I$, $r_i \geqq 0$, with $r_i(x_i - x_0) \to h$, and for every $\varepsilon > 0$, there exists the index $i_0 \in I$ such that

$$\left| r_i[f(x_i, s) - f(x_0, s)] - f'_Q(x_0, s)(h) \right| < \varepsilon \quad \forall i > i_0, \ \forall s \in \mathfrak{S},$$

We give one example. Later we shall consider the examples of uniform differentiability.

We remark that sometimes the derivative $F'_Q(x_0)(h)$ is called a differential.

EXAMPLE 6.1. Suppose that $X := \mathbb{R}^2$, $Q := S := \{(x,y) \in \mathbb{R}^2 \mid x^2 \leqq y \leqq 2x^2,\ x \geqq 0\}$, $F\colon S \to \mathbb{R}$ defined by $F(x,y) := x + \sqrt{(y-x^2)^3}$. In this case we have

$$K[Q:\ (0,0)] = \{(x,0) \in \mathbb{R}^2 \mid x \geqq 0\}$$

and

$$F'_Q((0,0))(h) = x$$

for every $h = (x,0) \in K[Q;\ (0,0)]$.

We observe that the previous example shows that the differential notion with respect to cone is more general than the differential on the direction (in the sense of GATEAUX). Namely, in our example the operator F does not have the derivative with respect to the cone at the point $\Theta_{\mathbb{R}^2}$.

It is clear that the derivative defined above is a patriculer case of a general scheme of the limit with respect to a prescribed set. In the sequel we will need some properties of the derivative with respect to cone.

(a) The derivative $F'_Q(x_0)$ is positive homogeneous mapping.

(b) The set of all mappings $F\colon S \to G$ for which there exists $F'_Q(x_0)$ forms a linear space.

(c) Let $h \in X$, $h \neq \Theta_X$, be a fixed element and

$$Q := \{\lambda x_o + (1-\lambda)h \mid 0 \leqq \lambda \leqq 1\}.$$

If there exists a derivative in the direction (in the sense of GATEAUX):

$$F'(x_o)(h) := \lim_{\alpha \downarrow 0} \frac{F(x_o + \alpha h) - F(x_o)}{\alpha},$$

then the derivative $F'_Q(x_o)$ also exists and the equality

$$F'_Q(x_o)(h) = F'(x_o)(h)$$

holds. If $Q \subseteq X$, $h \in X$ and $\varepsilon > 0$ are fixed and

$$x_o + \alpha h \in Q \qquad \forall \alpha \in (0,\varepsilon)$$

and if $F'_Q(x_o)(h)$ exists, then the derivative $F'(x_o)(h)$ also exists and the equality

$$F'_Q(x_o)(h) = F'(x_o)(h)$$

occurs.

(d) Let $f:[0,1] \to X$ be a continuous mapping with the properties:

$$f(0) = x_o, \quad f(t) \in Q \qquad \forall\, t \in (0,1],$$

$$\lim_{t \downarrow 0} \frac{f(t) - f(0)}{t} = h \in X.$$

Then if $F'_Q(x_o)(h)$ exists, there also exists the derivative

$$\frac{d}{dt} F(f(t))\Big|_{t=0}$$

and the equality

$$F'_Q(x_o)(h) = \frac{d}{dt} F(f(t))\Big|_{t=0}$$

holds.

(e) Let $X := (X, \|\cdot\|)$ be a normed linear space and $h \in K[Q; x_o]$ with $\|h\| = 1$. If $F'_Q(x_o)(h)$ exists, then for every $P \subseteq Q$ with the property

$$(*) \qquad \lim_{\substack{x \to x_o \\ x \in P}} \frac{x - x_o}{\|x - x_o\|} = h$$

there also exists the limit

$$\lim_{\substack{x \to x_0 \\ x \in P}} \frac{F(x) - F(x_0)}{\|x - x_0\|}$$

and these values are equal for every P from Q for which (*) is satisfied.

(f) If X,Y are normed linar spaces and $Q \subseteq S \subseteq X$ are given subsets and if the FRECHET differential*) $DF(x_0)$ of the operator $F: S \to Y$ at the point $x_0 \in S$ exists, then there also exists the derivative $F'_Q(x_0)$ and they are equal, $DF(x_0) = F'_Q(x_0)$.

(g) If the topological linear spaces X,Y,Z, the subsets Q,S such that $Q \subseteq S \subseteq X$, the point $x_0 \in S$ and the mappings $F: S \to Y$, $\Phi: F(S) \to Z$ are given, suppose that

$$F'_Q(x_0)(h) \quad \text{and} \quad \Phi'_{F(Q)}(F(x_0))(F'_Q(x_0)(h))$$

exist. Then we have

$$(\Phi \cdot F)'_Q(x_0)(h) = \Phi'_{F(Q)}(F(x_0))(F'_Q(x_0)(h)).$$

7. THE DEFINITION AND THE CHARACTERIZATION OF THE OPTIMIZING PROBLEM

The topological linear spaces X,Y,Z with the subsets $D \subseteq X$, $C \subseteq Y$ and a fixed element $z_0 \in Z$ are given. Let $\mathbb{T} := (\mathbb{T}, \leq)$ be a real semiordered topological linear space. We consider the mappings $f: D \to \mathbb{T}$, $F: D \to Y$, $G: D \to Z$ and the sets

$$(7.1) \qquad \begin{aligned} Q &:= \{x \in D \mid F(x) \in C\}, \\ M &:= \{x \in D \mid G(x) = z_0\}. \end{aligned}$$

The optimizing problem, more precisely, the minimizing problem, is defined in the following way:

*) In the sequel we will denote by $DF(x_0)$ the FRECHET differential of the operator F at a point x_0.

The determination of the elements $\bar{x} \in A := Q \cap M$ such that

(7.2) $$f(\bar{x}) = \min_{x \in A} f(x)$$

is required.

The function f is the t a r g e t f u n c t i o n while Q and M are restrictions (of the inequality type, respectively of the equality type). Here we had a global minimum. If instead of (7.2) we take

(7.3) $$f(x^*) = \min_{x \in A \cap V(x^*)} f(x)$$

where $V(x^*)$ is a neighbourhood of x^*, then we have the local minimization problem. The elements $\bar{x}$, x^* which verify (7.2) respectively (7.3) are called m i n i m a l respectively l o c a l m i n i m a l elements.

Sometimes in applications we have more restrictions. Put

(7.4) $$Y := \prod_{i \in I} Y_i, \qquad Z := \prod_{j \in J} Z_j$$

where Y_i and Z_j are real topological linear spaces and I,J are sets of indices. We denote

$$C := \prod_{i \in I} C_i, \quad C_i \subseteq Y_i,$$

$$z_o := (z_j)\ j \in J, \quad z_j \in Z_j,$$

(7.5) $$F := (F_i)\ i \in I, \quad F_i : D \to Y_i,$$

$$G := (G_j)\ j \in J, \quad G_j : D \to Z_j,$$

$$Q_i := \{x \in D \mid F_i(x) \in C_i\}, \; Q := \bigcap_{i \in I} Q_i,$$

$$M_j := \{x \in D \mid G_j(x) = z_j\}, \; M := \bigcap_{j \in J} M_j.$$

With these notations we have the following characterization, generally enough, of the optimal elements.

THEOREM 7.1. If $x_o \in A$ is a minimal element for f on A and the set

$$P := \{x \in D \mid f(x) < f(x_o)\}$$

is open in $\mathbb{T}$, then we have

(7.6) $K(P;\, x_o) \cap K(Q;\, x_o) \cap K[M;\, x_o] = \emptyset .$

PROOF. The proof is similar to that of the case when D:=X, see for example LAURENT [1972, Theorem 1.4.1]. We suppose that there exists an element h_o from the left side set of (7.6). Using Definition 5.1 and the property §4.(g), from

$$h_o \in K(P;\, x_o) \cap K(Q;\, x_o) = K(P\cap Q;\, x_o)$$

we see that there exists a neighbourhood $\mathcal{V}(h_o)$ of the point h_o and there exists a positive number $\varepsilon > 0$, such that

$$x_o + \eta h \in P\cap Q \qquad \forall \eta \in (0,\varepsilon),\ \forall h \in \mathcal{V}(h_o).$$

On the other hand from Definition 5.2, from $h \in K[M;\, x_o]$ we see that in the neighbourhood $\mathcal{V}(h_o)$ there exists a vector h_ε and a positive number $\eta_\varepsilon \in (0,\varepsilon)$ such that

$$x_o + \eta_\varepsilon h_\varepsilon \in M.$$

Thus the point $x_o + \eta_\varepsilon h_\varepsilon$ satisfies all conditions imposed by restrictions. Further, from $x_o + \eta_\varepsilon h_\varepsilon \in P$ results

$$f(x_o + \eta_\varepsilon h_\varepsilon) < f(x_o)$$

contradicting that x_o is a minimal element. □

By property §4.(a) we come to the conclusion that (7.6) yields a necessary and sufficient condition for a local minimum in x_o.

Now, let the target function f has the derivative $f'_D(x_o)(h)$ for all $h \in K[D;\, x_o]$, $x_o \in A$ where

(7.7) $$A := (\bigcap_{i\in I} Q_i) \cap (\bigcap_{j\in J} M_j) ,$$

Then we have from above theorem (see also [Theorem 21])*)

COROLLARY 7.2. If the set $\{t \in T \mid t < \theta_T\}$ is open and if $x_o \in A$ is a local minimal point for f on A, then there is not a vector $h \in K[A;\, x_o]$ which satisfy the inequality

(7.8) $$f'_D(x_o)(h) < \theta_T .$$

*) In the sequel we will denote by [·] some resultats from the paper HOFFMANN & KOLUMBAN [1974].

But in the practice many a time is difficult to find a cone $K[D; x_0]$ and then we cannot use the inequality (7.8). Thus in the cones method in optimization problems one of the fundamental question is the finding of the cones mentioned above and the studying of relations between these cones.

8. METHODS OF FINDING THE ADMISSIBLE AND ADHERENT CONES

Let X,Y,Z be topological linear spaces and $\mathbb{T}:=(\mathbb{T}, \leqq)$ a semiordered topolcgical linear space. Further let D be a given subset from X, $x_0 \in D$ the point and $F: D \to Y$, $f: D \to \mathbb{T}$ given mappings. For fixed $x_0 \in D$ we denote

$$
\begin{aligned}
&P := \{x \in D \mid f(x) < f(x_0)\}, \\
(8.1)\qquad &Q := \{x \in D \mid f(x) \leqq f(x_0)\}, \\
&M := \{x \in D \mid F(x) = F(x_0)\}.
\end{aligned}
$$

We shall quote some known results.

THEOREM 8.1. [Theorem 20]. Let $K[Q; x_0]$ be a convex cone and let $f'_Q(x_0)$ be a convex functional. If an element $\tilde{h} \in K[Q; x_0]$ exists such that $f'_Q(x_0)(\tilde{h}) < \theta_{\mathbb{T}}$ (SLATER's condition), then

$$(8.2)\qquad K(Q; x_0) \subseteq \{h \in K(D; x_0) \mid f'_Q(x_0)(h) < \theta_{\mathbb{T}}\}.$$

DEFINITION 8.2. Let X and Y be a topological linear spaces and $Q \subseteq D \subseteq X$, $x_0 \in D$, $K(Q; x_0) \neq \emptyset$. Further, for an operator $F: D \to Y$ there exists the derivative (in the sense of GATEAUX) $F'(x_0)(h)$ for every $h \in K(Q; x_0)$. We say that F is u n i f o r m l y d i f f e r e n t i a b l e o n $K(Q; x_0)$ if for every $\tilde{h} \in K(Q; x_0)$ and for every neighbourhood $\mathcal{W}$ of $F'(x_0)(\tilde{h})$, a neighbourhood $\mathcal{V}(\tilde{h})$ and a number $\varepsilon > 0$ exist such that

$$x_0 + \alpha h \in Q, \quad \frac{1}{\alpha}[F(x_0 + \alpha h) - F(x_0)] \in \mathcal{W} \quad \forall h \in \mathcal{V}(\tilde{h}), \forall \alpha \in (0, \varepsilon).$$

We emphasize that the above definition refers to the differentiation in the sense of GATEAUX.

EXAMPLE 8.i. [Lemma 15]. Let X,Y be normed linear spaces, $x_0 \in \mathring{Q}$ and let F be a FRECHET-differentiable mapping at x_0. Then F is uniformly differentiable on $K(Q; x_0) := X$.

EXAMPLE 8.ii. [Lemma 16]. Let X,Y be normed linear spaces and Q,D $(Q \subseteq D)$, be subsets from X, $x_0 \in \hat{Q}$ being prescribed. Further, let $F: D \to Y$ be a mapping which in a neighbourhood $\mathcal{V}(x_0)$ is LIPSCHITZ-limited. Then if there exists the differential $F'(x_0)(h)$ for every $h \in K(Q; x_0)$, then F is uniformly differentiable on $K(Q; x_0)$.

THEOREM 8.3. [Theorem 17]. If $\{t \in T \mid t < \theta_T\}$ is an open set in T, and if $f: D \to T$ is uniformly differentiable with respect to the directions from $K(D; x_0)$, then

(8.3) $\quad K(D; x_0) \supseteq \{h \in K(D; x_0) \mid f'(x_0)(h) < \theta_T\}$.

Combining Theorems 8.1 and 8.3 the conditions under which the equality sign would be valid in (8.2) and (8.3) could be obtained, which is of great importance in practice.

As for as the adherent cone is concerned, we have the following results.

THEOREM 8.4. [Theorem 18]. If the derivative $F'_M(x_0)$ of the operator F exists and $x_0 \in M$, then

(8.4) $\quad K[M; x_0] \subseteq \{h \in K[D; x_0] \mid F'_M(x_0)(h) = \theta_Y\}$.

THEOREM 8.5. [Theorem 19]. Let X,T be BANACH spaces and $\mathcal{V}(x_0)$ a neighbourhood of the point $x_0 \in D \subseteq X$. Let the operator F have the following properties:

1° $F'_D(x)$ exists for all $x \in D \cap \mathcal{V}(x_0)$ and $F'_D(x): K[D; x_0] \to T$ is linear and continuous mapping for all $x \in D \cap \mathcal{V}(x_0)$.

2° F'_D is a continuous operator in $D \cap \mathcal{V}(x_0)$,

3° $F'_D(x_0)$ maps the entire space X onto the entire space Y,

4° $K[D; x_0] \subseteq X$,

5° For every h from $\{h \in K[D; x_0] \mid F'_D(x_0)(h) = \theta_Y\}$ there exists a sequence $(x_i)_{i \in N}$ of points from M and a sequence (r_i) $i\in\mathbb{N}$ of positive numbers, so that $x_i \to x_o$, $r_i(x_i - x_o) \to h$ $(i\in\mathbb{N})$. Then

$$(8.5) \qquad K[M;\ x_o] = \{h\in K[D;\ x_o] \mid F_D'(x_o)(h) = 0_{\mathbb{T}}\}.$$

Notice that the latest theorem requires stronger conditions, for example 4^o. Theorem 8.5 generalises the well-known result of L.A.LUSTERNIK regarding the case when X,Y are BANACH spaces, $\mathbb{T} := \mathbb{R}$, $F: X\to Y$ is FRECHET-differentiable operator and $f: X\to\mathbb{R}$ is a functional whose FRECHET-derivative is continuous. For a long time it has been the only paper of that kind.

The implication of the " $\subseteq$ " type for the adherent cone (of an arbitrary set) can be given under quite weak conditions. More precisely, see [Theorem 12], let X,Y be topological spaces, A an arbitrary subset from X, $x_o\in D\supseteq A$, $K[A;\ x_o]\neq\emptyset$ and $F: D\to Y$ the given operator. Let the derivative $F_A'(x_o)$ exist. Then

$$(8.6) \qquad K[A;\ x_o] \subseteq (F_A'(x_o))^{-1}\, K[F(A);\ F(x_o)].$$

However, it is difficult to give implication of "$\supseteq$" type, even when F is differentiable (in any sense), particularly for sets having an empty interior, as for example the set M from (8.1). That is why the formulas of (8.5) type can be studied under quite strong conditions. The second obstacle in finding these formulas entails from the kind of the set D and its adherent cone. It seems that when these sets are not subspaces of X, the study of the above-mentioned implications requires the particular notions and definitions, which will be studied in further text. Another problem is the one quoted in Remark 5.5. In the next chapter we shall see that the formulas of the mentioned type condition the efficiency of the application of the cones method in the optimization problems. The following two sections deal with this topic.

9. TANGENTIALLY CONNECTED SETS AND TANGENTIALLY SEPARABLE FUNCTIONS

Let X, Y be linear topological spaces, D a subset of X, x_o a fixed point in X and $K[D; x_o]$ the adherent cone.

DEFINITION 9.1. A set D is *tangentially connected* at a point $x_o \in D$ if for every vector $h_o \in K[D; x_o]$ and any neighbourhood $\mathcal{V}(h_o)$ of h_o there is a positive number $\varepsilon_o = \varepsilon_o(h_o, \mathcal{V}(h_o))$ such that for every $\varepsilon \in (0, \varepsilon_o)$ the set

$$\Delta_\varepsilon := D \cap \{x_o + \eta h \mid \forall \eta \in (0, \varepsilon), \forall h \in \mathcal{V}(h_o)\}$$

is nonvoid and connected.

EXAMPLE 9.i. Let D be a convex set from X. Then

$$K[D; x_o] = \text{adh}\ (\bigcup_{\lambda > 0} \{\lambda(z - x_o) \mid z \in D\}).$$

Let $h \in K[D; x_o]$ and we put $\varepsilon_o = 1$. For every $\varepsilon \in (0,1)$, any $\eta_1, \eta_2 \in (0, \varepsilon)$ and any $h_1, h_2 \in X$ with

$$x_o + \eta_1 h_1 \in D, \quad x_o + \eta_2 h_2 \in D,$$

we may define a function φ by

$$(9.1) \quad \varphi(t) := (1 - t)(x_o + \eta_1 h_1) + t(x_o + \eta_2 h_2), \quad t \in [0,1].$$

The function φ is continuous and

$$\varphi(0) = x_o + \eta_1 h_1 \in D, \quad \varphi(1) = x_o + \eta_2 h_2 \in D.$$

The convexity of D implies

$$\varphi(t) \in D \qquad \forall t \in [0,1].$$

Taking into account the continuity of φ it follows that the set D is tangentially connected at a point x_o.

EXAMPLE 9.ii. Let $X := R^2$ and let

$$D := \{x = (\xi, \eta) \in R^2 \mid \xi^2 \leq \eta \leq 2\xi^2, \ \xi \geq 0\}.$$

Here we have

$$K[D; \theta_{R^2}] = \{(\xi, \eta) \in R^2 \mid \xi \geq 0, \ \eta = 0\}.$$

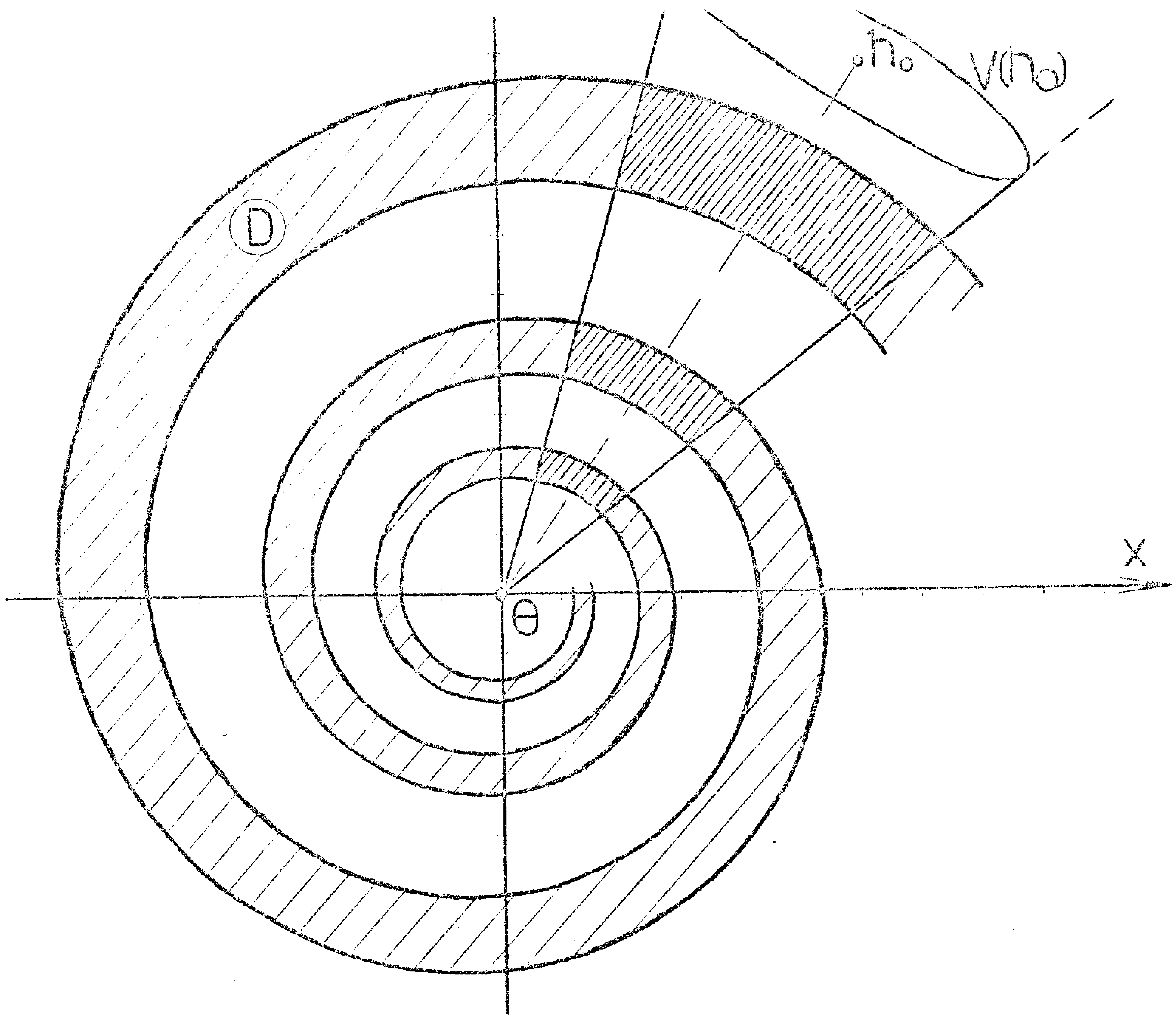

Fig. 9.2.

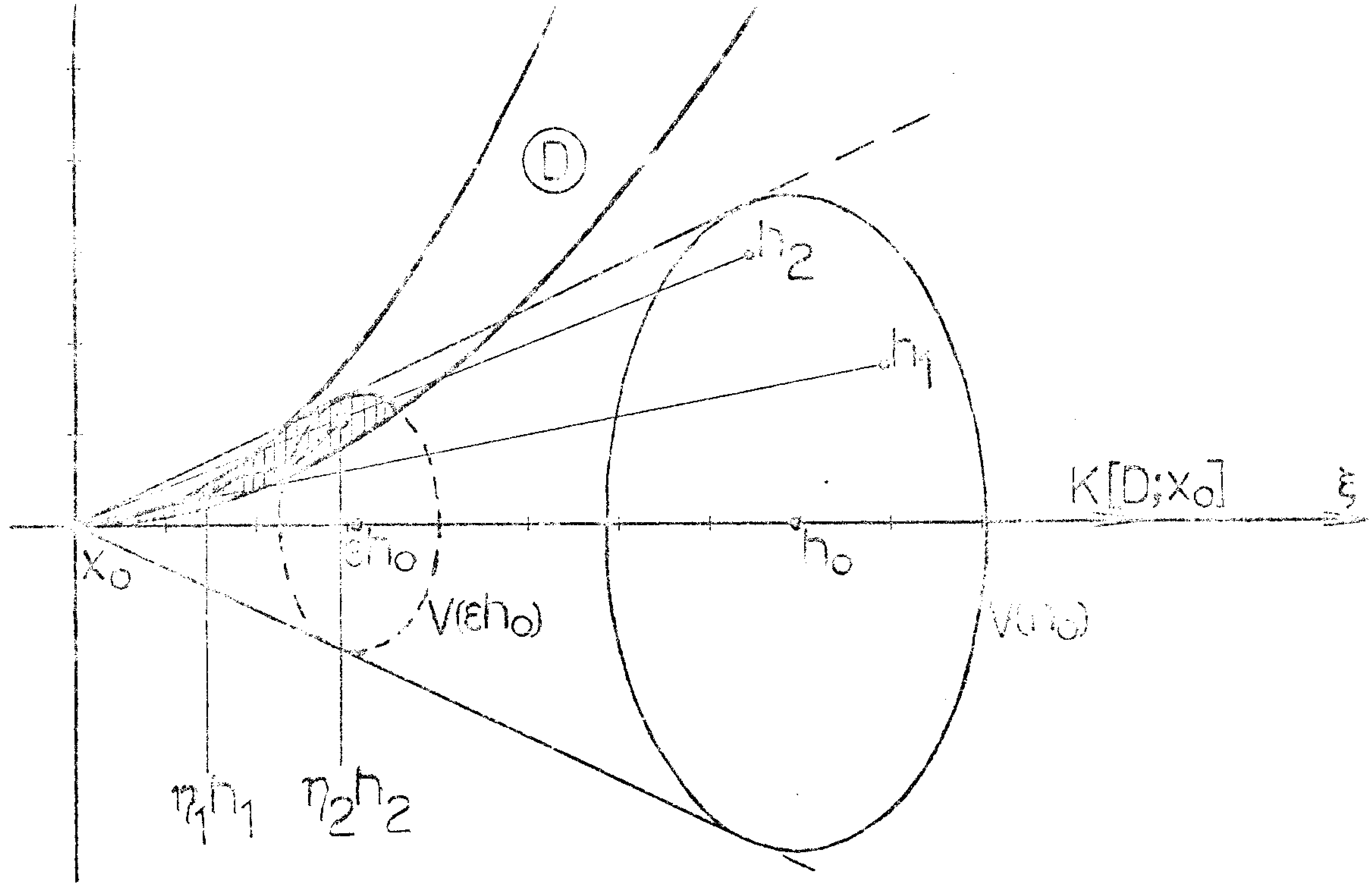
D
h_2
h_1
$K[D;x_o]$
ξ
x_o
εh_o
h_o
$V(\varepsilon h_o)$
$V(h_o)$
$\eta_1 h_1$
$\eta_2 h_2$

Fig. 9.1.

Similarly to the preceding example it may be shown that the set D is tangentially connected at a point $\Theta_{\mathbb{R}^2}$. We note here that D is not convex (in usual sense), Fig.9.1.

Fig. 9.1.

EXAMPLE 9.iii. Let $X := \mathbb{R}^2$ and

$$D := \{x := (\varphi, \rho) \in \mathbb{R}^2 \mid e^{-2\varphi} \leqq \rho \leqq e^{-\varphi}, \varphi \geqq 0\}.$$

Here we have

$$K[D; \Theta_{\mathbb{R}^2}] = \mathbb{R}^2.$$

The set D is not tangentially connected at a point $\Theta_{\mathbb{R}^2}$, Fig.9.2. We note that in last example the set D is connected (in usual sense). This shows that the notion of connectedness and tangential connectedness are independent.

Fig. 9.2.

DEFINITION 9.2. Let D be a set from X such that $K[D; x_0] \neq \emptyset$ and let $F: D \to Y$ be an operator. The operator F is t a n g e n t i a l l y s e p a r a b l e at a point $x_0 \in D$ if for every neighbourhood $\mathcal{V}(h_0)$, $h_0 \in K[D; x_0]$, and for any $\varepsilon > 0$ for which on the set

$$D \cap \{x_0 + \eta h \mid \forall \eta \in (0, \varepsilon), \forall h \in \mathcal{V}(h_0)\}$$

we have

$$F(x_0 + \eta h) \neq F(x_0), \tag{9.2}$$

there is a neighbourhood $\mathcal{W}(h_0) \subset \mathcal{V}(h_0)$ and there is a functional $y^* \in Y^*$, $y^* \neq \Theta_Y^*$ (which may depend on h_0 and $\mathcal{W}(h_0)$) as well as a positive number $\varepsilon^* (\leqq \varepsilon)$, such that on the set

$$D \cap \{x_0 + \eta h \mid \forall \eta \in (0, \varepsilon^*),\ \forall h \in \mathcal{W}(h_0)\}$$

the following is satisfied

(9.3) $$y^*[F(x_0 + \eta h)] \neq y^*[F(x_0)] .$$

EXAMPLE 9.iv. Let X and Y be BANACH spaces and let $F: X \to F(X) \subseteq Y$ be a linear continuous operator. Then F is a tangentially separable at any point $x_0 \in X$. Indeed, in this case $K[X; x_0] = X$ and on the basis of (9.2) we may write

(9.4) $$F(h) \neq \theta_Y \qquad \forall h \in \mathcal{V}(h_0)$$

where $\mathcal{V}(h_0)$ is an arbitrary convex neighbourhood of a point $h_0 \in X$. As $\theta_X \notin \mathcal{V}(h_0)$, the neighbourhood $\mathcal{V}(h_0)$ may be separated from θ_X, that is, there is a functional $x^* \in \in X^*$, $x^* \neq \theta_X^*$, so that

(9.5) $$x^*(h) > 0 \qquad \forall h \in \mathcal{V}(h_0).$$

On the other hand, if we put

$$L := \{x \in X \mid F(x) = \theta_Y\} =: \mathrm{Ker}\, F$$

and if X/L denotes the quotient-space, then F is invertible on X/L. Let Φ be the inverse of F on X/L. Then for every $h \in X/L$ we may write $h = \Phi(F(h))$ and by means of (9.5) we conclude that

(9.6) $$x^*[\Phi(F(h))] > 0 \qquad \forall h \in \mathcal{V}(h_0).$$

If Φ^* is the conjugate of Φ then the functional

(9.7) $$y^* := \Phi^* x^* \in Y^*$$

satisfies the inequality

$$y^*[F(h)] > 0 \qquad \forall h \in \mathcal{V}(h_0).$$

On the basis of (9.6) we see that the functional (9.7) can play the role of the functional which is sought.

EXAMPLE 9.v. If F acts from X in $\mathbb{R}$ i.e., if F is a functional, then we may take $y^* := 1$. In this case in Definition 9.2 we consider $\mathcal{W}(h_0) = \mathcal{V}(h_0)$ and $\varepsilon^* = \varepsilon$.

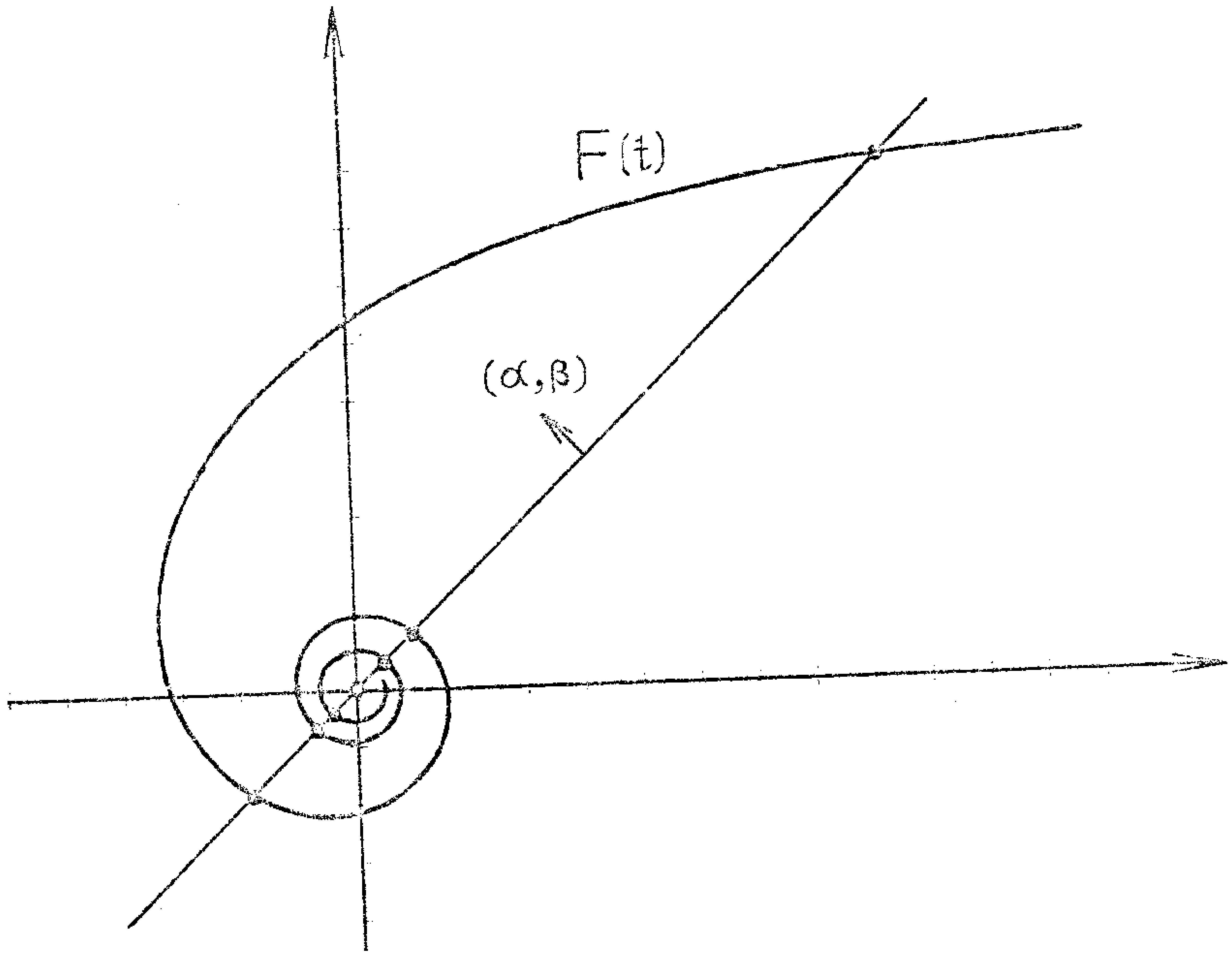

Fig. 9.3.

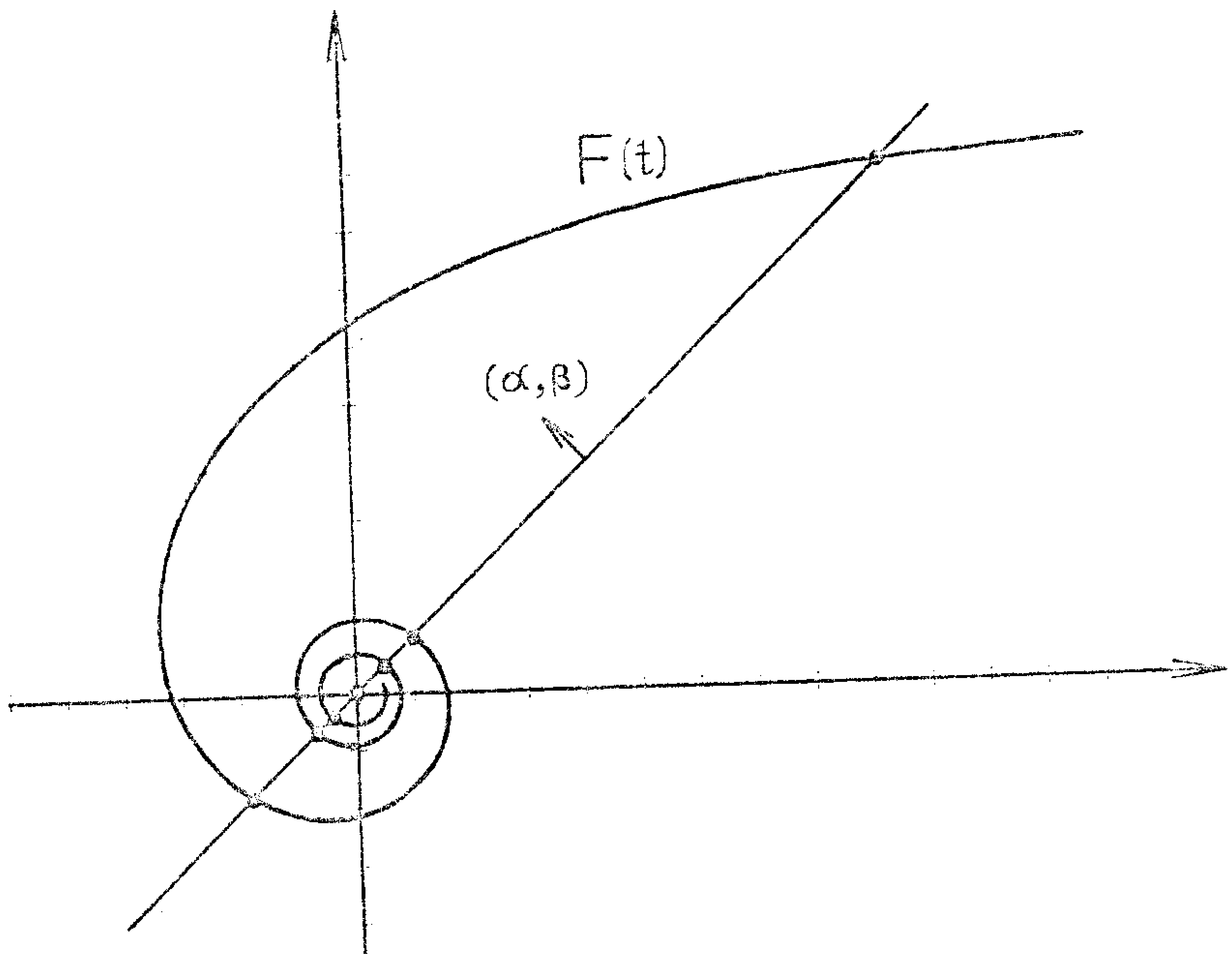

Fig. 9.3.

(10.1) $$M_o := \{ x \in X \mid F(x) = \theta_Y \}$$

the equality

(10.2) $$K[M_o;\ x_o] = \text{Ker } DF(x_o) = \{ h \in X \mid DF(x_o)(h) = \theta_Y \}$$

holds. This result plays a crucial role in the establishment of the multipliers rule, i.e., for the establishment of the necessary conditions for the optimum. In the case when $D \neq X$, x_o being a boundary point of D, the FRECHET differential $DF(x_o)$ cannot be defined; hence we cannot speak on an equality of the type (10.2).

We note here that if $x_o \in \overset{\circ}{D}$ and assuming that the extension of operator $F: D \to Y$ on X has the FRECHET differential $DF(x_o)$, then as follows from §6.(f), there also exists the derivative $F'_D(x_o)$ (with respect to the cone $K[D;\ x_o]$) and they are equivalent (on $K[D;\ x_o]$). But even in this case, for the formulas of inclusions of " $\supseteqq$ " type, stronger properties are required. For this reason we will study the cones of the sets defined by (8.1) in the case of tangential properties considered in §9.

DEFINITION 10.1. A operator $F: X \to Y$ is called s o l i d if any set from X with a nonvoid interior is transformed into a set from Y also with a nonvoid interior.

THEOREM 10.2. Let $D \subseteqq X$ be a given set tangentially connected at a point $x_o \in D$ and let $\text{int}(K[D;\ x_o]) \neq \emptyset$. Further let $F: D \to Y$ be a continuous operator which is tangentially separable at a point x_o and $F'_D(x_o): K[D;\ x_o] \to Y$ be solide. Then for the set

(10.3) $$M := \{ x \in D \mid F(x) = F(x_o) \}$$

we have

(10.4) $$K[M;\ x_o] \supseteqq \{ h \in \text{int}(K[D;\ x_o]) \mid F'_D(x_o)(h) = \theta_Y \}.$$

PROOF. We denote by T the set from the right hand side of (10.4). Let us suppose that the assertion is not true. Then there is a point $h_o \in T$, $h_o \notin K[M;\ x_o]$. This means that $\varepsilon_o > 0$ and that a neighbourhood $\mathcal{V}^{\varepsilon}(h_o)$ of h_o, exists so that

$$F'_D(x_o)(h_o) = \vartheta_Y,$$

$$x_o + \eta h \notin M \qquad \forall \eta \in (0,\varepsilon), \ \forall h \in V(h_o).$$

But the last inclusion implies

(10.5) $$F(x_o + \eta h) \neq F(x_o) \quad \forall \eta \in (0,\varepsilon_o), \forall h \in V(h_o).$$

On the basis of this, from the property of tangential separation of the operator F at x_o, we conclude that there exists a neighbourhood $W(h_o) \subset V(h_o)$, functional $y^* \in Y^*$, $y^* \neq \vartheta_Y^*$, and a positive number $\varepsilon^* \in (0, \varepsilon_o)$ such that on the set

$$D \cap \{ x_o + \eta h \mid \forall \eta \in (0, \varepsilon^*), \forall h \in W(h_o) \}$$

we have

(10.6) $$y^*[F(x_o + \eta h) - F(x_o)] \neq 0.$$

On the other hand by the tangential connectedness of the set D, a positive number $\varepsilon_1 = \varepsilon_1(h_o)$ can be attached to the neighbourhood $W_1(h_o)$ so that for every $\tilde{\varepsilon} \in (0, \varepsilon_1)$ the set

$$\Lambda_{\tilde{\varepsilon}} := D \cap \{ x_o + \eta h \mid \forall \eta \in (0,\tilde{\varepsilon}), \ \forall h \in W_1(h_o) \}$$

is nonvoid and convex. If $\bar{\varepsilon} = \min\{\varepsilon_1, \varepsilon^*\}$ then the set

$$\Lambda_{\bar{\varepsilon}} := D \cap \{ x_o + \eta h \mid \forall \eta \in (0,\bar{\varepsilon}), \ \forall h \in W_1(h_o) \}$$

is convex and the inequality (10.6) occurs. We will show that on the set $\Lambda_{\bar{\varepsilon}}$ the left-hand side expression from (10.6) preserves a constant sign. Indeed, assuming the contrary let

$$z_1 = x_o + \eta_1 h_1 \in \Lambda_{\bar{\varepsilon}}, \quad z_2 = x_o + \eta_2 h_2 \in \Lambda_{\bar{\varepsilon}}$$

be two points for which we have

(10.7) $$y^*[F(z_1) - F(x_o)] < 0 < y^*[F(z_2) - F(x_o)].$$

By means of the tangential connectedness of the set $\Lambda_{\bar{\varepsilon}}$ we can consider an auxilary continuous function φ with the following properties

(a) $$\varphi : [0,1] \longrightarrow \Lambda_{\bar{\varepsilon}},$$

(b) $$\varphi(0) = z_1, \qquad \varphi(1) = z_2$$

where z_1 and z_2 are the points considered above. Let Ψ

be another auxiliary function defined by

$$\psi(t) := y^*[F(\varphi(t)) - F(x_o)] \qquad t \in [0,1]$$

which is continuous and for which we have

$$\psi(0) = y^*[F(z_1) - F(x_o)] < 0,$$
$$\psi(1) = y^*[F(z_2) - F(x_o)] > 0.$$

Herefrom we conclude that there exists a value $t = \tau \in (0,1)$ such that $\psi(\tau) = 0$. But $\varphi(\tau) = x_o + \eta_\tau h_\tau = z_\tau \in \Lambda_\xi$.

In consequence, the equality

$$y^*[F(z_\tau) - F(x_o)] = 0$$

is valid and we notice a contradiction with (10.6).

To settle the ideas, we can consider that on the set Λ_ξ we have

(10.8) $$y^*[F(x_o + \eta h) - F(x_o)] > 0.$$

Indeed, let $h \in \mathrm{int}(K[D;\ x_o]) \cap \mathcal{W}_1^0(h_o)$ be an arbitrary point. It results from the inclusion $h \in \mathrm{int}(K[D;\ x_o])$ that there exists a generalized sequence $(\eta_i)\ i \in I$ of positive numbers and a generalized sequence $(h_i)\ i \in I$ of elements from X with $\eta_i \to 0$, $h_i \to h$ and $x_o + \eta_i h_i \in D$.

From (10.8) it follows

$$y^*[F'_D(x_o)(h)] = y^*\Big(\lim_{i \in I} \frac{1}{\eta_i}\,[F(x_o + \eta_i h_i) - F(x_o)]\Big) \geqq 0.$$

Hence we have for all $h \in \mathrm{int}(K[D;\ x_o]) \cap \mathcal{W}_1(h_o)$

(10.9) $$y^*[F'_D(x_o)(h)] \geqq 0.$$

On the other hand, from the inclusion $h_o \in T$ we get

(10.10) $$y^*[F'_D(x_o)(h)] = y^*(\theta_Y) = 0.$$

But from the fact that $F'_D(x_o)$ is a solide mapping it results that the set

$$\{F'_D(x_o)(h) \mid \forall h \in \mathrm{int}(K[D;\ x_o]) \cap \mathcal{W}_1(h_o)\}$$

has a nonempty interior. Hence, by relations (10.9) and (10.10) we conclude $y^* = \theta_Y^*$ which contradicts the assumtion that y^* is nontrivial. □

REMARK 10.3. The set $\operatorname{int}(K[D;\ x_o])$ cannot generally be replaced by $K[D;\ x_o]$; according to the following

EXAMPLE 10.i. Let $X:=\mathbb{R}^2$, $D:=\{(x,y)\in\mathbb{R}^2 \mid x\geqq 1, y\geqq 0\}$, $F((x,y)) := x^2+y^2$, $x_o:=(x,y)_o:=(1,0)$, and we consider the set

$$M:=\{(x,y)\in D \mid F((x,y)) = F((1,0)) = 1\} = \{(1,0)\}.$$

Then we have

$$K[M;\ x_o] = \{(0,0)\},$$
$$K[D;\ x_o] = \{h:=(\xi,\eta)\in\mathbb{R}^2 \mid \xi\geqq 0,\ \eta\geqq 0\}.$$

Further

$$F_D'(x_o)(h) = 2\xi + 0\cdot\eta = 2\xi .$$

The mapping $F_D'(x_o) : K[D;\ x_o]\to\mathbb{R}$ is solid. The vectors $z:=(0,\eta)$ with $\eta\geqq 0$ belong to cone $K[D;\ x_o]$ and satisfy the equality $F_D'(x_o)(z) = 0$. But for $\eta>0$ these vectors do not belong to the cone $K[M;\ x_o]$.

REMARK 10.4. If $D:=X$ then we have $\operatorname{int}(K[D;\ x_o]) = K[D;\ x_o] = X$ which is always occurring, and we have (HOFFMANN & KOLUMBAN [1974]):

$$K[M;\ x_o]\subseteq\{h\in X \mid F_D'(x_o)(h) = \theta_Y\}.$$

On the other hand from (10.4) in this case we obtain

$$K[M;\ x_o]\supseteq\{h\in X \mid F_D'(x_o)(h) = \theta_Y\}.$$

Finally it results

$$K[M;\ x_o] = \{h\in X \mid F_D'(x_o)(h) = \theta_Y\}.$$

CHAPTER III

THE EXTREMUM OF A FUNCTIONAL

11. ON NECESSARY CONDITIONS

Formulas with inclusion of "$\supseteq$" type for admissible and adherent cones are of interest for expressing the cones of union or intersection by means of the corresponding cones of sets of participants. Another application of such formulas is in determination of necessary conditions for the extremum.

Let (S_i) $i \in I$ be an arbitrary set of sets from linear space X. Besides inclusions given in § 5 (g), the following can be considered, too

$$\text{(i) } K(\bigcap_{i\in I} S_i;\, x_o) \supseteq \bigcap_{i\in I} K(S_i;\, x_o), \quad \text{(ii) } K[\bigcup_{i\in I} S_i;\, x_o] \subseteq \bigcup_{i\in I} K[S_i:\, x_o],$$

$$\text{(iii) } K(\bigcup_{i\in I} S_i;\, x_o) \supseteq \bigcup_{i\in I} K(S_i;\, x_o), \quad \text{(iv) } K[\bigcap_{i\in I} S_i;\, x_o] \supseteq \bigcap_{i\in I} K[S_i;\, x_o].$$

Using formulas § 5 (d),(e) as well as DE MORGAN's formulas, the study of implications (ii) and (iii) could be reduced to the study of inclusions (i) and (iv) respectively. So, for example, if we put $T_i := \complement_X S_i$, i.e., $S_i = \complement_X T_i$ in (ii) and proceed to complements, we have

$$\text{(ii)} \Rightarrow K[\bigcup_{i\in I}(\complement_X T_i);\, x_o] \subseteq \bigcup_{i\in I} K[\complement_X T_i;\, x_o]$$

$$\Rightarrow K[\complement_X(\bigcap_{i\in I} T_i);\, x_o] \subseteq \bigcup_{i\in I} K[\complement_X T_i;\, x_o]$$

$$\Rightarrow \complement_X K[\complement_X(\bigcap_{i\in I} T_i);\, x_o] \supseteq \bigcap_{i\in I} \complement_X K[\complement_X T_i;\, x_o]$$

$$\Rightarrow K(\bigcap_{i\in I} T_i;\, x_o) \supseteq \bigcap_{i\in I} K(T_i;\, x_o)$$

$$\Rightarrow \text{(i)}.$$

In that sense the inclusions (i) and (ii), (iii) and (iv) are dual as well as $(*)$ and $(^*_*)$, $(_*{}^*{}_*)$ and $(^{**}_{**})$ from § 5 (g).

Therefore out of the 8 mentioned inclusions, it is sufficient to study four, which are independent i.e., non-dual. We shall consider the inclusions (i) and (iv).

Comparing (i)-(iv) with those from §5 (g) we conclude that when the rigorous sign holds in (*), then (i) is not valid. This case can occur only if the set of indices I is infinite. Further, since in (**) the equality sign must not be valid even in the case when I is a finite and therefore (iv) can be wrong even when I is a finite. We shall quote a characteristic example, quite ubiquitous in the practice.

EXAMPLE 11.1. Let X be a HILBERT space and let

$$Q_1 := \{ x \in X \mid \|x + a\| \leq \|a\| \},$$

$$Q_2 := \{ x \in X \mid \|x - a\| \leq \|a\| \},$$

a ($\neq \theta_X$) being a given fixed point.

Since the functional $\rho := \|\cdot + a\|$ is differentiable in directions, on the basis of §6 (c) we can write

$$\rho'_{Q_1}(\theta_X)(h) = \rho'(\theta_X)(h) := \lim_{\alpha \downarrow 0} \frac{\rho(\theta_X + \alpha h) - \rho(\theta_X)}{\alpha}$$

$$= \lim_{\alpha \downarrow 0} \frac{\|\alpha h + a\| - \|a\|}{\alpha}$$

$$= \lim_{\alpha \downarrow 0} \frac{\|\alpha h + a\|^2 - \|a\|^2}{\alpha(\|\alpha h + a\| + \|a\|)}$$

$$= \lim_{\alpha \downarrow 0} \frac{\alpha^2 \|h\|^2 + 2\alpha \langle a, h \rangle}{\alpha(\|\alpha h + a\| + \|a\|)}$$

$$= \frac{\langle a, h \rangle}{\|a\|} .$$

Analogously, if we put $\varphi := \|\cdot - a\|$ we get

$$\varphi'_{Q_2}(\theta_X)(h) = \varphi'(\theta_X)(h) = \frac{\langle -a, h \rangle}{\|a\|} .$$

Since Q_1 and Q_2 are convex sets, on the basis of §5 (f) and Theorem 8.1 we get

$$K[Q_1; \theta_X] = \operatorname{adh} K(Q_1 : \theta_X) =$$

$$= \{ h \in X \mid \rho'_{Q_1}(\theta_X)(h) \leqq 0 \}$$

$$= \{ h \in X \mid \langle a,h \rangle \leqq 0 \}.$$

Analogously,

$$K[Q_2 ; \theta_X] = \{ h \in X \mid \langle a,h \rangle \geqq 0 \}.$$

Therefore, it follows

$$K[Q_1 ; \theta_X] \cap K[Q_2 ; \theta_X] = \{ h \in X \mid \langle a,h \rangle = 0 \}$$

while, on the other hand $Q_1 \cap Q_2 = \theta_X$ and the cone $K[Q_1 \cap \cap Q_2 ; \theta_X]$ is an empty set; so that (iv) is not valid.

Now, Theorem 7.1 gives the necessary condition for minimum of the functional f on the set $A := (\bigcap_{i \in I} Q_i) \cap \cap (\bigcap_{j \in J} M_j)$. This condition is expressed by equality (7.6) wherein the cones $K(P; x_o)$, $K(\bigcap_{i \in I} Q_i ; x_o)$ and $K[\bigcap_{j \in J} M_j ; x_o]$ explicitly appear. Equality (7.6) can be substituted by a more general

$$K(P; x_o) \cap K[A; x_o] = \emptyset . \tag{11.1}$$

This is the most general form of the necessary condition for the extremum of the functional, obtained in the present work. In the case when

$$K[A; x_o] \supseteqq (K(\cap Q_i ; x_o)) \cap (K[\cap M_j ; x_o]), \tag{11.2}$$

from (11.1) clearly follows (7.6). But our aim is to substitute (11.1) by a formula where the cones of each set-restriction would explicitly appear. This is motivated by the fact that the cones of some sets-restrictions could be, sometime, more easily determined then the cone $K[A;x_o]$. In this case if the operator $F: D \to Y$ and a set

$$Q := \{ x \in D \mid F(x) \leqq F(x_o) \}$$

satisfy conditions from Theorems 8.1 and 8.3, then we get

$$K(Q; x_o) = \{ h \in K(D; x_o) \mid F'_Q(x_o)(h) = \theta_Y \}. \tag{11.3}$$

Analogously, if the set D and the operator $G: D \to Z$ satisfy the conditions from Theorems 8.4 and 10.2, then we have

$$\begin{aligned} \{ h \in K[D; x_o] \mid G'_M(x_o)(h) = \theta_Z \} \supseteqq K[M; x_o] \supseteqq \\ \supseteqq \{ h \in \operatorname{int}(K[D; x_o]) \mid G'_D(x_o)(h) = \theta_Z \} \end{aligned} \tag{11.4}$$

In Example 10.i, quoted in §10, we had $F'_M(x_o)(h) = F'_D(x_o)(h) = 2\xi$ and $K[M; x_o] = \{(0,0)\}$. On the other hand we have

$$\{h \in K[D; x_o] \mid F'_M(x_o)(h)=0\} = \{h:=(\xi,\eta)\in \mathbb{R}^2 \mid \xi \geq 0, \eta \geq 0, 2\xi=0\} = \{(\xi,\eta)\in \mathbb{R}^2 \mid \xi=0,\ \eta\geq 0\}$$

and

$$\{h \in \mathrm{int}(K[D; x_o]) \mid F'_D(x_o)(h)=0\} = \{h:=(\xi,\eta)\in \mathbb{R}^2 \mid \xi > 0, \eta > 0, 2\xi=0\} = \emptyset.$$

This example shows that in the general case in (11.4) neither inclusion reduces to equality.

On the basis of (11.4) various conditions can be studied, enabling the writing of an exact formula for cone $K[M; x_o]$.

THEOREM 11.1. Let the set D and the operator G satisfy all conditions from Theorems 8.4 and 10.2 and let the equality

$$\text{(11.5)} \quad \mathrm{adh}(\{h \in \mathrm{int}(K[D; x_o]) \mid G'_D(x_o)(h) = \theta_Z\}) = \{h \in K[D; x_o] \mid G'_D(x_o)(h) = \theta_Z\}$$

hold. Then

$$\text{(11.6)} \quad K[M; x_o] = \{h \in K[D; x_o] \mid G'_D(x_o)(h) = \theta_Z\}.$$

PROOF. According to Definition 6.1, the derivatives $G'_D(x_o)(h)$ and $G'_M(x_o)(h)$ represent relative limit-values with respect to sets D and M respectively. Since $D \supseteq M$ and since $G'_D(x_o)(h)$ exists, $G'_M(x_o)(h)$ exists also and is equal to the former, (see for example ALJANČIĆ [1968, pp. 69-70]). On the basis of (11.4) and (11.5) and using the above statement equality (11.6) may be inferred. □

Let us mention a simple, but frequent case: (11.5) is fulfilled when for example, the set on the right hand side in (11.4) is convex, having a nonempty interior, (see for example LAURENT [1972, Proposition 1.3.5]).

THEOREM 11.2. (i) Let the following conditions be fulfilled:

1° The set D ($\subseteq X$) is tangentially connected at a point x_o and $\mathrm{int}(K[D; x_o]) \neq \emptyset$,

2° For every $j \in J$ the operator G_j is continuous and tangentially separable at a point x_o,

3° For every $j \in J$ the operator $G'_{jD}(x_o) : K[D;x_o] \to Z_j$ is solide.

Then we have

$$K[\bigcap_{j\in J} M_j;\ x_o] \supseteq \bigcap_{j\in J}\{h \in \operatorname{int}(K[D;\ x_o] \mid G'_{jD}(x_o)(h) = \theta_{Z_j}\}. \tag{11.7}$$

(ii) If, besides the quoted conditions, following equality holds

$$\begin{aligned}&\operatorname{adh}(\bigcap_{j\in J}\{h \in \operatorname{int}(K[D;x_o])\mid G'_{jD}(x_o)(h) = \theta_{Z_j}\}) = \\ &= \bigcap_{j\in J}\{h \in K[D;\ x_o] \mid G'_{jD}(x_o)(h) = \theta_{Z_j}\},\end{aligned} \tag{11.8}$$

then we have

$$K[\bigcap_{j\in J} M_j;\ x_o] = \bigcap_{j\in J} K[M_j;\ x_o]. \tag{11.9}$$

PROOF. Put $Z := \prod_{j\in J} Z_j$ and $G := (G_j)\ j \in J$. Then we get

$$M := \bigcap_{j\in J} M_j = \{x \in D \mid G(x) = G(x_o)\}.$$

It is easy to see that all conditions for Theorem 10.2 are fulfilled; and the proof of (i) immediately stems from this theorem. The proof of (ii) results on the basis of Theorem 11.1. □

When we are not formulas of the type (11.3) and (11.6), then we shall be satisfied by a formula (11.2).

Let x_o be a point of the minimum of functional f on a set A, already defined. In §5 (c) we saw that for the set-restriction Q_i for which $x_o \in \mathring{Q}_i$, we have $K(Q_i;\ x_o) = X$. Therefore this cone is not essential for decomposition of the cone $K[A;\ x_o]$ through the cone of component subsets.

Let

$$I_o := \{i \in I \mid x_o \in \bar{Q}_i \setminus \mathring{Q}_i\}.$$

We endeavour to determine the conditions under which the following will be valid

$$K[A;\ x_o] \supseteq (\bigcap_{i\in I_o} K(Q_i;\ x_o)) \cap (\bigcap_{j\in J} K[M_j;\ x_o]). \tag{11.10}$$

The set of restrictive conditions Q_i and M_j (defined by (7.5)) satisfying the inclusion (11.10) in HOFFMANN & KOLUMBAN [1974] is said to constitute a *regular* set-restriction in the optimization problem which is under consideration. Such a result was obtained in the above mentioned paper ([Example 25]). The properties of tangential connectedness and separatness enable us to give various results for a regularity of a set-restriction. Here we will not insist on such formulas.

In practice, the sets-restrictions are frequently finite. That is why it will be of interest to mention that when I is a finite set of indices and $J := 1$, the set-restriction is regular. The proof is easily derived on the basis of the mentioned fact that the formula (i) in §11 is valid. Indeed, let $h_o \in (\bigcap_{i \in I} K(Q_i; x_o)) \cap K[M; x_o]$. Then on the basis of the Definition 5.1, for every $i \in I$ there is a neighbourhood $\mathcal{V}^{Q(i)}(h_o)$ and there is a positive number $\varepsilon^{(i)}$, such that for all numbers $\eta^{(i)} \in (0, \varepsilon^{(i)})$ we have $x_o + \eta^{(i)} h \in Q_i$. Put $\mathcal{V}^{Q*}(h_o) := \bigcap_{i \in I} \mathcal{V}^{Q(i)}(h_o)$ and $\varepsilon^* := \min_{i \in I} \varepsilon^{(i)}$. Then we get

$$x_o + \eta h \in \bigcap_{i \in I} Q_i \qquad \forall h \in \mathcal{V}^{Q*}(h_o), \forall \eta \in (0, \varepsilon^*).$$

On the other hand, from $h_o \in K[M; x_o]$, on the basis of Definition 5.2 we conclude that there exists a vector $h_{\varepsilon^*} \in \mathcal{V}^{Q*}(h_o)$ and there exists a number $\eta_{\varepsilon^*} \in (0, \varepsilon^*)$ such that $x_o + \eta_{\varepsilon^*} h_{\varepsilon^*} \in M$, i.e., $x_o + \eta_{\varepsilon^*} h_{\varepsilon^*} \in (\bigcap_{i \in I} Q_i) \cap M$. But, this is equivalent to $h_o \in K[(\bigcap_{i \in I} Q_i) \cap M; x_o]$.

Now, let X be a topological linear space and $\mathbb{T} := (\mathbb{T}, \leqq)$ a semiordered topological linear space such that $\{t \in \mathbb{T} \mid t < \Theta_{\mathbb{T}}\}$ is an open set in $\mathbb{T}$. A set A ($\subset D$) and a point $x_o \in A$ are given, D ($\subseteq X$) being any super-set of A. The following definition will be necessary.

DEFINITION 11.3. A continuous operator $f: D \to \mathbb{T}$ is called *h-uniformly differentiable with respect to a set* A (precisely with respect to the cone $K[A; x_o]$), if for every $h \in K[A; x_o]$ and for every neighbourhood $\mathcal{W}(f'_A(x_o)(h))$, there exists an index

$i_0 := i_0(h, \mathcal{W}(\theta_{\mathbb{T}}))$ such that for every pair of sequences (x_i) $i \in I$, (r_i) $i \in I$ with $x_i \in A$, $r_i \geqq 0$, $x_i \to x_0$, $r_i(x_i - x_0) \to h$, we have

$$r_i[f(x_i) - f(x_0)] \in \mathcal{W}(f'_A(x_0)(h)) \quad \forall i > i_0.$$

THEOREM 11.4. Let f be a h-uniformly differentiable with respect to the set A and let $f'_A(x_0)$: $K[A; x_0] \to \mathbb{T}$ be a continuous operator. Let $x_0 \in A$ be a local minimal element for f on A. Then we have

$$\min_{h \in K[A;x_0]} f'_A(x_0)(h) = \theta_{\mathbb{T}}. \tag{11.11}$$

PROOF. We suppose that the assumption of the theorem is not true. Then there exists $h_0 \in K[A; x_0]$ such that for some $t_0 \in \mathbb{T}$, $t_0 > \theta_{\mathbb{T}}$ we have

$$f'_A(x_0)(h_0) = - t_0 < \theta_{\mathbb{T}}.$$

Using the continuity of $f'_A(x_0)$ we may find a neighbourhood $\mathcal{V}(h_0)$ such that

$$f'_A(x_0)(h) < -\frac{t_0}{2} \qquad \forall h \in \mathcal{V}(h_0).$$

On the basis of h-uniform differentiability of f, we conclude that for h_0 and for $t_0/4$ there exists an index $i_0 \in I$ such that for every pair of sequences (x_i) $i \in I$, (r_i) $i \in I$ with $x_i \in A$, $r_i \geqq 0$, $x_i \to x_0$, $r_i(x_i - x_0) \to h_0$, we have

$$\begin{aligned} r_i[f(x_i) - f(x_0)] &< f'_A(x_0)(h_0) + \frac{t_0}{4} \\ &< -\frac{t_0}{2} + \frac{t_0}{4} \\ &< \theta_{\mathbb{T}} \end{aligned}$$

for every $i > i_0$. But the last inequality shows that

$$f(x_i) < f(x_0) \qquad \forall i > i_0,$$

which is contradictory to the fact that x_0 is a local minimal point. □

On the basis of the tangential properties, in the following section we shall give the principle of extremum which presented a necessary condition for extremum, also.

12. THE RULE OF THE MULTIPLIERS

Let us consider now the optimization problem defined in §7. The point $x_o \in A$ must be determined such that

$$f(x_o) = \min_{x \in A} f(x).$$

THEOREM 12.1. Let $x_o \in A$ be a minimal element and let the conditions from Theorem 10.2 be satisfied for D and F. We suppose that $f'_D(x_o)$ and $G'_D(x_o)$ are existing and continuous. Then the following system is incompatible

$$\begin{aligned} &f'_D(x_o)(h) < 0,\\ &F'_D(x_o)(h) = \theta_Y,\\ (12.1)\qquad &G(x_o) + G'_D(x_o)(h) \in \overset{\circ}{C},\\ &h \in \operatorname{int}\,(K[D;\ x_o]). \end{aligned}$$

PROOF. We suppose that a vector $h \in \operatorname{int}(K[D;\ x_o])$, $h \neq \theta'_X$ with $F'_D(x_o)(h) = \theta_Y$, exists. Then by Theorem 10.2, more precisely by (10.4), we have $h \in K[M;\ x_o]$, where M is given by (10.3). Consequently a generalized sequence (x_i) $i \in I$ of the points from D such that $F(x_i) = F(x_o) = \theta_Y$ and a generalized sequence (r_i) $i \in I$ of positive numbers such that $x_i \to x_o$, $r_i(x_i - x_o) \to h$, exist. On the other hand from the inclusion $G(x_o) + G'_D(x_o)(h) \in \overset{\circ}{C}$ results that there exists the index i_1 such that for every $i > i_1$ we have

$$G(x_o) + r_i\{G(x_i) - G(x_o)\} \in \overset{\circ}{C}.$$

But on the basis of the convexity of the set C we have

$$G(x_i) = (1 - \frac{1}{r_i})G(x_o) + \frac{1}{r_i}\left\{G(x_o) + r_i(G(x_i) - G(x_o))\right\} \in \overset{\circ}{C}$$

and for $i > i_1$ we deduce

$$x_i \in \{x \in D \mid F(x) = \theta_Y\} \cap \overset{\circ}{C} \subset A.$$

If h also satisfies the inequality $f'_D(x_o)(h) < 0$, then there exists the index $i_o (> i_1)$, such that for every $i > i_o$ we have

$$r_1\{f(x_1) - f(x_0)\} < 0,$$

which shows that x_0 cannot be an optimal point.

DEFINITION 12.2. Let S be an arbitrary convex set. If for every $h_1, h_2 \in K[D; x_0]$ and all $\alpha \in [0,1]$, for the operator $\mathcal{G}$ we have

$$\alpha\, \mathcal{G}(h_1) + (1-\alpha)\, \mathcal{G}(h_2) - \mathcal{G}(\alpha h_1 + (1-\alpha) h_2) \in S$$

respectively

$$\mathcal{G}(\,h_1 + (1-\alpha)h_2) - \alpha\, \mathcal{G}(h_1) - (1-\alpha)\, \mathcal{G}(h_2) \in S$$

then we say that $\mathcal{G}$ is c o n v e x respectively c o n c a v e on the set S.

THEOREM 12.3. Put

$$S := \bigcup_{\lambda \geq 0} \{\lambda(u - G(x_0)) \mid u \in C\}$$

and let $x_0 \in A$ be an optimal element. We suppose that the following properties are fulfilled:

1^o The set D is tangentially connected at a point x_0 and the set $\mathrm{int}(K[D; x_0])$ is nonempty and convex,

2^o The operator F is continuous and tangentially separable at a point x_0 and the derivative $F'_D(x_0)$ is a solide mapping from X onto Y,

3^o The operators $f'_D(x_0)$ and $G'_D(x_0)$ exist, they are continuous, $f'_D(x_0)$ is convex on $\mathbb{R}$ and $G'_D(x_0)$ is concave on S.

Then there exists a nonnegative number $\mu\ (\geq 0)$ and the functionals $y^* \in Y^*$ and $z^* \in Z^*$, which, when $\mu = 0$ are not equal to zero-functionals Θ^*_Y respectively Θ^*_Z, so that the inequalities

$$\mu f'_D(x_0)(h) + y^*[F'_D(x_0)(h)] + z^*[G'_D(x_0)(h)] \geq 0,$$

$$z^*(z) \leq z^*[G(x_0)]$$

for every $h \in \mathrm{int}(K[D; x_0])$ and for all $z \in C$ hold.

PROOF. We consider the sets

$$P := \{(r, F'_D(x_0)(h), z) \in \mathbb{R}\times Y\times Z \mid h \in \mathrm{int}(K[D; x_0]),$$

$$r \geq f'_D(x_0)(h),\ G'_D(x_0)(h) - z \in S\},$$

(12.1)

$$Q := \{(\varrho, \theta_Y, \zeta) \in \mathbb{R}\times Y\times Z \mid \varrho < 0,$$
$$\zeta \in \{\lambda(u - G(x_o)) \mid \lambda > 0,\ u \in \mathring{C}\}\}$$

If $(r_1, F'_D(x_o)(h_1), z_1)$ and $(r_2, F'_D(x_o)(h_2), z_2)$ are two elements of P, then by convexity of $f'_D(x_o)$ we have

$$f'_D(x_o)(\theta h_1 + (1-\theta)h_2) \leqq \theta f'_D(x_o)(h_1) + (1-\theta)f'_D(x_o)(h_2)$$
$$\leqq \theta r_1 + (1-\theta) r_2$$
$$r_1, r_2 \in \mathbb{R}, \quad 0 \leqq \theta \leqq 1.$$

On the other hand, taking into account that $G'_D(x_o)$ is concave on S as well as by convexity of the set S, we have

$$\theta G'_D(x_o)(h_1) - \theta z_1 = \theta s_1 \in S,$$
$$(1-\theta) G'_D(x_o)(h_2) - (1-\theta) z_2 = (1-\theta) s_2 \in S$$
$$s_1, s_2 \in S, \quad 0 \leqq \theta \leqq 1.$$

Further we have

$$G'_D(x_o)(\theta h_1 + (1-\theta)h_2) - (\theta z_1 + (1-\theta)z_2) =$$
$$= G'_D(x_o)(\theta h_1 + (1-\theta)h_2) - \theta G'_D(x_o)(h_1) -$$
$$- (1-\theta) G'_D(x_o)(h_2) + (\theta s_1 + (1-\theta)s_2) =$$
$$= s_3 + (\theta s_1 + (1-\theta)s_2)$$
$$S,$$

because $s_3 \in S$. Since $\theta h_1 + (1-\theta)h_2 \in \operatorname{int}(K[D; x_o])$, we draw the conclusion that the set P is convex. It is easily seen that $\mathring{P} \neq \emptyset$ and that Q also is a convex set. Let us show that $P \cap Q = \emptyset$ occurs too. Indeed, assume contrary and for $h_1 \in \operatorname{int}(K[D; x_o])$ let

$$(r, F'_D(x_o)(h_1), z) = (\varrho, \theta_Y, \zeta) \in P \cap Q,$$

i.e.,

$$r = \varrho, \quad F'_D(x_o)(h_1) = \theta_Y, \quad z = \zeta.$$

Then we have

we have

$$f'_D(x_o)(h_1) \leqq r = \varrho < 0$$

i.e.,

$$G'_D(x_o)(h_1) - \lambda_1(u_1 - G(x_o)) = \lambda_2(u_2 - G(x_o)),$$
$$\lambda_1 > 0, \quad \lambda_2 \geqq 0, \quad u_1 \in \overset{\circ}{C}, \quad u_2 \in C.$$

So we conclude

$$G(x_o) + G'_D(x_o)(h_1) = G(x_o) + \lambda_1(u_1 - G(x_o)) + \lambda_2(u_2 - G(x_o)) \in$$
$$\in \overset{\circ}{S}.$$

From the last relation it follows that a positive number $\alpha >$ > 0 can be found such that

$$G(x_o) + G'_D(x_o)(\alpha h_1) \in \overset{\circ}{C}.$$

But then the equality $F'_D(x_o)(\alpha h_1) = \theta_Y$ is also satisfied as well as the inequality $f'_D(x_o)(\alpha h_1) < 0$. This shows that the vector αh_1 satisfies all conditions (12.1), which contradicts the assumption that x_o is an optimal point. Consequently, the sets P and Q from (12.2) can be separated, i.e., there exists a linear functional $(\mu, y^*, z^*) \in$ $\in (\mathbb{R} \times Y \times Z)^*$ such that

(12.3) $\mu r + y^*[F'_D(x_o)(h)] + z^*[G'_D(x_o)(h)] \geqq 0 \geqq \mu\varrho + z^*(\zeta)$

for all $(r, F'_D(x_o)(h), z) \in P$ and for all $(\varrho, \theta_Y, \zeta) \in Q$.
From the right side of (12.3) we conclude that $\mu \geqq 0$, since in the contrary case for $|\varrho|$ sufficiently great the inequality is no more true. Similarly from the second part of the inequality from (12.3) we deduce that

$$z^*(z) \leqq z^*[G(x_o)]$$

because the hyperplane of the separation passes through the common vertex of the cones P and Q. This is justified by the fact that if we suppose the reverse inequality, then for $\lambda > 0$ sufficiently great, from $z^*[\lambda(u - G(x_o))] > 0$ we come at a contradiction.

Among the numerous particular cases we mention here only those which are contained in the papers HOFFMANN & KOLUMBAN [1974], LEMPIO [1972], DUBOVIZKI & MILUTIN [1965], LAURENT [1972].

13. ON SUFFICIENT CONDITIONS

Let X, $\mathbb{T}$, $x_0 \in A \subseteq D \subseteq X$ be the symbols appearing in 11, and let A be a tangentially connected set at a point x_0.

DEFINITION 13.1. The derivative $f'_A(x_0)$ is called s i g n- r e g u l a r at a point x_0, if for every $h \in \in K[A;\ x_0]$ and every pairs of sequences (x_i) $i \in I$, (r_i) $i \in I$, $x_i \in A$, $r_i \geq 0$, with $x_i \to x_0$, $r_i(x_i - x_0) \to h$, there exists an index $i_0 := i_0(h) \in I$ such that

$$r_i[f(x_i) - f(x_0)] \geq f'_A(x_c)(h) \qquad \forall i > i_0.$$

EXAMPLE 13.i. Let A be tangentially connected at a point x_0 and $f\colon D \to \mathbb{T}$ be a differentiable with respecr to A and tangentially separable at x_0. Then either $f'_A(x_0)$ or $-f'_A(x_0)$ is sign-regular at x_0. Particulatly, in Example 9.v. a criteria for tangential separability of a functional $f\colon$ $\colon X \to \mathbb{R}$ was given by (9.2). It is a classical result that a convex functional f, defined on a convex set D from the normed linear space, is differentiable (in the sense of GATEAUX). Since $f'(x_0)$ is tangentially separable, we see that $f'(x_0)(h)$ is sign-regular at a point x_0.

THEOREM 13.2. Let the following conditions be satisfied:

1^o There exists a point $x_0 \in A$ such that $\operatorname{int}(K[A;$ $;\ x_0])$ is convex, nonvoid and

$$\min_{h \in K[A;x_0]} f'_A(x_0)(h) = \vartheta_{\mathbb{T}},$$

2^o $f'_A(x_0)$ is sign-regular at x_0,

3^o f is convex on a cone $K[A;x_0]$,

4^o The set A is tangentially connected at x_0,

5^o There exists a neighbourhood $\mathcal{V}(x_0)$ of x_0 such that

$$(*) \qquad A \cap \mathcal{V}(x_0) \subseteq A \cap (x_0 + K[A;\ x_0]).$$

Then f attains a local minimum at a point x_0.

If instead of $(*)$ we suppose

$$(\ast\ast) \qquad A \subseteq x_0 + K[A;\ x_0],$$

then f has a global minimum (on the set A) at x_0.

PROOF. Let us assume that the statement of the theorem is not true. Then for a point $\tilde{x} \in A \cap \mathcal{V}(x_0)$ the following is valid: $f(\tilde{x}) < f(x_0)$. On the basis of (*) we have

$$\tilde{h} := \tilde{x} - x_0 \in K[A;\, x_0].$$

Having in view 2^0, the index $i_0 \in I$ exists such that for every pair of sequences (x_i) $i \in I$, (r_i) $i \in I$, $x_i \in A$, $r_i \geq 0$, with $x_i \to x_0$, $r_i(x_i - x_0) \to \tilde{h}$, we have

$$f_A'(x_0)(\tilde{h}) < r_i[f(x_i) - f(x_0)] \qquad \forall i > i_0. \tag{13.1}$$

On the other hand, due to the continuity of f we may find a neighbourhood $\mathcal{V}(\tilde{x})$ of $\tilde{x}$, such that for $t > 0$, we have

$$f(x) < f(x_0) - t \qquad \forall x \in \mathcal{V}(\tilde{x}) \cap K[A;\, x_0].$$

Using convexity of $K[A;\, x_0]$ we may find an index $\tilde{i} \in I$ such that

$$x_i =: \alpha_i x_0 + (1-\alpha_i)x \qquad \forall i > \tilde{i}, \quad \forall x \in \mathcal{V}(\tilde{x}) \cap K[A;x_0],\ \alpha_i \in (0,1).$$

On the basis of convexity of f we have

$$f(x_i) = f(\alpha_i x_0 + (1-\alpha_i)x) \leq f(x_0) + (1-\alpha_i)f(x)$$

for every $i > \tilde{i}$ and every $x \in \mathcal{V}(\tilde{x}) \cap K[A;x_0]$. If we put $i^* := \min\{i_0, \tilde{i}\}$, we get

$$r_i[f(x_i) - f(x_0)] \leq r_i[\alpha_i f(x_0) + (1-\alpha_i)f(x) - f(x_0)] =$$
$$= r_i(1-\alpha_i)[f(x) - f(x_0)] \leq r_i(1-\alpha_i)[f(x_0) - t - f(x_0)] =$$
$$= - r_i(1-\alpha_i)t.$$

But, combining this with (13.1) we conclude that

$$f_A'(x_0)(\tilde{h}) < - r_i(1-\alpha_i)t \qquad \forall i > i^*,$$

which is contrary to (*). The proof will be finished by means of the Theorem 11.4.

14. CONVEX PROBLEMS

Let A be a convex set. On the basis of §5 (f), we conclude that $K(A;\, x_0)$ and $K[A;\, x_0]$ are convex. Then $K(A;x_0) \supseteq A$ and $K[A;x_0] \supseteq A$ are valid also.

Let $f\colon X \to \mathbb{R}$ be a convex functional, X being a nor-

med space. It is known that $f'(x_0)$ is a positive-homogeneous subaditive functional. The classical result is:
$\frac{1}{\alpha}[f(x_0+\alpha h)-f(x_0)] - f'(x_0)(h) \geq 0$ for every x_0 and every h from X, (see for example DEMIANOV & RUBINOV [1968, p. 26], compare also with ROBERTS & VARBERG [1973, p.98]). If f, besides the already said, satisfies the LIPSCHITZ condition, then f is uniformly differentiable (in the sense of Definition 8.2) and $f'(x_0)$ is a sublinear functional, (The hierarchy of different notions of convexity may be seen, for example, in MITRINOVIĆ [1970], ROBERTS & VARBERG [1973]). It is easy to see that the convex functional, by a convenable interpretation, satisfies for example, the conditions from the Definitions 8.2, 11.3 as well as the conditions from Theorems in §8 - §14. The consequence is that the necessary conditions for extremum are sufficient also. Moreover, the solution of the optimization problem is uniqae, while the local extremum is simultaneously a global extremum. If the functional f is convex, then the set $\{x \in D \mid f(x) \leq \theta_\pi\}$ is convex, D being convex too. The norms, subnorms as well as numerous other in the practice, are convex. Because we frequently deal with convexity, the Optimization problems in such cases are known as c o n v e x p r o b l e m s.

Convex sets (considered, for example, in normed space) are separable by means of hyperplanes. More precisely, if P and Q are convex and disjoint sets, provided at least one of them has a nonempty interior, they are separable. The nonempty interior requirement may be substituted by a compactness-condition of at least one set (or bicompactness in a local convex space). A linear functional $\ell \in X^{\#}$ corresponds to every hyperplane H. H is closed iff $\ell \in X^{*}$. If H "passes" through $x_0 \in X$, i.e., $x_0 \in H$, then

(*) $$H = \{x \in X \mid \ell(x) = \ell(x_0)\}.$$

On the basis of the foregoing we may say that $\ell \in X^{*}$ and $c \in \mathbb{R}$ are foudable for separable convex sets P, Q ($\dot{P} \neq \emptyset$), so that

(☆) $$\ell(x) < c \quad \forall x \in P \quad \text{and} \quad \ell(x) \geq c \quad \forall x \in Q.$$

Therefore,

$$(**)\qquad \sup_{x\in P} \ell(x) \leq \inf_{x\in Q} \ell(x)$$

holds.

If in the optimization problem from §11 the cones $K(P; x_0)$ and $K[A; x_0]$ are convex, on the basis of (**) and (11.1) it follows

$$(14.1)\qquad \sup_{h\in K(P;x_0)} \ell(h) = \min_{h\in K[A;x_0]} \ell(h) = 0.$$

The condition (**) is of paramount importance. We may even say that the entire classical optimization theory has developed from the existence of functional $\ell \in X^*$, having the properties (*) and (**). If the shape of the functional $\ell \in X^\#$ is known, which is a frquent occarrence in practice, the optimization problem, by means of (14.1) may be transferred to the linear programming problem. In connection with (*) we stress that we often have explicit formulas for the set L^* defined by

$$(14.2)\qquad L^\#\{S; x_0\} := \{\ell \in X^\# | \forall x \in S, \ell(x) \leq \ell(x_0)\} .$$

$L^\#$ is said to represet a set of s u p p o r t f u n c t i o n a l s on S at a point $x_0 \in S$. Some authors attribute this name to the set $-L^\#$. When S is a cone, $-L^\#\{S; x_0\}$ is called sometimes a d u a l c o n e (with respect to S), (see for example DUBOVIZKI & MILUTIN [1965], GIRSANOV [1970]). If D is convex set, f a convex functional and $S := \{x \in D | f(x) \leq f(x_0)\}$, then the set $L^\#\{S; x_0\}$ is equal to

$$(14.3)\qquad L^\#\{f; x_0\} := \{\ell \in X^\# | \forall x \in X, \ell(x) \leq f(x), \ell(x_0) = f(x_0)\}$$

which represent a set of s u p p o r t e d linear functionals on convex functional f at a point x_0. The Theorem HAHN-BANACH sees to it that $L^\#\{f; x_0\}$ is nonempty. If we deal with the continuous functionals, instead $L^\#$ we write L^*. For a convex functional f we have $L^*\{f; x_0\} = L^*\{f'(x_0); x_0\}$, (see for example DEMIANOV & RUBINOV [1968, pp. 28-29]). If X is a linear space, then $L^*\{f; x_0\}$ is convex, bounded and weak sequentially closed in X^*. If, besides the quoted, X is local convex space, then $L^*\{f; x_0\}$ is weak compact in X^*.

We often have to deal with "two degree" optimization, i.e., the target-functional f is given with an optimization process, so

(14.4) $$f(x) := \max_{s \in \mathfrak{S}} \Phi(x,s).$$

Let $\Phi : D \times \mathfrak{S} \to \mathbb{R}$ satisfy the conditions of Definition 6.1 (iii) and let $\Phi(\cdot,s)$ be continuous for every $s \in \mathfrak{S}$, and let $\Phi(x,\cdot)$ be upper semicontinuous for every $x \in D$ as well as $\Phi_A'(x_o)(h) : \mathfrak{S} \to \mathbb{R}$ be upper semicontinuous. Then f is well defined by (14.4) and we have

(14.5) $$f_A'(x_o)(h) = \max_{s \in \mathfrak{S}(x_o)} \Phi(x,s)$$

where

(14.6) $$\mathfrak{S}(x_o) := \{ s \in \mathfrak{S} \mid f(x_o) = \Phi(x_o,s) \}$$

is a compact set, (see [Theorem 9]).

Particularly, if $f : D \to \mathbb{R} \cup \{+\infty\}$ is a convex and continuous in a neighbourhood $\mathcal{V}(x_o)$, there exists $f_D'(x_o)$ and

(14.7) $$f_D'(x_o)(h) = \max_{\ell \in L^*\{f;x_o\}} \ell(h)$$

holds, (see [Korollar 11]). Using (14.5) and (14.7) we can often find the mentioned derivative.

EXAMPLE 14.i. Let $X := (X, \|\cdot\|)$ be the BANACH space and let f be a subnorm. Then we have

$$L^*\{f; x_o\} = \{ \ell \in X^* \mid \|\ell\|_f = 1, \ \ell(x_o) = f(x_o) \},$$

where $\|\ell\|_f := \sup_{f(x) \le 1} |\ell(x)|$.

Especially, if $f := \|x\|$, then

$$L^*\{f; x_o\} = \{ \ell \in X^* \mid \|\ell\| = 1, \ \ell(x_o) = \|x_o\| \}.$$

EXAMPLE 14.ii. Let $X := C[D]$, D being a closed and bounded set in $\mathbb{R}^n$. Taking into account that the conjugat space of $C[D]$ may be equalize with a linear space $NV_o(D)$ of so called normalized functions of bounded variation on D, (see for example ALJANČIĆ [1968, p. 239]), we have the following results.

1° If $f(x) := \|x\| := \max_{t \in E} |x(t)|$, $E \subseteq D$, then

$$L^*\{f;\ x_0\} = \{\mu \in NV_0(D) \mid \mu \geqq 0 \ \text{ on } E_1(x_0),\ \mu = 0 \ \text{ on } E_3(x_0),$$
$$\mu \leqq 0 \ \text{ on } E_2(x_0),\ \mu(E_1(x_0)) - \mu(E_2(x_0)) = 1\}$$

where

$$E_1(x_0) := \{t \in E \mid x_0(t) = f(x_0)\},$$
$$E_2(x_0) := \{t \in E \mid x_0(t) = -f(x_0)\},$$
$$E_3(x_0) := D \setminus (E_1(x_0) \cup E_2(x_0)).$$

For the derivative we have

$$f'(x_0)(h) = \max\{\max_{t \in E_1(x_0)} h(t),\ \max_{t \in E_2(x_0)} -h(t)\}.$$

2^o If $f(x) := \max_{t \in E} x(t)$, $E \subseteq D$, then

$$L^*\{f;x_0\} = \{\mu \in NV_0(D) \mid \mu \geqq 0 \ \text{ on } E,\ \mu(E) = 1,\ \mu(D \setminus E(x_0)) = 0\}$$

where

$$E(x_0) := \{t \in E \mid x_0(t) = f(x_0)\}.$$

The directional derivative is

$$f'(x_0)(h) = \max_{t \in E(x_0)} h(t).$$

EXAMPLE 14.iii. Let $X := L^p(E)$, $1 < p < +\infty$, where E is any bounded and measurable set in $\mathbb{R}^n$. Let $f(x) := \|x\| := \|x\|_{L^p(E)}$. Then $L^*\{f;\ x_0\}$, $x_0 \neq \Theta_{L^p}$, is formed from one element*):

$$L^*\{f;\ x_0\} = \left\{ \left(\frac{|x_0(t)|}{\|x_0\|} \right)^{p-1} \operatorname{sign} x_0(t) \right\}.$$

For the directional derivative we have

$$f'(x_0)(h) = \frac{1}{\|x_0\|^{p-1}} \int_E |x_0(t)|^{p-2} x_0(t) h(t)\, dt.$$

If $X := L^1(E)$ and $f(x) := \|x\|_1$, then we have

$$L^*\{f;\ x_0\} = \{\mu \in L^\infty(E) \mid \mu(t) = \operatorname{sign} x_0(t) \ \text{ on } E_1(x_0),$$
$$\mu(t) = \tau(t) \ \text{ on } E_2(x_0)\}$$

where

---

*) This situation we have whenever X is a reflexive space.

$$E_1(x_o) := \{ t \in E \mid |x_o(t)| > 0 \},$$
$$E_2(x_o) := \{ t \in E \mid |x_o(t)| = 0 \},$$

and $\tau(t)$ is an arbitrary measurable function on $E_2(x_o)$ such that $|\tau(t)| \leq 1$ on $E_2(x_o)$.

To calculate the derivative we use the formula

$$f'(x_o)(h) = \max_{\mu \in L^*\{f; x_o\}} \int_E \mu(t) h(t) \, dt.$$

EXAMPLE 14.iv. Let $X := \ell^p$, $1 < p < +\infty$, and let $f(x) := \|x\| := \|x\|_{\ell^p}$. Then we have

$$L^*\{f; x_o\} = \left\{ \left(\frac{|x_o|}{\|x_o\|^{p-1}} \right)^{p-1} \text{sign } x_o \right\}, \quad x_o \neq \theta_{\ell^p},$$

and

$$f'(x_o)(h) = \frac{1}{\|x_o\|^{p-1}} \sum_{\nu=1}^{+\infty} |\xi_\nu|^{p-2} \xi_\nu h_\nu$$

with

$$x_o := (\xi_\nu) \in \ell^p, \quad h := (h_\nu) \in \ell^p, \quad (|\xi_\nu|^{p-2} \xi_\nu) \in \ell^q,$$
$$p^{-1} + q^{-1} = 1.$$

In fine we note that frequently the directional derivative $f'(x_o): X \to \mathbb{R}$ is a linear and continuous functional, i.e., this derivative is the gradient of f. The norm of a BANACH space X has the gradient iff the unit sphere $S := \{ x \in X \mid \|x\| = 1 \}$ is smooth. If grad $f(x_o)$ is continuous (in variable x, i.e., in topology of X^*) then there exist the FRECHET differential of f and we have

$$Df(x_o) = \text{grad } f(x_o).$$

CHAPTER IV

A P P L I C A T I O N S

In this part we give some examples for the application of the results described in the previous chapters. Primarily we shall observe some cases of the One-sided approximation.

15. ON BEST ONE-SIDED APPROXIMATION

1. INTRODUCTION. Let X be a set formed with $n+1$ points of the real axis and $f: X \to \mathbb{R}$ be the restriction, at X, of a polynomial of degree n. Evidently, there are polynomials P and Q of degree n, such that

(15.1) $\qquad P(x) \geqq 0, \quad Q(x) \geqq 0 \qquad \forall x \in X$

and

(15.2) $\qquad f := P - Q \quad$ on X.

Professor T.POPOVICIU has proposed the following problem: to study the existence and the uniqueness of a pair (P^*, Q^*) of polynomials of degree n which is minimal, i.e.,

(15.3) $\qquad P \geqq P^* \geqq 0, \quad Q \geqq Q^* \geqq 0 \qquad$ on X

for every pair (P,Q) which satisfies (15.2); if the problem has a solution, let this minimal pair be determined. Further, let a similar problem be solved in the case when X contains $n+2$ points.

From (15.1), (15.2) and (15.3) we have

$$(P^* \geqq 0 \wedge P^* - f = Q^* \geqq 0) \Rightarrow (P^* \geqq 0 \wedge P^* \geqq f) \Rightarrow P^* \geqq \max\{0, f\} := f_+$$

and

$$(Q^* \geqq 0 \wedge f + Q^* = P^* \geqq 0) \Rightarrow (Q^* \geqq 0 \wedge Q^* \geqq -f) \Rightarrow Q^* \geqq \max\{0, -f\} := (-f)_+ .$$

If we put

(15.4) $$\rho(\phi) := \max_{x \in X} \phi(x)$$

for every $\phi \in C[X]$, then we want to find the polynomials P^* and Q^* with properties

(15.5) $$\rho(P^* - f_+) = \min_{P \in \mathcal{P}_n^{f_+}} \rho(P - f_+)$$

and

(15.6) $$\rho(Q^* - (-f)_+) = \min_{Q \in \mathcal{P}_n^{(-f)_+}} (Q - (-f)_+),$$

where

$$\mathcal{P}_n^{f_+} := \{ p \in \mathcal{P}_n \mid \forall x \in X, \quad p(x) \geqq f_+(x) \},$$

$$\mathcal{P}_n^{(-f)_+} := \{ p \in \mathcal{P}_n \mid \forall x \in X, \quad p(x) \geqq (-f)_+(x) \}.$$

Therefore, we observe that the above problem is a specimen of the best one-sided approximation in uniform norm.

Inspired by POPOVICIU's problem our aim in this part is a study of several problems of best one-sided approximation. Some of our theorems generalize the known results about the same topic, and several new results are incorporated into a presentation of the previously known theory.

ii. STATEMENT OF THE PROBLEM; EXISTENCE OF THE BEST APPROXIMATION. Let X be a compact set of $\mathbb{R}$ and let $C[X]$ be normed by means of

(15.7) $$p(f) := \max_{x \in X} |f(x)| \qquad f \in C[X].$$

The norm p is the measure of the closeness of the approximation. For a fixed element $f \in C[X]$ we write

(15.8) $$C_B := \{ g \in C[X] \mid \forall x \in X, \ g(x) \leqq f(x) \}.$$

Let $\mathcal{H} \subseteq C[X]$ be a subspace (closed), and

$$\mathcal{H}_B := C_B \cap \mathcal{H} = \{ g \in \mathcal{H} \mid \forall x \in X, \ g(x) \leqq f(x) \}.$$

The purpose is to study the elements $g_* \in \mathcal{H}_B$ such that

(15.9) $$p(f - g_*) = \min_{g \in \mathcal{H}_B} p(f - g),$$

in which case we say that $g_* = g_*(f; \mathcal{H}; X)$ is the element of best one-sided approximation from below on X of the function f by elements from $\mathcal{H}$. In the same way the elements of the best one-sided approximation from above $g^* = g^*(f; ;\mathcal{H};X)$ may be defined. The following theorems deal with approximation from below; an analogous result holds for approximation from above. In our case $p(f-g) = \max_{x \in X} (f(x)- - g(x))$, but in order to put in evidence the similarity with the unconstrained uniform approximation, we shall frequently use the notation introduced in (15.7). It will be supposed that $d := E(f;\mathcal{H}_B;X) := \inf_{g \in \mathcal{H}_B} p(f-g) > 0$.

The proofs about one-sided approximation may be performed by using the fact that the one-sided approximation is a particular case of some problems of uniform approximation with constraints (see for example TAYLOR [1968]). But our aim here is to use the cones-method. In the second part of this chapter we generalize the problem in the following manner: instead of an interval $[a,b]$ we consider an arbitrary compact set, the T-space on $[a,b]$ being substituted by an I-set of functions defined on this set.

First we get two results regarding the existence problem in the case when X contains only a finite number of points as well as when $\mathcal{H}$ is a finite dimensional space. We note that in this case any $f: X \to \mathbb{R}$ is continuous on X. It is well known that the proof of the existence of a best approximation essentially depends on the fact that the best nucus function on a compact set attains its extreme values.

THEOREM 15.1. Let $X := \{x_1,\dots,x_m\}$, $x_i \in \mathbb{R}$ and let $\mathcal{H} \subseteq C[X]$ be given. Then for every $f: X \to \mathbb{R}$ there exists at least one element $g_*(f;\mathcal{H};X)$.

PROOF. We consider the sets

$$M_i := \{ g \in C[X] \mid 0 \leq f(x_i) - g(x_i) \leq c \}$$

$$i = 1,\dots,m$$

where c is an arbitrary sufficiently large positive constant. The sets $Q := \bigcap_{i=1,\dots,m} M_i$ and $Q \cap \mathcal{H}$ are compact in $\mathcal{H}$, and $d = \inf_{g \in Q \cap \mathcal{H}} p(f-g)$. The function φ defined by

$$\varphi(g) := p(f-g) = \max_{i=1,\dots,m} |f(x_i) - g(x_i)|, \qquad g \in C[X],$$

being continuous on $C[X]$, attains its minimum at least at a point $g_* \in Q \cap \mathcal{H}$, and we have

$$\varphi(g_*) = p(f-g_*) = \min_{g \in Q \cap \mathcal{H}} p(f-g) = \min_{g \in \mathcal{H}_B} p(f-g). \quad \square$$

THEOREM 15.2. If X ($\subset \mathbb{R}$) is a compact and $\mathcal{H}$ ($\subseteq C[X]$) is finitly dimensional, then for every $f \in C[X]$ there exists at least one element $g_*(f;\mathcal{H};X)$.

PROOF. Let $\mathcal{H} := \text{span}\{g_1,\dots,g_n\}$ where $g_1,\dots,g_n$ is any basis, $g_i \in C[X]$ for $i=1,\dots,n$. If we put

$$C_B := \{g \in C[X] \mid \forall x \in X,\ g(x) \leq f(x)\},$$

then the set $V := C_B \cap \mathcal{H}$ is convex and closed in $\mathcal{H}$. The functional p, defined by (15.7), is continuous on $\mathcal{H}$. Now, let us consider

$$M := \{g \in C[X] \mid p(f-g) \leqq d + \delta\},$$

where δ is an arbitrary positive number. Taking into account that

$$M \cap \mathcal{H}_B \neq \emptyset, \qquad \inf_{g \in M \cap \mathcal{H}_B} p(f-g) = \inf_{g \in \mathcal{H}_B} p(f-g) = d,$$

we see that it is enough to seek the minimum of $\varphi(g) = p(f-g)$ on $M \cap \mathcal{H}_B$. The set $M \cap \mathcal{H}_B$ is compact in $\mathcal{H}$, because it is closed and bounded in $\mathcal{H}$. The continuity of $p(f-g)$ with respect to g, implies that there is at least one element $g_* \in M \cap \mathcal{H}_B$ such that

$$p(f-g_*) = \min_{g \in M \cap \mathcal{H}_B} p(f-g) = \min_{g \in \mathcal{H}_B} p(f-g). \quad \square$$

One observes that the above theorems follow from a more general situation. Namely, if Q is a closed subset from $C[X]$ such that its intersections with every subset $\{g \in C[X] \mid p(f-g) \leqq \rho\}$ are compact, then $f \in C[X]$ admits at least one element of best approximation by means of elements from Q. In particular, this is true if Q is a closed subset from a finite dimensional space $\mathcal{H}$.

iii. CHARACTERIZATION, General case.- The following important theorem is a modification of the characterization theorem for approximation with restricted range due to TAYLOR [1969, Theorem 3.2]. In this section X is a compact set containing at least n distinct points, $\mathcal{H}$ ($\in C[X]$) is an arbitrary space (closed with respect to norm p).

For a fixed $g \in \mathcal{H}_B$ we put

$$E_+(g) := \{x \in X \mid f(x) - g(x) = p(f-g)\},$$
(15.10)
$$C_-(g) := \{x \in X \mid f(x) - g(x) = 0\}$$

and

$$A(g) := E_+(g) \cup C_-(g),$$
(15.11)
$$\sigma(x) := \begin{cases} 1 & \text{if } x \in E_+(g) \\ -1 & \text{if } x \in C_-(g). \end{cases}$$

Points in $E_+(g)$ are called (plus) extremal points and those in $C_-(g)$ are called points of the (lower) contact. The points in $A(g)$ are called critical points (with respect to the function $e := f - g$. Analogous symbols can be introduced for approximation from above. If $g_* = g_*(f; \mathcal{H}; X)$, then $d_* := p(f - g_*)$ is called the deviation and $e_* := f - g_*$ is called the deviation-function. In the sequel we will symbolize by $g_o < f$ the fact that $g_o \in C[X]$ satisfies the SLATER condition, i.e., that $g_o(x) - f(x)$ for all $x \in X$.

THEOREM 15.3. If there exists an element $g_o \in \mathcal{H}$ satisfying $g_o < f$, then a necessary and sufficient condition which must be fulfilled by $g_* \in \mathcal{H}$ such that g_* is an element of the best one-sided approximation from f, namely $g_* = g_*(f; \mathcal{H}; X)$, is that there does not exist $g \in \mathcal{H}$ such that

(15.12) $\sigma(x)\, g(x) > 0 \qquad \forall\, x \in A(g_*)$,

where $\sigma := \sigma_{g_*}$.

PROOF. The functional $\varphi : C[X] \to \mathbb{R}$ defined by $\varphi(g) := p(f - g)$ is continuous and convex. Therefore, the sets

$$Q := \{g \in C[X] \mid p(f - g) < p(f - g_*)\},$$
$$C_B := \{g \in C[X] \mid \forall x \in X,\ f(x) - g(x) \geq 0\}$$

are convex. Since $g_0 \in \overset{\circ}{C}_B$ we have $\overset{\circ}{C}_B \cap \mathcal{H} \neq \emptyset$. In this way the conditions of the Theorem 7.1 are fulfilled. Thus $g_*(f; \mathcal{H}; X)$ satisfies

(15.13) $\quad K(Q;\, g_*) \cap K(C_B;\, g_*) \cap K[\mathcal{H};\, g_*] = \emptyset.$

On the basis of the Example 14.ii, we have

(15.14) $K(Q;\, g_*) = \{g \in C[X] \mid \forall x \in E_+(g_*),\ g(x)\cdot \mathrm{sign}(f(x)-g_*(x)) > 0\}$,

(15.15) $K(C_B;\, g_*) = \{g \in C[X] \mid \forall x \in C_-(g_*),\ \ g(x) < 0\}$,

(15.16) $K[\mathcal{H};\, g_*] = \mathcal{H}$.

In our case $\mathrm{sign}\,(f(x) - g_*(x)) = +1$ for all $x \in E_+(g_*)$ and (15.13) reduces to

$$\{g \in C[X] \mid \forall x \in E_+(g_*),\ \ g(x) > 0\} \cap$$
$$\cap \{g \in C[X] \mid \forall x \in C_-(g_*),\ \ g(x) < 0\} \cap \mathcal{H} = \emptyset,$$

which proves our assertion. □

iv. CHARACTERIZATION. The case with a finite dimensional space.- First we note that in this case the set $A(g_*)$ defined in 15.iii., contains only finitly many points. On a real line let X be a compact set containing at least n distinct points, let $\mathcal{H}$ $(\subset C[X])$ have the dimension n. We may write $\mathcal{H} = \mathrm{span}\,\{g_1, \dots, g_n\}$ where $g_1, \dots, g_n$ is a basis of $\mathcal{H}$, $g_i \in C[X]$. Then we may transpose our problem to the same topic in $\mathbb{R}^n$. This is motivated by taking into account the linear biunivoque correspondence between $\mathcal{H}$ and $\mathbb{R}^n$:

$$g = \sum_{k=1}^{n} \alpha_k g_k \in \mathcal{H} \leftrightarrow \alpha = (\alpha_1, \dots, \alpha_n) \in \mathbb{R}^n.$$

Therefore to the set C_B, the set

(15.17) $\quad V_B := \{\alpha \in \mathbb{R}^n \mid \forall x \in X,\ \sum_{k=1}^{n} \alpha_k\, g_k(x) \leqq f(x)\}$

corresponds. Further we define the continuous application $a : X \to \mathbb{R}^n$ by

$$a(x) = (\,g_1(x), \dots, g_n(x)\,) \qquad \forall x \in X$$

and put

$$a(A(g)) := \{a(x) \mid x \in A(g)\}.$$

The following characterization theorem is valid.

THEOREM 15.4. Let us assume that there exists an element $g_0 \in \mathcal{H}$ such that $g_0 < f$, and let f be a given element in $C[X]$, $f \notin \mathcal{H}$. Then each of the following is a necessary and sufficient condition for $g_* \in \mathcal{H}$ to be an element of the best one-sided approximation, $g_* = g_*(f; \mathcal{H}; X)$:

(a) There does not exist $g \in \operatorname{span}\{g_1, \dots, g_n\}$ such that $\sigma(x)g(x) > 0$ for all $x \in A(g_*)$;

(b) The vector $\theta_{\mathbb{R}^n}$ belongs to the set

$$\operatorname{co}\left(\{\sigma(x)(g_1(x), \dots, g_n(x)) \mid \forall x \in A(g_*)\}\right);$$

(c) There exist the points $x_1, \dots, x_m$ in $A(g_*)$, $m \leq n+1$, and positive numbers γ_i such that a linear (and continuous) functional $\ell \in C[X]^*$ defined on $C[X]$ of the "point evaluation" type

$$\ell(f) := \ell\left(\begin{matrix}\gamma_1, \dots, \gamma_m \\ x_1, \dots, x_m\end{matrix}; f\right) := \sum_{i=1}^{m} \sigma(x_i)\, \gamma_i\, f(x_i), \quad f \in C[X]$$

satisfies $\ell \in \mathcal{H}^{\perp}$, i.e., $\ell(g) = 0$ for all $g \in \mathcal{H}$.

PROOF. We shall show that: $(g_* \in g_*(f; \mathcal{H}; X) \Leftrightarrow (a) \Leftrightarrow (b) \Leftrightarrow (c)$. In view of Theorem 15.3 the first part is established. To show that $(a) \Leftrightarrow (b)$, let $\langle \cdot , \cdot \rangle$ denote the inner product in $\mathbb{R}^n$. Then the set defined in (15.17) may be written in the form

$$V_B = \{\alpha \in \mathbb{R}^n \mid \forall x \in X,\ \langle \alpha, a(x) \rangle \leq f(x)\}$$

and our problem reduces to finding a point $\alpha_* \in \mathbb{R}^n$ for which

$$\varphi(\alpha_*) = \min_{\alpha \in V_B} \varphi(\alpha)$$

where

$$\varphi(\alpha) := p\left(f - \sum_{k=1}^{m} \alpha_k g_k\right).$$

By using the formulas (15.14)-(15.16) we find

$$\begin{aligned} K(Q; g_*) &= \{g \in C[X] \mid \forall x \in E_+(g_*),\ g(x) > 0\} = \\ &= \{\alpha \in \mathbb{R}^n \mid \forall x \in E_+(g_*),\ \langle \alpha, a(x) \rangle > 0\} = \\ &= \{\alpha \in \mathbb{R}^n \mid \forall \beta \in a(E_+(g_*)),\ \langle \alpha, \beta \rangle > 0\}; \end{aligned}$$

$$\begin{aligned} K(C_B; g_*) &= \{g \in C[X] \mid \forall x \in C_-(g_*),\ g(x) < 0\} = \\ &= \{\alpha \in \mathbb{R}^n \mid \forall x \in C_-(g_*),\ \langle \alpha, -a(x) \rangle > 0\} = \\ &= \{\alpha \in \mathbb{R}^n \mid \forall \beta \in -a(C_-(g_*)),\ \langle \alpha, \beta \rangle > 0\}; \end{aligned}$$

$$K[\mathcal{H}; g_*] = \mathbb{R}^n.$$

By means of the Theorem 15.3 we reach the fact that the system of linear inequalities in $\mathbb{R}^n$:

$$\langle \alpha , \beta \rangle > 0 \qquad \forall \beta \in a(E_+(g_*)),$$
$$\langle \alpha , \beta \rangle > 0 \qquad \forall \beta \in -a(C_-(g_*)),$$

in unknown α, is inconsistent, i.e., there is no any solution. Since

$$-a(C_-(g_*)) = \{ \sigma(x_i)a(x_i) \mid \forall x_i \in C_-(g_*) \},$$

this system may be written in the form

$$(15.18) \qquad \langle \alpha , \beta \rangle < 0$$

with

$$\beta \in \{ \sigma(x_i)a(x_i) \mid \forall x \in A(g_*) \} =: \{\beta\} .$$

The set $\{\beta\}$ of n-tuples, is compact in $\mathbb{R}^n$. It is known (see for example CHENEY [1966, p.19]) that the inconsistent property of the (15.18) is equivalent to

$$(15.19) \qquad \theta_{\mathbb{R}^n} \in \mathrm{co}\ (\{\beta\})$$

which is equivalent to (b). The chain will be completed by showing (b)$\Leftrightarrow$(c). In order to show this, it is sufficient to show that (15.19) is equivalent to (c). On account of CARATHEODORY Theorem, there exist at most $n+1$ points $\beta_1, \dots, \beta_m$ and positive coefficients ρ_i such that

$$(15.20) \qquad \theta_{\mathbb{R}^n} = \sum_{i=1}^{m} \rho_i \beta_i$$

with $\beta_i \in \{\beta\}$, $\sum_{i=1}^{m} \rho_i = 1$, $\rho_i > 0$, $1 \leq m \leq n+1$. If $x_i \in$ $\in a(E_+(g_*))$ we have

$$\rho_i \beta_i = \rho_i a(x_i) = \sigma(x_i)\ \rho_i\ a(x_i), \quad \forall x_i \in E_+(g_*)$$

as well as, if $x_i \in -a(C_-(g_*))$, we have

$$\rho_i \beta_i = -\rho_i\ a(x_i) = \sigma(x_i)\ \rho_i\ a(x_i), \quad \forall x_i \in C_-(g_*).$$

Thus, from (15.20) we have

$$\theta_{\mathbb{R}^n} = \sum_{i=1}^{m} \sigma(x_i)\ \rho_i a(x_i) = \sum_{i=1}^{m} \rho_i\ \sigma(x_i) \big(g_1(x_i), \dots, g_n(x_i) \big)$$

which is equivalent to

$$(15.21) \qquad \sum_{i=1}^{m} \sigma(x_i)\ \rho_i\ g_k(x_i) = 0, \qquad k = 1, \dots, n.$$

By multiplying every equation in (15.21) with arbitrary numbers α_k, $k = 1, \dots, n$ and by adding up these equalities we obtain

$$\sum_{i=1}^{m} \sigma(x_i)\, \rho_i \left[\sum_{k=1}^{n} \alpha_k\, g_k(x_i) \right] = 0, \qquad x_i \in A(g_*),$$

that is

$$\ell(g) = \sum_{i=1}^{m} \sigma(x_i)\, \rho_i\, g(x_i) = 0 \qquad \forall g = \sum_{k=1}^{n} \alpha_k\, g_k.$$

We note that $E_+(g_*) \neq \emptyset$. Indeed, let us assume the contrary. The inequality $1 \leqq m$ implies $C_-(g_*) \neq \emptyset$, i.e., we must have $A(g_*) = C_-(g_*)$ and

$$0 = \sum_{i=1}^{m} \sigma(x_i)\, \rho_i\, g(x_i) = -\sum_{i=1}^{m} \rho_i\, g(x_i)$$

for every $x_i \in A(g_*)$. In other words, for the element $\tilde{g} :=$ $:= g - g_o$

$$\sum_{i=1}^{m} \rho_i (\tilde{g} - g_o)(x_i) = \sum_{i=1}^{m} \rho_i \bigl(f(x_i) - g_o(x_i)\bigr) > 0$$

holds; because $g_*(x_i) = f(x_i)$ for all $x_i \in A(g_*)$, which asserts that there is an element $\tilde{g}$ in $\mathcal{H}$ such that $\ell(\tilde{g}) =$ $= 0$. □

In connection with the uniqueness we note that the conditions which we have imposed, do not ensure the uniqueneness of the best approximation element for each $f \in C[X]$. In the following paragraph we shall see that in the case when $\mathcal{H}$ is an I-space, the best approximation has a unique solution as well as that the I-property is also a necessary condition for the uniqueness.

16. THE BEST ONE-SIDED APPROXIMATION WITH I-FAMILIES

Let $\mathcal{H}$ ($\subseteq C[X]$) be an I-subspace of dimension n, i.e., $\mathcal{H}$ is a linear I-set of the order n on X. Let $\ell :$ $: C[X] \to \mathbb{R}$ be a non-zero linear continuous functional defined by

(16.1) $$\ell(f) := \ell\begin{pmatrix} \gamma_1, \dots, \gamma_{n+1} \\ x_1, \dots, x_{n+1} \end{pmatrix}; f) =: \sum_{i=1}^{n+1} \gamma_i\, f(x_i).$$

If $\ell \in \mathcal{H}^\perp$, then we have

(16.2) $$\gamma_i = (-1)^{n+1-i}\, \gamma_{n+1}\, V\begin{pmatrix} g_1, & \dots & , g_n \\ x_1, \dots, x_{i-1}, x_{i+1}, \dots, x_{n+1} \end{pmatrix} :$$

$$: V\begin{pmatrix} g_1, \dots, g_n \\ x_1, \dots, x_n \end{pmatrix},$$

$i = 1,\dots,n+1,$

and the prevronskian at the denominator differs from zero. Moreover, we have

(16.3) $$\ell(f) = \gamma_{n+1}\, V\binom{g_1,\dots,g_n,f}{x_1,\dots,x_{n+1}} : V\binom{g_1,\dots,g_n}{x_1,\dots,x_n}.$$

In his work T.POPOVICIU 1959 established that the unique functional of the form (16.1), from $\mathcal{H}^{\perp}$, was defined by (16.3). Likewise he noted that in the case when $X:=[a,b]$, the prevronskian $V\binom{g_1,\dots,g_n}{x_1,\dots,x_n}$ may be assumed positive. Therefore, in this case (16.2) implies the well-known alternate property of the coefficients $\gamma_1,\dots,\gamma_{n+1}$. This property remains valid even if X is an arbitrary compact set on $\mathbb{R}$ and if $\mathcal{H}$ is an I-set of $I_n\{[a,b]\}$ type and $[a,b]$ is an arbitrary interval which contains the set X.

The following theorem deals with uniqueness of solution of the best one-sided approximation. The proof of this may be performed by using similar arguments with those from the unconstrained case (see for example LAURENT [1972, p.96]).

THEOREM 16.1. Let us suppose that there exists an element $g_0 \in \mathcal{H}$ such that $g_0 < f$. In order that a unique element $g_* = g_*(f;\mathcal{H};X)$ may exist for every $f \in C[X]$, it is necessary and sufficient that the subspace $\mathcal{H}$ be interpolatory on X.

We note that in the I-case, if the functional which is considered in Theorem 15.4 (c), is from $\mathcal{H}^{\perp}$, we must have $m = n+1$. Indeed, let us assume $m \leq n$. Then there exists an element $\tilde{g} \in \mathcal{H}$ defined by conditions $\tilde{g}(x_i) = \gamma_i$, $i=1,\dots,m$. Then we have

$$\ell(\tilde{g}) = \sum_{i=1}^{m} \gamma_i\, \tilde{g}(x_i) = \sum_{i=1}^{m} \gamma_i^2 > 0,$$

which contradicts our hypothesis. But this means that in the I-case there exist exactly $n+1$ points which form the set $A(g_*)$ of critical points from X.

17. THE ALTERNATORY CASE

If $\mathcal{H}$ is of the type $I_n\{[a,b]\}$, $[a,b]\supset X$, then the quotient of prevronskians in (16.2), has a constant sign and all coefficients γ_i, $i=1,\dots,n+1$ are different from zero. Therefore, from (16.2) we conclude that γ_i have alternating sign. Finally the fact that the set $A(g_*)$ contains exactly $n+1$ distinct points, enables us to assert that the deviation-function $e_* = f - g_*$ has exactly $n+1$ points in X at which e takes alternatively the values d_* and 0. This situation is called t h e c a s e w i t h a l t e r n a n c e. Further we extend the Theorem 15.4 as

THEOREM 17.1. Let $\mathcal{H} := \operatorname{span}\{g_1,\dots,g_n\} \subset C[a,b]$ be an I-space of the type $I_n\{[a,b]\}$, $[a,b]\supset X$, and suppose that trere exists an element $g_0\in\mathcal{H}$ satisfying the SLATER condition on X. Then each of the following is a necessary and sufficient condition for $g_*\in\mathcal{H}$ to be $g_* = g_*(f;\mathcal{H};X)$:

(a) There does not exist $g\in\mathcal{H}$ such that $\sigma(x)\cdot g(x) > 0$ for all $x\in A(g_*)$;

(b) The vector $\theta_{\mathbb{R}^n}$ belongs to the set

$$\operatorname{co}\left(\{\sigma(x)(g_1(x),\dots,g_n(x)) \mid \forall x\in A(g_*)\}\right);$$

(c) There exists a linear functional $\ell\in C[X]^*$ of the form (16.3) and $\ell\in\mathcal{H}^\perp$, where $(x_1,\dots,x_{n+1})\in\Delta_{n+1}(X)\cap\cap(E_+(g_*)\cup C_-(g_*))$, $\sigma(x_i)\gamma_i > 0$ for $i=1,\dots,n+1$ and $\gamma_i\gamma_{i+1} < 0$;

(d)(Alternanca) There exists the $n+1$-tuple $(x_1,\dots,x_{n+1})\in\Delta_{n+1}(X)\cap(E_+(g_*)\cup C_-(g_*))$ such that $\sigma(x_{i+1}) = -\sigma(x_i)$ for $i=1,\dots,n$.

In this case we have the following consequences.

COROLLARY 17.2. Let $(x_1,\dots,x_{n+1})\in\Delta_{n+1}(X)$ be the "critical" set from X with respect to $e_* = f - g_*$. Then the coefficients α_k of $g_* = g_*(f;\mathcal{H};X)$ satisfy

$$(17.1)\qquad \sum_{k=1}^{n}\alpha_k\, g_k(x_i) = f(x_i) - \vartheta_i\, d_*,\qquad i=1,\dots,n+1$$

where ϑ_i are the coordinates of the "alternating" vectors

$\vartheta := (1,0,1,0,\dots) \in \mathbb{R}^{n+1}$ or $\bar{\vartheta} := (0,1,0,1,\dots) \in \mathbb{R}^{n+1}$.

The system of coefficients

$$\lambda_1, \mu_1, \lambda_2, \mu_2, \dots, \lambda_r, \mu_s$$

or

$$\mu_1, \lambda_1, \mu_2, \lambda_2, \dots, \mu_s, \lambda_r$$

with $\lambda_i > 0$ and $\mu_j < 0$ and $r+s = n+1$ corresponds in the functional ℓ to the system of points $x_1,\dots,x_{n+1}$. We select one of the above systems such that $d_* > 0$.

Similarly, if $(x_1',\dots,x_{n+1}') \in \Delta_{n+1}(X)$ are critical points with respect to deviation-function $e^* := f - g^*$ where $g^* = g^*(f;\mathcal{H};X)$, then the coefficients of g^* satisfy

(17.2) $\qquad \sum_{k=1}^{n} \alpha_k g_k(x_i') = f(x_i') - \eta_i d^*, \qquad i=1,\dots,n+1$

where η_i, $i=1,\dots,n+1$ are the coordinates of vector $\eta := (0,-1,0,-1,\dots) \in \mathbb{R}^{n+1}$ or $\bar{\eta} := (-1,0,-1,0,\dots) \in \mathbb{R}^{n+1}$ respectively. Now, if g_* is $g_*(f;\mathcal{H};[a,b])$, then in KAMMERER[1959, p.12] it is shown that this is equivalent to the fact that $-g_*$ is the $g^*(-f;\mathcal{H};[a,b])$, i.e., $g^*(f;\mathcal{H};[a,b]) = -g_*(-f;\mathcal{H};[a,b])$. This may be extended to the I-case with an arbitrary compact set X ($\subset [a,b]$). Therefore in the case when the knots $x_1,\dots,x_{n+1}$ are elements from $\Delta_{n+1}(X)$ we can assert the following: if for the vector the system (17.1) furnishes the element g_*, then from the some system one finds, with the vector η and for a certain system $(x_1',\dots,x_{n+1}') \in \Delta_{n+1}(X)$, the element g^*. We remark that generally these two systems of knots are not the same.

COROLLARY 17.3. If n is even, then the points x_1 and x_{n+1} are critical points of the same kind, i.e., both belong to $E_+(g_*)$ or to $C_-(g_*)$. For n odd the same points are critical of the different nature.

PROOF. Let us suppose $n = 2k \in \mathbb{N}$ and $r+s = 2k+1$. According to the alternatory property, it follows $r = k+1$, $s = k$ or $r = k$, $s = k+1$. In the first case, x_1 and x_{n+1} belong to $E_+(g_*)$, while in the second case these points are in $C_-(g_*)$. If $n = 2k-1 \in \mathbb{N}$ and $r+s = 2k$, we have

$r = s = k$. But this means that x_1 and x_{n+1} are not of the same kind. Moreover, in both cases the number of contact points in $C_-(g_*)$ is $s = [\frac{n}{2}] + 1$. □

COROLLARY 17.4. Let $X := [a,b]$, $\mathcal{H} :=$ span $\{g_1,\ldots,g_n\}$ be an I-set of the type $I_n\{[a,b]\}$ (i.e., $\mathcal{H}$ is a T-space on $[a,b]$), and let $g_1 := 1$. Then, if f is non-polynomial with respect to $\mathcal{H}$, the function $e_* := f - g_*(f;\mathcal{H};[a,b])$ has the end-points $\underline{a}$ and $\underline{b}$ in its set of alternance $A(g_*)$.

PROOF. Suppose that x_{i_0} is extremal, i.e., $e_*(x_{i_0}) = p(f - g_*)$, and x_{i_0+1} is a contact point, $e_*(x_{i_0+1}) = 0$. Let us suppose thatthere is an extremal point x_0 between x_{i_0} and x_{i_0+1} (the same study when x_0 is a contact point). The function e_* differs on $[x_{i_0}, x_0]$ from the constant function $p(f - g_*)$. This is motivated by the fact that the number of extremal points is finite. Therefore we find a positive number c so that the function $e_* - c = f - (g_* + c)$ vanishes at three points from $[x_{i_0}, x_{i_0+1}]$ and it results that the above function has at least $n+2$ roots on $[a,b]$. But this contradicts the fact that $g_* + c \in \mathcal{H}$ and that $\mathcal{H} + \{f\}$ is an I_{n+1}-space. Further, assume that the end-point $\underline{a}$ is not critical, i.e., let $0 < e_*(a) < p(f - g_*)$. Then we can find a positive number c for which the function $e_* - c = f - (g_* + c)$ has a root on $[a, x_1]$ as well as a root on $[x_1, x_2]$. This means that $e_* - c$ has at least $n+1$ roots on $[a,b]$, and must be $f = g_* + c$. But this is a contradiction and the proof is complete. □

The above Corollary extends to the well-known result by T.POPOVICIU [1939] in the case $X := [a,b]$, $\mathcal{H}$ being the space of algebraic polynomials of the degree $n-1$.

18. COMPUTATION

Let $X := \{x_1,\ldots,x_{n+1}\}$ be fixed and ℓ be the functional considered in Theorem 15.4 (c).

Put

$$\gamma_i := \begin{cases} \lambda_i & \text{if } \gamma_i > 0, \\ \mu_j & \text{if } \gamma_i < 0, \end{cases}$$

and

$$I := \{ i \in \{1,\dots,n+1\} \mid \gamma_i > 0 \}, \quad J := \{ i \in \{1,\dots,n+1\} \mid \gamma_i < 0 \}.$$

Then we have

$$\begin{aligned} \ell(f) = \ell(f - g_*) &= \sum_{i\in I} \lambda_i (f(y_i) - g_*(y_i)) + \sum_{j \in J} \mu_j (f(z_j) - g\ (z_j)) \\ &= \sum_{i \in I} \lambda_i (f(y_i) - g_*(y_i)) \\ &= d_* \sum_{i \in I} \lambda_i , \end{aligned}$$

$$y_i \in E_+(g_*), \quad z_j \in C_-(g_*).$$

Therefore, if $I \neq \emptyset$ and $d_* > 0$, we see that $\ell(f) > 0$. In this manner we conclude that by selecting the sign of γ_{n+1} in (16.3) so that $\ell(f) > 0$, by means of (16.2) we get the possibility to divide the index-set $\{1,\dots,n+1\}$ into the subsets I and J. Instead of ℓ we shall consider the normalized functional $\ell_* := \ell / \sum_{i \in I} \lambda_i$ which satisfies $\ell_*(f) = d_*$. If ℓ_* is well-defined the deviation $d_* := \ell_*(f)$ is known. The familiarity with the sets I and J is equivalent to the fact that the subsets $E_+(g_*)$ and $C_-(g_*)$ are known. The coefficients α_k of $g_* = \sum_{k=1}^{n} \alpha_k g_k$ satisfy the system of equations

(18.1) $$\sum_{k=1}^{n} \alpha_k g_k(x_i) = \varphi(x_i), \qquad i=1,\dots,n+1$$

where

(18.2) $$\varphi(x_i) := f(x_i) - \frac{1 - \sigma(x_i)}{2} d_* .$$

Thus g_* is the element which interpolates the function φ on the set X. On the other hand, from the construction of φ as well as by taking into account that every element from $\mathcal{H}$ is determined by n distinct points, we may write

(18.3) $$g_* = L(\mathcal{H}; x_1,\dots,x_{i-1},x_{i+1},\dots,x_{n+1}; \varphi).$$

We note that the sign $\sigma(x_i)$ may be determined as follows: if (18.1) is considered as a system in the unknowns $\alpha_1,\dots,\alpha_n, d_*$ then we have

$$d_* = V\begin{pmatrix} g_1,\dots,g_n,f \\ x_1,\dots,x_{n+1} \end{pmatrix} : V\begin{pmatrix} g_1,\dots,g_n, \frac{1+\sigma(\cdot)}{2} \\ x_1, \quad \dots \quad ,x_{n+1} \end{pmatrix},$$

or

$$d_* = V\binom{g_1,\dots,g_n,f}{x_1,\dots,x_{n+1}} : \sum_{i=1}^{n+1} \frac{1+\sigma(x_i)}{2} D_i$$

where D_i are the cofactors in the developing of the denominator by the elements of the last row. Since the numerator has a constant sign and the denominator is a linear combination of the values $(1+\sigma(x_i)/2$, we may determine the sign of coefficients D_i such that d_* has a positive and minimal value. Indeed, put

$$\sigma(x_i) = \text{sign}\left(V\binom{g_1,\dots,g_n,f}{x_1,\dots,x_{n+1}} : D_i^{-1}\right),$$

$$i = 1,\dots,n+1.$$

The above method depends on a functional ℓ which enables us to find d_* and $\sigma(x_i)$, $i=1,\dots,n+1$. Using the method of POPOVICIU of the consideration of two I-space: $\mathcal{H} := \text{span}\{g_1,\dots,g_n\}$ and $\mathcal{W} := \text{span}\{g_1,\dots,g_n,g_{n+1}\}$, g_{n+1} being selected in a convenable manner, we shall give a more elegant solution. Indeed, let

$$u := L(\mathcal{W};\ x_1,\dots,x_{n+1};\ f) := \sum_{k=1}^{n+1} a_k\, g_k$$

and

$$v := L(\mathcal{W};\ x_1,\dots,x_{n+1};\ f) := \sum_{k=1}^{n+1} b_k\, g_k,$$

and let us consider

$$h := u - d\cdot v,$$

where d is a real number. If $d =: d_*$ is selected such that the coefficient of g_{n+1} is zero, then the element

$$h_* := u - d_* v = \sum_{k=1}^{n} (a_k - d\ b_k)\, g_k$$

belongs to $\mathcal{H}$ and satisfies the system (18.1). Therefore we have $h_* = g_*(f;\mathcal{H};X)$. From the above remarks and by taking into account(see for example ELENA POPOVICIU [1972, p.34]) that

$$a_{n+1} = [\mathcal{W};\ x_1,\dots,x_{n+1};\ u],$$

$$b_{n+1} = [\mathcal{W};\ x_1,\dots,x_{n+1};\ v],$$

if we write $d_* = a_{n+1}/\,b_{n+1}$, we can find h_*. Since the generalized LAGRANGE operator is linear (see for example ELENA POPOVICIU [1972, p.27]), we may write

(18.4) $h_* = g_*(f;\mathcal{H};X) = L(\mathcal{W};\ x_1,\dots,x_{n+1};\ \ell)$

where ℓ is given by (18.2). We note that in (18.4) the coefficient of g_{n+1} is zero. It is of interest to remark that from equalities $f(x_i) = u(x_i)$, $i=1,\dots,n+1$ as well as from $\ell \in \mathcal{H}^{\perp}$, we conclude that

$$\ell(f) = d_* = \frac{1}{b_{n+1}}[\mathcal{W};\ x_1,\dots,x_{n+1};\ f]$$

is a divided difference (of order n with respect to the I-segment $\mathcal{H} \subset \mathcal{W}$).

The function defined by (18.3) is symmetric with respect to the knots. Taking into account the relationship between ℓ and f, we shall represent the element g_* in a more convenient form. Namely, let us denote

$$\delta_k := f(x_k) - L(\mathcal{H};\ x_1,\dots,x_{k-1},x_{k+1},\dots,x_{n+1};\ f)(x_k)$$

$$i = 1,\dots,n+1.$$

We may write (see for example ELENA POPOVICIU [1972, p.45])

$$\delta_k = \frac{V\begin{pmatrix} g_1,\dots,g_{n+1} \\ x_1,\dots,x_{n+1}\end{pmatrix}}{V\begin{pmatrix} g_1, & \dots & ,g_n \\ x_1,\dots,x_{k-1}, & x_{k+1},\dots, & x_{n+1}\end{pmatrix}}[\mathcal{W};\ x_1,\dots,x_{n+1};\ f].$$

Let $\theta_1,\dots,\theta_{n+1}$ be non-negative numbers so that

(18.5) $$\sum_{k=1}^{n+1} \theta_k = 1.$$

Put

(18.6) $$h = \sum_{k=1}^{n+1} \theta_k\, L(\mathcal{H};\ x_1,\dots,x_{k-1},x_{k+1},\dots,x_{n+1};\ f)$$

where θ_k must be determined in order that $h = g_*(f;\mathcal{H};X)$. On the basis of the equalities

$$L(\mathcal{H};\ x_1,\dots,x_{k-1},x_{k+1},\dots,x_{n+1};\ f)(x_i) = \begin{cases} f(x_i) & \text{if } i \neq k, \\ f(x_k) - \delta_k & \text{if } i=k,\end{cases}$$

$$i = 1,\dots,n+1,$$

we may write

$$h(x_i) = \sum_{k=1}^{n+1} \theta_k\, f(x_i) - \theta_i\,\delta_i = f(x_i) - \theta_i\,\delta_i,$$

$$i = 1,\dots,n+1.$$

From this it follows

(18.7) $$f(x_i) - h(x_i) = \theta_i\,\delta_i, \qquad i=1,\dots,n+1.$$

Let us suppose that $|\delta_i| \neq 0$, $i=1,\dots,n+1$. (For example, this condition is fulfilled when $\mathcal{H}$ is an I-space on X and f is non polynomial with respect to $\mathcal{H}$ on X). Since $h = g_*$, we have

$$f(x_i) - h(x_i) = \begin{cases} d_* & \text{if } x_i \in E_+(g_*), \\ 0 & \text{if } x_i \in C_-(g_*), \end{cases}$$

and from (18.6) we conclude that

$$\theta_i = \begin{cases} \dfrac{d_*}{|\delta_i|} = \dfrac{d}{\delta_i} & \text{if } i \in I, \\ 0 & \text{if } i \in J. \end{cases}$$

Thus (18.5) implies

$$1 = \sum_{k=1}^{n+1} \theta_k = \sum_{i \in I} \frac{d_*}{\delta_i} = d_* \sum_{i \in I} \frac{1}{\delta_i} ,$$

that is

$$d_* = 1 / \sum_{i \in I} \frac{1}{\delta_i} ,$$

and finally

$$\theta_k = \begin{cases} \left(\dfrac{1}{\delta_k}\right) / \sum_{k \in I} \dfrac{1}{\delta_k} & \text{for } k \in I, \\ 0 & \text{for } k \in J. \end{cases} \tag{18.8}$$

Therefore g_* is given by (18.6) and (18.8). For the unconstrained case, a similar result was obtained by MOTZKIN & & SHARMA [1966, Theorem 2]. In the case with alternance we have $\delta_i \, \delta_{i+1} < 0$, $i=1,\dots,n$. This implies the remarks from the preceding paragraph. Likewise, a similar representation may be given when X contains a finite number $m > n+1$ of distinct points.

Further we present an algorithm for finding the solution $g_*(f;\mathcal{H};X)$. In what follows $\mathcal{H}$ is an I-space. The algorithm is the analogue of the one-point REMES exchange algorithm for uniform unconstrained approximation (see for example RICE [1964, p.173]). Our algorithm starts with a set $X^{(o)}$ of n+1 distinct points $x_1^{(o)},\dots,x_{n+1}^{(o)}$ of X. Putting $g^{(o)} := g_*(f;\mathcal{H};X^{(o)})$, the deviation-function $f - g^{(o)}$ is examined over X and $g^{(o)} := \sum_{k=1}^{n} \alpha_k g_k$ is compared with f. From the inclusion

$$\mathcal{H}_B(o) = \{g \in \mathcal{H} \mid \forall x \in X^{(o)},\ g(x) \le f(x)\} \supset \{g \in \mathcal{H} \mid \forall x \in X,\ g(x) \le f(x)\} = \mathcal{H}_B$$

it follows

$$p^{(o)}(\phi) := \max_{i=1,\dots,n+1} |\phi(x_i^{(o)})| \leq \max_{x \in X} |\phi(x)| = p(\phi)$$

for every $\phi \in C[X]$. In particular

$$d^{(o)} = p^{(o)}(e^{(o)}) = \min_{g \in \mathcal{H}_B(o)} p^{(o)}(f - g^{(o)}) \leq \min_{g \in \mathcal{H}_B} p(f - g^{(o)}),$$

i.e.,

(18.9) $$d^{(o)} \leq p(e^{(o)}).$$

There is always at least a point $x_q \in X \setminus X^{(o)}$ such that

(18.10) $$e^{(o)}(x_q) = \max_{x \in X \setminus X^{(o)}} e^{(o)}(x) := M^{(o)}$$

as well as a point $x_{q'} \in X \setminus X^{(o)}$ such that we have

(18.11) $$e^{(o)}(x_{q'}) = \min_{x \in X \setminus X^{(o)}} e^{(o)}(x) := m^{(o)}.$$

By means of (18.9)-(18.10) we see that we have

(18.12) $$M^{(o)} = p(e^{(o)}).$$

If we have simultaneously

(18.13) $$m^{(o)} \geq 0 \quad \text{and} \quad d^{(o)} = p(e^{(o)}),$$

then the fact $g^{(o)} \in \mathcal{H}_B$ and because $e^{(o)}$ has $n+1$ critical points,(all points of $X^{(o)}$ $(\subset X)$), enable us to assert that $g^{(o)} = g_*(f;\mathcal{H};X)$. If it is not the case, then at least one relation from (18.13) is violated. The point with the greatest deviation from $X^{(o)}$ is substituted by the point above described point in way which will be shown. On the basis of (18.9) and (18.10) we conclude that it is possible to have simultaneously

(18.14) $$m^{(o)} < 0 \quad \text{and} \quad d^{(o)} < p(e^{(o)}).$$

Let x_q or $x_{q'}$ be one of the points satisfying(18.10) and (18.11) respectively. Setting

(18.15) $$x_o^{(1)} := \begin{cases} x_q & \text{if} \quad M^{(o)} - d^{(o)} > |m^{(o)}|, \\ x_{q'} & \text{if} \quad M^{(o)} - d^{(o)} \leq |m^{(o)}|, \end{cases}$$

and

$$(18.16) \qquad X^{(1)} := \{x_1^{(1)}, \ldots, x_{n+1}^{(1)}\}$$

where

$$x_i^{(o)} := \begin{cases} x_i^{(o)} & \text{for } i=1,\ldots,n+1,\ i \neq i_o \\ x_o^{(1)} & \text{for } i=i_o, \end{cases}$$

we have prepared the next step in the algorithm, i.e., from $X^{(o)}$ we proceed to $X^{(1)}$. If $x_o^{(1)} = x_q$, then $x_o^{(1)} \in E_+(g^{(1)})$ and if $x_o^{(1)} = x_{q'}$, then $x_o^{(1)} \in C_-(g^{(1)})$. Similarly with the uniform non-restricted approximation (see for example LAURENT [1972, p.151]), it may be proved that the sequences $(d^{(\tau)})\ \tau \in \mathbb{N}$, $(m^{(\tau)})\ \tau \in \mathbb{N}$, $(M^{(\tau)})\ \tau \in \mathbb{N}$, $(g^{(\tau)})\ \tau \in \mathbb{N}$, where $d^{(\tau)}$, $m^{(\tau)}$, $M^{(\tau)}$, $g^{(\tau)}$ are the elements which appear in the τ^{th}-step, are convergent when $\tau \to +\infty$.

In the case with alternance it is easy to exchange a point $x_i^{(o)} \in X^{(o)}$ by $x_o^{(1)}$. Indeed, let $x_{i_o-1}^{(o)}$ and $x_{i_o}^{(o)}$ be two points from $X^{(o)}$ so that $x_o^{(1)} \in (x_{i_o-1}^{(o)}, x_{i_o}^{(o)})$. Since we know the sign of the coefficient corresponding to $x_o^{(1)}$, which appears in the functional ℓ, we exchange with preservation of the alternance, one of the points $x_{i_o-1}^{(o)}$, $x_{i_o}^{(o)}$ by $x_o^{(1)}$.

19. THE CONNECTION BETWEEN ONE-SIDED AND UNCONSTRAINED APPROXIMATION

Let $\mathcal{H} := \text{span}\{g_1,\ldots,g_n\} \subseteq C[X]$ be of the type $I_n\{[a,b]\}$, $[a,b] \supseteq X$, and let $g_* = g_*(f;\mathcal{H};X)$, $g^* = g^*(f;\mathcal{H};X)$, $\bar{g} = \bar{g}(f;\mathcal{H};X)$ be the elements of the one-sided best approximation from below, above, and of unconstrained best approximation respectively. Then we have

THEOREM 19.1. Let g_*, g^*, $\bar{g}$ be the elements described as above. In order that a positive constant c exists such that

(19.1) $g_* + c = \bar{g} = g^* - c,$

it is necessary and sufficient that $g_1 = 1$.

PROOF. Let $g^* = g_* + 2c$ where $c > 0$, and let

$A(g_*) := E_+(g_*) \cup C_-(g_*)$, $A(g^*) := E_-(g^*) \cup C_+(g^*)$ be the sets of critical points relative to g_* and g^* respectively. Then $C_+(g^*) \cap C_-(g_*) = \emptyset$. Indeed, if we assume that the intersection is nonvoid, then must $C_+(g^*) = C_-(g_*)$. Therefore the elements g^* and g_* have $[\frac{n}{2}] + 1$ contact-points, which implies $g^* = g_*$ (see for example LAZAREVIĆ [1971]). It follows that we have

(19.2) $E_-(g^*) = C_-(g_*), \quad E_+(g_*) = C_+(g^*).$

Taking into account (19.2) and Corollary 17.2 we conclude that g_* and g^* may be obtained from (17.1) or (17.2) by means of one of the pairs (ϑ, η) or $(\bar{\vartheta}, \bar{\eta})$ of alternating vectors. Evidently $d^* = d_* = c$. From (18.3) we may write, with the pair (ϑ, η),:

$$g_* = g_*(f; \mathcal{H}; X) = -\frac{\begin{vmatrix} V\begin{pmatrix} g_1, \ldots, g_n \\ x_1, \ldots, x_n \end{pmatrix} & \begin{matrix} g_1 \\ \vdots \\ g_n \end{matrix} \\ f(x_1)-c \quad f(x_2) \quad f(x_3)-c \quad f(x_4) \quad \ldots & 0 \end{vmatrix}}{V\begin{pmatrix} g_1, \ldots, g_n \\ x_1, \ldots, x_n \end{pmatrix}}$$

and

$$g^* = -g_*(-f; \mathcal{H}; X) = \frac{\begin{vmatrix} V\begin{pmatrix} g_1, \ldots, g_n \\ x_1, \ldots, x_n \end{pmatrix} & \begin{matrix} g_1 \\ \vdots \\ g_n \end{matrix} \\ -f(x_1) \quad -f(x_2)-c \quad -f(x_3) \quad -f(x_4)-c \quad \ldots & 0 \end{vmatrix}}{V\begin{pmatrix} g_1, \ldots, g_n \\ x_1, \ldots, x_n \end{pmatrix}},$$

(see for example ELENA POPOVICIU [1972, p.34]). Substituting these values in the equality $g^* = g_* + 2c$, we obtain

$$(19.3) \qquad \begin{vmatrix} V\begin{pmatrix} g_1, \ldots, g_n \\ x_1, \ldots, x_n \end{pmatrix} & \begin{matrix} g_1 \\ \vdots \\ g_n \end{matrix} \\ 2c \quad 2c \quad \ldots \quad 2c & 2c \end{vmatrix} = 0.$$

Because $V\binom{g_1,\dots,g_n}{x_1,\dots,x_n} \neq 0$, in the determinant in (19.3) the rows are linearly dependent: there are the constants λ_k such that

(19.4) $$\sum_{k=1}^{n} \lambda_k\, g_k(x_i) = 2c, \qquad i = 1,\dots,n+1.$$

From the fact that $\mathcal{H}$ is I-space, the equality (19.4) remains valid for every $x \in X$. But, this implies that $\mathcal{H}$ contains the constant-functios, i.e., we may write $g_1 := 1$.

Now, let $g_1 := 1$ and $g_* = g_*(f;\mathcal{H};X)$. Then $g_* + d_* \in \mathcal{H}$. Further we have that the function $f - (g_* + d_*)$ takes the values $-d_*$ and 0 at $n+1$ distinct points. Therefore $g_* + d_* = g^* = g^*(f;\mathcal{H};X)$. It is clear that $\bar{g} = \frac{1}{2}\cdot(g_* + g^*)$.

The case when $X := [a,b]$ was investigated in LEWIS [1973], where the reader finds many references.

We note that in general case, even if $g_1 := 1$, the following inequality

$$\frac{1}{d_*} + \frac{1}{d^*} \leqq \frac{1}{\bar{d}}$$

is valid, KAMMERER [1959].

20. THE APPROXIMATION WITH POSITIVE ELEMENTS

Let X be a set on $\mathbb{R}$ containing at least $n+1$ distinct points and $f := \sum_{k=1}^{n} a_k\, g_k$ be a given element from $\mathcal{H}$; $\mathcal{H}$ is assumed to be an I-space of the type $I_n\{[a,b]\}$, $[a,b] \supseteqq X$. By $\mathcal{H}^{\perp}$ we denote the subset of $\mathcal{H}$ which contains all non-negative elements on X. Here we wish to find a pair (P^*, Q^*) of elements from $\mathcal{H}^{\perp}$ which satisfy the following minimum property

(20.1) $$P \geqq P^* \geqq 0, \quad Q \geqq Q^* \geqq 0 \qquad \text{on } X$$

and

(20.2) $$f := P^* - Q^* \qquad \text{on } X.$$

If we use the same reason as in § 15, we have

(20.3) $$p(f_+ - P^*) = \min_{P \in \mathcal{H}^{f_+}} p(f_+ - P)$$

and

$$p((-f)_+ - Q^*) = \min_{Q \in \mathcal{H}(-f)_+} p((-f)_+ - Q), \tag{20.4}$$

where

$$\mathcal{H}^{f_+} := \{g \in \mathcal{H}^{\perp} \mid \forall x \in X,\ \ g(x) \geqq f_+(x)\},$$

$$\mathcal{H}(-f)_+ := \{g \in \mathcal{H}^{\perp} \mid \forall x \in X,\ \ g(x) \geqq (-f)_+(x)\}.$$

Thus we have

$$P^* = g^*(f_+; \mathcal{H}; X), \quad Q^* = g^*((-f)_+; \mathcal{H}; X). \tag{20.5}$$

Evidently, if f does not change its sign on X, then the solution is trivial. If $f_+ = f$ on X, then we have $P^* = f$ and $Q^* = 0$. If $f_- = f$ on X, then $P^* = 0$ and $Q^* = -f$.

We remark that the solution of the best approximation of this kind (which is described as above), has a solution even if X contains only n distinct points. In this case, if $X := \{x_1, \ldots, x_n\}$, from (18.3) and (20.5) we get

$$P^* = L(\mathcal{H}; x_1, \ldots, x_n; f_+),$$

$$Q^* = L(\mathcal{H}; x_1, \ldots, x_n; (-f)_+).$$

In the case when $X := \{x_1, \ldots, x_{n+1}\}$, the solution is given by (18.6). If $\mathcal{H} := \mathcal{P}_n$ we have the problem proposed by T.POPOVICIU.

Finally we note that if either P^* or Q^* is determined, the other may be found according to (20.2).

21. THE BEST ONE-SIDED APPROXIMATION WITH WM-SYSTEMS

We continue to use the notations as in the preceding sections. Let $g_1, \ldots, g_n$ be a WM-system (Weak Markoff system) on X and let $\mathcal{H} := \operatorname{span}\{g_1, \ldots, g_n\}$. We have

THEOREM 21.1. The element $g_* := \sum_{k=1}^{n} \alpha_k g_k \in \mathcal{H}$ is $g_* = g_*(f; \mathcal{H}; X)$ iff the following conditions are fulfilled:

There exists the r-tuple $(x_1, \ldots, x_r) \in \Delta_r(X)$,

$1 \leq r \leq n+1$ and the set of indices $k_1,\dots,k_{r-1}$, such that

$$1^\circ \qquad e_*(x_{k_i}) = \frac{1+\sigma(x_{k_i})}{2} d_*,$$

$$2^\circ \qquad V\begin{pmatrix} g_{k_1},\dots & \dots,g_{k_{r-1}} \\ x_1,\dots,x_{i-1},x_{i+1},\dots,x_r \end{pmatrix} > 0, \quad i=1,\dots,r,$$

$$3^\circ \qquad V\begin{pmatrix} g_{k_1},g_{k_2},\dots,g_{k_{r-1}},g_k \\ x_1,\dots & \dots,x_{r-1}, x_r \end{pmatrix} = 0 \qquad \forall k=1,\dots,n.$$

PROOF. We have seen in §15.iii and §15.iv that, inespective of additional properties of the space $\mathcal{H}$, we have

$$(g_* = g_*(f;\mathcal{H};X)) \Longleftrightarrow \theta_{\mathbb{R}^n} \in \mathrm{co}\,(\{\sigma(x)(g_1(x),\dots$$
$$\dots,g_n(x)) \mid \forall x \in A(g_*)\}) =: C$$

where the sign-function $\sigma(\cdot)$ and the set $A(g_*)$ corresponding to the deviation-function $e_* := f - g_*$, are given by (15.10) and (15.11). We will prove that

$$(1^\circ \wedge 2^\circ \wedge 3^\circ) \Longleftrightarrow \theta_{\mathbb{R}^n} \in C.$$

Let r be the minimal number of n-tuples from the above convex hull C needed to represent the vector $\theta_{\mathbb{R}^n}$. Thus there exist positive numbers $\lambda_1,\dots,\lambda_r$ and a r-tuple $(x_1,\dots,x_r) \in \Delta_r(X)$ such that

$$\sum_{i=1}^{r} \lambda_i = 1$$
$$\sum_{i=1}^{r} \lambda_i \sigma(x_i)(g_1(x_i),\dots,g_n(x_i)) = \theta_{\mathbb{R}^n},$$

i.e.,

$$\sum_{i=1}^{r} \lambda_i = 1$$

(21.1) $$\sum_{i=1}^{r} \lambda_i \sigma(x_i)\, g_k(x_i) = 0, \qquad k = 1,\dots,n.$$

($r \leq n+1$ is proved similarly as in the proof of the CARATHEODORY Theorem, see for example CHENEY [©1966, pp. 17-18]). Since the system (21.1) of $n+1$ equations in r unknowns

$_1, \ldots, {}_r$, allows a nontrivial solution. It follows that there exists a subsystem of (21.1):

$$\begin{aligned} &\sum_{i=1}^{r} \lambda_i = 1 \\ (21.2) \qquad &\sum_{i=1}^{r} \lambda_i\, \sigma(x_i)\, g_k(x_i) = 0, \qquad k = k_1, \ldots, k_{r-1} \end{aligned}$$

wherefrom the mentioned solution is obtained. Due to the nontriviality of the solution, the equation $\sum_{i=1}^{r} \lambda_i = 1$ belongs obligatly to (21.2).

The principal determinant of (21.2):

$$D = \begin{vmatrix} 1 & \cdots & 1 \\ \sigma(x_1) g_{k_1}(x_1) & \cdots & \sigma(x_r) g_{k_1}(x_r) \\ \vdots & & \\ \sigma(x_1) g_{k_{r-1}}(x_1) & & \sigma(x_r) g_{k_{r-1}}(x_r) \end{vmatrix} =$$

$$= \left(\prod_{i=1}^{r} \sigma(x_i)\right) \begin{vmatrix} \sigma(x_1) & \cdots & \sigma(x_r) \\ g_{k_1}(x_1) & \cdots & g_{k_1}(x_r) \\ \vdots & & \\ g_{k_{r-1}}(x_1) & & g_{k_{r-1}}(x_r) \end{vmatrix},$$

i.e.,

$$(21.3) \qquad D = \left(\prod_{i=1}^{r} \sigma(x_i)\right) V \begin{pmatrix} \sigma(\cdot), g_{k_1}, \ldots, g_{k_{r-1}} \\ x_1, \ldots \quad \ldots, x_r \end{pmatrix}$$

is not zero. On the other hand, solving this system, we have

$$(21.4) \qquad D \lambda_\nu = D_\nu, \qquad \nu = 1, \ldots, r$$

where $D_\nu =$

$$= \begin{vmatrix} 1 & \cdots & 1 & 1 & 1 & \cdots & 1 \\ \sigma(x_1) g_{k_1}(x_1) & \cdots & \sigma(x_{\nu-1}) g_{k_1}(x_{\nu-1}) & 0 & \sigma(x_{\nu+1}) g_{k_1}(x_{\nu+1}) & \cdots & (x_r) g_{k_1}(x_r) \\ \vdots & & & & & & \\ \sigma(x_1) g_{k_{r-1}}(x_1) & & \sigma(x_{\nu-1}) g_{k_{r-1}}(x_{\nu-1}) & 0 & \sigma(x_{\nu+1}) g_{k_{r-1}}(x_{\nu+1}) & & (x_r) g_{k_{r-1}}(x_r) \end{vmatrix},$$

i.e.,

$$(21.5)\qquad D_\nu = (-1)^\nu \sigma(x_\nu)\left(\prod_{i=1}^{r}\sigma(x_i)\right) V\begin{pmatrix} g_{k_1}, & \cdots & & \cdots, g_{k_{r-1}} \\ x_1, & , x_{\nu-1}, & x_{\nu+1}, & , x_r \end{pmatrix}.$$

Since we deal with WM-system, the prevronskian in (21.5) must be nonnegative. Moreover, on the basis of $D \neq 0$, from (21.4) we conclude

$$(21.6)\qquad D_\nu > 0, \qquad \nu = 1,\ldots,r,$$

which implies 1^0 and 2^0. If the first equation in (21.2) is substituted by any equation from (21.1), excluding the first, we obtain a homogeneous system of r equations (in unknowns $\lambda_1,\ldots,\lambda_r$). From nontriviality of these solutions, 3^0 results. We remark that, taking into account the nature of λ_i, from (21.4) and (21.6), $D > 0$ results, as well as, on the basis of (21.3):

$$(21.7)\ \sigma(x_\nu) = (-1)^\nu \ \operatorname{sign}\left(\left(\prod_{i=1}^{r}\sigma(x_i)\right) V\begin{pmatrix} g_{k_1}, & \cdots & & \cdots, g_{k_{r-1}} \\ x_1, & , x_{\nu-1}, & x_{\nu+1}, & , x_r \end{pmatrix}\right)$$

Now, from 2^0 and 3^0 it follows thatthe rank of matrix $\|\sigma(x_i)g_k(x_i)\|$, $(x_1,\ldots,x_r) \in \Delta_r(X)$, $k = 1,\ldots,n$, is $r-1$. Thus its columns (totalling to r) are linearly dependent, i.e., there exist the numbers $\gamma_1,\ldots,\gamma_r$ with $\sum_{i=1}^{r}|\gamma_i| \neq 0$, such that

$$\sum_{i=1}^{r}\gamma_i \begin{Vmatrix} \sigma(x_i)g_1(x_i) \\ \vdots \\ \sigma(x_i)g_n(x_i) \end{Vmatrix} = \begin{Vmatrix} 0 \\ \vdots \\ 0 \end{Vmatrix} =: \theta_{\mathbb{R}^n} .$$

If we divide this equality by $\sum_{i=1}^{r}|\gamma_i|$ and write $\mu_i := \gamma_i / \sum_{i=1}^{r}|\gamma_i|$, we get

$$\sum_{i=1}^{r}\mu_i = 1$$

$$(21.8)\qquad \sum_{i=1}^{r}\mu_i \begin{Vmatrix} \sigma(x_i)g_1(x_i) \\ \vdots \\ \sigma(x_i)g_n(x_i) \end{Vmatrix} = \theta_{\mathbb{R}^n} .$$

But (21.8) is of the same form as (21.1) and we may repeat the discussion of solvability as in (21.1) except that we do not have a priori assumption that all $_i$ are positive. But, on the basis of (21.7) we conclude that $\sigma(x_i)\sigma(x_{i+1}) < 0$. From this and from 2^0 and (21.5) it follows that

$$\text{sign } D_\nu = \text{sign } D_{\nu+1}, \quad \nu=1,\dots,r.$$

Similarly from (21.3) and 2^0 it follows $D \neq 0$. On the basis of (21.4) we conclude that sign μ_i = sign μ_{i+1}, $i=1,\dots,r-1$. Evidently, we may take $\mu_i > 0$, $i=1,\dots,r$, and (21.8) shows that $\Theta_{\mathbb{R}^n}$ belongs to the set C. ◻

The case with unconstrained approximation of this kind is given in MICCHELLI [1974].

It is easy to see that it is possible to get the similar results in different cases. For example, the set $\{1,x,\dots,x^{n-1},(x-\xi_1)_+^{n-1},\dots,(x-\xi_r)_+^{n-1}\}$ constitutes a WM-system on [0,1]. Thus the approximation in space of s p l i n e functions may be obtained similarly as above. The approximation in I_{n-k}-spike or in every set which may be related to TP-kernels, can be studied as above.

Any two norms defined on a finite-dimensional linear space, are equivalent. Hence we have similar results if we consider the cases in different normed spaces. For example if we consider the psace L^p, then using the formulas of §14.(Example 14.iii), we may calculate the necessary cones and for one-sided approximation we may obtain the results, as for example, BOJANIC & DeVORE [1966], DeVORE [1968].

R E Z I M E

Problemi ekstremuma bili su predmet proučavanja već
u najstarijim etapama razvoja Matematike. U novije vreme
mnogi problemi ekstremuma bili su objedinjeni u okviru Va-
rijacionog računa koji je dobio široku primenu u mnogim
naučnim oblastima. Proučavanje ekstremuma funkcionala, mo-
že se reći, počinje sa ČEBIŠEVljevom teorijom aproksimacija.
Po uzoru na LAGRANGEovu teoriju multiplikatora razvila se
Teorija ekstrmuma funkcionala sa ograničenjima. Za mnoge
probleme optimizacije dobijeni su potrebni uslovi za ekst-
remum u obliku PONTRJAGINovog principa maksimuma koji je
imao jedan nov karakter koji je različit od klasične teo-
rije Varijacionog računa. Proučavanje ekstremuma glatkih
ra-funkcionala na skupovima sa delimično glatkim rubovima ra-
zvilo se četrdesetih godina našeg veka u vidu Matematičkog
programiranja. Dvadesetak godina docnije dobijeni su potre-
bni uslovi za ekstremum u analitičkoj formi. Pomoću tih re-
zultata mogli su da se sintetizuju svi do tada poznati re-
zultati teorije ekstremuma (videti na pr. ДУБОВИЦКИЙ &
& МИЛЮТИН [1965]). U najnovije vreme dato je više mo-
dela za dobijanje potrebnih uslova za ekstremum. Navodimo
samo neke: ПОНТРЯГИН & БОЛТЯНСКИЙ & ГАМКРЕЛИДЗЕ &
& МИЩЕНКО [1961], РУБИНШТЕЙН [1963], ПШЕНИЧНИЙ
[1965], ДУБОВИЦКИЙ & МИЛЮТИН [1965], NEUSTADT [1966]
i drugi. Medjutim, još uvek se mnogi problmi optimizacije
proučavaju izolovano. Jedna jedinstvena i opšta teorija ko-
ja bi predstavljala sintezu različitih partikularnih rezul-
tata još uvek ne postoji. Jedan od ozbiljnih razloga leži
u poznatoj razlici izmedju potrebnih i dovoljnih uslova za
ekstremum.

U ovom radu izložen je metod konusa u teoriji opti-
mizacije. Ovaj metod pokazao se efikasnijim od mnogih dru-
gih i čini se da će se u tom smislu još intenzivnije razvi-

jati. Pojmovi t a n g e n c i j a l n e p o v e z a n o s t i skupa i t a n g e n c i j a l n o r a z d v o j i v o g operatora (videti §9.) omogućili su nam da dobijemo nove rezultate o uslovima za ekstremum. Dobijena je tešnja veza izmedju potrebnih i dovoljnih uslova. Smatramo da t a n g e n c i j a l n a svojstva koja smo dobili u §9. i §10. omogućuju "potpunu afirmaciju" metode konusa u problemima optimizacije.

Ideja o povezivanju Analize sa konusima nije nova. Još u Starom veku APOLONIJE i DANDELEN proučavaju osobine krivih linija drugoga reda pomoću geometrije konusa. CARATHÉODORY [1907] primenom Geometrije i konusa ispituje osobine analitičkih funkcija. POPOVICIU [1945], KREIN [1951], KARLIN & SHAPLEY [1953], ROGOSINSKI [1958], ZIGLER [1966] i drugi razradili su i objasnili svu teoriju T-sistema pomoću geometrijskih metoda, preciznije pomoću teorije konusa. Jedna iscrpna studija geometrijskih metoda data je u KARLIN [1968]. Pojam konveksnosti višega reda koji uvodi T. POPOVICIU u svojim radovima, omogućava jednu potpunu klasifikaciju funkcija (realnih) prema odredjenim geometrijskim osobinama koje upravo su i potrebne u izgradjivanju teorije konusa i primene ovih u teoriji optimizacije.

Prvi kapitol ove teze uvodnog je karaktera. Daju se definicije i pojmovi koji se koriste u radu a prilagodjeni su za primenu konusa. Različiti fenomeni interpolacija objedinjeni su preko pojmova t o t a l n o - p o z i t i v n i h j e z g a r a koji se pridružuju interpolacionim sistemima. Uslovi jedinstvenosti rešenja interpolacionog problema izražavaju se na prirodan način pomoću totalne pozitivnosti pridruženog jezgra. Date su još definicije koje su vezane za pojam konveksnosti višega reda i utvrdjena šema označavanja i notacija.

Drugi kapitol posvećen je konusima. Specijalnu ulogu dobili su m o g u ć i i a d h e r e n t n i konusi. Za skup Q iz linearnog prostora X, definiše se mogući konus $K(Q;\, x_0)$ sa vrhom u tački x_0 :

$$K(Q;\, x_0) := \{ h \in X \mid \exists\, \mathcal{V}(h),\ \exists\, \varepsilon > 0,\ \forall y \in \mathcal{V}(h),\ \forall \eta \in (0,\varepsilon) : x_0 + \eta y \in Q \}$$

i adherentni konus $K[Q;\ x_0]$ sa vrhom u istoj tački:

$$K[Q;\ x_0]:=\{h\in X \mid \forall \vartheta(h),\ \forall \varepsilon>0,\ \exists y_\varepsilon\in \vartheta(h),$$
$$\exists \eta_\varepsilon\in(0,\varepsilon)\ :\ x_0+\eta_\varepsilon y_\varepsilon\in Q\},$$

(vidi Def. 5.1 i 5.2).

Neka su X, Y, Z topološki linearni prostori u kojima su $D\subseteq X$, $C\subseteq Y$ i $z_0\in Z$ dati. Neka je dalje $\mathbb{T}:=(\mathbb{T},\leqq)$ dati realan poluuredjen topološki linearni prostor. Data su takodje preslikavanja $f\colon D\to\mathbb{T}$, $F\colon D\to Y$, $G\colon D\to Z$ i uočimo skupove

$$Q:=\{x\in D\mid F(x)\in C\},$$
$$M:=\{x\in D\mid G(x)=z_0\}.$$

Tada imamo jedan og glavnih polaznih rezultata:Ako je za tačku $x_0\in A:=Q\cap M$,

$$f(x_0)\leqq f(x),\qquad \forall x\in A,$$

onda važi

$$K(P;\ x_0)\cap K(Q;\ x_0)\cap K[M;\ x_0]=\emptyset \tag{22.1}$$

gde je

$$P:=\{x\in D\mid f(x)<f(x_0)\}.$$

Ako u problemu optimizacije imamo više restrikcija, što je u praksi čest slučaj, možemo staviti u (22.1) $Q=\bigcap_{i\in I} Q_i$ i $M=\bigcap_{j\in J} M_j$ gde su I, J dati skupovi indeksa. Ovim se odmah nameće pitanje valjanosti inkluzija oblika

$$(22.2)\quad K(\cap S_i;\ x_0)\subseteq\cap K(S_i;\ x_0),\qquad K[\cup S_i;\ x_0]\supseteq\cup K[S_i;\ x_0],$$
$$K(\cup S_i;\ x_0)\supseteq\cup K(S_i;\ x_0),\qquad K[\cap S_i;\ x_0]\subseteq\cap K[S_i;\ x_0]$$

gde su S_i dati skupovi. Takodje su od interesa inkluzije "dualne" ovima o čemu ćemo govoriti docnije. Radi efektivnog izračunavanja mogućih i adherentnih konusa tj. za iznalaženje analitičkih izraza za njih, (§8.), uveli smo više pojmova i definicija kao što su d i f e r e n c i r a n j e i i z v o d p o k o n u s u, (§6.), t a n g e n c i j a l n a p o v e z a n o s t s k u p a, t a n g e n c i j a l n o s e p a r a b i l a n o p e r a t o r, (§§9., 10.), k o n v e k s a n (k o n k a v a n) o p e r a t o r, (§ 12.) i druge. Sve ove definicije uvode se u ovom radu prvi put. Teoremom 10.2 generališe se LUSTERNIKov rezultat (vidi ЛЮСТЕРНИК [1934]) koji je dugo bio jedini rezultat te vrste i imao veliki uticaj na dalja proučavanja ekstremuma sa ograničenjima. Pomenuti naš rezultat obuhvata i slučajeve kada u problemu optimizacije restrikcije nemaju strukturu linearnog podprostora osnovnog prostora. Teorema se takodje odnosi na slučajeve kada operatori koji se posmatraju nisu FRECHET-diferencijabilni. Rezultat je dobijen korišćenjem izvoda po konusu. Diferencijabilnost po konusu kao i tangencijalna svojstva, (§§ 9.,10.), rasvetljavaju činjenicu nedovoljnosti egzistencije parcijalnih izvoda za diferencijabilnost što nam je poznato u Klasičnoj analizi.

U trećem kapitolu najpre se proučavaju inkluzije oblika

$$K(\cap S_i;\ x_o) \supseteqq \cap K(S_i;\ x_o),\quad K[\cup S_i;\ x_o] \subseteqq K[S_i;\ x_o],$$

(22.3)

$$K(\cup S_i;\ x_o) \supseteqq \cup K(S_i;\ x_o),\quad K[\cap S_i;\ x_o] \supseteqq \cap K[S_i;\ x_o].$$

Na osnovu (22.2) i (22.3) dobijeni su egzaktni analitički izrazi za moguće i adherentne konuse za mnoge skupove koji imaju odredenu ulogu u problemima optimizacije. Na osnovu pomenutih inkluzija analizirani su takodje potrebni i dovoljni uslovi za ekstremum; daje se pojam r e g u l a r n o s t i restrikcija u problemu optimizacije. U okviru potrebnih uslova za ekstremum u § 12. daje se pravilo multiplikatora. Ovde su dobijeni rezultati za koje smatramo da su najvredniji u ovoj tezi. U teoremama 12.1 i 12.3 sadržan je jedan princip maksimuma u najopštijem obliku i koji objedinjuje mnoge poznate rezultate te vrste. Naredni odeljak posvećen je dovoljnim uslovima za ekstremum a kapitol se završava studiranjem k o n v e k s n i h p r o b l e m a optimizacije u kojima potrebni uslovi su istovremeno i dovoljni. Napomenimo još da se iz rezultata koje smo dobili u 13. vidi da su konveksnost restriktivnih skupova i funkcije cilja vrlo strogi pri formulisanju dovoljnih uslova za ekstremum. Vrednost ovog odeljka upravo je u tome.

U poslednjem kapitolu daju se primene dobijenih rezultata u prethodna dva kapitola. Izabrani su problemi u kojima se može uspešno primeniti (i ilustrovati) metod konusa. Posmatrani su problemi uniformne obostrane i jednostrane najbolje aproksimacije u interpolacionim sistemima sa odredjenim svojstvima totalne pozitivnosti pridruženih jezgara. U § 19. data je veza izmedju pomenutih aproksimacija. Kapitol se završava primerima aproksimacija u MARKOVljevim sistemima. Rad ima 21 odeljak. U registru, na kraju, dat je pregled pojmova a u tekstu posebno je istaknuto 10(5) definicija. Rad sadrži takodje 20(14) Lema i Teorema, (4) Posledica, (5) Primedaba, 15(11) Primera, (3) Crteža. Brojevi u zagradi označavaju broj priloga autora. Bibliografija sadrži 47 bibliografskih navoda.

B I B L I O G R A P H Y

ALJANČIĆ, S.

1968 Uvod u Realnu i funkcionalnu analizu; Gradj. knjiga, Beograd.

BOJANIC, R. and R. DeVORE

1966 On Polynomials of Best One-Sided Approximation; L'Enseignement Math. t.XII, fasc.3, 139-164.

BOULIGAND, G.

1932 Introduction à la Géométrie Infinitésimale directe; Paris.

CHENEY, E. W.

1966 Introduction to Approximation Theory; McGraw-Hill Book Company, New York.

DeVORE, R.

1968 One-Sided Approximation of Functions; Jour. of Approx. Theory 1, 11-25.

ДУБОВИЦКИЙ, А. Я., и А. А. МИЛЮТИН

1963 Задачи на экстремум при наличии ограничений; ДАН СССР, 149, № 4, 759-762.

1965 Задачи на экстремум при наличии ограничения; Ж. Выч. мат. и мат. физ. 5. №3, 395-453.

ДЕМЬЯНОВ, В. Ф. и А. М. РУБИНОВ

1968 Приближенные методы решения экстремальных задач, Изд. Ленинградского университета.

ГИРСАНОВ, И. В.

1970 Лекции по математической теории экстремальных задач; Изд. Моск. Унив.

ГОНЧАРОВ, В. Л.
1954 Теория интерполирования и приближения функций; Гос. изд. Москва-Ленинград.

HOFFMANN, K.-H. and I. KOLUMBAN
1974 Verallgemeinerte Differenzierbarkeitsbegriffe und ihre Anwendung in der Optimierungstheorie; Computing 12, 17-41.

KAMMERER, W.
1959 Optimal Approximation of Functions; One-Sided Approximation and Extreme Preserving Approximation; Doctoral thesis, Univ. of Wisconsin, Madison.

KARLIN, S. J.
1968 Total Positivity, vol.I; Stanford Univ. Press, Stanford, California.

KARLIN, S. J. and W. J. STUDDEN
1966 Tchebycheff Systems: With Applications in Analysis and Statistics; Interscience Publishers. New York.London.Sydney.

KREIN, M. G.
1951 The Ideas of P.L. ČEBYŠEV and A.A. MARKOV in the Theory of Limiting Values of Integrals and Their Further Development; Amer. Math. Soc. Transl. Ser. 2.,12, 1-12 .

KARLIN, S.J. and L.S. SHAPLEY
1953 Geometry of Moment Spaces; Mem. Am. Math. Soc. № 12.

LAZAREVIĆ, I.
1971 Some Properties of n-Parameter Families of Functions; Univ. Beograd, Publ. Elektrotehn. Fak., Ser. Mat. Fiz. №357-380, 101-106.

1974 On Best One-Sided Approximation With Interpolatory Families; to apper, Revue D'Analyse Numerique et de la Theorie de L'Approximation, t.3., №2.

LAZAREVIĆ, I. and I. KOLUMBAN
1974 The Role of the Multipliers in Optimal Problems; Mathematica Balkanica 4., 127-135.

LAURENT, P. J.
1972 Approximation et Optimisation; Paris.

LEMPIO, F.
1972 Tangentialmannigfaltigkeiten und Infinite Optimierung; Habilitationsschrift, Hamburg.

LEWIS, J. T.
1973 Approximation With Convex Constraints; SIAM Review, vol.15., N 1., 193-217.

ЛЮСТЕРНИК, Л. А.
1934 Об условных экстремах функционалов; Мат. Сборник, т. 41. №3., 390-401.

MICCHELLI, C. A.
1974 Characterization of Chébyshev Approximation by Weak Markoff Systems; Computing 12/1, 1-8.

МИЛЮТИН, А. А.
1966 Задачи на экстремум при наличии ограничений; Докторска дисертация, М. ИПМ.АН.СССР.

MITRINOVIĆ, D. S.
1970 Analytic Inequalities; Springer-Verlag, Berlin-Heidelberg-New York.

MOLDOVAN, ELENA
1955 Asupra unei generalizări a notiunei de convexitate; Studii și cerc. St. Seria

Mat.(Cluj), VI nr.3-4., 65-73.

МОРОЗОВ, М.И.

1952 О некоторых вопросах равномерного приближения непрерывных функций посредством функций интерполяционных классов; Изв. Акад. Наук, Сер. Мат. 16., 75-100.

MOTZKIN, T. S. and A. SHARMA

1966 Next-to-Interpolatory Approximation on Sets With Multiplicities; Canadian Jour. of Math. XVIII, №6., 1196-1211.

NEMETH, A. B.

1970 Nonlinear Differential n-Parameter Families; Revue Romaine de Math. Pures et Appliquées; t.XV, № 1., 111-118.

NEUSTADT, L.W.

1966 An abstract variational theory with applications to a broad class of optimization problems; J.SIAM Control 4, 3.

POPOVICIU, ELENA

1972 Teoreme de medie din analiza matematica si legatura lor cu teoria interpolari; Editura dacia, Cluj.

POPOVICIU, T.

1933 Sur quelques proprietés des functions d'une ou de deux variables réelles; Mathematica (Cluj), VIII, 1-85.

1936 Notes sur les functions convexes d'ordre supérieur (II), Mathematica (Cluj), XII, 227-233.

1939 Notes sur les functions convexes d'ordre supérieur (VI); Revue Math. de L'Union Interbalkanique, II, fasc. III-IV, 1-10.

1945 Les Fonctions Convexes; Herman and Cie, Paris.

1959 Sur le reste dans certaines formules linéaires D'Approximation de L'Analyse; Mathematica (Cluj), vol. 1 (24), 95-142.

ПОНТРЯГИН, Л.С. & В.Г. БОЛТЯНСКИЙ & Р.В. ГАМКРЕЛИДЗЕ & Е.Ф. МИЩЕНКО
1961 Математическая теория оптимальных процессов; Физматгиз.

ПШЕНИЧНЫЙ, Б.Н.
1965 Выпуклое программирование в нормированных пространствах; Кибернетика 1, 5.

RICE, J. R.
1964 The Approximation of Functions, vol.I ; Addison-Wesley Publishing Company, INC., Reading, Massachusetts.Palo Alto. London.

ROBERTS, A. W. and D. E. VARBERG
1973 Convex Functions; Acad. Press, New York-London.

ROGOSINSKI, W.W.
1958 Moments of Non-negative Mass; Proc. Roy. Sci. (London), A, 245, 1-27.

РУБИНШТЕЙН, Г.Ш.
1963 Двойственные экстремальные задачи; ДАН СССР, 152.

TAYLOR, G. D.
1968 On Approximation by Polynomials Having Restricted Ranges; SIAM Jour. Numer. Anal. 5, 258-268.

1969 Approximation by Functions Having Restricted Ranges, III, ; Jour. Math. Anal. Appl. 27, 241-248.

TORNHEIM, L.
1950 On n-Parameter Families of Functions and

Associated Convex Functions; Trans. Amer. Math. Soc. 69, 457-467.

ZIEGLER, ZVI

1966 Generalized Convexity Cones; Pacific J. Math. 17, 561-580.

S U B J E C T   I N D E X