CONVEX CONES, SETS, AND FUNCTIONS

by

W. FENCHEL

Princeton University
Department of Mathematics

September 1953

CONVEX CONES, SETS, AND FUNCTIONS

by

W. FENCHEL

from notes by D. W. Blackett
of lectures at Princeton University,
Spring Term, 1951

Princeton University
Department of Mathematics
Logistics Research Project
sponsored by the
Office of Naval Research
(Contract N-6-ONR-27011)

September, 1953

(includes corrections)

TABLE OF CONTENTS

ACKNOWLEDGEMENT

The author wishes to express his gratitude to Professor A. W. Tucker for giving him the opportunity to write this report and for calling his attention to the problems dealt with in the final sections (pp. 105-137). The author is also indebted to Professors J. W. Green and H. W. Kuhn for critical remarks, and especially to Dr. D. W. Blackett for his valuable help in the preparation of this report.

PREFACE

The following notes contain a survey of those properties of convex cones, convex sets, and convex functions in finite dimensional spaces which are most frequently used in other fields. Emphasis is given to results having applications in the theory of games and in programming problems.

Chapters I and II center about the interaction of the two features of convexity in linear spaces and affine spaces: 1. the original definition of a convex set as a set containing all segments whose endpoints are in the set and 2. the existence of a support through every boundary point. The convex hull of a set is the set of all centroids of points in the given set, while its closure is the intersection of all halfspaces containing the set. This fact may be considered as the kernel of many of the applications of the concept of convexity. It indicates also the important (though not quite complete) self-duality of the theory. The projective and - it is believed - most general formulation of this duality is given at the end of Chapter II.

The first part of Chapter III deals with the well-known elementary properties of continuous convex functions. No differentiability assumptions are made, but the directional derivative which always exists is investigated and used rather extensively. The second part of the chapter contains recent investigations. By means of a suitable polarity, an involutory correspondence between convex functions is established and applied to a generalized convex programming problem. Finally the level sets of a convex function are studied and the existence of a convex function with given level sets is discussed.

Since the end of the last century numerous papers have dealt mainly or partially with convex sets or functions.

Many results have been discovered several times in different formulations - often adapted to particular applications in other fields. No attempt has been made in these notes to quote for each theorem the first paper in which it appears in the formulation chosen here. In fact most of the basic concepts and results can be traced back in one form or another to the very first papers on the subject. Short historical notes and references are gathered at the end of this report.

Chapter I

CONVEX CONES

§1. PRELIMINARIES

Let L^n be an n-dimensional Euclidean vector space with origin 0, vectors x,y, ..., inner product (x,y), norm $\|x\| = \sqrt{(x,x)}$, and metric $d(x,y) = \|x - y\|$. Identify the vector x with the n-tuple $\begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}$ of its coordinates with respect to a particular orthonormal basis of L^n.

Then $(x,y) = x'y = \sum_{i=1}^{n} x_i y_i$.

A subset M of L^n is called a <u>cone</u> if 0 is in M and $x \in M$ implies $\lambda x \in M$ for every non-negative real scalar λ. The particular cones consisting of a non-zero vector x and all its multiples λx $(\lambda \geq 0)$ are <u>rays</u>. A cone which contains at least one non-zero vector is therefore just the union of the rays it contains.

Since all non-trivial cones may be thought of as sets of rays, it is desirable to introduce a topology on these rays from the topology on L^n. This might be done by defining the angle

$$\phi(x,y) = \text{arc cos } \frac{x'y}{\|x\| \ \|y\|} \quad (0 \leqq \phi \leqq \pi)$$

as a metric on $L^n - 0$. This angle depends only on the rays (x) and (y) to which x and y belong. It may be thought of as the angle between the two rays. The proof that this angle is indeed a metric for the rays, in particular that it satisfies the triangle inequality, is not obvious. An equivalent metric is

$$[x,y] = \sqrt{2 - \frac{2x'y}{\|x\|\,\|y\|}} \, .$$

This new metric is the chord distance between the two points $\frac{x}{\|x\|}$ and $\frac{y}{\|y\|}$ on the unit sphere $\|z\| = 1$. That is

$$[x,y] = d\left(\frac{x}{\|x\|}, \frac{y}{\|y\|}\right).$$

Clearly $[x,y]$ depends only on the rays (x) and (y). $[x,y]$ also satisfies the defining conditions for a metric on the space of rays. The geometric description shows that the two metrics are topologically equivalent.

A <u>sequence of rays</u> (x^{ν}) is said to <u>converge</u> to a ray (x) if $[x^{\nu}, x] \rightarrow 0$. A ray (x) is called a <u>limit ray</u> of a cone M if there is a sequence of rays of the cone which are different from (x) and which converges to (x). A <u>closed cone</u> or a closed set of rays is a cone which contains all its limit rays. A cone is closed in this sense if and only if it is closed in the usual topology of L^n. A cone is <u>open</u> if and only if the complementary set of rays is a closed cone. This is equivalent to the definition:

> DEFINITION. M is open if and only if for every (x) in M there is an $\epsilon > 0$ such that all rays (y) with $[x,y] < \epsilon$ are in M.

The set of such rays (y) is called an <u>ϵ-neighborhood</u> of (x). An open cone as a set in L^n is an open set of L^n plus the origin. A ray (x) is called an <u>interior ray</u> of a cone M if M contains an ϵ neighborhood of (x) for some $\epsilon > 0$. A ray (x) such that the complementary cone to the cone M contains a neighborhood of (x) is called an <u>exterior ray</u> of M. A <u>boundary ray</u> of a cone M is a limit ray of M which is not an interior ray of M.

With any cone M there is associated a smallest linear subspace S(M) of L^n which contains M. This space may be defined as the intersection of all subspaces containing M. The dimension d(M) of the space S(M) is called the linear dimension of the cone M. In the theorems which follow S(M) will often play a more important role than L^n itself. For these results cones, open or closed relative to S(M), and interior, exterior, and boundary rays relative to S(M) will be considered rather than their counterparts in the topology of the full space L^n. They will be called for simplicity relative interior, relative exterior, and relative boundary rays.

§2. CONVEX CONES

A cone C is convex if the ray (x+y) is in C whenever (x) and (y) are rays of C. Thus a set C of vectors is a convex cone if and only if it contains all vectors

$$\lambda x + \mu y \ (\lambda, \mu \geq 0;\ x, y \in C).$$

The largest subspace s(C) contained in a convex cone C is called the lineality space of C and the dimension l(C) of s(C) is called the lineality of C.

LEMMA 1. If (x) is an interior ray of a convex cone C relative to S(C) and (y) is a boundary or interior ray of C relative to S(C), every ray $(\lambda x + \mu y)$, where λ and μ are positive real numbers, is an interior ray of C relative to S(C).

PROOF:

Case 1: (y) = (-x), that is $(\lambda x + \mu y) = (x)$ or (y). It will be shown that C = S(C). It may be assumed that $y \in C$.

Otherwise there is a $y^* \in C$ so close to y that $(x^*) = (-y^*)$ is in a neighborhood of (x) contained in C. Thus, x^* and y^* satisfy the assumptions of the lemma. Let $z \neq 0, \pm x$ be any vector in $S(C)$. Consider the plane P spanned by x and z. Now $C \cap P$ contains an angle around (x). In this angle there is a ray $(\bar{x})$ such that z is in the ~~acute~~ convex angle determined by $(\bar{x})$ and (y). Hence z is a linear combination of $\bar{x}$ and y with positive coefficients. Therefore z is in C and $C = S(C)$. The lemma follows in this case because every ray of $S(C)$ is relative interior to $S(C)$.

Case 2: $(y) \neq (-x)$ which implies $(\lambda x + \mu y) \neq (y)$.

2.1: $y \in C$. There is an $\eta > 0$ such that C contains an η-neighborhood of (x) relative to $S(C)$. It has to be shown that there is an $\epsilon > 0$ such that C contains an ϵ-neighborhood of $(\lambda x + \mu y)$ relative to $S(C)$. Consider first an arbitrary $\epsilon > 0$. Let (z) be any ray in the ϵ-neighborhood of $(\lambda x + \mu y)$. Put $z = \lambda x + \mu y + v$ and suppose z is normalized so that $\|z\| = \|\lambda x + \mu y\|$. Then

$$[\lambda x + \mu y + v, \lambda x + \mu y]^2 = \frac{\|v\|^2}{\|\lambda x + \mu y\|^2} < \epsilon^2,$$

hence

$$\|v\|^2 < \epsilon^2 \|\lambda x + \mu y\|^2 .$$

Consider now the vector $x + \frac{1}{\lambda}v$ for which $\lambda(x + \frac{1}{\lambda}v) + \mu y = z$. The distance of $(x + \frac{1}{\lambda}v)$ from (x) satisfies

$$[x + \tfrac{1}{\lambda}v, x]^2 = 2 - 2\frac{\|x\|^2 + \frac{1}{\lambda}v'x}{\|x + \frac{1}{\lambda}v\|\,\|x\|} < 2 - 2\frac{\|x\| - \frac{1}{\lambda}\|v\|}{\|x\| + \frac{1}{\lambda}\|v\|}$$

$$= \frac{\frac{4}{\lambda}\|v\|}{\|x\| + \frac{1}{\lambda}\|v\|} < \frac{4\|v\|}{\lambda\|x\|} .$$

This will be less than η^2 when $\epsilon < \frac{\lambda\|x\|\eta^2}{4\|\lambda x + \mu y\|}$. Hence $x + \frac{1}{\lambda}v$ and z are in C when ϵ is this small.

2.2: $y \notin C$. There is a sequence of vectors $y^\nu \in C$ tending to y. Since the bound found for ϵ remains greater than a positive constant when y varies in a bounded region, there is a fixed $\epsilon > 0$ such that the ϵ-neighborhood of $\lambda x + \mu y^\nu$ is in C for $\nu = 1, 2, \ldots$. Since $\lambda x + \mu y^\nu \to \lambda x + \mu y$ every vector z for which $[z, \lambda x + \mu y] < \epsilon$ will be in this neighborhood for sufficiently large ν. This completes the proof.

The following list gives some of the more important simple properties of convex cones.

1. The closure $\overline{C}$ of a convex cone C is convex.

This follows directly from the definition of convexity.

2. The interior of a convex cone C relative to S(C) is a convex cone.

This is a corollary of Lemma 1.

3. A convex cone has interior rays relative to S(C).

This follows because the set of vectors $v(\lambda) = \lambda_1 x^1 + \ldots + \lambda_d x^d$ (where $x^1, \ldots, x^d$ are fixed vectors of C which form a basis of S(C) and $\lambda_1, \ldots, \lambda_d$ are positive variables) form a set of rays in C which is open in S(C).

4. In every neighborhood of a relative boundary ray (z) of a convex cone C there is a ray exterior to C.

Let $(x) \neq (-z)$ be any relative interior ray of C. If N is a given neighborhood of (z) select some ray (w) in N such that $w = -\eta x + z, \eta > 0$. (w) is therefore a ray near (z) in the plane of (x) and (z) such that (z) is in the smaller angle between (x) and (w). If (w) were not an exterior ray of C, Lemma 1 would state that all rays $(\lambda x + \mu w)$ $(\lambda, \mu > 0)$ would be relative interior rays. In particular (z) $= (\eta x + w)$ would be a relative interior ray. Hence (w) is an exterior ray of C.

This property does not hold for cones in general as is shown by the example of the cone which is the whole space with exception of one ray.

5. A convex cone C and its ~~complement~~ exterior have the same boundary rays.

This is ~~merely a restatement~~ an immediate consequence of Property 4.

6. A convex cone which is everywhere dense in L^n is L^n.

This follows from Property 4.

§3. SUPPORTS

A closed half-space defined by a relation $x'u \leqq 0$ for a fixed $u \neq 0$ is called a <u>support</u> for a cone M if M is contained in this half-space.

THEOREM 1. If C is a convex cone and (z) a ray exterior to C, there is a support of C which does not contain (z).

To prove this theorem a vector u must be found such that $x'u \leqq 0$ for all x in C and $z'u > 0$. It will certainly be sufficient to show this for any closed convex cone, since a ray exterior to a cone is also exterior to the closure of the cone. Since the rays of a closed cone form a compact set, there is some ray (x^0) such that $[z,x^0] = \min [z,x]$. It can be assumed without loss of generality that $\|z\| = \|x^0\| = 1$.

Case 1. $[z,x^o] = \min\limits_{(x)\in C} [z,x] = 2.$ Then $x^o = -z$ and any vector u such that $x'u \leq 0$ defines a support for C for which $z'u > 0$.

Case 2. $[x,x^o] < 2$. Since $[z,x]$ is a monotone decreasing function of $z'x$ if z and x are unit vectors,

$$[z,x^o] = \min_{x\in C} [z,x] \text{ implies } z' \frac{x}{\|x\|} \leqq z'x^o \text{ for all } x \in C.$$

Because $x^o \in C$ and $x \in C$ implies that $(1-\theta)x^o + \theta x \in C$ $(0 \leqq \theta \leqq 1)$, it follows that

$$z' \frac{(1-\theta)x^o + \theta x}{\|(1-\theta)x^o + \theta x\|} \leqq z'x^o \text{ for any } 0 \leqq \theta \leqq 1 \text{ and any } x \in C.$$

Therefore

$$(z'x - z'x^o) \leqq z'x^o \, \frac{\sqrt{(1-\theta)^2 + \theta^2\|x\|^2 + 2(1-\theta)\theta x^{o\prime}x} - 1}{\theta} .$$

If θ tends to zero, the limiting relation

$$z'x - z'x^o \leqq z'x^o(x^{o\prime}x - 1)$$

is derived. (The right side is the derivative of the square root with respect to θ at $\theta = 0$.) Hence

$$z'x \leqq (z'x^o)(x^{o\prime}x) \text{ or}$$

$$x'(z - (z'x^o)x^o) \leqq 0 \text{ for all } x \text{ in } C.$$

Since z and x^o are linearly independent

$$z - (z'x^o)x^o \neq 0 .$$

Therefore the vector $u = z - (z'x^o)x^o$ defines a halfspace of support for C. Now $z'(z - (z'x^o)x^o) = 1 - (z'x^o)^2$ which is greater than zero since z and x^o are not opposite unit vectors. This completes the proof of Theorem 1.

COROLLARY 1. A convex cone which is not the whole of L^n has a support, i.e. it is in some half-space.

There must be at least one ray (z) not in C if $C \neq L^n$. If this is not an exterior ray then by Property 4 there is some other ray (z^1) which is an exterior ray. Theorem 1 says that C is contained in a halfspace not containing this exterior ray.

COROLLARY 2. If (z) is a boundary ray of C there is a supporting half-space, $x'u \leqq 0$, to C such that $z'u = 0$ that is z is on the boundary of this support.

Let $z^1, \ldots, z^t$ be a sequence of vectors exterior to C and converging to z. For each t there is a support such that $x'u^t \leqq 0$ for $x \in C$, $z^{t\prime}u^t \geqq 0$. The u^t may be assumed to be unit vectors and hence contain a subsequence which converges to some vector u. Now $x'u \leqq 0$ for all $x \in C$ and $z'u \geqq 0$. Since $z \in \bar{C}$, $z'u = 0$.

§4. THE CONVEX HULL AND THE NORMAL CONE

If M is a cone, the cone $\{M\}$ which is the intersection of all convex cones containing M is called the <u>convex hull</u> of M. The convex hull of M is the smallest convex cone containing M.

For any cone M, $\overline{\{M\}} \supset \{\bar{M}\}$ because $\overline{\{M\}}$ is a closed convex cone containing M and hence $\bar{M}$ and $\{\bar{M}\}$. The more interesting question is when $\overline{\{M\}} \subset \{\bar{M}\}$ that is when $\overline{\{M\}} = \{\bar{M}\}$. Examination of the possible two dimensional cones shows that $\overline{\{M\}} = \{\bar{M}\}$ if $d(M) \leq 2$. It will be proved later that if M consists of a finite number of rays or if M is closed and $l(\{M\}) = 0$ the equality also holds. That the equality does not hold in general is shown by the following example in L^3:

$$M = (\text{vectors } (x_1, x_2, x_3) \mid (x_1 - |x_3|)^2 + x_2^2 \leq x_3^2).$$

Here $\bar{M} = M$ and $\{\bar{M}\}$ is the open half-space defined by $x_1 > 0$ plus the line $x_1 = x_2 = 0$. On the other hand $\overline{\{M\}}$ is the closed half-space defined by $x_1 \geq 0$.

THEOREM 2. The closure $\overline{\{M\}}$ of the convex hull of a cone M is the intersection of all the supports of M.

The intersection I of all supports of M is a closed convex cone containing M. Therefore $I \supset \overline{\{M\}}$.

If (z) were a ray of I which was not in the closed convex cone $\overline{\{M\}}$, it would be an exterior ray to $\overline{\{M\}}$ and hence by Theorem 1 there would be a homogeneous hyperplane separating (z) from $\overline{\{M\}}$ and hence from M. The halfspace defined by this hyperplane which contained M would be a support of M which did not contain (z). Therefore $\overline{\{M\}} \supset I$.

The cone M^* formed by all vectors u such that $x'u \leq 0$ for every vector x in a cone M is called the <u>normal cone</u> of M; for, it consists of all outer normals of supports to M. Clearly M^* is convex and closed and hence $\overline{\{M\}}^* = M^*$.

If M is a subspace, M^* is its orthogonal complement.

THEOREM 3. $M^{**} = \overline{\{M\}}$

If $y \in M^{**}$, then $y'u \leqq 0$ for all u such that $x'u \leqq 0$ for all x in M. Therefore y is in the half space of support of M which is defined by $z'u \leqq 0$ for a particular u in M^*. Now as u ranges over M^*, this half space ranges over all supports of M. Therefore y is in the intersection of the supports of M and hence in $\overline{\{M\}}$ by Theorem 2. Since M^{**} is a convex closed cone, it follows that $M^{**} \supset \overline{\{M\}}$. Hence $M^{**} = \overline{\{M\}}$.

COROLLARY: If C is a closed convex cone $C^{**} = C$.

Because of this relation, the normal cone C^* is also called the <u>polar</u> <u>cone</u> of C when C is closed and convex.

THEOREM 4: For any two cones M and N

$$(M \cup N)^* = M^* \cap N^*$$

and

$$(M \cap N)^* \supset M^* \cup N^*$$

If $u'x \leqq 0$ for all $x \in M \cup N$ then $u'x \leqq 0$ for all $x \in M$ and for all x in N, and conversely. Hence

$$(M \cup N)^* = M^* \cap N^*$$

Substitution of M^* for M and N^* for N in this equation gives $(M^* \cup N^*)^* = M^{**} \cap N^{**}$. If the normal cone is now considered $M^* \cup N^* \subset (M^* \cup N^*)^{**} = (\ \overline{\{M\}} \cap \overline{\{N\}}\ )^* \subset (M \cap N)^*$.

COROLLARY: If C and D are convex cones,

$$(C + D)^* = C^* \cap D^* \text{ ‡}$$

and

$$(\bar{C} \cap \bar{D})^* = \overline{C^* + D^*} \ .$$

For general cones M and N, $\{M \cup N\} \supset M + N \supset M \cup N$. Since $(\{M \cup N\})^* = (M \cup N)^*$, $(M + N)^* = (M \cup N)^*$. Hence for convex cones $(C + D)^* = C^* \cap D^*$. Also $(C^* + D^*)^* = C^{**} \cap D^{**} = \bar{C} \cap \bar{D}$. Therefore $\overline{C^* + D^*} = (C^* + D^*)^{**} = (\bar{C} \cap \bar{D})^*$.

THEOREM 5. For any cone M,

$$d(M) + l(M^*) = n$$

and

$$l(\{M\}) + d(M^*) \leqq l(\{\bar{M}\}) + d(M^*) = n$$

From the definition of the normal cone, it follows that $s(M^*) \subset M^*$ implies $s(M^*)^* \supset M^{**} \supset M$. Now $s(M^*)^*$ is a subspace of dimension $n - l(M^*)$. Therefore $n - l(M^*) \geqq d(M)$. On the other hand $S(M) \supset M$. Hence $S(M)^* \subset M^*$. Since $S(M)^*$ is a subspace of dimension $n - d(M)$, it follows that $n - d(M) \leqq l(M^*)$. Hence $l(M^*) + d(M) = n$. Substitution of M^* for M in this relation gives

$$l(M^{**}) + d(M^*) = l\overline{(\{M\})} + d(M^*) = n.$$

Since $l(\{M\}) \leqq l\overline{(\{M\})}$ the theorem is proved.

---

‡ The sum M + N of two cones M and N is defined as the cone of all vectors $x + y$, $x \in M, y \in N$.

COROLLARY: For a closed convex cone C

$$l(C) + d(C^*) = n$$

and

$$l(C^*) + d(C) = n.$$

§5. THE CONVEX HULL AND POSITIVE LINEAR COMBINATIONS.

THEOREM 6. Any vector x of $\{M\}$ is of the form $x = \lambda_1 x^1 + \dots + \lambda_r x^r$ for $x^\rho \in M$ and $\lambda_\rho \geq 0$.

This follows immediately because the set of all such non-negative finite linear combinations is in $\{M\}$ and on the other hand these linear combinations do form a convex cone.

THEOREM 7. Any vector $x \neq 0$ in $\{M\}$ is a positive linear combination of linearly independent vectors in M. (This shows that any vector of $\{M\}$ can be expressed as a non-negative linear combination of some d(M) vectors of M where d(M) is the linear dimension of M.)

By Theorem 6, $x = \lambda_1 x^1 + \dots + \lambda_r x^r$ for some vectors x^ρ of M and some constants $\lambda_\rho \geq 0$. If the vectors $x^1, \dots, x^r$ are linearly dependent then there are some real numbers $\mu_1, \dots, \mu_r$ not all zero such that $\mu_1 x^1 + \dots + \mu_r x^r = 0$. It may be assumed that at least one μ_ρ is positive. Let τ be an index such that

$$\frac{\lambda_\tau}{\mu_\tau} = \min_{\substack{\rho \text{ such that} \\ \mu_\rho > 0}} \frac{\lambda_\rho}{\mu_\rho} \geq 0 .$$

Now

$$x = \lambda_1 x^1 + \ldots + \lambda_r x^r - \frac{\lambda_\tau}{\mu_\tau}(\mu_1 x^1 + \ldots + \mu_r x^r) =$$

$$(\lambda_1 - \frac{\lambda_\tau \mu_1}{\mu_\tau})x^1 + \ldots + (\lambda_r - \frac{\lambda_\tau \mu_r}{\mu_\tau})x^r .$$

Since $(\lambda_\rho - \frac{\lambda_\tau \mu_\rho}{\mu_\tau}) \geq 0$ for all ρ and $= 0$ for $\rho = \tau$, the expression above represents x as a non-negative linear combination of fewer than r vectors. Therefore if r is chosen minimal, $x^1, \ldots, x^r$ must be linearly independent. This proves the theorem.

LEMMA 2. If H is a supporting hyperplane to a cone M

$$\{M \cap H\} = \{M\} \cap H.$$

Now $\{M \cap H\} \subset \{M\}$ and $\{M \cap H\} \subset H$. Therefore $\{M \cap H\} \subset \{M\} \cap H$. Consider the union D of $\{M \cap H\}$ and the open half space determined by H which is a support for M. D is convex and it contains M hence $\{M\}$. On the other hand $D \cap H = \{M \cap H\}$. Therefore $\{M \cap H\} \supset \{M\} \cap H$.

LEMMA 3. If $s = s(\{M\})$ is the largest subspace contained in the convex hull $\{M\}$ of a cone M, then

$$\{M \cap s\} = s .$$

The proof is by induction on d -1 where d is the linear dimension of M, and l is the lineality of $\{M\}$ that is the dimension of s.

If $d = 1$, $s = S(M)$ so $M \cap s = M$. Therefore $\{M \cap s\} = \{M\}$. Since $s \subset \{M\} \subset S(M)$ and $s = S(M)$, $\{M \cap s\} = s$.

If $d > 1$, let H be a supporting hyperplane of M in the space S(M). By the preceding lemma $\{M \cap H\} = \{M\} \cap H$. Now $\{M \cap H\}$ is of dimension at most $d - 1$ and s is the largest subspace contained in $\{M \cap H\}$. The assumption $s = \{(M \cap H) \cap s\}$ therefore immediately yields $s = \{M \cap (H \cap s)\} = \{M \cap s\}$. This proves the lemma by induction.

THEOREM 8. Let M be a cone such that $\{M\} = S(M)$. Given any finite set V of vectors in M which contains at least one non-zero vector, there is a set W of at most $d = d(M)$ vectors in M such that the vectors of $V \cup W$ are linearly dependent with positive coefficients. Conversely, if there is a finite set of vectors in M which span S(M) and which are linearly dependent with positive coefficients, then the convex hull of the rays determined by these vectors and, hence, $\{M\}$ is S(M).

Let $y^1, \ldots, y^r$ be the vectors of V. Then by Theorem 7 the vector $-y^1 - \ldots - y^r$ is a non-negative linear combination of at most d vectors in M.

Suppose that $x^1, \ldots, x^r$ are vectors of a cone M which span S(M) and there exist constants $\mu_s > 0$ such that

$$\mu_1 x^1 + \ldots + \mu_r x^r = 0.$$

Let N denote the cone consisting of the rays (x^1), ..., (x^r). If $\{N\}$ is not $S(M)$, by Corollary 1 to Theorem 1, there is some half-space of support relative to the space $S(M)$ for N. Let such a half space be defined by the relation $x'u \leqq 0$ for a fixed vector $u \neq 0$. Then $x^{\rho'} u \leqq 0$ for $\rho = 1, \ldots, r$. Therefore $(\mu_1 x' + \ldots + \mu_r x^r)'u = 0$ implies $\mu_\rho x^{\rho'} u = 0$ and hence $x^{\rho'} u = 0$ for all ρ. Since the $x^{\rho'}$ span the whole space $S(M)$, this is impossible. This proves the last statement of the theorem.

COROLLARY: If for a cone M, $\{M\} = S(M)$, $d = d(M) > 0$, then there is a set of at most $d + 1$ non-zero vectors of M which are linearly dependent with positive coefficients. There is also in M a set of at most 2d vectors spanning $S(M)$ which are linearly dependent with positive coefficients.

This follows from Theorem 8 when V consists of one vector or d linearly independent vectors.

The following example shows that $d + 1$ is the best possible number in the first statement. Let $x^1, \ldots, x^d$ form a basis of a subspace of L^n. The cone M consisting of the rays (x^1), ..., (x^d), and $(-x^1 - \ldots - x^d)$ has $d(M) = d$ and contains no set of d vectors which are linearly dependent with positive coefficients. The cone consisting of the rays (x^1), ..., (x^d), $(-x^1)$, ..., $(-x^d)$ is an example showing that 2d is the best possible result for the second statement.

THEOREM 9. Let M be a cone and let $l > 0$ be the lineality of $\{M\}$. There is a set of at most $l+1$ non-zero vectors of M which are linearly dependent with positive coefficients. There is also a set of at most $2l$ vectors of M which span $s(\{M\})$ and which are linearly dependent with positive coefficients. If there is a set of vectors of M among which r are linearly independent and such that the set of vectors as a whole is linearly dependent with positive coefficients then $r \leqq l$ and the convex hull of the rays determined by these vectors is an r-dimensional subspace of $s(\{M\})$.

By Lemma 3 this reduces to Theorem 8 and its corollary applied to the cone $M \cap s$.

By means of the preceding results the former statements concerning the validity of $\overline{\{M\}} = \{\overline{M}\}$ will now be proved.

THEOREM 10. If M consists of a finite number of rays $\overline{\{M\}} = \{\overline{M}\}$.

If x is in $\overline{\{M\}}$ there are vectors $x^{\vartheta} = \lambda_{1\vartheta}x^{1\vartheta} + \ldots + \lambda_{r\vartheta}x^{r\vartheta}$, $\vartheta = 1, 2, \ldots$, in $\{M\}$ such that $x^{\vartheta} \longrightarrow x$ as $\vartheta \longrightarrow \infty$. Here the vectors $x^{\rho\vartheta} \in M$ and the vectors $x^{1\vartheta}, \ldots, x^{r\vartheta}$ can be assumed linearly independent because of Theorem 7. It can be assumed without loss of generality that all the vectors x, x^{ϑ}, and $x^{\rho\vartheta}$ are unit vectors. By replacing the sequence of x^{ϑ}'s by a subsequence of them, r can be made to be constant with respect to ϑ. A still finer subsequence can be chosen such that the unit vectors $x^{\rho\vartheta}$ can be made to converge to some unit vectors $\overline{x}^{\rho}$. Since there are only a finite number of rays in M this means that this subsequence can be assumed to have $x^{\rho\vartheta} = \bar{x}^{\rho}$ for all ϑ and ρ. Suppose therefore that the original sequence x^{ϑ} had been chosen so that r does not depend upon ϑ and $x^{\vartheta} = \lambda_{1\vartheta}\bar{x}^{1} + \ldots + \lambda_{r\vartheta}\bar{x}^{r}$. Consider the function $f(\mu_1, \ldots, \mu_r) = \|\mu_1\bar{x}^1 + \ldots + \mu_r\bar{x}^r\|^2$. On the sphere $\sum_{\rho=1}^{r}\mu_\rho^2 = 1$ this function has a positive minimum m since the $\bar{x}^\rho$ are linearly independent. Therefore $\|\lambda_1\bar{x}^1 + \ldots + \lambda_r\bar{x}^r\|^2 \geq m(\lambda_1^2 + \ldots + \lambda_r^2)$. Since $\|\lambda_{1\vartheta}\bar{x}^1 + \ldots + \lambda_{r\vartheta}\bar{x}^r\| = 1$, the $\lambda_{\rho\vartheta}$ are bounded by $\sqrt{\frac{1}{m}}$. Therefore there is a subsequence of the x^{ϑ} such that for each ρ, $\lambda_{\rho\vartheta} \to \lambda_\rho$ as $\vartheta \to \infty$ for some non-negative number λ_ρ. Therefore $x = \lambda_1\bar{x}^1 + \ldots + \lambda_r\bar{x}^r$. Hence x is in $\{M\}$.

THEOREM 11. If M is closed and $l(\{M\}) = 0$, then $\overline{\{M\}} = \{M\}$.

Let x be a vector in $\overline{\{M\}}$ and let x^{ϑ} be a sequence of vectors in $\{M\}$ which approach x. Then

$$x^{\vartheta} = \lambda_{1\vartheta}x^{1\vartheta} + \ldots + \lambda_{r\vartheta}x^{r\vartheta} \text{ for some } \lambda_{\rho\vartheta} \geq 0 \text{ and}$$

some $x^{\rho\vartheta} \in M$.

Here the $x^{\rho\vartheta}$ may be assumed to be unit vectors and r may be assumed less than or equal to d the dimension of S(M). As in the proof of Theorem 10 the sequence x^{ϑ} can be selected so that r does not depend on ϑ and $x^{\rho\vartheta} \longrightarrow \bar{x}^{\rho}$ as $\vartheta \longrightarrow \infty$. Since M is a closed cone and the $x^{\rho\vartheta}$ are unit vectors, the $\bar{x}^{\rho}$ are also unit vectors in M. If $y^1, \ldots, y^r$ are unit vectors in S(M) which are linearly independent with non-negative coefficients, the function $\|\mu_1 y^1 + \ldots + \mu_r y^r\| =$

$f(\mu_1, \ldots, \mu_r)$ for $\mu_\varrho \geq 0$ and $\sum_{\varrho=1}^{r} \mu_\varrho{}^2 = 1$ is positive and continuous. Hence it has a positive minimum $m(y^1, \ldots, y^r)$. Consider the function $m(z^1, \ldots, z^r)$ over sets of r unit vectors in M. Any relation $\mu_1 z^1 + \ldots + \mu_r z^r = 0$ with $\mu \geq 0$, $\sum_{\varrho=1}^{r} \mu_\varrho{}^2 = 1$ would contradict the hypothesis that $\{M\}$ contains no linear subspace (Theorem 9). Hence any r unit vectors of M are linearly independent with non-negative coefficients. Sets of r unit vectors of M range over a closed set in the product of r unit spheres because M is closed. Therefore $m(z^1, \ldots, z^r)$ has a positive minimum m. Hence $\|x^\vartheta\| = \|\lambda_{1\vartheta} x^{1\vartheta} + \ldots + \lambda_{r\vartheta} x^{r\vartheta}\| \geq m \sum_{\varrho=1}^{r} \lambda_{\varrho\vartheta}^2$. Since $x^\vartheta \longrightarrow x$, $\|x^\vartheta\|$ and, hence, $\sum_{\varrho=1}^{r} \lambda_{\varrho\vartheta}^2$ are bounded. Therefore a subsequence of the x^ϑ can be chosen so that $\lambda_{\varrho\vartheta} \longrightarrow \lambda_\varrho$ for $\vartheta \longrightarrow \infty$. With such a selection

$$x = \lambda_1 \bar{x}^1 + \ldots + \lambda_r \bar{x}^r \text{ so that } x \in \{M\}.$$

THEOREM 12. Let C be a closed convex cone which is not the whole space L^n and let H be the hyperplane which bounds a support to C defined by $x'u \leq 0$. Then $C \cap H = s(C)$ if and only if (u) is a relative interior ray of C^*.

Suppose that (u) is a relative interior ray of C^*. Put $d(C^*) = d^*$ and let $v^1, \ldots, v^{d-1}$ be vectors such that $u, v^1, \ldots, v^{d-1}$ form a basis for $S(C^*)$. Consider the vectors

$$u^1 = u + v^1, \ldots, u^{d-1} = u + v^{d-1}, u^d = u - v^1 - \ldots - v^{d-1}.$$

These also form a basis for $S(C^*)$.

Suppose that the selected vectors $v^1, \ldots, v^{d-1}$ are so short that $u^1, \ldots, u^d$ are in C^*. This is possible because (u) is relative interior to C^*. Now

$$u = \frac{1}{d}(u^1 + \ldots + u^d).$$

Suppose that $x \in C \cap H$. Then $x'u^1 \leqq 0, \ldots, x'u^d \leqq 0$ and $x'u = 0$. Hence $x'u^1 = 0, \ldots, x'u^d = 0$. Since the u^j span the subspace $S(C^*)$, x is in its orthogonal complement which contains $s(C)$ and has dimension $n-d(C^*)$. By the corollary to Theorem 5, $l = n-d(C^*)$. Therefore the two spaces coincide. This proves the sufficiency part of Theorem 12.

Suppose on the other hand that (u) is a relative boundary ray of C^*. A sequence $v^1, v^2, \ldots, v^\vartheta, \ldots$ can then be selected so that $x'v^\vartheta \leqq 0$ does not define a support of C but v^ϑ tends to u. This means that for every v^ϑ an $x^\vartheta \in C$ can be found such that $x^{\vartheta\prime} v^\vartheta > 0$. Since $w'x = 0$ for any $w \in S(C^*)$ and $x \in s(C)$, x^ϑ is not in $s(C)$. Write x^ϑ as $y^\vartheta + z^\vartheta$ where z^ϑ is in $s(C)$ and y^ϑ is in C but in the orthogonal complement of $s(C)$. Then $y^{\vartheta\prime} v^\vartheta = x^{\vartheta\prime} v^\vartheta > 0$. It may be assumed without loss of generality that $\| y^\vartheta \| = 1$. If only a suitable subsequence of the y^ϑ is considered, these y^ϑ will converge to some unit vector y. For this y, $y'u \geqq 0$, and hence $y'u = 0$. However, y is not in $s(C)$. This completes the proof of the theorem.

§6. EXTREME RAYS AND SUPPORTS

THEOREM 13. If C is a closed convex cone of dimension greater than one and C

is not $S(C)$ or a half-space of $S(C)$, C is the convex hull of its relative boundary rays.

The assumption that C is not a subspace or a half subspace means that $l = \dim s(C) \leqq d(C) - 2$. Since $s(C)$ is contained in every supporting hyperplane of C in the space $S(C)$ and since there is at least one such hyperplane because $C \neq S(C)$, every ray in $s(C)$ is a relative boundary ray of C. Let z be any vector in C which is not in $s(C)$. Since $l \leqq n-2$, there is a plane P in $S(C)$ which contains the vector z and intersects $s(c)$ only in the origin. The at most two dimensional cone $P \cap C$ contains z but no two opposite rays because $P \cap s(C) = 0$. Therefore it is a sector of less than 180° in the plane P. Hence z is a non-negative linear combination of boundary vectors of $P \cap C$. A boundary ray of $P \cap C$ is however a relative boundary ray of C. Therefore (z) is in the convex hull of the relative boundary rays of C. This proves Theorem 13.

DEFINITION: A ray (x) of a convex cone C is an extreme ray of C if x is not a positive linear combination of two linearly independent vectors of C.

Clearly this definition does not depend upon the choice of the representative vector x.

THEOREM 14. A closed convex cone C with $l(C) = 0$ is the convex hull of its extreme rays.

This is true for a one dimensional cone with $l(C) = 0$ because the one ray of the cone is necessarily an extreme ray.

Suppose the theorem has been proved for cones of dimension less than d. Let (x) be a relative boundary ray of the d dimensional closed convex cone C. Select a supporting hyperplane H containing (x). $C \cap H$ is a closed convex cone of dimension at most $d - 1$. By the induction hypothesis $C \cap H$ is the convex hull of its extreme rays. Since C is all on one side of H, an extreme ray of $C \cap H$ is also an extreme ray of C. Therefore every relative boundary ray of C is in the convex hull of the extreme rays of C. Theorem 13 therefore gives that C is in the convex hull of its extreme rays. This finishes the induction proof.

For the determination of the extreme rays of a particular cone it is helpful to note that any ray which is the only ray in the intersection of a supporting hyperplane and a convex closed cone is necessarily an extreme ray. It is not true, however, that for a general convex closed cone every extreme ray is the intersection of a supporting hyperplane and the cone. For example if in L^3, C is the convex hull of a circular cone D and a ray (x) such that both (x) and (-x) are outside D, the extreme rays which are at the juncture of the curved surface of the cone and the flat surface of the cone are not the intersection of the cone with any supporting plane. Any supporting plane which contains one of these two rays contains the whole two dimensional cone spanned by this ray and (x).

> DEFINITION: A support $x'u \leq 0$ of a convex cone C is an extreme support if u is not a positive linear combination of two linearly independent outer normal vectors of supports of C, in other words if (u) is an extreme ray of C^*.

THEOREM 15. A closed cone C with $d(C) = n$ is the intersection of its extreme supports.

This follows from Theorem 14 applied to C^* and Theorem 6.

DEFINITION: A cone is called polyhedral if it is the convex hull of a finite number of rays.

A subspace is a polyhedral cone.

It is obvious that a sum of polyhedral cones is polyhedral.

The polar of a polyhedral cone is the intersection of a finite number of halfspaces. For, let C be the convex hull of the rays (a^ρ), $\rho = 1, \ldots, r$; then C^* consists of all vectors u for which $u'a^\rho \leqq 0$, $\rho = 1, \ldots, r$. Hence C^* is the intersection of these halfspaces.

THEOREM 16. The polar of a polyhedral cone is polyhedral. In other words, a convex cone is polyhedral if and only if it is the intersection of a finite number of halfspaces.

Let C be the convex hull of the rays (a^ρ), $\rho = 1, \ldots, r$. Then C^* is the intersection of the halfspaces $u'a^\rho \leqq 0$. If (u^0) is an extreme ray of C^*, the vector u must satisfy n-1 linearly independent equations $u^{0\prime}a^\rho = 0$. For, otherwise there would be an at least two-dimensional neighborhood of (u^0) all of whose rays satisfy all the inequalities $u'a^\rho \leqq 0$

and (u^o) could not be extreme. Since there are only a finite number of systems of n-1 linearly independent equations $u'a \xi = 0$, C^* has only a finite number of extreme rays.

If $l(C^*) = 0$ that is $d(C) = n$ it follows from Theorem 14 that C^* is polyhedral. If $d(C) < n$ this, applied to C in S(C), yields that $C^* \cap S(C)$ is polyhedral. Now C^* is the sum of $C^* \cap S(C)$ and the subspace $s(C^*) = S(C)^*$, hence polyhedral.

§7. SYSTEMS OF LINEAR HOMOGENEOUS INEQUALITIES.

Various theorems on the solvability of systems of linear homogeneous inequalities are obtained by specializing some of the preceding results to polyhedral cones.

In this section the inequalities $x \geq 0$ or $x > 0$ for a vector x mean that the corresponding inequalities hold for each component. $x \geq 0$ means $x \geqq 0$ but $x \neq 0$.

Let A be an m by n matrix. Denote by ξ and x vectors in L^m and L^n respectively (both considered as column matrices). Let A be fixed, ξ and x variable. Then the following statements are valid:

I. One and only one of the two systems

$$Ax > 0$$

and

$$A'\xi = 0, \xi \geq 0$$

of linear inequalities has a solution.

II. One and only one of the two systems

$$Ax \geq 0$$

and

$$A'\xi = 0, \xi > 0$$

has a solution.

These statements may be interpreted geometrically either in L^m or L^n. In each of these spaces there are two mutually polar interpretations depending on whether ξ and x represent vectors or hyperplanes. The two most convenient interpretations are described in the following.

First interpretation:

Consider x as a normal vector of a hyperplane and the rows of A as vectors in L^n. The existence of a solution of $Ax > 0$ means that the cone M consisting of the rays determined by the row vectors of A has a supporting hyperplane whose intersection with M is the origin only. This is the case if and only if the lineality of $\{M\}$ is 0. On the other hand, this is equivalent with the non-existence of a non-trivial linear relation with non-negative coefficients between the rows of A, that is $A'\xi = 0$ and $\xi \geq 0$ imply $\xi = 0$ (Theorem 9). This yields I.

Let $d = d(M)$ be the linear dimension of M. Then d is just the rank of A. Suppose $Ax > 0$ has no solution, that is $l = l(\{M\}) > 0$. From Theorem 9 it then follows that there are $l + 1$ or less among the rows of A which are linearly dependent with positive coefficients. This together with I implies that the system $Ax > 0$ of m inequalities has a subsystem consisting of at most $l + 1$ inequalities which has no solution. Now $l \leq d$; hence: $Ax > 0$ has a solution if and only if every subsystem consisting of $d + 1$ of the inequalities has a solution.

Consider now the system $Ax \geq 0$. The existence of a solution means that M has a supporting hyperplane which does not contain the whole of M. This is the case if and only if

$\{M\}$ is not a subspace (Theorem 12). Now $\{M\}$ is a subspace if and only if there is a linear relation with positive coefficients between all the rows of A (Theorem 8). This yields II.

From Theorem 8 and its corollary it follows further that if $\{M\}$ is a subspace there are 2d or less rays in M such that their convex hull is the same subspace; hence: $Ax \geq 0$ has a solution if and only if every subsystem with rank d consisting of 2d inequalities has a solution.

Second interpretationL

Denote the closed positive orthant of L^m, that is the set of all $\xi \geqq 0$, by D. Consider ξ and the columns of A as vectors in L^m and let S be the subspace spanned by the column vectors of A. The orthogonal complement S^* of S consists of the solutions ξ of $A'\xi = 0$. The statements I and II then follow by substituting $C = S^*$ and $C = S$ in the following theorem:

> A closed convex cone C contains no point of D except the origin if and only if its polar cone C^* contains an interior point of D.

This is the case $k = m$ of

THEOREM 17. Let C be a closed convex cone, D the closed positive orthant, and E_k, $0 \leqq k \leqq m$, the subspace of all vectors whose first k components vanish. Then $C \cap D \subset E_k$ if and only if, for every $\varepsilon > 0$, the polar cone C^* contains a vector whose k first components are greater than a fixed positive constant γ and whose $m - k$ last components are greater than $-\varepsilon$.

If C is polyhedral the condition may be simplified to: $C^* \cap D$ contains a vector whose k first components are positive.

To prove the sufficiency consider an arbitrary vector $\xi \in C \cap D$. The polar Cone C^* is contained in the half-space $\xi'\eta \leqq 0$, η variable. For an $\eta \in C^*$ such that $\eta_1 > \gamma$, $\ldots, \eta_k > \gamma$, $\eta_{k+1} > -\varepsilon, \ldots, \eta_m > -\varepsilon$ it follows that

$$(\xi_1 + \ldots + \xi_k)\gamma - (\xi_{k+1} + \ldots + \xi_m)\varepsilon \leqq 0.$$

Since $\xi \geqq 0$, this can be valid for all $\varepsilon > 0$ only if $\xi_1 = \ldots = \xi_k = 0$; that is, if $\xi \in E_k$.

The necessity may be seen in the following way. From $C \cap D \subset E_k$ it follows that $(C \cap D)^* \supset E_k^*$. Obviously, E_k^* contains the vector $\xi = (\underbrace{1,\ldots,1}_{k},\underbrace{0,\ldots,0}_{m-k})$. Since $(C \cap D)^* = \overline{C^* + D^*}$ (Corollary to Theorem 4) there are vectors $\eta^i \in C^*$, $\zeta^i \in D^*$, $i = 1,2,\ldots,$ such that $\eta^i + \zeta^i \longrightarrow \xi$. Now $\zeta^i \leqq 0$, since $\zeta^i \in D^*$. Hence, given $0 < \varepsilon < 1/2$, it follows that $\eta^i > \xi - \varepsilon$ for sufficiently large i. This is the statement of the theorem with $\gamma = 1/2$. If C is polyhedral, $C^* + D^*$ is closed (Theorem 10). Hence there are vectors $\eta \in C^*$ and $\zeta \in D^*$ such that $\eta + \zeta = \xi$, and the vector $\eta = \xi - \zeta \geqq \xi$ satisfies the requirement for every $\varepsilon > 0$.

Consider again an m by n matrix A. Let $m = k + l$ with fixed non-negative integers k and l. Write

$$A = \begin{pmatrix} B \\ \Gamma \end{pmatrix}, \quad \xi = \begin{pmatrix} \eta \\ \zeta \end{pmatrix}$$

where the matrices B, Γ, η, ζ are k by n, l by n, k by 1, and l by 1 respectively. Then the following statements hold:

III. One and only one of the two systems

$$Bx > 0, \qquad \Gamma x \geqq 0$$

and

$$B'\eta + \Gamma'\zeta = 0, \ \eta \geqq 0, \ \zeta \geqq 0$$

has a solution.

IV. One and only one of the two systems

$$Bx \geq 0, \ \Gamma x \geqq 0$$

and

$$B'\eta + \Gamma'\zeta = 0, \ \eta > 0, \ \zeta \geqq 0$$

has a solution.

V. One and only one of the two systems

$$Bx > 0, \ \Gamma x \geqq 0, \ x \geqq 0$$

and

$$B'\eta + \Gamma'\zeta \leqq 0, \ \eta \geq 0, \ \zeta \geqq 0$$

has a solution.

VI. One and only one of the two systems

$$Bx \geq 0, \ \Gamma x \geqq 0, \ x \geqq 0$$

and

$$B'\eta + \Gamma'\zeta \leqq 0, \ \eta > 0, \ \zeta \geqq 0$$

has a solution.

To prove these statements apply Theorem 17 to the following polyhedral cones C: the subspace of all vectors $\binom{\eta}{\zeta}$ satisfying $B'\eta + \Gamma'\zeta = 0$ (III), the subspace of all vectors $\begin{pmatrix} Bx \\ \Gamma x \end{pmatrix}$ (x unrestricted) (IV), the cone of all vectors $\begin{pmatrix} \eta \\ \zeta \end{pmatrix}$ satisfying $B'\eta + \Gamma'\zeta \leq 0$ (V), and the cone of all vectors $\begin{pmatrix} Bx \\ \Gamma x \end{pmatrix}$, $x \geq 0$ (VI).

Theorems on systems of infinitely many inequalities may also be obtained. Let a^α denote a vector in L^n depending on the index α which may run through any set. Let M be the cone consisting of all rays (a^α). (For instance, α may be a real variable. Then the point a^α might describe a curve in L^n for which M would be the cone projecting this curve from the origin.) As an example take the following generalization of statement I which is derived in the same way as I using the first interpretation above:

> The system of inequalities $x'a^\alpha > 0$ has no solution if and only if there are finitely many among the vectors a^α which are linearly dependent with positive coefficients.

Let b be a vector with the property that $x'b \leq 0$ for every x which satisfies all the inequalities $x'a^\alpha \leq 0$. Geometrically this means that b is contained in all supports of M, hence $b \in \overline{\{M\}}$. If in particular $\{M\} = \overline{\{M\}}$ which is the case if α runs through a finite set (Theorem 10) or if M is closed and $\{M\}$ has lineality 0 (Theorem 11), then b is in $\{M\}$ and, hence, b is a positive linear combination of at most n of the vectors a^α (Theorem 7). In the general case b is a limit of such linear combinations (generalization of a theorem of Farkas).

Chapter II

CONVEX SETS

§1. LINEAR COMBINATIONS OF POINT SETS

The cones of Chapter I were always considered to be in an n-dimension Euclidean vector space L^n. In a vector space the origin or zero vector is necessarily distinguished and its coordinate representation is invariant under a change of the coordinate basis of the space.

Convex sets, however, are more naturally thought of in an n-dimensional affine space A^n. If a particular coordinate system has been chosen a point is described by the n-tuple $x = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}$ of its coordinates. Denote the point with coordinates x by $\hat{x}$. If t is a fixed n-tuple and T is a non-singular n x n matrix,

$$x \longrightarrow \bar{x} = T(x - t)$$

is a transformation of the representation of A^n in terms of the coordinates x_i into a representation in terms of coordinates $\bar{x}_i$. For A^n all allowable coordinate transformations are of this type.

In terms of particular coordinates the expression $x = \sum_{\rho=1}^{r} \lambda_\rho x^\rho$ (λ_ρ real) represents a point $\hat{x}$ which is a "<u>linear</u>

combination" of the points $\hat{x}^\rho$. If $\lambda_\rho \geq 0$ ($\rho = 1, \ldots, r$), $\hat{x}$ is called a non-negative linear combination of the $\hat{x}^\rho$. If $\lambda_\rho > 0$ ($\rho = 1, \ldots, r$), $\hat{x}$ is a positive linear combination. These definitions are not independent of the choice of coordinates, for if $\bar{x} = T(x-t)$

$$\sum_{\rho=1}^{r} \lambda_\rho \bar{x}^\rho = \sum_{\rho=1}^{r} \lambda_\rho T(x^\rho - t) = T(\sum_{\rho=1}^{r} \lambda_\rho x^\rho) - \sum_{\rho=1}^{r} \lambda_\rho Tt$$

$$= T(\sum_{\rho=1}^{r} \lambda_\rho x^\rho - t) + (1 - \sum_{\rho=1}^{r} \lambda_\rho)Tt$$

$$= \overline{\sum_{\rho=1}^{r} \lambda_\rho x^\rho} + (1 - \sum_{\rho=1}^{r} \lambda_\rho)Tt.$$

This shows that if the coordinates $\bar{x}$ instead of x are used the linear combination of $\hat{x}^1, \ldots, \hat{x}^r$ with coefficients $\lambda_1, \ldots, \lambda_r$ may be a point which is different from $\hat{x}$. It should be noted that this difference depends on $\sum_{\rho=1}^{r} \lambda_\rho$ and t but not on the points $\hat{x}^\rho$. In the particular case that $t = 0$, that is the change of coordinates does not shift the origin, $\sum_{\rho=1}^{r} \lambda_\rho \bar{x}^\rho = \overline{\sum_{\rho=1}^{r} \lambda_\rho x^\rho}$. This is also the case whenever $\sum_{\rho=1}^{r} \lambda_\rho = 1$. These linear combinations with $\sum_{\rho=1}^{r} \lambda_\rho = 1$, for which the resulting point is independent of the choice of coordinates, are particularly important as the following example shows.

A line through the points with coordinates x^0 and x^1 is just the set of all points represented by

$$x = \lambda_0 x^0 + \lambda_1 x^1 = (1-\theta)x^0 + \theta x^1 \quad (\lambda_0 + \lambda_1 = 1,\ \theta = \lambda_1).$$

Those points on this line with $0 \leqq \theta \leqq 1$ form the segment between the points $\hat{x}^0$ and $\hat{x}^1$.

The points $\hat{x}^0, \ldots, \hat{x}^p$ are defined to be <u>linearly dependent</u> if

$$\mu_0 x^0 + \ldots + \mu_p x^p = 0$$

for some real numbers μ_π with

$$\mu_0 + \ldots + \mu_p = 0 \text{ and } \mu_0^2 + \ldots + \mu_p^2 > 0.$$

If μ_0 is one of the non-zero μ_π

$$x^0 = \lambda_1 x^1 + \ldots + \lambda_p x^p \text{ where } \lambda_\pi = \frac{-\mu_\pi}{\mu_0} \text{ and } \sum_{\pi=1}^{p} \lambda_\pi = 1.$$

Therefore the point $\hat{x}^0$ is expressed as a linear combination of the other points in a fashion which is independent of the choice of coordinates.

Equivalently the points $\hat{x}^0, \ldots, \hat{x}^p$ are linearly dependent if and only if

$$\operatorname{rank} \begin{pmatrix} 1 & \cdots\cdots & 1 \\ x_1^0 & \cdots\cdots & x_1^p \\ \vdots & & \vdots \\ x_n^0 & & x_n^p \end{pmatrix} < p .$$

That two points are linearly dependent means they are the same point. Three points are linearly dependent precisely when they are collinear. Similarly four points are coplanar if and only if they are linearly dependent.

A <u>p-flat</u> is defined to be all points with coordinates $x = \lambda_0 x^0 + \dots + \lambda_p x^p$ where $\lambda_0 + \dots + \lambda_p = 1$ and $\hat{x}^0, \dots, \hat{x}^p$ are linearly independent points. Note that a p-flat is a p-dimensional affine space. Similarly a <u>p-simplex</u> is the set of points with coordinates

$$x = \lambda_0 x^0 + \dots + \lambda_p x^p \text{ where } \lambda_0 + \dots + \lambda_p = 1, \lambda_\rho \geqq 0 \ (\rho = 0, \dots, p),$$

and $\hat{x}^0, \dots, \hat{x}^p$ are linearly independent.

Although all the proofs that follow are affine proofs, it is desirable for conceptual clarification occasionally to introduce a projective interpretation. Identify the point $\begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}$ of A^n with the point $\begin{bmatrix} \lambda \\ \lambda x_1 \\ \vdots \\ \lambda x_n \end{bmatrix}$ of the projective space P^n.

With this identification A^n may be thought of as the "finite" portion of P^n. (The "hyperplane at infinity" consists of the projective points with first coordinate 0.) It is now seen that the points $\hat{x}^0, \dots, \hat{x}^p$ of A^n are linearly dependent if and only if the projective points with coordinates $\begin{bmatrix} 1 \\ x_1^0 \\ \vdots \\ x_n^0 \end{bmatrix}, \dots, \begin{bmatrix} 1 \\ x_1^p \\ \vdots \\ x_n^p \end{bmatrix}$ are linearly dependent that is

$$\operatorname{rank} \begin{pmatrix} 1 & \cdots\cdots & 1 \\ x_1^0 & \cdots\cdots & x_1^p \\ \vdots & & \vdots \\ x_n^0 & & x_n^p \end{pmatrix} < p.$$

If M and N are sets of points in A^n, $M+N$ is defined to be the set of all points $\widehat{x+y}$ for $\hat{x}$ in M and $\hat{y}$ in N. Since $\widehat{x+y}$ may be different from $\hat{x}+\hat{y}$, $M+N$ must be expected to vary with the choice of coordinates. However $\overline{x+y}$ always differs from $\bar{x}+\bar{y}$ by $(1 - (1+1))Tt$. Therefore $M+N$ is determined up to a translation.

The set of points with coordinates λx where $\hat{x}$ is some point in M is denoted by λM.

Relative to a fixed coordinate system the following rules of calculation are valid:

1) $(M+N) + O = M + (N+O)$
2) $M+N = N+M$
3) $\lambda(\mu M) = (\lambda\mu)M$
4) $\lambda(M+N) = \lambda M + \lambda N$
5) $(\lambda+\mu)M \subset \lambda M + \mu M$.

It is not true in general that $(\lambda+\mu)M = \lambda M + \mu M$, for if $\mu = -\lambda \neq 0$, $(\lambda+\mu)M$ consists of only the origin while $\lambda M + \mu M$ contains more points if M has at least two points. It is true, however, that $(\lambda+\mu)M = \lambda M + \mu M$ if M is a flat and $\lambda+\mu \neq 0$ or if $\lambda \geqq 0$, $\mu \geqq 0$, and M is a convex (see below) set.

The previous calculation with linear combinations of points shows that a sum $\sum_{\rho=1}^{r} \lambda_\rho M^\rho$ is independent of the choice of coordinates if $\sum_{\rho=1}^{r} \lambda_\rho = 1$, otherwise it is determined up to a translation.

The distinction between the points of A^n and their coordinate n-tuples is not important for the properties which follow. Therefore the point $\hat{x}$ will be identified with its coordinate n-tuple x from now on.

§2. CONVEX SETS AND THE CONVEX HULL OF A SET

A set M is called <u>convex</u> if M contains every segment joining a pair of points from M. Expressed in terms of coordinates this means that $(1-\theta)x + \theta y$ $(0 \leqq \theta \leqq 1)$ represents a point in M whenever x and y are in M.

An example of a convex set is the "ellipsoid" of all points x such that $Q(x,x) \leqq 1$, where $Q(x,x) = \sum_{i,j=1}^{n} a_{ij}x_i x_j$ is a positive semidefinite quadratic form.

With the notation $Q(x,y) = \sum_{i,j=1}^{n} a_{ij}x_i y_j$,

$$
(1) \qquad \begin{aligned} &Q(\lambda x+\mu y, \lambda x+\mu y) = \\ &\lambda^2 Q(x,x) + 2\lambda\mu\, Q(x,y) + \mu^2 Q(y,y) \geqq 0 \end{aligned}
$$

for all real λ, μ. For $\lambda = -\mu = 1$ this yields

$$
2Q(x,y) \leqq Q(x,x) + Q(y,y).
$$

Use of this in (1) when $\lambda = 1 - \theta$, $\mu = \theta$, $0 \leqq \theta \leqq 1$, gives

$$
Q((1-\theta)x + \theta y, (1-\theta)x + \theta y) \leqq (1-\theta)\, Q(x,x) + \theta\, Q(y,y).
$$

This shows that $Q((1-\theta)x + \theta y, (1-\theta)x + \theta y) \leqq 1$ whenever $Q(x,x) \leqq 1$, $Q(y,y) \leqq 1$. Hence the ellipsoid is convex.

Certain properties of convex sets will now be listed.

1. If the sets M_α are convex $\bigcap_\alpha M_\alpha$ is also convex.

This follows immediately from the definition of convexity.

2. If the sets $M_\rho (\rho=1, \ldots, r)$ are convex, then $\sum_{\rho=1}^{r} \lambda_\rho M_\rho$ is convex.

If x and y are in $\sum_{\rho=1}^{r} \lambda_\rho M_\rho$, $x = \sum_{\rho=1}^{r} \lambda_\rho x^\rho$ and $y = \sum_{\rho=1}^{r} \lambda_\rho y^\rho$ for some x^ρ and y^ρ in M. Now

$$(1-\theta)x + \theta y = \sum_{\rho=1}^{r} \lambda_\rho ((1-\theta)x^\rho + \theta y^\rho).$$

Therefore $\sum_{\rho=1}^{r} \lambda_\rho M_\rho$ is convex if the sets M_ρ are.

3. If M is convex and $N^1, \ldots, N^r$ are any sets such that $N^\rho \subset M$, then

$\sum_{\rho=1}^{r} \lambda_\rho N^\rho \subset M$, if $\sum_{\rho=1}^{r} \lambda_\rho = 1$ and $\lambda_\rho \geqq 0$ $(\rho=1, \ldots, r)$.

If $r = 2$, $\lambda_1 x^1 + \lambda_2 x^2 \subset M$ $(\lambda_1, \lambda_2 \geqq 0, \lambda_1 + \lambda_2 = 1)$ for all x^1 in $N_1 \subset M$ and x^2 in $N_2 \subset M$ because M is convex. Hence $\lambda_1 N_1 + \lambda_2 N_2 \subset M$. Assume the property has been proved for $r = s - 1 \geqq 2$. Now

$$\lambda_1 N_1 + \lambda_2 N_2 + \ldots + \lambda_s N_s = \frac{\lambda_1 N_1 + \ldots + \lambda_{s-1} N_{s-1}}{\lambda_1 + \ldots + \lambda_{s-1}}(1-\lambda_s) + \lambda_s N_s$$

if $\sum_{\rho=1}^{s} \lambda_\rho = 1$ and $\lambda_s \neq 1$. This last condition may be assumed without loss in generality. By the induction assumption

$$\frac{\lambda_1}{\lambda_1 + \ldots + \lambda_{s-1}} N_1 + \ldots + \frac{\lambda_{s-1}}{\lambda_1 + \ldots + \lambda_{s-1}} N_{s-1} \subset M \text{ if } N_\rho \subset M.$$

From the case of $r = 2$ it follows that $\lambda_1 N_1 + \ldots + \lambda_s N_s \subset M$ if

$N_\rho \subset M$, $\sum_{\rho=1}^{s} \lambda_\rho = 1$.

The <u>convex hull</u> $\{M\}$ of a set M is defined to be the intersection of all convex sets containing M. By Property 1, it is the smallest convex set containing M.

4. If $N_1, \ldots, N_r$ are sets such that $N_\rho \subset M$ (M any set), then $\sum_{\rho=1}^{r} \lambda_\rho N_\rho \subset \{M\}$ if $\lambda_\rho \geq 0$ ($\rho = 1, \ldots, r$) and $\sum_{\rho=1}^{r} \lambda_\rho = 1$.

This is an immediate consequence of Property 3 and the definition of $\{M\}$.

A point $x = \sum_{\rho=1}^{r} \lambda_\rho x^\rho$ ($\lambda_\rho \geq 0$, $\sum_{\rho=1}^{r} \lambda_\rho = 1$) is called a <u>centroid</u> of the points x^ρ.

5. The convex hull $\{M\}$ of a set M consist of all centroids of all finite sets of points from M.

That all such centroids are in $\{M\}$ follows from Property 4. To prove the reverse inclusion, it is sufficient to show that the set of centroids is convex. Suppose $x = \sum_{\rho=1}^{r} \lambda_\rho x^\rho$ and $y = \sum_{\sigma=1}^{s} \mu_\sigma y^\sigma$ for some x^ρ and y^σ in M. Then

$(1-\theta)x + \theta y = (1-\theta) \sum_{\rho=1}^{r} \lambda_\rho x^\rho + \theta \sum_{\sigma=1}^{s} \mu_\sigma y^\sigma$ and hence $(1-\theta)x + \theta y$ is a centroid of $x^1, \ldots, x^r, y^1, \ldots, y^s$.

6. If $z \in \{M\}$, z is a centroid of linearly independent points of M. (A set of linearly independent points in M contains at most n+1 points.)

Suppose $z = \lambda_0 x^0 + \ldots + \lambda_r x^r \left(\sum_{\rho=1}^{r} \lambda_\rho = 1, \lambda_\rho \geqq 0 \right)$ where $x^0, \ldots, x^r$ are linearly dependent, that is there are real numbers μ_ρ such that $\sum_{\rho=0}^{r} \mu_\rho x^\rho = 0$, $\sum_{\rho=0}^{r} \mu_\rho = 0$ and $\sum_{\rho=0}^{r} \mu_\rho^2 \neq 0$.

Let $\frac{\lambda_\tau}{\mu_\tau} = \underset{\mu_\rho \geqq 0}{\mathrm{Min}} \frac{\lambda_\rho}{\mu_\rho}$. Then $z = \sum_{\rho \neq \tau} \left(\lambda_\rho - \frac{\lambda_\tau}{\mu_\tau} \mu_\rho \right) x^\rho$ and $\lambda_\rho - \frac{\lambda_\tau}{\mu_\tau} \mu_\rho \geqq 0$. Repetition of this procedure proves Property 6 for any particular z.

7. If M and N are convex sets

$$\{M \cup N\} = \bigcup_{0 \leqq \theta \leqq 1} ((1-\theta)M + \theta N)$$

This follows because every point of $\{M \cup N\}$ is a centroid of a point from M and a point from N.

8. If M_0 is any set and

$$M_{i+1} = \bigcup_{0 \leqq \theta \leqq 1} ((1-\theta)M_i + \theta M_i) \quad (i = 0, 1, 2, \ldots),$$

then $\{M_0\} = M_k$ where k is the smallest integer such that 2^k is greater than or equal to n+1.

This is a corollary of Property 6.

§3. METRIC AND TOPOLOGY

If a particular coordinate system has been chosen, the definition

$$d(x,y) = \sqrt{(x_1-y_1)^2 + \dots + (x_n-y_n)^2}$$

gives a Euclidean metric on $A^n(x_1, \dots, x_n)$. This metric is not an invariant of A^n for $d(x,y)$ is invariant only under orthogonal transformations of coordinates. In general

$$d(x,y) = \sqrt{Q(x_1', \dots, x_n', y_1', \dots, y_n')}$$

where x_i' and y_i' are new coordinates and Q is a positive definite quadratic form. While this metric is not an invariant of A^n, the uniform topology it defines is. From here on it will be assumed that A^n has this topology. It is convenient to consider A^n metrized with a particular Euclidean metric. This is no actual restriction of generality, but it allows simple geometric interpretation of the theorems.

9. If M_1 is a non-empty open set of A^n and λ_1 is a non-zero real number, $\lambda_1 M_1 + \dots + \lambda_r M_r$ is an open set for any sets M_ϱ ($\varrho=2, \dots, r$) and for λ_ϱ ($\varrho=2, \dots, r$) any real numbers.

If M_1 is open and $\lambda_1 \neq 0$, $\lambda_1 M_1$ is also open. Now $\lambda_1 M_1 + N = \bigcup_{x \in N} (\lambda_1 M_1 + x)$. Since $\lambda_1 M_1 + x$ is open when M is open, $\lambda_1 M_1 + N$ is open. Let $N = \lambda_2 M_2 + \dots + \lambda_r M_r$.

10. If $M_1, \dots, M_r$ are closed sets and $M_2, \dots, M_r$ are bounded, $\lambda_1 M_1 + \dots + \lambda_r M_r$ is closed.

Suppose z is a limit point of $\lambda_1 M_1 + \ldots + \lambda_r M_r$. Then there is a sequence $x^{\vartheta} = \lambda_1 x^{1\vartheta} + \ldots + \lambda_r x^{r\vartheta}$ $(x^{\rho\vartheta} \in M_\rho)$ such that x^{ϑ} converges to z as $\vartheta \to \infty$. Since $M_2, \ldots, M_r$ are closed and bounded, it may be assumed that $x^{\rho\vartheta}$ converges to some x^{ρ} in M_ρ for $\rho = 2, \ldots, r$. $x^{\vartheta} - (\lambda_2 x^{2\vartheta} + \ldots + \lambda_r x^{r\vartheta})$ must also converge, so $\lambda_1 x^{1\vartheta}$ converges to some point $\lambda_1 x^1$ of $\lambda_1 M_1$. Therefore $z = \lambda_1 x^1 + \ldots + \lambda_r x^r$.

If M is any set and U is the open unit sphere with center at the origin of the coordinates, $M + \epsilon U$ is the ϵ-neighborhood of M. If M is convex, this neighborhood is also convex. If M is closed and $\overline{U}$ is the closed unit sphere $M + \epsilon \overline{U}$ is a closed ϵ-neighborhood of M.

11. If C is a convex set $\overline{C}$ is also convex.

This is true because if $x^{\vartheta} \to x$ and $y^{\vartheta} \to y$, the points of the line segment joining x and y are limit points of the points on the segments joining x^{ϑ} to y^{ϑ}.

Let S(M) denote the intersection of all the flats containing a set M. This is just the flat with any maximal set of linearly independent points of M as a basis. The dimension d(M) of S(M) is called the <u>linear dimension</u> of M.

A point is called a <u>relative interior point</u> of M if it is interior to M relative to the topology of S(M). (Note that if M is a point, that is d(M) = 0, this point is a relative interior point of M.) A boundary point of M is called a <u>relative boundary point</u> if it is a boundary point relative to S(M). Since points of S(M) are exterior to M relative to S(M) if and only if they are exterior to M relative to A^n, no distinction need be made between exterior points and relative exterior points.

12. If C is a convex set with $d(C) > 0$, then every point of C is a limit point of C.

If $d(C) > 0$ and x is any point in C, there must be some other point y in C. x is a limit point of points on the segment joining x and y. Since this segment must be in C, x is a limit point of C.

13. A convex set C has relative interior points.

Let $d = d(C)$ and suppose $x^0, \ldots, x^d$ are linearly independent points of C which span S(C). The d-simplex spanned by $x^0, \ldots, x^d$ has interior points relative to S(C) and hence C does also because C contains this simplex.

14. If x is a relative interior point of a convex set C and z is in $\bar{C}$, all points of the segment joining x to z with the possible exception of z are relative interior points of C. If z is a relative boundary point of C, the points on the line through x and z which are separated from x by z are exterior points of C.

Suppose $y = (1-\theta)x + \theta z$ for $0 \leqq \theta < 1$. Let $U_x(\epsilon)$ be the open sphere of radius ϵ with center x. If x is relative interior to C, there is some $\epsilon > 0$ such that $U_x(\epsilon) \cap S(C) \subset C$.

Let z^{ϑ} be a sequence of points of C converging to z. The set $((1-\theta)U_x(\epsilon) + \theta z^{\vartheta})$ is an open sphere of radius $(1-\theta)\epsilon$ and center $(1-\theta)x + \theta z^{\vartheta}$. From the convexity of C $((1-\theta)U_x(\epsilon) + \theta z^{\vartheta}) \cap S(C)$ is contained in C. Since $z^{\vartheta} \longrightarrow z$, $(1-\theta)x + \theta z^{\vartheta} \longrightarrow y$. Therefore for ϑ sufficiently large y is interior to $(1-\theta)U_x(\epsilon) + \theta z^{\vartheta}$. Hence y is relative interior to C. This proves the first statement of Property 14.

Suppose z is a relative boundary point of C and $y = (1-\theta)x + \theta z$ $(\theta > 1)$. Now $z = \frac{1}{\theta}y + (1 - \frac{1}{\theta})x$ so that, if y were not an exterior point of C, z would be in the relative interior of C by the first part of Property 14. This contradiction proves the second statement.

15. If C is convex, the relative interior of C is convex.

This is a corollary of Property 14.

16. If C is convex and everywhere dense in S(C), C = S(C).

This is because a convex set C with no exterior points in S(C) can have no relative boundary points and hence is S(C) itself.

§4. PROJECTING AND ASYMPTOTIC CONES, s-CONVEXITY

A ray $\overrightarrow{px}(x \neq p)$ consists of all points $(1-\theta)p + \theta x$ for $\theta \geq 0$. The <u>projecting cone</u> $P_p(M)$ of a set M from a point p is defined to be $\bigcup_{x \in M} \overrightarrow{px}$. (If M = p, set $P_p(M) = p$.) Note that $P_p(M)$ need not be closed when M is closed. For example if

M is an (n-1)-flat and p is a point outside M, $P_p(M)$ is an open half-space through p plus the point p.

17. If C is convex, $P_p(C)$ is convex (for any p).

This is a direct consequence of the definition of $P_p(C)$.

DEFINITION: A set C is called s-convex if for every point p not in C, $s(\overline{P_p(C)}) \cap C$ is empty.

18. An s-convex set C has the property that if $x \in C$ and $y \in \overline{C}$, $p = (1-\theta)x + \theta y$ is in C for $0 < \theta < 1$.

$s(\overline{P_p(C)})$ contains the line xy and hence $s(\overline{P_p(C)}) \cap C$ is non-empty. This shows that p is in C.

Property 18 shows that a s-convex set is convex. Clearly closed and relatively open convex sets are s-convex. On the other hand an open triangle with one point of the boundary adjoined is convex but not s-convex.

DEFINITION: If M is any set and p any fixed point, the set of rays $\overrightarrow{px}$ which are the limit of a sequence of rays $\overrightarrow{px^\nu}$ where $x^\nu \in M$ and $x^\nu \longrightarrow \infty$ is called the asymptotic cone $A_p(M)$ of M with vertex p.

19. $A_p(M)$ is closed for any M and p.

An ordinary diagonal process shows that a limit ray of $A_p(M)$ is a limit ray of rays $\vec{px}$, $x \in M$, $x \longrightarrow \infty$.

20. For any set M and any points p and q

$$A_q(M) = A_p(M) + (q-p).$$

This follows because the convergence of $\vec{px^\vartheta}$ to $\vec{px}$ as $x^\vartheta \longrightarrow \infty$ implies that $\vec{qx^\vartheta}$ converges to $\vec{qx} = \vec{px} + (q-p)$, and conversely.

21. If M is any set and p is any point

$$A_p(M) = \bigcap_{q \in A^n} [\overline{P_q(M)} + (p-q)]$$

By definition $A_q(M) \subset \overline{P_q(M)}$. Therefore by Property 20 $A_p(M) \subset \overline{P_q(M)} + (p-q)$ for every point q. Suppose $\vec{px} \notin A_p(M)$. There is then a neighborhood $N_\epsilon(\vec{px})$ of rays emanating from p such that $N_\epsilon(\vec{px})$ (as a point set) has a bounded intersection with M. A point q in $N_\epsilon(px)$ can therefore be selected so that $(N_\epsilon(\vec{px}) + q - p) \cap M$ is empty. For this q, $x \notin \overline{P_q(M)} + (p-q)$. This completes the proof of Property 21.

22. For a convex set C and any point p, $A_p(C)$ is convex.

This may be regarded as a corollary of Properties 17 and 21.

23. If C is an s-convex set and p is any point of C, $A_p(C)$ is the set of all rays contained in C.

Denote by A_p' the cone consisting of all rays emanating from p and contained in C. Obviously $A_p' \subset A_p(C)$. Let $(\overrightarrow{px})$ be a ray of $A_p(C)$. Then there is a sequence of points $x^{\vartheta} \in C$ such that $x^{\vartheta} \longrightarrow \infty$ and $(\overrightarrow{px^{\vartheta}}) \rightarrow (\overrightarrow{px})$. Since the segments px^{ϑ} are in C, $(\overrightarrow{px}) \subset \bar{C}$. From the Property 18 it follows that $(\overrightarrow{px}) \subset C$ if $p \in C$. Hence $A_p' = A_p(C)$.

> COROLLARY: If C is an arbitrary convex set and p is a relative interior point of C, $A_p(C)$ is the set of all rays emanating from p and contained in C.

Apply 23 to the relative interior of C.

Consider the cone $A_p(C)$ (C convex) as a cone of the linear space of the vectors with initial point p. $A_p(C)$ then contains a largest subspace $s(A_p(C))$ with dimension $l(A_p(C))$ (the lineality of $A_p(C)$). This subspace considered in A^n is the largest flat in $A_p(C)$ containing p.

> 24. If C is an s-convex set, C is the union of l-flats parallel to $s(A_p(C))$, that is
>
> $$C = s(A_p(C)) + [C \cap s(A_p(C))^*].$$

By Property 23, $A_q(C)$ for q any point of C simply consists of all rays $\overrightarrow{qx}$ contained in C. Therefore C contains $s(A_q(C)) = s(A_p(C)) + (q-p)$. Hence C is just

$$\bigcup_{q \in C} \left(s(A_p(C)) + (q-p) \right).$$

If C is a convex set in three space and $l(A_p(C)) = 1$, Property 24 says that C is a cylinder.

§5. BARRIERS AND NORMAL CONES

Any <u>oriented (n-1) flat</u> F may be described as the set of all points x such that $x'u = u_o$ where u is a vector in the positive normal direction to the flat. If $\sup_{x \in M} x'u < u_o$, F is called a <u>bound</u> of M and the set M is said to be bounded in the direction u and to be in the "negative" half-space of F. If $\sup_{x \in M} x'u = u_o$, F is called a <u>supporting flat</u> for M and the negative half-space of F (the points with $x'u \leqq u_o$) is called a <u>support</u> of M. Note that if u and u_o define a supporting flat for M, u and $u_o + \epsilon$ ($\epsilon > 0$) define a bound of M in the direction u. A flat which is either a bound or a supporting flat of M is called a <u>barrier</u> of M.

> 25. If M is any set and p is a fixed point, all vectors from p which are positive normal vectors of barriers of M through p form a closed convex cone $N_p(M)$, the normal cone of M at p. This cone is in the linear space of vectors with origin p. If the projecting cone $P_p(M)$ is interpreted as being in the same space $N_p(M) = P_p(M)^*$.

This equation just states that all the barriers of M containing p are supporting hyperplanes of $P_p(M)$.

Property 25 is of particular interest when M is convex and p is a relative boundary point of M. $P_p(M)$ is not the whole space because the ray $\overrightarrow{p(p-x)}$ contains no points of M if p + x is in M (Property 14). This cone has a supporting hyperplane, and hence M has a supporting flat through p.

> 26. If M is any set, the vectors from the coordinate origin which are positive normals to barriers of M form a convex cone $B_\theta(M) \subset (A_\theta(M))^*$. If M is convex $\overline{B_\theta(M)} = A_\theta(M)^*$.

If $x'u \leqq u_o$ and $x'v \leqq v_o$ for all x in M, $x'(\lambda u + \mu v) \leqq \lambda u_o + \mu v_o$ $(\lambda \geqq 0, \mu \geqq 0)$ for all x in M. Therefore if u and v are in $B_\theta(M)$, $\lambda u + \mu v$ $(\lambda \geqq 0, \mu \geqq 0)$ is also. This shows that $B_\theta(M)$ is a convex cone. If the flat defined by $x'u = u_o$ is a barrier for M, the flat of points x such that $x'u = \text{Max } (u^o, p'u)$ is a barrier for $M \cup A_p(M)$. Hence the hyperplane of vectors y with $y'u = 0$ in the linear space with origin p is a supporting hyperplane of the cone $A_p(M)$. Therefore if $u \in B_\theta(M)$, $u \in (A_\theta(M))^*$.

Suppose now that M is convex. Let $s = s(A_\theta(M))$ and $l = l(A_\theta(M))$. To prove $\overline{B_\theta(M)} = (A_\theta(M))^*$, it is sufficient to show that if a ray is not in $B_\theta(M)$, it is not a relative interior ray of $(A_\theta(M))^*$. By Property 24, the relative interior of M is the union of l-flats parallel to s. If the n-1-flat defined by $x'u = u_o$ is not a barrier for M, the structure of M shows that there is a point y in $M \cap s^*$ (s^* is the orthogonal complement of s) such that $y'u > u_o$. For a u not in $B_\theta(M)$, such a y may be selected for each u_o. From these y's a sequence which tends to infinity may be chosen so that the rays $\overrightarrow{\theta y}$ (or (y)) converge to a ray $\overrightarrow{\theta z}$ (or (z)) in $A_\theta(M)$. Since (y) is in s^* for each y, (z) is also in s^*. In Chapter I it was shown that the supporting hyperplane to a convex cone C corresponding to a relative interior ray of the polar cone C^* intersects C only in s(C). Therefore (u) is not a relative interior ray of $A_\theta(M)$ and $\overline{B_\theta(M)} = (A_\theta(M))^*$.

That this equation cannot be strengthened to $B_\theta(C) = A_\theta(C)^*$ for C convex is shown by the following example. In the x_1, x_2 plane let C consist of all points such that $x_2 \geqq e^{x_1}$. Then $B_\theta(C)$ is the half-open quadrant defined by $x_1 \geqq 0$, $x_2 < 0$. $A_\theta(C)$ is the closed quadrant given by $x_1 \leqq 0$, $x_2 \geqq 0$. Therefore $(A_\theta(M))^*$ is not $B_\theta(M)$ but its closure.

Note that if coordinates with a different origin θ' had been used for A^n, the set $B_{\theta'}(M)$ would be a translate of $B_\theta(M)$. More precisely

$$B_{\theta'}(M) = B_\theta(M) + (\theta' - \theta).$$

Property 26 shows that $B_\theta(M)$ determines $A_\theta(M)$, but the example above demonstrates that $A_\theta(M)$ does not determine $B_\theta(M)$ uniquely.

§6. SEPARATION THEOREMS

27. If C and D are closed convex sets with an empty intersection and C is bounded, there is a support H of C such that $D \cap H$ is empty. There is also a support H' of D such that $C \cap H'$ is empty.

Since D is closed there is a point $p(x)$ in D such that the minimum of the distance from points of D to a fixed point x is attained at $p(x)$. Because C is compact, there is a point q of C such that the distance from q to $p(q) = p$ is less than or equal to the distance from any point x of C to any point y of D. Let H be the half-space of points with

$$x'(p-q) \leqq q'(p-q).$$

The oriented flat which bounds this halfspace passes through q and has $p-q$ as normal vector. If x is some point in C different from q the segment from x to q is in C. The shortest distance from this segment to p is either $\|x - p\|$, the length of the altitude from p of the triangle (p,q,x), or $\|q - p\|$. By assumption the last of these three possibilities must be the case. For this to happen, however, the vector $x - q$ must make an obtuse or right angle with $p - q$. Therefore C is in H. If H' is the halfspace defined by

$$x'(q-p) \leqq p'(q-p) \text{ or } x'(p-q) \geqq p'(p-q)$$

an analogous argument shows that D is in H'. Since $H \cap H' = \emptyset$, $H \cap D = H' \cap C = \emptyset$ and H and H' are the desired supports.

28. If C and D are convex sets such that no common point is relative interior to both C and D, there is in $S(C \cup D)$ a $(d(C \cup D)-1)$-dimensional hyperplane separating C and D. (i.e. there is a vector u and a number u_o such that $x'u \leqq u_o$ for all x in C and $x'u \geqq u_o$ for all x in D.)

The theorem for the closures of C and D implies the theorem for C and D. Therefore assume C and D are closed. Suppose x is a point in $C \cap D$ relative interior to C and y is a point in $C \cap D$ relative interior to D. By Property 14, $(1-\theta)x + \theta y$ $(0 < \theta < 1)$ is a point which is relative interior to both C and D. By the hypothesis of the theorem this is impossible. Hence it can be assumed that $C \cap D$ contains ~~only~~ no relative boundary points of one of the sets (say C). If in particular C consists of a single point p, $C \cap D$ is empty since p is relative interior to C. This case when C is a point disjoint from D is covered by the last theorem. Assume therefore that $d(C) > 0$. Define

$$C_{\vartheta} = \left((1 - \tfrac{1}{\vartheta})C + \tfrac{1}{\vartheta}p\right) \cap \overline{U_p(\vartheta)} \quad (\vartheta = 1, 2, \ldots,)$$

where p is a fixed point and $\overline{U_p(\vartheta)}$ is the closed sphere of radius ϑ and center p. C_{ϑ} is just a linear contraction with center p of the part of C near p. Choose p as a relative interior point of C. Then C_{ϑ} is in the relative interior of C by Property 14. Therefore $C_{\vartheta} \cap D$ is empty. Theorem 27 asserts that there is a hyperplane defined by $x'u^{\vartheta} = u_o^{\vartheta}$ such that $x'u^{\vartheta} \leqq u_o^{\vartheta}$ for $x \in C_{\vartheta}$ and $x'u^{\vartheta} \geqq u_o^{\vartheta}$ for $x \in D$. In particular $p'u^{\vartheta} \leqq u_o^{\vartheta} \leqq q'u^{\vartheta}$ (q any point of D). Suppose the vectors u^{ϑ} had all been normalized to length one. Then a subsequence of the ϑ could be selected so that the corresponding u^{ϑ} converge to a vector u and the corresponding u_o^{ϑ} converge to a number u_o. For u and u_o, $x'u \geqq u_o$ for all x in D and $x'u \leqq u_o$ for all x in the relative interior of C. It immediately follows that $x'u \leqq u_o$ for every x in C.

If D is just a single relative boundary point of C, Theorem 28 states that there is a supporting hyperplane of C through this point.

29. For any set M, $\overline{\{M\}} = I_s = I_{bd} = I_{br}$ where I_s is the intersection of all the supports of M, I_{bd} is the intersection of the half-spaces on the same side of a bound as M, and I_{br} is the intersection of all half-spaces on the same side of a barrier as M.

Clearly $\overline{\{M\}} \subset I_{br} \subset I_s \subset I_{bd}$. If p is not in $\overline{\{M\}}$, by Theorem 27 there is a support of M (defined by $x'u \leqq u_o$) which does not contain p. For ϵ sufficiently small, the hyperplane defined by $x'u = u_o + \epsilon$ is a bound of M which separates M from p. Therefore $p \notin I_{bd}$ and $\overline{\{M\}} = I_{br} = I_s = I_{bd}$.

30. $\overline{\{M\}} \supset \{\bar{M}\}$ for any set M and $\overline{\{M\}} = \{\bar{M}\}$ if M is bounded.

That $\overline{\{M\}} \supset \{\bar{M}\}$ is obvious. If x is in $\overline{\{M\}}$, $x = \lim_{\nu \to \infty} \sum_{\rho=0}^{r} \lambda_{\rho\nu} x^{\rho\nu}$ ($\lambda_{\rho\nu} \geqq 0$, $\sum_{\rho=1}^{r} \lambda_{\rho\nu} = 1$, $x^{\rho\nu} \in$ M) for a fixed $r \leqq d(M)$ because of Property 6. Since M is bounded a subsequence may be selected so that $x^{\rho\nu} \longrightarrow x^{\rho}$, $\lambda_{\rho\nu} \longrightarrow \lambda_{\rho}$. Therefore

$$x = \sum_{\rho=0}^{r} \lambda_{\rho} x^{\rho} .$$

Since $x^{\rho} \in \bar{M}$, $x \in \{\bar{M}\}$ and $\overline{\{M\}} = \{\bar{M}\}$.

§7. CONVEX HULL AND EXTREME POINTS

31. If M is any set and H is any supporting hyperplane $\{M \cap H\} = \{M\} \cap H$.

Clearly $\{M \cap H\} \subset \{M\} \cap H$. If I is the interior of the support of M bounded by H, the convex set $I \cup \{M \cap H\} \supset M$. Therefore $I \cup \{M \cap H\} \supset \{M\}$ and

$$\{M \cap H\} = \{((I \cup \{M \cap H\}\ ) \cap H\} = (I \cup \{M \cap H\}) \cap H \supset \{M\} \cap H.$$

32. If C is a closed convex set which is neither a flat nor a half-flat, then C is the convex hull of its relative boundary points.

Let p be any relative interior point of C. It is sufficient to show that there is in $S(C)$ a line L through p which has no other point in common with $A_p(C)$. For, then $C \cap L$ is bounded and L contains two relative boundary points such that p is on the segment determined by these points. If $d(A_p(C) < d(C)$, there is clearly a line through p which has no other point in common with $A_p(C)$. If $d(A_p(C)) = d(C)$, $A_p(C)$ is neither a flat (because C would equal $A_p(C)$ and be a flat) nor a half-flat (because C would be a half flat). This means $l(A_p(C)) \leqq d(A_p(C)) - 2$. Since in $S(A_p(C)) = S(C)$ there is a support to $A_p(C)$ which has only $s(A_p(C))$ in common with $A_p(C)$, there is a line L in $S(C)$ with the required property.

A point of a convex set is called an <u>extreme point</u> if it is not interior to any segment in the convex set, that is, it is not the centroid of other points of the convex set.

33. A closed bounded convex set is the convex hull of its extreme points.

This is obvious in one dimension. If C is n dimensional and p is a boundary point of C, there is a supporting hyperplane H of C passing through p. Now the extreme points of the n-1 dimensional closed bounded convex set $C \cap H$ are extreme points of C. This is because any segment not in H containing a point of H as an interior point would have to pierce H, i.e. have points on both sides of H. From the theorem in n-1 dimensions, it follows that p is a centroid of extreme points in $C \cap H$. Therefore the relative boundary of C is in the convex hull of the extreme points of C. By Theorem 31, C itself must be in the convex hull of its extreme points. This completes the inductive proof of Theorem 33.

§8. POLARITY IN THE PROJECTIVE SPACE

DEFINITION: A point set C in the projective space is called p-convex if it has the following properties:

1) C is not the entire projective space but not empty.
2) C is connected.
3) Through every point not in C there is a hyperplane which has no points in common with C.

A hyperplane set Γ in the projective space is called p-convex if it has the following properties:

1) Γ does not contain all hyperplanes of the projective space but is not empty.

2) Γ is connected.

3) In every hyperplane not in Γ there is a point which is in no hyperplane of Γ.

Let C be p-convex and choose any hyperplane outside C as the plane at infinity. Then C is an s-convex point set in the affine space. For let x and y be any two points in C, and suppose there were a point z on the finite segment xy which is not in C. Then there would be a hyperplane through z which does not meet C. This hyperplane (together with the plane at infinity) would separate x and y in contradiction to the assumption that C is connected. Hence, C is convex. Let p be any point not in C. There is a hyperplane through p not intersecting C. Now this hyperplane bounds a support to $\overline{P_p(C)}$, hence it contains $s(\overline{P_p(C)})$. This proves the s-convexity of C.

Conversely, every s-convex point set in the affine space is p-convex in the projective space obtained by adjoining the plane at infinity. For, the points at infinity do not belong to C and they are in a hyperplane which does not intersect C. C is obviously connected. Through every exterior point of C there is a bound to C. Through every point $y \in \overline{C}$ but not in C there is a supporting hyperplane which has no points in common with C. This is true because $\overline{P_y(C)}$ has a supporting hyperplane which intersects $\overline{P_y(C)}$ only in $s(\overline{P_y(C)})$, and $s(\overline{P_y(C)}) \cap C$ is empty.

34. If C is a p-convex set, the set Γ of all hyperplanes which have no point in common with C is p-convex.

35. If Γ is a p-convex hyperplane set, the set C of all points which are in no hyperplane of Γ is p-convex.

PROOF: The two statements are duals of each other; hence it is sufficient to prove one. Let C be given and denote by Γ the set of all hyperplanes not intersecting C. Since C is not empty Γ does not contain all hyperplanes. Choose one of the hyperplanes of Γ, as the plane at infinity. Every other hyperplane of Γ then is a barrier to C. Since the barriers form a convex set and since there are barriers which are arbitrarily far away, Γ is connected. Every hyperplane which is not in Γ contains a point of C and no hyperplane through this point is in Γ.

Obviously the set of all those points which are in no hyperplane of Γ is exactly the original point set C. Hence the sets C and Γ determine each other in this simple way.

Consider any such pair of sets C, Γ and apply any correlation $\xi = Ax$. Then $\Gamma^* = AC$ and $C^* = A'^{-1}\Gamma$ form another pair of the same kind. If the correlation is involutory, that is if $A = \pm A'$, we have

$$C^{**} = C .$$

In the case $A = A'$, C^* is called the polar body of C with respect to the quadric $x'Ax = 0$. By means of the bilinear equation $x'Ax^* = 0$ the polar body C^* of C is determined as follows: For each fixed point $x \in C$ this is the equation of a hyperplane in Γ^*, and C^* consists of all points x^* which are on no such hyperplane.

Let

$$A = \begin{pmatrix} -1 & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \dots & 1 \end{pmatrix}$$

and choose $x_0 = 0$ as the plane at infinity. Then the bilinear equation is

$$x_1 x_1^* + \ldots + x_n x_n^* - x_0 x_0^* = 0.$$

The origin corresponds to the plane at infinity. In the euclidean space, putting $x_0 = x_0^* = 1$, we have the polarity with respect to the unit sphere. To a bounded convex set C with the origin as an interior point corresponds a C^* with the same properties. If C is open C^* is closed and conversely. The closures of C and C^* obviously determine each other, and this gives Minkowski's polarity for convex bodies.

Let C be a closed convex cone whose vertex is the origin. Then C^* is the polar cone of C in the former sense, if it is defined by means of $\Gamma^* = AC$. Otherwise the origin has to be added.

Replace now n by n+1, denote the homogeneous coordinates by $x_0, \ldots, x_n, z$, and consider

$$A = \begin{pmatrix} 0 & 0 & \ldots & 0 & -1 \\ 0 & 1 & \ldots & 0 & 0 \\ \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & \ldots & 1 & 0 \\ -1 & 0 & \ldots & 0 & 0 \end{pmatrix} .$$

The corresponding bilinear equation is

$$-x_0 z^* - x_0^* z + x_1 x_1^* + \ldots + x_n x_n^* = 0.$$

This is the polarity with respect to the paraboloid of revolution $2z = x_1^2 + \ldots + x_n^2$ if the inhomogeneous coordinates are interpreted as rectangular coordinates. The infinite point of the z-axis, that is, the point with all

$x^{\vartheta} = 0$ and $z = 1$, corresponds to the plane at infinity, $x_0 = 0$. To all other points at infinity correspond hyperplanes parallel to the z axis. The origin corresponds to the hyperplane $z^* = 0$. If a convex set C has an asymptotic cone which contains the positive z-axis,* the polar set C^* has the same property. For a closed convex cone C whose vertex is the origin and which contains the positive z-axis the polar set C^* is a half-cylinder generated by open half-lines whose end points make up a closed convex set in $z = 0$. This polarity is especially useful in treating convex functions.

* add: but not the negative z-axis

CHAPTER III

CONVEX FUNCTIONS

1. DEFINITIONS AND ELEMENTARY PROPERTIES

DEFINITION: Let D be a convex set of $A^n(x_1,\ldots,x_n)$. A real-valued function $f(x)$ defined for x in D is said to be convex in D if

$$f((1-\theta)x + \theta y) \leqq (1-\theta)f(x) + \theta f(y)$$

for $0 \leqq \theta \leqq 1$ and x and y in D. If $<$ is always valid for $0 < \theta < 1$ and x and y distinct points in D, $f(x)$ is said to be strictly convex in D. A function $f(x)$ is called concave (strictly concave) if $-f(x)$ is convex (strictly convex).

If $f(x)$ is a function defined in the set D of A^n, the set of all points in $A^{n+1}(x_1,\ldots,x_n,z)$ such that $x = (x_1,\ldots,x_n)$ is in D and $z \geqq f(x)$ will be denoted by $[D,f]$.

For each of the properties listed below the domains of the functions are always assumed to be convex unless a contrary assumption is explicitly made.

1. The function $f(x)$ is convex in the set D if and only if the set $[D,f]$ is convex.

If $f(x)$ is convex in D and (x, z_o) and (y, z_1) are points of $[D, f]$,

$$(1-\theta)z_o + \theta z_1 \geqq (1-\theta)f(x) + \theta f(y) \geqq f((1-\theta)x + \theta y).$$

This means the point $((1-\theta)x + \theta y, (1-\theta)z_o + \theta z_1)$ is in $[D, f]$. The proof of the reverse implication is even more obvious.

2. If $f(x)$ is convex in D and $x = My + b$ where M is an n by m matrix and b is a vector of A^n, then $f(My+b)$ is convex in the inverse image of D, that is in the set of all $y = (y_1, \ldots, y_m)$ for which $My + b \in D$.

This is true because

$$f(M((1-\theta)y^o + \theta y^1) + b) = f((1-\theta)(My^o + b) + \theta(My^1 + b)).$$

3. If $f_\rho(x)$, $\rho = 0, 1, \ldots, r$, are convex functions in D and $\lambda_\rho \geqq 0$, the function $\sum_{\rho=0}^{r} \lambda_\rho f_\rho(x)$ is also convex in D.

This follows from the rules for adding inequalities.

4. If $f(x)$ is convex in D, $x^\rho \in D$, $\lambda_\rho \geqq 0$, and $\sum_{\rho=0}^{r} \lambda_\rho = 1$,

$$f\left(\sum_{\rho=0}^{r} \lambda_\rho x^\rho\right) \leqq \sum_{\rho=0}^{r} \lambda_\rho f(x^\rho).$$

The definition of convexity says that this is true if $r = 1$. If $\lambda_o = 1$ the statement is trivial. Suppose $\lambda_o < 1$. Because of $1 - \lambda_o = \sum_{\rho=1}^{r} \lambda_\rho$, Property 4 for $r - 1$ and 1 yields

$$f\left(\sum_{\rho=0}^{r} \lambda_\rho x^\rho \right) = f\left(\lambda_o x^o + (1-\lambda_o) \sum_{\rho=1}^{r} \frac{\lambda_\rho}{1-\lambda_o} x^\rho \right)$$

$$\leqq \lambda_o f(x^o) + (1-\lambda_o)\, f\left(\sum_{\rho=1}^{r} \frac{\lambda_\rho}{1-\lambda_o} x^\rho \right)$$

$$\leqq \lambda_o f(x^o) + (1-\lambda_o) \sum_{\rho=1}^{r} \frac{\lambda_\rho}{1-\lambda_o} f(x^\rho)$$

$$= \sum_{\rho=0}^{r} \lambda_\rho f(x^\rho).$$

Property 4 follows by induction.

> 5. A function $f(x)$ is both convex and concave in D if and only if it is linear in D.

The sufficiency of the condition is obvious. If $f(x)$ is both convex and concave in D, Property 4 applied to f and $-f$ yields

$$(*) \qquad f\left(\sum_{\rho=0}^{r} \lambda_\rho x^\rho \right) = \sum_{\rho=0}^{r} \lambda_\rho f(x^\rho)$$

for $\lambda_\rho \geqq 0$, $\sum_{\rho=0}^{r} \lambda_\rho = 1$. If r equals the linear dimension of D and the points x^ρ are linearly independent, (*) shows that f is linear in the simplex with vertices x^ρ.

Suppose now that $x = \sum_{\rho=0}^{r} \mu_\rho x^\rho$, $\sum_{\rho=0}^{r} \mu_\rho = 1$, is any point of D. Application of (*) to the point x and the centroid $\frac{1}{r+1} \sum_{\rho=0}^{r} x^\rho$ of the simplex gives

$$f\left(\sum_{\rho=0}^{r}\left(\frac{1-\theta}{r+1}+\theta\mu_\rho\right)x^\rho\right)$$
$$=(1-\theta)\,f\left(\frac{1}{r+1}\sum_{\rho=0}^{r}x^\rho\right)+\theta f\left(\sum_{\rho=0}^{r}\mu_\rho x^\rho\right)$$

for $0 \leqq \theta \leqq 1$. Since the point $\sum_{\rho=0}^{r}\left(\frac{1-\theta}{r+1}+\theta\mu_\rho\right)x^\rho$ is in the simplex with vertices x^ρ for some θ sufficiently small, it follows from (*) that

$$f\left(\sum_{\rho=0}^{r}\left(\frac{1-\theta}{r+1}+\theta\mu_\rho\right)x^\rho\right)=\sum_{\rho=0}^{r}\left(\frac{1-\theta}{r+1}+\theta\mu_\rho\right)f(x^\rho).$$

Therefore,

$$\sum_{\rho=0}^{r}\left(\frac{1-\theta}{r+1}+\theta\mu_\rho\right)f(x^\rho)=\frac{1-\theta}{r+1}\sum_{\rho=0}^{r}f(x^\rho)+\theta f\left(\sum_{\rho=0}^{r}\mu_\rho x^\rho\right)$$

and

$$f\left(\sum_{\rho=0}^{r}\mu_\rho x^\rho\right)=\sum_{\rho=0}^{r}\mu_\rho f(x^\rho).$$

6. If $f_\nu(x)$, $\nu = 1,2,\ldots,$ are convex functions in D and $f_\nu(x)$ converges pointwise to $f(x)$, $f(x)$ is also convex in D.

This is because the inequality defining convexity in D for $f(x)$ is the limit of the corresponding inequalities for $f_\nu(x)$.

7. If $f_\alpha(x)$, where α runs through any set, are convex functions in D, the set of all points x of D at which $\sup_\alpha f_\alpha(x)$ is finite is convex and $\sup_\alpha f_\alpha(x)$ is a convex function in this set.

Define $g(x) = \sup_{\alpha} f_{\alpha}(x)$ where this supremum is finite. Let x and y be any two points for which $\sup_{\alpha} f_{\alpha}$ is finite. Then

$$f_{\alpha}((1-\theta)x + \theta y) \leqq (1-\theta)f_{\alpha}(x) + \theta f_{\alpha}(y)$$

$$\leqq (1-\theta)g(x) + \theta g(y).$$

This shows that $\sup_{\alpha} f_{\alpha}((1-\theta)x + \theta y)$ is finite and $g((1-\theta)x + \theta y) \leqq (1-\theta)g(x) + \theta g(y)$.

8. If $f(x)$ is convex in D and $\varphi(t)$ is a monotone increasing convex function over an interval which contains the values of $f(x)$, $\varphi(f(x))$ is convex in D.

From the convexity of f and the monotone character of φ, and from the convexity of φ

$$\varphi(f((1-\theta)x + \theta y)) \leqq \varphi((1-\theta)f(x) + \theta f(y))$$

$$\leqq (1-\theta)\varphi(f(x)) + \theta\varphi(f(y)).$$

9. If $f(x)$ is convex in D and D' is a compact set in the relative interior of D, $f(x)$ is bounded above in D'.

Cover D' with a finite number of closed simplexes contained in D. Every point x of D' is a centroid of the vertices of any simplex which contains it. By Property 4, $f(x)$ is less than or equal to the maximum of $f(x^i)$

as x^i ranges over the vertices of a simplex containing x. Since the number of simplices is finite, $f(x)$ is bounded above in D'.

10. If $f(x)$ is convex in D, it is bounded below in every bounded subset of D.

Let x^o be a fixed point relative interior to D. Select a positive number δ so small that, in the flat spanned by D, the closed sphere K about x^o with radius δ is in the relative interior of D. For an arbitrary point x in D, denote by y that point of the line joining x and x^o which does not separate x and x^o and which is a distance δ from x^o. This definition insures that $y \in K \subset D$. From the convexity of f it follows that

$$f(x^o) \leq \frac{\delta}{\rho+\delta} f(x) + \frac{\rho}{\rho+\delta} f(y)$$

where ρ denotes the distance $\|x-x^o\|$. Hence

$$\delta f(x) \geqq (\rho+\delta) f(x^o) - \rho f(y).$$

Since K is compact and relative interior to D, $f(y)$ is bounded above (Property 9). Hence $f(x)$ is bounded below for ρ bounded.

11. If $f(x)$ is a convex function in D which attains a maximum value at a relative interior point of D, then $f(x)$ is constant in D.

Suppose $f(x)$ has a maximum at a relative interior point x^o. If x is any point of D, for some sufficiently small positive η the point $y = (1+\eta)x^o - \eta x$ is also in D.

Because $f(x) \leqq f(x^o)$ and $f(y) \leqq f(x^o)$,

$$f(x^o) = f\left(\frac{\eta}{1+\eta}\,x + \frac{1}{1+\eta}\,y\right) \leqq \frac{\eta}{1+\eta}\,f(x) + \frac{1}{1+\eta}\,f(y) \leqq f(x^o).$$

Hence $f(x) = f(x^o)$.

This argument also shows that a convex function cannot have a local maximum in a relatively open neighborhood unless $f(x)$ is constant in that neighborhood. If this does happen, the next property shows that this constant value must be an absolute minimum of $f(x)$.

> 12. If $f(x)$ is convex in D, $f(x)$ has at most one local minimum. If there is such a minimum, it is an absolute minimum and is attained on a convex set.

Suppose there is a local minimum at x^o. For any point x of D,

$$f(x^o) \leqq f((1-\theta)x^o + \theta x) \leqq (1-\theta)f(x^o) + \theta f(x)$$

if θ is a sufficiently small positive number. Hence $f(x) \geqq f(x^o)$ and $f(x^o)$ is the absolute minimum of f. If x^o and x^1 are two points at which $f(x)$ attains its minimum value μ,

$$\mu \leqq f((1-\theta)x^o + \theta x^1) \leqq (1-\theta)f(x^o) + \theta f(x^1) = \mu.$$

Hence f also attains its minimum at $(1-\theta)x^o + \theta x^1$.

> 13. Let $f(x)$ be a convex function defined in a set D which contains a flat F. If there exists a (non-homogeneous) linear function

$\ell(x)$ in A^n such that $f(x) \leqq \ell(x)$ in F, then $f(x) - \ell(x)$ is constant in F and in every flat which is a translate of F and is in the relative interior of D.

The function $g(x) = f(x) - \ell(x)$ is convex in D and non-positive throughout F. If x^o is a fixed point in F and x is any other point in F, the points $x^\lambda = (1-\lambda)x^o + \lambda x$ are in F for all λ. If $\lambda > 1$ the convexity of $g(x)$ implies that

$$g(x) \leqq (1 - \tfrac{1}{\lambda})g(x^o) + \frac{1}{\lambda}g(x^\lambda) \leqq (1 - \tfrac{1}{\lambda})g(x^o).$$

Letting $\lambda \longrightarrow \infty$ gives the relation

$$g(x) \leqq g(x^o).$$

If $\lambda < 0$, it follows from the convexity of $g(x)$ that

$$g(x^o) \leqq \frac{1}{1-\lambda}g(x^\lambda) + \frac{\lambda}{\lambda - 1}g(x) \leqq \frac{\lambda}{\lambda - 1}g(x).$$

Letting $\lambda \longrightarrow -\infty$ shows that

$$g(x) \geqq g(x^o).$$

Hence $g(x)$ is constant over F. Suppose that for the vector v the translate, $F' = F + v$, of F is relative interior to

D. Select $\lambda > 1$ so large that the point $x^{o} + \frac{\lambda}{\lambda - 1}v$ is in D. With x and x^{λ} as before, the definition of convexity applied to the points $x^{o} + \frac{1}{\lambda}v$, $x^{o} + v$, and x^{λ} gives

$$g(x+v) \leqq (1 - \tfrac{1}{\lambda})g(x^{o} + \tfrac{\lambda}{\lambda - 1}v) + \tfrac{1}{\lambda}g(x^{\lambda}) \leqq (1 - \tfrac{1}{\lambda})g(x^{o} + \tfrac{\lambda}{\lambda - 1}v).$$

By Property 9 g is bounded above in a neighborhood of $x^{o} + v$. Hence, $(1 - \frac{1}{\lambda})g(x^{o} + \frac{\lambda}{\lambda - 1}v)$ is uniformly bounded above for all sufficiently large λ. Thus $g(x+v)$ is bounded above for $x \in F$, that is to say g is bounded above in F'. That g is constant on F' follows from the first part of the theorem applied to the function g in the flat F'.

Property 13 is also a consequence of Chapter II, Property 24 applied to the set $[D,f]$.

14. Let $f(x)$ be convex in D and let p be a relative interior point of D. Assume that $f(x)$ is linear on each of finitely many (finite or infinite) segments in D which have linearly independent directions and which have p as a common interior point. Then $f(x)$ is linear over the convex hull of these segments.

There is a (non-homogeneous) linear function $\ell(x)$ in A^n which is identical with $f(x)$ along the segments. Hence, the convex function $g(x) = f(x) - \ell(x)$ vanishes on the segments. Now every point x of the convex hull of the segments may be written

$$x = \sum_{\rho=0}^{r} \lambda_\rho x^\rho , \qquad \lambda_\rho \geqq 0, \qquad \sum_{\rho=0}^{r} \lambda_\rho = 1,$$

with points x^ρ belonging to the segments. Hence, by Property 4

$$g(x) \leqq \sum_{\rho=0}^{r} \lambda_\rho g(x^\rho) = 0.$$

But p is a relative interior point of the convex hull and $g(p) = 0$. Therefore (Property 11) $g(x)$ is identically zero in the convex hull of the segments.

DEFINITION. A function $f(x)$ defined in a cone D with the origin as vertex is said to be positively homogeneous (of degree 1) in D if $f(\lambda x) = \lambda f(x)$ for every $x \in D$ and all $\lambda \geqq 0$.

15. A positively homogeneous function $f(x)$ in a convex cone D is convex in D

if and only if

$$f(x+y) \leqq f(x) + f(y)$$

for every x and y in D.

Convexity of $f(x)$ implies

$$\tfrac{1}{2}f(x+y) = f(\tfrac{1}{2}x + \tfrac{1}{2}y) \leqq \tfrac{1}{2}f(x) + \tfrac{1}{2}f(y).$$

On the other hand this inequality implies for $0 \leqq \theta \leqq 1$ that

$$f((1-\theta)x + \theta y) \leqq f((1-\theta)x) + f(\theta y) = (1-\theta)f(x) + \theta f(y).$$

Important examples of positively homogeneous convex functions are the support functions of point sets in A^n.

DEFINITION. Let M be an arbitrary point set in A^n. Denote by $B(M)$ the convex cone with the origin as vertex consisting of all vectors ξ such that M is bounded in the direction ξ (Chapter II, Section 5). The function

$$h_M(\xi) = \sup_{x \in M} x'\xi$$

defined in $B(M)$ is called the support function of M.

That $h_M(\xi)$ is positively homogeneous in $B(M)$ is clear. That it is convex follows from Property 7.

Obviously, $h_M(\xi) \leqq h_N(\xi)$ in $B(N)$ if $M \subset N$.

If $\|\xi\| = 1$, $h_M(\xi)$ is the distance from the origin to the supporting flat of M with positive normal vector ξ. Thus, $h_M(\xi)$ determines all the supports of M. The converse holds because $h_M(\xi)$ is positively homogeneous. Therefore M and $\overline{\{M\}}$, the closure of the convex hull of M, have the same support function. Also two sets M and N have the same support function if and only if $\overline{\{M\}} = \overline{\{N\}}$.

Let M be a point set with the support function $h_M(\xi)$ and λ a real number. Then the set λM has the support function $\lambda h_M(\xi)$ defined in $B(M)$ if $\lambda \geqq 0$, and the support function $-\lambda h_M(-\xi)$ defined in $-B(M)$ if $\lambda < 0$.

If M and N be point sets with the support functions $h_M(\xi)$ and $h_N(\xi)$, the set $M + N$ has the support function

$$h_{M+N}(\xi) = h_M(\xi) + h_N(\xi)$$

defined in $B(M) \cap B(N)$. This follows because

$$\sup_{x+y \in M+N} (x+y)'\xi = \sup_{\substack{x \in M \\ y \in N}} (x'\xi + y'\xi) = \sup_{x \in M} x'\xi + \sup_{y \in N} y'\xi .$$

2. CONTINUITY AND DIFFERENTIABILITY OF CONVEX FUNCTIONS OF ONE VARIABLE

The case of a convex function $\varphi(t)$ over a convex set D of A^1 $(-\infty<t<\infty)$ will now be considered. Here D must be an interval (open, closed, or half-open, possibly unbounded). If $x \neq y$ and $\theta \neq 0$ or 1, the inequality used in Section 1 to define convex functions is equivalent to

$$\varphi(t_2) \leqq \frac{t_3-t_2}{t_3-t_1}\varphi(t_1) + \frac{t_2-t_1}{t_3-t_1}\varphi(t_3)$$

for any three points $t_1 < t_2 < t_3$ of D. If $x = y$ or $\theta = 0$ or 1 the inequality of Section 1 is valid for all functions. Hence the present inequality is no weaker than the previous one.

16. If $\varphi(t)$ is convex in D

$$\frac{\varphi(t_2)-\varphi(t_1)}{t_2-t_1} \leqq \frac{\varphi(t_3)-\varphi(t_1)}{t_3-t_1} \leqq \frac{\varphi(t_3)-\varphi(t_2)}{t_3-t_2}$$

for $t_1 < t_2 < t_3$. Conversely, if one of these inequalities is satisfied for all $t_1 < t_2 < t_3$ in D, the function

$\varphi(t)$ is convex in D.

The first inequality of Property 16 follows from the defining inequality above by subtraction of $\varphi(t_1)$ from both sides and division by $t_2 - t_1$. Reversal of the steps proves the opposite implication. Similarly the second inequality of Property 16 is also equivalent to the defining inequality.

Property 16 shows that $\frac{\varphi(t+h)-\varphi(t)}{h}$ is monotone decreasing as $h \longrightarrow +0$. Hence, the right hand derivative

$$\varphi'_+(t) = \lim_{h \longrightarrow +0} \frac{\varphi(t+h)-\varphi(t)}{h}$$

exists and is either finite or $-\infty$. Similarly the left hand derivative

$$\varphi'_-(t) = \lim_{h \longrightarrow +0} \frac{\varphi(t-h)-\varphi(t)}{-h}$$

exists and is either finite or $+\infty$. From Property 16 it also follows for an interior point t of D and a sufficiently small $\varepsilon > 0$ that

$$\varphi'_+(t-\varepsilon) \leqq \varphi'_-(t) \leqq \varphi'_+(t) \leqq \varphi'_-(t+\varepsilon).$$

Since $\varphi'_+ < \infty$ and $\varphi'_- > -\infty$, both derivatives are finite at any interior point of D. This implies the continuity of φ in the interior of D. Furthermore, at any point where one of the derivatives is continuous the two derivatives agree, i.e. $\varphi(t)$ has an ordinary derivative. Since both derivatives are monotone increasing functions, they have at most a denumerable number of jump discontinuities. For a sufficiently small fixed $h \neq 0$, $\frac{\varphi(t+h)-\varphi(t)}{h}$ is continuous in an arbitrary closed interval interior to D. Therefore $\varphi_+'(t)$ is the limit of a decreasing sequence of continuous functions and consequently is upper semicontinuous. Similarly $\varphi_-'(t)$ is lower semicontinuous. The combination of semicontinuity and monotoneity shows that $\varphi_+'(t)$ is continuous from the right and $\varphi_-'(t)$ is continuous from the left.

These facts may be summarized as follows:

17. If $\varphi(t)$ is a convex function in an interval D, at every interior point of D it is continuous and has finite one-sided derivatives $\varphi_-'(t)$ and $\varphi_+'(t)$. These derivatives are monotone increasing functions which have identical values everywhere except for an at most denumerable number of points where they both have jumps. The value of $\varphi_-'(t)$ at a jump is the left hand limit, while the value of $\varphi_+'(t)$ is the right hand limit.

18. If $\varphi(t)$ is a convex function in D, $\varphi''(t)$ exists everywhere in D

except on a set of Lebesgue measure zero. Where it exists it is non-negative.

This follows from the Lebesgue theorem that a monotone function has a derivative almost everywhere.

19. A function $\varphi(t)$, which is continuous in an interval D and is twice differentiable in the interior of D, is convex in D if $\varphi''(t) \geqq 0$ for all t in the interior of D.

According to 16 it only has to be shown that

$$\frac{\varphi(t_3)-\varphi(t_2)}{t_3-t_2} - \frac{\varphi(t_2)-\varphi(t_1)}{t_2-t_1} \geqq 0$$

for any $t_1 < t_2 < t_3$. This is true, because repeated application of the Theorem of the Mean shows that the left hand side, apart from a positive factor, equals some value of φ''.

20. Under the same assumptions as in Property 19 $\varphi(t)$ is strictly convex if and only if $\varphi''(t) \geqq 0$ for all t in the interior of D, but is not identically zero in any (non-trivial) subinterval of D.

Property 20 is equivalent to the fact that $\varphi(t)$ is convex but not strictly convex if and only if it is convex in D and linear on some subinterval of D. Here the latter condition is obviously sufficient. That it is necessary is seen in the following way. If $\varphi(t)$ is convex but not strictly convex there are values t_o and t_1 such that

$$\psi(\theta) = \varphi((1-\theta)t_o + \theta t_1) - (1-\theta)\,\varphi(t_o) - \theta\,\varphi(t_1) \leqq 0$$

for $0 \leqq \theta \leqq 1$ and $\psi(\theta_o) = 0$ for some θ_o, $0 < \theta_o < 1$. This means that the convex function $\psi(\theta)$ has a maximum at θ_o and is, therefore, constant (Property 11).

The behavior of a convex function $\varphi(t)$ at the endpoints of its domain D may be described in the following way: $\varphi(t)$ is monotone either in the whole of D or in each of two complementary subintervals of D separated by a point at which $\varphi(t)$ is a minimum (Property 12). Hence, as t approaches an endpoint e of D, $\varphi(t)$ has a finite or infinite limit. If e is finite, $\lim_{t \to e} \varphi(t) > -\infty$ because of Property 10. If e belongs to D, the convexity of $\varphi(t)$ implies $\lim_{t \to e} \varphi(t) \leqq \varphi(e)$. On the other hand, it is easy to see that $\varphi(e)$ may be given an arbitrary value satisfying this inequality without violating the convexity of $\varphi(t)$. It is often convenient to redefine D and $\varphi(t)$ in the following manner: If e is an endpoint of D belonging to D, change the value of φ at e, if necessary, so that $\varphi(e) = \lim_{t \to e} (t)$. If e is a finite endpoint of D not belonging to D, and if $\lim_{t \to e} \varphi(t)$ is finite, adjoin e to D and define $\varphi(e) = \lim_{t \to e} \varphi(t)$. By these inessential changes a convex function $\varphi(t)$ is obtained which is continuous in the whole interval D of definition and $\varphi(t) \longrightarrow \infty$ as t approaches a finite end-point of D which is not in D. For a function with these properties the set $[D, \varphi]$ is closed. Conversely if the convex set $[D, \varphi]$ is closed, φ is such a function.

3. CONTINUITY PROPERTIES OF CONVEX FUNCTIONS OF SEVERAL VARIABLES

21. Suppose $f(x)$ is a convex function over a convex set D of A^n and D' is a compact convex set in the relative interior of D. Let $\delta > 0$ be such that the closed relative δ-neighborhood $D'' = D' + \delta \overline{U}$ of D' is also in the relative interior of D. Here $\overline{U}$ denotes the closed unit sphere of that subspace through the origin which is a translate of the minimal flat containing D. Let M and m be numbers such that $m \leqq f(x) \leqq M$ in D'' (Properties 9 and 10). Under these conditions

$$\left| f(x+y) - f(x) \right| \leqq \frac{M - m}{\delta} \|y\|$$

for any $x \in D'$ and any vector y for which $x + y \in D''$.

If $y = 0$, the statement is trivial. If $y \neq 0$, consider the function $f(x+ty)$ of the real variable t for fixed $x \in D'$, $y \in S(D) - x$. This is a convex function at least in the interval $-\frac{\delta}{\|y\|} \leqq t \leqq \frac{\delta}{\|y\|}$. From Property 16 it follows that for $0 < t \leqq \frac{\delta}{\|y\|}$

$$\frac{f(x)-f(x-\frac{\delta}{\|y\|}y)}{\delta}\|y\| \leqq \frac{f(x+ty)-f(x)}{t} \leqq \frac{f(x+\frac{\delta}{\|y\|}y)-f(x)}{\delta}\|y\| .$$

Hence

$$\left|\frac{f(x+ty) - f(x)}{t}\right| \leqq \frac{M - m}{\delta}\|y\| .$$

If $\|y\| \leqq \delta$, the value one may be substituted for t. The inequality obtained is obviously also valid when $\|y\| > \delta$ provided $x + y \in D''$.

The inequality in 21 shows that f satisfies a uniform Lipschitz condition in D'. Hence a uniformly bounded family of convex functions over the domain D'' is equicontinuous in D'. From this follows

> 22. If a set of convex functions over a relatively open convex set D is uniformly bounded in every compact subset of D, a sequence of functions may be selected from this set so that the sequence converges in D to a convex function. Moreover, this convergence is uniform in any compact subset of D.

An immediate consequence of 21 is

> 23. If $f(x)$ is convex in D, it is continuous in the relative interior of D.

The behavior of a convex function at the boundary of its domain is essentially described by

> 24. If $f(x)$ is convex in D

and y is a relative boundary point of D

$$\underline{\lim}_{x \to y} f(x) > -\infty .$$

If $y \in D$

$$\underline{\lim}_{x \to y} f(x) \leqq f(y).$$

The first statement follows from Property 10 and the second is true because

$$\begin{aligned}\underline{\lim}_{x \to y} f(x) &\leqq \lim_{\theta \to 1} f((1-\theta)x^{o} + \theta y) \\ &\leqq \lim_{\theta \to 1} \big((1-\theta)f(x^{o}) + \theta f(y)\big) \\ &= f(y)\end{aligned}$$

is valid for any fixed $x^{o} \in D$.

Let

$$f(x_1,x_2) = \frac{x_1^2 + x_2^2}{2x_2}$$

for $x_2 > 0$ and define $f(0,0)$ to be an arbitrary non-negative number. Then f is convex over the half-plane $x_2 > 0$ plus the origin. Now $\underline{\lim}_{x \to 0} f(x) = 0$ while $\overline{\lim}_{x \to 0} f(x) = +\infty$. This example shows that "$\underline{\lim}$" in 24 cannot be replaced by "lim" and that the inequality cannot be strengthened.

25. Let $f(x)$ be convex in a relatively open convex set $\tilde{D}$. Denote by D

A function obtained in the way described in 25 has the properties given in the

> DEFINITION: A convex function $f(x)$ defined in a convex set D is called closed if $\lim_{x \to y} f(x) = \infty$ for every relative boundary point y of D which is not in D, and $\underline{\underline{\lim}}_{x \to y} f(x) = f(y)$ for every relative boundary point y of D which is in D.

A closed convex function may be obtained from any convex function by removing the relative boundary points of its domain and then extending the function in the way described in 25.

> 26. If $f(x)$ in D is a closed convex function $\lim_{x \to y} f(x) = f(y)$ where y is any point in D and x approaches y along a segment belonging to D.

Letting x approach y along the segment from x^o to y is the same as allowing θ to approach one from below in the expression $(1-\theta)x^o + \theta y$. Since

$$f((1-\theta)x^o + \theta y) \leqq (1-\theta)f(x^o) + \theta f(y),$$

$$\overline{\lim_{\theta \to 1}}\, f((1-\theta)x^o + \theta y) \leqq f(y).$$

On the other hand, $f(y) = \underline{\underline{\lim}}_{x \to y} f(x)$. This proves the statement.

> 27. A convex function $f(x)$ in D is closed if and only if the set $[D,f]$ of A^{n+1} is closed.

the set obtained by adding to $\widetilde{D}$ all relative boundary points y for which $\underline{\lim}_{x \to y} f(x) < \infty$.
Define

$$f(y) = \underline{\lim}_{x \to y} f(x)$$

for y in D but not in $\widetilde{D}$ and x in $\widetilde{D}$. With these definitions D is convex and $f(x)$ is a convex function in D.

If y^o and y^1 are any two points of D, there are sequences x^{oi} and x^{1i}, $i = 1,2,\ldots,$ of points from $\widetilde{D}$ such that $x^{oi} \to y^o$, $x^{1i} \to y^1$ and

$$\lim_{i \to \infty} f(x^{oi}) = \underline{\lim}_{x \to y^o} f(x),$$

$$\lim_{i \to \infty} f(x^{1i}) = \underline{\lim}_{x \to y^1} f(x).$$

Now for $0 \leqq \theta \leqq 1$

$$f((1-\theta)x^{oi} + \theta x^{1i}) \leqq (1-\theta)f(x^{oi}) + \theta f(x^{1i}).$$

Hence

$$\underline{\lim}_{x \to (1-\theta)y^o + \theta y^1} f(x) \leqq \underline{\lim}_{i \to \infty} f((1-\theta)x^{oi} + \theta x^{1i})$$

$$\leqq (1-\theta)f(y^o) + \theta f(y^1) < \infty.$$

This shows that $(1-\theta)y^o + \theta y^1 \in D$ and that

$$f((1-\theta)y^o + \theta y^1) \leqq (1-\theta)f(y^o) + \theta f(y^1) .$$

Suppose the set $[D,f]$ in A^{n+1} is closed. Let y be a relative boundary point of D and $x^i \in D$ a sequence of points converging to y such that $\lim\limits_{i \to \infty} f(x^i) = \underline{\lim}_{x \to y} f(x)$. If this $\underline{\lim}$ is finite, the sequence of points $(x^i, f(x^i))$ in $[D,f]$ converges to the point $(y, \underline{\lim}_{x \to y} f(x))$ in $[D,f]$. This means that $y \in D$ and $f(y) \leqq \underline{\lim}_{x \to y} f(x)$. From 24 it now follows that $f(y) = \underline{\lim}_{x \to y} f(x)$. Conversely, suppose the function is closed. Consider any sequence of points (x^i, z_i) in $[D,f]$ which converges to a point (y,z). Since $z_i \geqq f(x^i)$, $z \geqq \underline{\lim}_{x \to y} f(x)$. This implies $y \in D$ and $z \geqq f(y)$, that is $(y,z) \in [D,f]$.

4. DIRECTIONAL DERIVATIVES AND DIFFERENTIABILITY PROPERTIES

28. If $f(x)$ is convex in D, the "directional derivative"

$$f'(x;y) = \lim_{t \to +0} \frac{f(x+ty) - f(x)}{t}$$

exists and is either finite or $-\infty$ for any x in D and any vector y such that $x + y$ is in the projecting cone $P_x(D)$. For a fixed x, $f'(x;y)$ is either finite for all y in the translate $P_x(D) - x$ of the projecting cone, or it is $-\infty$ for all relative interior vectors y of $P_x(D) - x$. When $f'(x;y)$ is finite in $P_x(D) - x$, $f'(x;y)$ is a positively homogeneous, convex function in

$P_x(D) - x$. If x is relative interior to D, the cone $P_x(D) - x$ is a subspace and $f'(x;y)$ is finite for all y in this subspace.

The above limit is the right hand derivative at $t = 0$ of the function $f(x+ty)$ of t which is defined and convex at least in some interval $0 \leq t < b$. Thus the limit exists and is $< \infty$ (Property 17). If x is relative interior to D, $f(x+ty)$ is defined and convex in some interval containing $t = 0$ in its interior and, hence, $f'(x;y)$ is finite.

If $\lambda > 0$,

$$\frac{f(x+\lambda ty) - f(x)}{t} = \lambda \frac{f(x+\lambda ty) - f(x)}{\lambda t} .$$

Hence

$$(*) \qquad f'(x; \lambda y) = \lambda f'(x;y)$$

for $\lambda > 0$. This equation is clearly also valid for $\lambda = 0$. If $f'(x;y) = -\infty$ for a particular y, it must be infinite on the ray generated by y in $P_x(D) - x$. If y^o and y^1 are in $P_x(D) - x$,

$$\frac{f(x+t(y^o+y^1)) - f(x)}{t} = \frac{f(\frac{1}{2}(x+2ty^o) + \frac{1}{2}(x+2ty^1)) - f(x)}{t}$$

$$\leq \frac{f(x+2ty^o) - f(x)}{2t} + \frac{f(x+2ty^1) - f(x)}{2t} .$$

If $f'(x;y)$ is $-\infty$ on any ray (y^o), it now follows that it must be $-\infty$ on every ray which is strictly between (y^o) and any other ray (y^1) of $P_x(D) - x$. In particular

$f'(x;y) = -\infty$ in the whole relative interior of $P_x(D) - x$. If $f'(x;y)$ for the x considered is $-\infty$ for no y, the above inequality gives $f'(x;y^o+y^1) \leq f'(x;y^o) + f'(x;y^1)$. This combined with equation (*) shows that $f'(x;y)$ is a positively homogeneous, convex function of y in the cone $P_x(D) - x$ (Property 15).

That $f'(x;y)$ need not be $-\infty$ on all relative boundary rays of $P_x(D) - x$ when it is $-\infty$ on the relative interior rays is shown by the following example: Let D be a closed strip of a plane and let $f(x)$ be a convex function over D with its graph half of a circular cylinder. If x is a boundary point of D, $f'(x;y) = -\infty$ in any direction from x into the interior of D but is finite in the two directions along the edge of D.

29. If $f(x)$ is convex in D

$$f(x) \geqq f(x^o) + f'(x^o;x-x^o)$$

for all x and x^o in D. If $f(x)$ is positively homogeneous and convex in a convex cone D

$$f(y) \geqq f'(x;y)$$

for all x and y in D.

If x^o and x are in D, $f(x^o+t(x-x^o))$ is a convex function of t in an interval including $0 \leqq t \leqq 1$. Hence for $t > 0$

$$\frac{f(x^o+t(x-x^o)) - f(x^o)}{t} \geqq f'(x^o;x-x^o)$$

because the left hand side decreases as t decreases (Property 16). Substitution of one for t gives

$$f(x^o+(x-x^o)) - f(x^o) \geqq f'(x^o;x-x^o),$$

that is the first statement of 29. If $f(x)$ is positively homogeneous

$$f(x^o) + f(x-x^o) \geqq f(x^o + (x-x^o)).$$

This and substitution of y for $x - x^o$ gives the second statement from the first.

30. If $f(x)$ is convex in D, the supporting hyperplanes of the set $[D,f(x)]$ which contain a fixed point $(x^o,f(x^o))$ are identical with the supporting hyperplanes of the corresponding set $[P_{x^o}(D),\ f(x^o) + f'(x^o;x-x^o)]$ for the function $f(x^o) + f'(x^o;x-x^o)$ of x.

The set $[P_{x^o}(D),\ f(x^o) + f'(x^o;x-x^o)]$ is a convex cone in A^{n+1} with vertex $(x^o,f(x^o))$. This follows easily from the facts that $P_{x^o}(D)$ is a convex cone and that $f'(x^o;y)$ is positively homogeneous in y. Hence every supporting hyperplane of this set goes through $(x^o,f(x^o))$. Furthermore

$$[D,f(x)] \subset [P_{x^o}(D),\ f(x^o) + f'(x^o;x-x^o)]$$

because of Property 29, and the inclusion $D \subset P_{x^o}(D)$. Every supporting hyperplane of the set $[P_{x^o}(D),\ f(x^o) + f'(x^o;x-x^o)]$ is thus a supporting hyperplane of $[D,f(x)]$ through $(x^o,\ f(x^o))$.

To prove the converse consider a supporting hyperplane of $[D,f(x)]$ which is not parallel to the z-axis and which contains $(x^o,f(x^o))$. Its equation may be written

$$z = f(x^o) + (x-x^o)'u$$

with some vector $u \neq 0$ in A^n. Now

$$f(x) \geqq f(x^o) + (x-x^o)'u \quad \text{for all} \quad x \in D.$$

Hence, replacement of x by $x^o + t(x-x^o) \in D$ for $0 < t \leqq 1$ gives

$$f(x^o+t(x-x^o)) \geqq f(x^o) + t(x-x^o)'u,$$

$$\frac{f(x^o+t(x-x^o)) - f(x^o)}{t} \geqq (x-x^o)'u$$

and

$$f(x^o) + f'(x^o;x-x^o) \geqq f(x^o) + (x-x^o)'u.$$

Since $f'(x^o;y)$ is positively homogeneous in y, the last inequality holds for all $x \in P_{x^o}(D)$. This means that a supporting hyperplane of $[D,f(x)]$ through $(x^o,f(x^o))$ and not parallel to the z-axis is also a supporting hyperplane of $[P_{x^o}(D), f(x^o) + f'(x^o;x-x^o)]$.

A supporting hyperplane of $[D,f(x)]$ through $(x^o,f(x^o))$ which is parallel to the z-axis has an equation of the form $(x-x^o)'u = 0$. For this u

$$(x-x^o)'u \leqq 0 \quad \text{for all} \quad x \in D$$

Clearly this inequality also holds for all $x \in P_{x^o}(D)$

because every $x \in P_{x^o}(D)$ may be written $x = x^o + \lambda(x^1-x^o)$ with $x^1 \in D$, $\lambda \geqq 0$. This means that the given hyperplane parallel to the z-axis is a supporting hyperplane of $[P_{x^o}(D),\ f(x^o) + f'(x^o;x-x^o)]$.

31. Let $f(x)$ be convex in D, and let x^o be an arbitrary point of D. Then there is a supporting hyperplane to $[D,f(x)]$ which contains the point $(x^o,f(x^o))$ and which is not parallel to the z-axis if and only if $f'(x^o;y)$ is finite for all y in $P_{x^o}(D) - x^o$.

Suppose $f'(x^o;x^1-x^o)$ is finite for some x^1 relative interior to $P_{x^o}(D)$. The ray in A^{n+1} with initial point $(x^o,f(x^o))$ and direction determined by the vector $(x^1-x^o,\ f'(x^o;x^1-x^o))$ is a relative boundary ray of the convex cone $C = [P_{x^o}(D),\ f(x^o) + f'(x^o;x-x^o)]$. Hence, in the minimal flat $S(C)$ containing this cone there is a supporting hyperplane H of C which contains this ray. If H were parallel to the z-axis, its intersection with the hyperplane $z = 0$ would be a supporting flat of $P_{x^o}(D)$. On the other hand it would contain the relative interior point x^1 of $P_{x^o}(D)$, but this is impossible. Now H can be extended to a supporting hyperplane in A^{n+1} of C not parallel to the z-axis. From 30 it follows that H also supports $[D,f(x)]$. The converse follows from the inequality

$$f'(x^o;x-x^o) \geqq (x-x^o)'u > -\infty$$

obtained in the proof of 30 for any supporting hyperplane

$$z = f(x^o) + (x-x^o)'u$$

of $[D,f(x)]$ which is not parallel to the z-axis.

Now let $f(x)$ be convex in an n-dimensional convex set D, and let x^o be a fixed interior point of D. Consider the function on the line $x = x^o + ty$ where y is an arbitrary fixed vector in A^n. In some interval $f(x^o+ty)$ is a convex function of t whose right hand derivative at $t = 0$ is $f'(x^o,y)$ and whose left hand derivative at $t = 0$ is $-f'(x^o;-y)$. Hence $f(x^o+ty)$ is differentiable at $t = 0$ if and only if $-f'(x^o;-y) = f'(x^o;y)$ that is if

$$f'(x^o;\lambda y) = \lambda f'(x^o;y)$$

for arbitrary real λ. Therefore the partial derivatives of $f(x)$ exist if and only if for all real λ

$$f'(x^o;\lambda u^i) = \lambda f'(x^o;u^i)$$

where the u^i, $i = 1,\dots,n$, denote the unit vectors $(0,\dots,0,1,0,\dots 0)$. The partial derivatives have the values

$$\frac{\partial f}{\partial x_i} = f'(x^o;u^i).$$

If they exist, $f'(x^o;y)$ is linear on every coordinate axis. From Property 14 it then follows that $f'(x^o;y)$ is linear over the whole y-space. Hence

$$f'(x^o;dx) = \sum_{i=1}^{n} \frac{\partial f}{\partial x_i} dx_i$$

is the total differential of $f(x)$.

32. Let $f(x)$ be convex in an n-dimensional convex set D. Let x^o

be an interior point of D and suppose that $f'(x^o;y)$ is a linear function of y. Then $f(x)$ is differentiable at $x = x^o$.

This statement is equivalent with the following: To every $\varepsilon > 0$ there is a $\delta > 0$ such that

$$|f(x^o+tu) - f(x^o) - tf'(x^o;u)| \leqq \varepsilon t$$

for all unit vectors u and $0 < t \leqq \delta$. From Property 29 and the definition of $f'(x^o;y)$, it follows that for each fixed vector y there is a $\delta(y)$ such that

$$(*) \qquad 0 \leqq f(x^o+ty) - f(x^o) - tf'(x^o;y) \leqq \varepsilon t$$

for $0 < t \leqq \delta(y)$. Apply this to the vectors y^i, $i = 1,\ldots,2^n$, all of whose coordinates have the value ± 1 and put $\delta = \min\limits_i \delta(y^i)$. Then $(*)$ is valid for each $y = y^i$ and $0 < t \leqq \delta$. Now for any fixed t in this interval $f(x^o+ty) - f(x^o) - tf'(x^o;y)$ is a convex function of y since $f'(x;y)$ is linear in y. Hence $(*)$ is valid for all y in the convex hull of the points y^i (Properties 4 and 29), in particular for all unit vectors u. This proves the statement.

33. Let $f(x)$ be convex in a relatively open convex set D of dimension d, and let y be a fixed vector parallel with the minimal flat containing D. Then $f'(x;y)$ is an upper semicontinuous function of x in D. The ordinary derivative of $f(x)$ in the direction y exists everywhere in D with the possible exception of a subset of d-measure zero. Where it exists the derivative is a continuous function of x.

In every compact subset of D the function $f'(x;y)$ of x is the limit of a decreasing sequence of continuous functions $\frac{f(x+t_i y) - f(x)}{t_i}$ where $t_i > 0$, $t_i \longrightarrow 0$. Hence $f'(x;y)$ is upper semicontinuous. The ordinary derivative of $f(x)$ in the direction y exists at a point x if and only if $f'(x;y) = -f'(x;-y)$. Now $f'(x;y) + f'(x;-y) \geq 0$, since $f'(x;y)$ is positively homogeneous and convex in y. Hence the set of points at which the derivative does not exist is the set of x at which $f'(x;y) + f'(x;-y) > 0$. Thus this set is measurable. Its intersection with a line parallel to y contains at most a denumerable number of points, (Property 17). Therefore the set has d-measure zero. Since $f'(x;y)$ is upper semicontinuous for every fixed y, $-f'(x;-y)$ is lower semicontinuous and, hence, $f'(x;y)$ is continuous in the set at which $f'(x;y) + f'(x;-y) = 0$.

34. If $f(x)$ is convex in an open convex set D, it is differentiable with continuous partial derivatives everywhere in D except for a set of measure zero.

Apply 33 to each of the unit vectors $u^i = (0,\ldots,0,1,0,\ldots,0)$ on the coordinate axes instead of to y. For each $i = 1,\ldots,n$, there is a set of measure zero at which $\frac{\partial f}{\partial x_i}$ does not exist. The union U of these sets has measure zero. At every x in D but outside U all partial derivatives exist, that is $f'(x;y)$ is linear in y and $f(x)$ differentiable (Property 32). The continuity of the partial derivatives is an immediate consequence of 33.

35. if $f(x)$ is a twice differentiable function in an open donvex set D, $f(x)$ is convex in D if and only if the quadratic form

$$\sum_{i,j=1}^{n} f_{ij}(x) y_i y_j, \quad f_{ij}(x) = \frac{\partial^2 f}{\partial x_i \partial x_j},$$

is positive semidefinite for every x in D.

From the fact that $f(x)$ is convex if and only if it is convex on every straight segment in D and from Properties 18 and 19 it follows that $f(x)$ is convex if and only if

$$\left[\frac{d^2 f(x+ty)}{dt^2}\right]_{t=0} = \sum_{i,j=1}^{n} f_{ij}(x) y_i y_j \geqq 0$$

for all $x \in D$ and all y.

A sufficient condition that a function $f(x)$, twice differentiable in an open convex set D, is strictly convex in D is that $\Sigma f_{ij}(x) y_i y_j$ is positive definite. It is even sufficient that the form is positive semidefinite for all x in D and the determinant $\det f_{ij}(x)$ is not identically zero on any segment in D.

5. CONJUGATE CONVEX FUNCTIONS

In Chapter II, Section 8, polarity with respect to the paraboloid

$$2z = x_1^2 + \dots + x_n^2$$

in $A^{n+1}(x_1, \dots, x_n, z)$ was described. This polarity will now be used to define an involutory correspondence between closed convex functions.

For the sake of brevity, a flat in A^{n+1} will be called <u>vertical</u> or <u>non-vertical</u> according as it is or is not parallel with the z-axis. The polar hyperplane to a point (x,z) of A^{n+1} has the equation $\zeta + z = x'\xi$, where (ξ, ζ) are variables in the space $A^{n+1}(\xi_1, \dots, \xi_n, \zeta)$. Let $f(x)$ in C be a closed convex function. To each point

pg 78

(x,z) in $[C,f]$ let correspond the closed upper half-space $\zeta \geqq x'\xi - z$ bounded by the polar hyperplane to the point. The intersection of all these half-spaces for (x,z) in $[C,f]$ is a closed convex set $[C,f]^*$ in A^{n+1}. Since $x'\xi - f(x) \geqq x'\xi - z$ for all $(x,z) \in [C,f]$, it is sufficient to consider the half-spaces

$$\zeta \geqq x'\xi - f(x), \quad x \in C.$$

Hence, $[C,f]^*$ is the set $[\Gamma, \varphi]$ for the function

$$\zeta = \varphi(\xi) = \sup_{x \in C} (x'\xi - f(x))$$

defined in the projection Γ in the ζ-direction of $[C,f]^*$ on the hyperplane $\zeta = 0$. This function is convex and closed since $[C,f]^*$ is convex and closed. A point ξ is in Γ if and only if the function $x'\xi - f(x)$ is bounded above for $x \in C$.

The set $[\Gamma, \varphi]$ may also be obtained from $[C,f]$ in a dual way. A non-vertical hyperplane has an equation of the form $z = x'\xi - \zeta$ with (x,z) variable. Its pole is the point (ξ, ζ). If and only if this hyperplane is a barrier to $[C,f]$, we have $f(x) \geqq x'\xi - \zeta$ for all $x \in C$, that is $(\xi, \zeta) \in [\Gamma, \varphi]$. Thus, $[\Gamma, \varphi]$ is the set of the poles of all non-vertical barriers to $[C,f]$. Since there exist such barriers (Propositions 28 and 31), $[\Gamma, \varphi]$ is not empty.

If $g(x)$ is a closed concave function defined in the convex set D, let $[D,g]$ denote the closed convex set of all points (x,z) in A^{n+1} such that $x \in D$ and $z \leqq g(x)$. To a point (x,z) in $[D,g]$ let correspond the closed lower half-space $\zeta \leqq x'\xi - z$ bounded by the polar hyperplane of the point. The intersection of all these half-spaces is the set $[\Delta, \psi]$ for the closed concave function

$$\zeta = \psi(\xi) = \inf_{x \in D} (x'\xi - g(x))$$

defined in the set Δ of all points ξ for which $x'\xi - g(x)$ bounded below in D. As in the convex case, $[\Delta, \psi]$ is the set of the poles of all non-vertical varriers to $[D, g]$.

DEFINITION: Let $f(x)$ in C be convex and closed. Then the closed convex function

$$\varphi(\xi) = \sup_{x \in C} (x'\xi - f(x))$$

defined in the set Γ of all points ξ for which $x'\xi - f(x)$ is bounded above for x in C is called the conjugate function of $C, f(x)$.

Let $g(x)$ in D be concave and closed. Then the closed concave function

$$\psi(\xi) = \inf_{x \in D} (x'\xi - g(x))$$

defined in the set Δ of all points ξ for which $x'\xi - g(x)$ is bounded below for x in D is called the conjugate function of $D, g(x)$.

From what has been said it follows that equivalent definitions of the conjugate functions are

$$\varphi(\xi) = \sup_{(x,z) \in [C,f]} (x'\xi - z), \quad \psi(\xi) = -\sup_{(x,z) \in [D,g]} (-x'\xi + z).$$

These show that $\varphi(\xi)$ is the support function of the point set $[C, f]$ for the argument $(\xi, -1)$ and that $-\psi(\xi)$ is the support function of $[D,g]$ for $(-\xi,1)$.

From the above remarks the following geometrical interpretations of the conjugates of convex and concave functions are immediately derived:

> 36. Let $f(x)$ in C be convex (or concave) and closed, and let $\varphi(\xi)$ in Γ be its conjugate. Then Γ consists of all ξ such that $[C,f]$ is bounded in the direction of the vector $(\xi,-1)$ (or $(-\xi,1)$), and $-\varphi(\xi)$ is the z-intercept of the supporting hyperplane of $[C,f]$ with the normal vector $(\xi,-1)$ (or $(-\xi,1)$).

As already mentioned, the correspondence defined above between closed convex or concave functions is involutory:

> 37. If $\varphi(\xi)$ in Γ is the conjugate of the closed convex (or concave) function $f(x)$ in C, then $f(x)$ in C is the conjugate of $\varphi(\xi)$ in Γ.

Let $f^*(x)$ in C^* be the conjugate of $\varphi(\xi)$, Γ. From the preceding statements it follows that $[C^*,f^*]$ is the intersection of all supports to $[C,f]$ whose bounding hyperplanes are non-vertical. Thus the statement $[C^*,f^*] = [C,f]$ follows from the

> LEMMA: A closed convex set M in A^{n+1} having supports bounded by non-vertical hyperplanes is the intersection of all these supports.

Since M is the intersection of all its supports, the statement is that the supports bounded by vertical hyperplanes may be omitted without changing the intersection. A point (ξ^0, ζ_0) not in M is outside some bound or support of M. It has to be shown that there is such a bound or support bounded by a non-vertical hyperplane. Let H be a barrier of M such that (ξ^0, ζ_0) is separated from M by H but (ξ^0, ζ_0) is not on H. If H is non-vertical there is nothing to prove. Suppose H is vertical and let H' be a non-vertical barrier of M. The hyperplanes H and H' divide the space A^{n+1} into four wedges, one of which contains M but not (ξ^0, ζ_0). Now turn H about the intersection of H and H' away from the wedge containing M, but so little that H remains in the wedges adjacent to that wedge containing M and that (ξ^0, ζ_0) is still separated from M. The hyperplane obtained bounds a bound or support with the required properties.

38. Let $f(x)$ in C and $\varphi(\xi)$ in Γ be conjugate closed convex functions. Then

$$x'\xi \leqq f(x) + \varphi(\xi) \quad \text{for } x \in C,\ \xi \in \Gamma .$$

To every $x \in C$ for which $f'(x,y)$ is finite for all y for which it is defined, there is at least one $\xi \in \Gamma$ such that equality is valid, and dually. For concave functions the inequality is reversed.

The inequality follows immediately from the definition of the conjugate function. The statement concerning the equality sign is a consequence of Propositions 31 and 36.

In general there is no simple relation between the properties of the domains C and Γ of two conjugate functions. To a point x in C correspond all points ξ of Γ with the property that through the point $(x,f(x))$ there is a supporting hyperplane to $[C,f]$ with the normal direction $(\xi,-1)$, and dually. Thus, the correspondence between the sets depends strongly on the behavior of the function $f(x)$. But there is one very simple direct relation between C and Γ which will play a role in the following:

If one of the sets is bounded in the direction η, the asymptotic cone of the other one contains the ray with the direction η.

This is seen in the following way: Suppose that C is bounded in the direction η. Then $[C,f]$ is bounded in the direction $(\eta,0)$ and like every set $[C,f]$, in some direction $(\xi,-1)$. Since the directions in which a set is bounded form a convex cone, $[C,f]$ is bounded in all directions $(\xi+\rho\eta,-1)$, $\rho \geqq 0$. Hence Γ contains the half-line $\xi+\rho\eta$, $\rho \geqq 0$.

In the remaining part of this section, only convex functions are considered. The corresponding results for concave functions are obtained by rather obvious changes following from the fact that $C,-f$ and $-\Gamma, -\varphi$ are conjugate if C,f and Γ,φ are conjugate. More generally:

39. Let $f(x)$ in C be a closed convex function and $\varphi(\xi)$ in Γ its conjugate function. Then for any real $\lambda \neq 0$ the conjugate function of $\lambda f(x)$ in C is $\lambda\varphi(\frac{\xi}{\lambda})$ in λC.

This follows from the relations

$$\sup_{x \in C} (x'\xi - \lambda f(x)) = \lambda \sup_{x \in C} (\frac{x'\xi}{\lambda} - f(x)),$$

for $\lambda > 0$, $\xi \in \lambda \Gamma$ and

$$\sup_{x \in C} (x'\xi - \lambda f(x)) = \lambda \inf_{x \in C} (\frac{x'\xi}{\lambda} - f(x)),$$

for $\lambda < 0$, $\xi \in \lambda \Gamma$.

Other obvious consequences of the definition of conjugate functions are the following:

> 40. Let $f(x)$ in C be a closed convex function and $\varphi(\xi)$ in Γ its conjugate function. Then the conjugate function of $f(x) + k$ in C, k a constant, is $\varphi(\xi) - k$ in Γ. The conjugate function of $f(x-v)$ in $C + v$, v a constant vector, is $\varphi(\xi) + v'\xi$ in Γ.

The first statement is clear and the second follows from

$$\sup_{x \in C+v} (x'\xi - f(x-v)) = \sup_{x-v \in C} ((x-v)'\xi - f(x-v) + v'\xi.$$

Now let $f_1(x)$ in C_1 and $f_2(x)$ in C_2 be closed convex functions, where C_1 and C_2 have points in common. Then $f_1(x) + f_2(x)$ is a convex function defined in $C_1 \cap C_2$. It is easily seen that this function is closed. To prove it, let y be a relative boundary point of $C_1 \cap C_2$. If $y \in C_1 \cap C_2$, we have $f_1(x) \longrightarrow f_1(y)$, $f_2(x) \longrightarrow f_2(y)$ as x approaches y on any segment in $C_1 \cap C_2$ (Proposition 26).

Hence $f_1(x) + f_2(x) \longrightarrow f_1(y) + f_2(y)$ under the same condition. This implies that $\underline{\lim}_{x \longrightarrow y} (f_1(x) + f_2(x)) < \infty$ as x approaches y arbitrarily and (again by Proposition 26) that this $\underline{\lim}$ is $f_1(y) + f_2(y)$. If y is not in $C_1 \cap C_2$, we have either $f_1(x) \longrightarrow \infty$ or $f_2(x) \longrightarrow \infty$ as $x \longrightarrow y$, and hence $f_1(x) + f_2(x) \longrightarrow \infty$, since f_1 and f_2 are bounded below in a neighborhood of y.

41. Let $f_1(x)$ in C_1 and $f_2(x)$ in C_2 be closed convex functions with the conjugates $\varphi_1(\xi)$ in Γ_1 and $\varphi_2(\xi)$ in Γ_2. Assume that $C_1 \cap C_2$ is not empty. Denote by $\varphi(\xi)$ in Γ the conjugate of the function $f_1(x) + f_2(x)$ in $C_1 \cap C_2$. Then

$$(*) \qquad [\Gamma, \varphi] = \overline{[\Gamma_1, \varphi_1] + [\Gamma_2, \varphi_2]},$$

$$\Gamma_1 + \Gamma_2 \subset \Gamma \subset \overline{\Gamma_1 + \Gamma_2},$$

and, for $\xi \in \Gamma_1 + \Gamma_2$,

$$\varphi(\xi) = \inf_{\substack{\xi^1 \in \Gamma_1,\ \xi^2 \in \Gamma_2 \\ \xi^1 + \xi^2 = \xi}} \left(\varphi_1(\xi^1) + \varphi_2(\xi^2)\right)$$

To prove the first statement, it will be shown that the conjugate of the function φ on the set Γ <u>defined</u> by (*) is $f_1(x) + f_2(x)$ in $C_1 \cap C_2$. According to an observation made in connection with the definition of the conjugate function, $f_1(x)$ and $f_2(x)$ are the support functions of the sets $[\Gamma_1, \varphi_1]$ and $[\Gamma_2, \varphi_2]$ with $(x,-1)$ as argument. Now, the set $[\Gamma_1, \varphi_1] + [\Gamma_2, \varphi_2]$, and therefore its closure, is bounded in all directions $(x,-1)$ in which both $[\Gamma_1, \varphi_1]$ $[\Gamma_2, \varphi_2]$ are bounded, and conversely. Hence, the support function of $\overline{[\Gamma_1, \varphi_1] + [\Gamma_2, \varphi_2]}$ taken for $(x,-1)$ is defined in $C_1 \cap C_2$ and equals $f_1(x) + f_2(x)$ (see the end of Section 1 of this Chapter). The two last statements of 41 follow from the fact that $[\Gamma_1, \varphi_1] + [\Gamma_2, \varphi_2]$ consists of all points (ξ, ζ) for which $\xi = \xi^1 + \xi^2$, $\xi^1 \in \Gamma_1$, $\xi^2 \in \Gamma_2$ and $\zeta = \zeta_1 + \zeta_2$, $\zeta_1 \geqq \varphi_1(\xi^1)$, $\zeta_2 \geqq \varphi_2(\xi^2)$.

For the application of this result which will be made in Section 6 it is important to have sufficient conditions in order that the inf in the statement of Proposition 41 may be replaced by min. This may obviously be done if $[\Gamma_1, \varphi_1] + [\Gamma_2, \varphi_2]$ is closed, which will be the case if C_1 and C_2 have points in common which are relative interior to both sets; that is, if C_1 and C_2 cannot be separated by a hyperplane of $S(C_1 \cup C_2)$ in the sense of the Separation Theorem 28, Chapter II, Section 6. However, this condition is not necessary. Necessary and sufficient conditions in terms of C_1, f_1 and C_2, f_2 are rather complicated and will not be formulated here. To the extent that the question is of importance it will be discussed in Section 6 in a slightly different and more intuitive formulation.

Let $C_\alpha, f_\alpha(x)$ where α runs through any set, be closed convex functions. Let $C \subset \bigcap_\alpha C_\alpha$ be the set of those points x at which $\sup_\alpha f_\alpha(x)$ is finite and define $f(x) = \sup_\alpha f_\alpha(x)$ for $x \in C$. According to Proposition 7, C is convex, and $f(x)$ is a convex function in C. This follows also from the relation

$$[C,f] = \bigcap_\alpha [C_\alpha, f_\alpha],$$

which shows in addition that $f(x)$ is closed in C.

42. Let C_α, $f_\alpha(x)$ be closed convex functions and Γ_α, $\varphi_\alpha(x)$ their conjugates. Assume that the set C in which $\sup_\alpha f_\alpha(x) < \infty$ is not empty and put $f(x) = \sup_\alpha f_\alpha(x)$ for $x \in C$. Denote by Γ, $\varphi(\xi)$ the conjugate of $C, f(x)$. Then

$$[\Gamma, \varphi] = \overline{\{\bigcup_\alpha [\Gamma_\alpha, \varphi_\alpha]\}} \quad ,$$

$$\{\bigcup_\alpha \Gamma_\alpha\} \subset \Gamma \subset \overline{\{\bigcup_\alpha \Gamma_\alpha\}} \quad ,$$

and, for $\xi \in \{\bigcup_\alpha \Gamma_\alpha\}$,

$$\varphi(\xi) = \inf \sum_{i=0}^{n} \lambda_i \, \varphi_{\alpha_i}(\xi^{\alpha_i}),$$

where

$$\xi^{\alpha_i} \in \Gamma_{\alpha_i}, \quad \lambda_i \geqq 0, \quad \sum_{i=0}^{n} \lambda_i = 1, \quad \sum_{i=0}^{n} \lambda_i \xi^{\alpha_i} = \xi \; ;$$

that is, for a given ξ the inf has to be taken over all representations of ξ as a centroid of $n + 1$ points taken from any $n+1$ of the sets Γ_α.

First observe that $[C,f] = \bigcap_\alpha [C_\alpha, f_\alpha]$. Thus the polar hyperplanes of the points of $[C,f]$ are on the one hand the non-vertical barriers of $[\Gamma, \varphi]$ and on the other hand the common non-vertical barriers of the sets $[\Gamma_\alpha, \varphi_\alpha]$. Hence, the sets $[\Gamma, \varphi]$ and $\bigcup_\alpha [\Gamma_\alpha, \varphi_\alpha]$ have the same supports bounded by non-vertical hyperplanes. From the above Lemma it now follows that $[\Gamma, \varphi]$ is the closure of the convex hull of $\bigcup_\alpha [\Gamma_\alpha, \varphi_\alpha]$. From Proposition 6, Chapter II, Section 2, we have that every point (ξ, ζ_c) of $\{\bigcup_\alpha [\Gamma_\alpha, \varphi_\alpha]\}$ is a centroid of at most $n + 2$ points $(\xi^{\alpha_i}, \zeta_{\alpha_i})$, $\xi^{\alpha_i} \in \Gamma_{\alpha_i}$, $\zeta_{\alpha_i} \geqq \varphi_{\alpha_i}(\xi^{\alpha_i})$, $i = 0,1,\ldots,n+1$. Thus

$$\xi = \sum_{i=0}^{n+1} \lambda_i \xi^{\alpha_i}, \quad \zeta_o \geqq \sum_{i=0}^{n+1} \lambda_i \varphi_{\alpha_i}(\xi^{\alpha_i}),$$

with $\lambda_i \geqq 0$, $\sum_{i=0}^{n+1} \lambda_i = 1$. This shows that $\{\bigcup_\alpha \Gamma_\alpha\} \subset \Gamma \subset \overline{\{\bigcup_\alpha \Gamma_\alpha\}}$ and that $\varphi(\xi)$ is an inf of the form in 42, but where ξ is the centroid of $n + 2$ points. That $n + 1$ points are sufficient is seen in the following way: The $n + 2$ points $(\xi^{\alpha_i}, \zeta_{\alpha_i})$ are the vertices of a (possibly degenerate) simplex in A^{n+1}. The vertical line through the point (ξ, ζ_o) intersects the simplex in a segment containing this point. That point (ξ, ζ_{min}) of this segment for which ζ is minimum is on some face of the simplex and, hence is a centroid of at most $n + 1$ of the points $(\xi^{\alpha_i}, \zeta_{\alpha_i})$. Since $\zeta_{min} \leqq \zeta_o$, in the expression for $\varphi(\xi)$ the original representation of ξ may be replaced by the new one as the centroid of $n + 1$ points. This completes the proof of 42.

43. With the notations of 42 assume that the set C is bounded and that $f(x) \geqslant a$ in C, where a is a constant. Then if $\varepsilon > 0$ is given, $n + 1$ functions $f_{\alpha_i}(x)$, $i = 0,1,\ldots,n$, may be chosen from the functions $f_\alpha(x)$ such that

$$\sum_{i=0}^{n} \lambda_i f_{\alpha_i}(x) > a - \varepsilon, \quad x \in \bigcap_i C_{\alpha_i},$$

for suitable $\lambda_i \geqq 0$, $\sum_{i=0}^{n} \lambda_i = 1$.

Since $[C,f]$ is closed, C is bounded, and f is bounded below, it follows that $f(x)$ has a minimum z_o. Then $z = z_o$ is a supporting hyperplane of $[C,f]$, and so $\varphi(0) = - z_o$. The assumption that C is bounded implies further that Γ is the whole ξ-space and that consequently, $\Gamma = \{\bigcup_\alpha \Gamma_\alpha\}$. In particular, the expression for $\varphi(\xi)$ in 42 may be applied for $\xi = 0$, giving

$$\varphi(0) = \inf \sum_{i=0}^{n} \lambda_i \varphi_{\alpha_i}(\xi^{\alpha_i}) = - z_o,$$

where $\xi^{\alpha_i} \in \Gamma_{\alpha_i}$, $\lambda_i \geqq 0$, $\sum_{i=0}^{n} \lambda_i = 1$, $\sum_{i=0}^{n} \lambda_i \xi^{\alpha_i} = 0$.

Hence there are $n + 1$ points $\xi^{\alpha_i} \in \Gamma_{\alpha_i}$ and $\lambda_i \geqq 0$, $\sum_{i=0}^{n} \lambda_i = 1$ such that

$$\sum_{i=0}^{n} \lambda_i \xi^{\alpha_i} = 0, \qquad \sum_{i=0}^{n} \lambda_i \varphi_{\alpha_i}(\xi^{\alpha_i}) < - z_o + \varepsilon .$$

For the corresponding functions $f_{\alpha_i}(x)$, $x \in \bigcap_i C_{\alpha_i}$,

Proposition 38 gives

$$\sum_{i=0}^{n} \lambda_i f_{\alpha_i}(x) \geqq x' \sum_{i=0}^{n} \lambda_i \xi^{\alpha_i} - \sum_{i=0}^{n} \lambda_i \varphi_{\alpha_i}(\xi^{\alpha_i}) > z_o - \varepsilon \geqq a - \varepsilon,$$

which is the desired result.

If closed convex functions C_α, $f_\alpha(x)$ are given, the question arises under which conditions $f(x) = \sup_\alpha f_\alpha(x)$ does exist, i.e. is finite for some x. This is the case if and only if the sets $[C_\alpha, f_\alpha]$ have a common point, which in turn is the case if and only if the sets $[\Gamma_\alpha, \varphi_\alpha]$ have a common non-vertical barrier. There will be such a common non-vertical barrier if $\{\bigcup_\alpha [\Gamma_\alpha, \varphi_\alpha]\}$ is not the whole space A^{n+1} while $\{\bigcup_\alpha \Gamma_\alpha\}$ is the whole A^n, i.e., if $\{\bigcup_\alpha \Gamma_\alpha\}$ has no barrier. The latter part of the condition is satisfied if the asymptotic cones $A_o(C_\alpha)$ of the sets C_α have no common ray, for the existence of a common barrier to the sets Γ_α implies the existence of a common ray of the cones $A_o(C_\alpha)$. (Compare the remark following 38.) To ensure that $\{\bigcup_\alpha [\Gamma_\alpha, \varphi_\alpha]\}$ is not the whole space it is sufficient to assume that there is a fixed hyperplane $z = x'\xi^o - \zeta_o$ such that any $n + 1$ of the sets $[C_\alpha, f_\alpha]$ have a common point below this hyperplane. Then the point (ξ^o, ζ_o) cannot belong to $\{\bigcup_\alpha [\Gamma_\alpha, \varphi_\alpha]\}$. If it did, it would be the centroid of $n + 1$ points $(\xi^{\alpha_i}, \zeta_{\alpha_i})$ taken from certain $n + 1$ sets $[\Gamma_{\alpha_i}, \varphi_{\alpha_i}]$, $i = 0,1,\ldots,n$. In other words there would be numbers $\lambda_i \geqq 0$, $\sum_{i=0}^{n} \lambda_i = 1$, such that

$$\xi^o = \sum_{i=0}^{n} \lambda_i \xi^{\alpha_i}, \quad \zeta_o = \sum_{i=0}^{n} \lambda_i \zeta_{\alpha_i} > \sum_{i=0}^{n} \lambda_i \varphi_{\alpha_i}(\xi^{\alpha_i}).$$

From 42 applied to the $n + 1$ functions $C_{\alpha_i}, f_{\alpha_i}(x)$, $i = 0,1,\ldots,n$, it would now follow that $z = x'\xi^o - \zeta^o$ is a barrier to $\bigcap_i [C_{\alpha_i}, f_{\alpha_i}]$ which contradicts the assumption. Thus the following theorem is proved:

> 44. Let f_α in C_α be closed convex functions. Assume that the asymptotic cones of the sets C_α have no common ray and that there is a fixed non-vertical hyperplane below which any $n + 1$ of the sets $[C_\alpha, f_\alpha]$ have at least one point in common. Then all the sets $[C_\alpha, f_\alpha]$ have

a common point, ~~and~~ in other words $\sup_{\alpha} f_{\alpha}(x)$ is finite for at least one x.

In the special case where all f_{α} are identically zero (and hence the sets C_{α} are closed) the existence of a hyperplane with the required property is obvious (any hyperplane $z = z_o > 0$ will suffice) and 44 becomes Helly's Theorem:

> 45. Let C_{α} be closed convex sets in A^n. Assume that the asymptotic cones of the sets C_{α} have no common ray and that any $n + 1$ of the sets have a common point. Then all sets have a common point.

Obviously the assumption that the asymptotic cones of the C_{α} have no common ray may be replaced by the usual one: There are sets among the C_{α} which have a non-empty, bounded intersection.

Finally some special cases and applications of conjugate convex functions will be mentioned.

Let $f(x)$ be identically zero in a closed convex set C. The conjugate function

$$\varphi(\xi) = \sup_{x \in C} x' \xi = h_C(\xi)$$

is the support function of C, and Γ is the cone of those directions ξ in which C is bounded. This implies that every support function is closed. Conversely, let $\varphi(\xi)$ be defined, positively homogeneous, convex, and closed in a convex cone Γ. Then $[\Gamma, \varphi]$ is a cone with the origin as vertex, and hence all non-vertical supporting hyperplanes to $[\Gamma, \varphi]$ pass through the origin. This means that the conjugate $f(x)$ of $\varphi(\xi)$ is identically zero in some convex set C (which must be closed since $f(x)$ is closed in C). Hence

> 46. A function $\varphi(\xi)$ defined in a convex cone Γ is the support function of some point set if and only if it is positively homogeneous, convex, and closed in Γ.

In the particular case $f_\alpha(x) \equiv 0$ the $\varphi(\xi)$ of Proposition 42 is the support function of the intersection $C = \bigcap_\alpha C_\alpha$ expressed in terms of the support functions $\varphi_\alpha(\xi)$ of the sets C_α. Because of the homogeneity of the functions φ_α the expression may here be written

$$h_C(\xi) = \inf \sum_{i=0}^{n} h_{C\alpha_i}(\xi^{\alpha_i})$$

where $\xi^{\alpha_i} \in \Gamma_{\alpha_i}$, $\sum_{i=0}^{n} \xi^{\alpha_i} = \xi$; that is, the inf has to be taken over all representations of ξ as a sum of $n + 1$ points taken from any $n + 1$ of the sets Γ_α.

Consider again an arbitrary convex function $f(x)$ closed in a convex set C. Denote its conjugate by $\Gamma, \varphi(\xi)$. The supporting hyperplane $z = x'\xi^0 - \varphi(\xi^0)$ to $[C,f]$ with normal direction $(\xi^0,-1)$, $\xi^0 \in \Gamma$, intersects $[C,f]$ in a (possibly empty) closed convex set. Let $C(\xi^0)$ denote the projection of this set on the hyperplane $z = 0$. Thus, x is in $C(\xi^0)$ if and only if $(x,f(x))$ is in the hyperplane $z = x'\xi^0 - \varphi(\xi^0)$; that is, if

$$f(x) = x'\xi^0 - \varphi(\xi^0).$$

Interpreted dually, x is in $C(\xi^0)$ if and only if there is a supporting hyperplane to $[\Gamma, \varphi]$ having normal direction $(x,-1)$ and passing through $(\xi^{o}, \varphi(\xi^{o}))$. In particular, $C(\xi^{o})$ is empty if and only if there is no non-vertical supporting hyperplane to $[\Gamma, \varphi]$ through $(\xi^{o}, \varphi(\xi^{o}))$. This is the case only if ξ^{o} is a relative boundary point of Γ at which the directional derivative $\varphi'(\xi^{o};\eta)$ is infinite. Dually, to a given $x^{o} \in C$ there corresponds a subset of Γ with the analogous properties. This set will be denoted by $\Gamma(x^{o})$. Obviously, $x^{o} \in C(\xi^{o})$ implies $\xi^{o} \in \Gamma(x^{o})$, and conversely.

The directional derivative $f'(x^{o};y)$ as a function of y is convex but not necessarily closed in its domain $P_{x^{o}}(C) - x^{o}$. But if it is not closed it may be made so by the unessential changes described in connection with Proposition 25. Then we may speak of its conjugate function, which is identically zero since $f'(x^{o};y)$ is positively homogeneous. To find the domain of the conjugate, consider first the convex function $f(x^{o}) + f'(x^{o};x-x^{o})$, $x \in P_{x^{o}}(C)$, or, if necessary, the function obtained by closing it. From Proposition 30 it follows that the conjugate of this function is the linear function

$$\varphi(\xi) = x^{o\prime}\xi - f(x^{o}), \qquad \xi \in \Gamma(x^{o}).$$

Application of Proposition 40 now shows that the conjugate of $f'(x^{o};y)$ (or of the function obtained by closing it) is defined in $\Gamma(x^{o})$.

Let $x^{o} \in C$ be such that $f'(x^{o};y)$ is finite. Denote by $\ell(f', x^{o})$ the linearity of the cone $M = [P_{x^{o}}(C) - x^{o}, f'(x;y)]$; that is, the maximum number of linearly independent directions in which $f(x)$ is differentiable at x^{o}. Then

$$\ell(f', x^o) + d(\Gamma(x^o)) = n,$$

where $d(\Gamma(x^o))$ is the dimension of $\Gamma(x^o)$. To prove this, observe that if the cone M is laid off from the point $(0,1)$, its normal cone M^* intersects the hyperplane $z = 0$ in $\Gamma(x^o)$. Hence $d(M^*) = 1 + d(\Gamma(x^o))$ and, by the corollary to Theorem 5, Chapter I, Section 4, $\ell(M) + d(M^*) = n + 1$.

Suppose now that C is open and that $f(x)$ is differentiable in C. Then for every $x^o \in C$, $\Gamma(x^o)$ consists of one point ξ^o whose coordinates are the partial derivatives of f at the point x^o. Hence, there is a one-valued mapping $x \longrightarrow \xi$ of C onto Γ determined by

$$(*) \qquad \xi_i = \frac{\partial f}{\partial x_i}, \quad i = 1, 2, \ldots, n.$$

If, moreover, $\varphi(\xi)$ satisfies the same conditions as $f(x)$, i.e. if $f(x)$ is strictly convex, the mapping is one-to-one and, because of the involutory character of the conjugate relation, the inverse mapping must be given by $x_i = \dfrac{\partial \varphi}{\partial \xi_i}$.

This leads to the following procedure for the computation of the conjugate of a smooth convex function: Let $f(x)$ be strictly convex, closed, and differentiable in an open convex set C. Then the domain Γ of the conjugate function φ is determined as the image of C by the mapping $(*)$. By solving $(*)$, the x_i are found as functions of the ξ_i and substituted in

$$\varphi(\xi) = x'\xi - f(x)$$

to give φ in terms of ξ.

6. A GENERALIZED PROGRAMMING PROBLEM

Let $f(x)$ in C be a closed convex function and $g(x)$ in D a closed concave function. Consider the following extremum problem:

> PROBLEM I: To find a point x^o in $C \cap D$ such that $g(x) - f(x)$ as a function in $C \cap D$ has a maximum at x^o.

If $g(x) - f(x) \geqq 0$ in $C \cap D$ this problem, stated geometrically, is to find the maximum vertical chord of the convex set $[C,f] \cap [D,g]$ in A^{n+1}. If $f(x) \equiv 0$ in C, Problem I reduces to a programming problem, viz. to maximize $g(x)$ under the condition $x \in C$.

Denoting by $\varphi(\xi)$ in Γ and $\psi(\xi)$ in Δ the conjugates of $C, f(x)$ and $D, g(x)$ respectively, consider the similar problem:

> PROBLEM II: To find a point ξ^o in $\Gamma \cap \Delta$ such that $\varphi(\xi) - \psi(\xi)$ as a function in $\Gamma \cap \Delta$ has a minimum at ξ^o.

If $\varphi(\xi) - \psi(\xi) \geqq 0$ in $\Gamma \cap \Delta$ this stated geometrically is to find the minimum vertical segment joining the sets $[\Gamma, \varphi]$ and $[\Delta, \psi]$ in A^{n+1}.

These two problems are connected by

> 47. Let the function $f(x)$ in C be convex and closed, $\varphi(\xi)$ in Γ its conjugate. Let further $g(x)$ in D be concave and closed, $\psi(\xi)$ in Δ its

conjugate. If the sets $C \cap D$ and $\Gamma \cap \Delta$ are non-empty, then $g(x) - f(x)$ is bounded above, $\varphi(\xi) - \psi(\xi)$ is bounded below, and

$$\sup_{x \in C \cap D} (g(x)-f(x)) = \inf_{\xi \in \Gamma \cap \Delta} (\varphi(\xi)-\psi(\xi)).$$

We shall give two proofs. The first and more formal proof is based on Proposition 41 applied to the functions f and $-g$ (instead of f_1 and f_2). Let $\chi(\xi)$ be the conjugate of $f(x) + (-g(x))$ in $C \cap D$. From 41 and 39 it follows that $\chi(\xi)$ is defined in a set containing $\Gamma + (-\Delta)$. Since Γ and Δ have points in common, $\Gamma - \Delta$ contains the origin. Hence, $\chi(0)$ is defined and, again by 41 and 39,

$$\chi(0) = \inf_{\substack{\xi^1 \in \Gamma,\, \xi^2 \in -\Delta \\ \xi^1 + \xi^2 = 0}} (\varphi(\xi^1) - \psi(-\xi^2))$$

$$= \inf_{\xi \in \Gamma \cap \Delta} (\varphi(\xi) - \psi(\xi)).$$

On the other hand, the very definition of the conjugate of $f(x) - g(x)$, taken for $\xi = 0$, yields

$$\chi(0) = \sup_{x \in C \cap D} (g(x) - f(x)).$$

This proves the statement.

A second proof, more geometrical and more elementary, is based on the interpretation given in Proposition 36 of the conjugate of a convex function. It does not give 47 in its full generality but, on the other hand, it allows an intuitive discussion of the existence of the extremum values in question.

If $\xi \in \Gamma \cap \Delta$, there exist supports $z \geqq x'\xi - \varphi(\xi)$ and $z \leqq x'\xi - \psi(\xi)$ of [C,f] and [D,g] respectively. Since $-\varphi(\xi)$ and $-\psi(\xi)$ are the z-intercepts of the supporting hyperplanes, $\varphi(\xi) - \psi(\xi)$ is the vertical width of the strip bounded by these hyperplanes, taken with a sign in the usual way. Now, for $\xi \in \Gamma \cap \Delta$,

$$f(x) \geqq x'\xi - \varphi(\xi), \quad x \in C,$$

$$g(x) \leqq x'\xi - \psi(\xi), \quad x \in D,$$

which gives,

$$g(x) - f(x) \leqq \varphi(\xi) - \psi(\xi), \quad x \in C \cap D.$$

(If $g(x) - f(x) \geqq 0$ in $C \cap D$, this simply means that $[C,f] \cap [D,g]$ is contained in the strip.) Hence, the left side is bounded above, the right side is bounded below, and

$$(1) \qquad \sup_{x \in C \cap D} \big(g(x) - f(x)\big) \leqq \inf_{\xi \in \Gamma \cap \Delta} (\varphi(\xi) - \psi(\xi)).$$

Denote by μ the value of the left side of inequality (1). Then

$$g(x) \leqq f(x) + \mu, \quad x \in C \cap D.$$

Thus, the only points (x,z) which are common to the sets [D,g] and $[C, f+\mu]$, if any, are those for which $z = g(x) = f(x) + \mu$. These points are obviously relative boundary points of both sets. Therefore, the Separation Theorem 28, Chapter II, Section 6 may be applied, and consequently, there is in the smallest flat S containing both sets a hyperplane h which separates [D,g] and $[C, f+\mu]$

in the sense of that theorem. The normals to S through the points of h form a hyperplane H of A^{n+1} with the following properties: H does not contain both sets, $[D,g]$ is contained in one of the closed half-spaces bounded by H, and $[C,f+\mu]$ is contained in the other closed half-space bounded by H.

Suppose first that there is a non-vertical separating hyperplane h in S. Then H too is non-vertical and its equation is of the form $z = x' \xi^0 - \zeta_0$. Now, the distance of the two sets being zero, H is a supporting hyperplane to both $[D,g]$ and $[C,f+\mu]$, and thus, by Propositions 36 and 40,

$$\zeta_0 = \psi(\xi^0) = \varphi(\xi^0) - \mu .$$

Together with (1) this shows that $\min(\varphi(\xi) - \psi(\xi))$ exists and that

$$(2) \qquad \sup_{x \in C \cap D} (f(x)-g(x)) = \min_{\xi \in \Gamma \cap \Delta} (\varphi(\xi)-\psi(\xi)).$$

Suppose now that there is no non-vertical hyperplane in S which separates $[D,g]$ and $[C,f+\mu]$. Let h be a vertical separating hyperplane and denote by h_0 its intersection with $z = 0$. By projection parallel to the z-axis $[D,g]$, $[C,f+\mu]$, and h are projected into D, C, and h_0 respectively, and h_0 separates C and D. This shows that the present case occurs only if C and D have no points in common which are relative interior to both sets. Hence, we may conclude that if $C \cap D$ contains points relative interior to both sets, the minimum problem has a solution and (2) is valid.

The preceding, together with the dual argument, leads to the following theorem:

48. With the notations of 47 suppose that C and D have points in common which are relative interior to both sets and that Γ and Δ satisfy the same condition. Then $g(x)-f(x)$ has a maximum in $C \cap D$, $\varphi(\xi)-\psi(\xi)$ has a minimum in $\Gamma \cap \Delta$, and

$$\max_{x \in C \cap D}(g(x)-f(x)) = \min_{\xi \in \Gamma \cap \Delta}(\varphi(\xi)-\psi(\xi)).$$

It may be mentioned without proof that if the directional derivatives $f'(x;y)$ and $g'(x;y)$ are uniformly bounded for $x \in C \cap D$ and all y for which they are defined, there is a non-vertical hyperplane separating $[D,g]$ and $[C,f+\mu]$ even if C and D have no points in common which are relative interior to both sets. Hence, if this condition and the corresponding condition for φ and ψ are satisfied, the conclusions of 48 are valid.

A continuous function whose domain is closed and can be divided into finitely many subsets in each of which the function is linear, will be called a <u>piecewise linear function</u>. Observe that if such a function is bounded above (below), it has a maximum (minimum) since it cannot approach its least upper (greatest lower) bound asymptotically. Consequently, if the functions f, g and, thus, φ, ψ are piecewise linear, and if the assumptions of Proposition 47 are made, the conclusions of 48 hold.

From the definitions of the conjugate functions it is clear that Proposition 47 is equivalent to either of the two following statements:

49. Under the assumptions of 47

$$\inf_{\xi \in \Gamma\cap\Delta} \sup_{x \in C} (x'\xi - f(x) - \psi(\xi))$$

$$= \sup_{x \in C \cap D} \inf_{\xi\in\Delta} (x'\xi - f(x) - \psi(\xi)),$$

and

$$\sup_{x \in C \cap D} \inf_{\xi\in\Gamma} (\varphi(\xi) + g(x) - x'\xi)$$

$$= \inf_{\xi \in \Gamma\cap\Delta} \sup_{x \in D} (\varphi(\xi) + g(x) - x'\xi).$$

If Problems I and II have solutions, as is the case under the assumptions of 48 or if the functions involved are piecewise linear, the outer inf and sup in the immediately preceding equations may be replaced by min and max respectively.

The pair of Problems I and II is equivalent to each of the two following saddle value problems:

PROBLEM III: Let $f(x)$ be convex and closed in C and let $\psi(\xi)$ be concave and closed in Δ. Put

$$F(x, \xi) = x'\xi - f(x) - \psi(\xi).$$

To find an $x^{\circ} \in C$ and a $\xi^{\circ} \in \Delta$ such that

$$F(x, \xi^{\circ}) \leqq F(x^{\circ}, \xi^{\circ}) \leqq F(x^{\circ}, \xi)$$

for all $x \in C$ and all $\xi \in \Delta$.

PROBLEM III': Let $g(x)$ be concave and closed in D and let $\varphi(\xi)$ be convex and closed in Γ. Put

$$\varphi(\xi, x) = \varphi(\xi) + g(x) - x'\xi .$$

To find an $x^{o} \in D$ and a $\xi^{o} \in \Gamma$ such that

$$\varphi(\xi, x^{o}) \geqq \varphi(\xi^{o}, x^{o}) \geqq \varphi(\xi^{o}, x)$$

for all $\xi \in \Gamma$ and all $x \in D$.

Consider Problem III. Denote the conjugates of C, f and Δ, ψ by Γ, φ and D, g respectively. From the definitions of the conjugate functions we have

(3) $$F(x, \xi) \leqq \varphi(\xi) - \psi(\xi)$$

for $x \in C$, $\xi \in \Gamma \cap \Delta$, and

(4) $$F(x, \xi) \geqq g(x) - f(x)$$

for $x \in C \cap D$, $\xi \in \Delta$.

Suppose Problem I has a solution $x^{o} \in C \cap D$ and Problem II has a solution $\xi^{o} \in \Gamma \cap \Delta$. Put

$$g(x^{o}) - f(x^{o}) = \varphi(\xi^{o}) - \psi(\xi^{o}) = \mu .$$

Then (3) and (4) give

$$F(x, \xi^{o}) \leqq \mu, \quad x \in C,$$

$$F(x^{o}, \xi) \geqq \mu, \quad \xi \in \Delta .$$

Hence, $F(x^{o}, \xi^{o}) = \mu$ and

$$F(x, \xi^0) \leqq F(x^0, \xi^0) \leqq F(x^0, \xi)$$

for $x \in C$, $\xi \in \Delta$.

Suppose now Problem III has a solution $x^0 \in C$, $\xi^0 \in \Delta$. From $F(x, \xi^0) \leqq F(x^0, \xi^0)$ for $x \in C$ it follows that $x'\xi^0 - f(x)$ attains a maximum at x^0. This implies that $\xi^0 \in \Gamma$ and that the maximum value is $\varphi(\xi^0)$. Hence,

$$F(x^0, \xi^0) = \varphi(\xi^0) - \psi(\xi^0).$$

Analogously, $F(x^0, \xi) \geqq F(x^0, \xi^0)$ for $\xi \in \Delta$ yields $x^0 \in D$ and

$$F(x^0, \xi^0) = g(x^0) - f(x^0).$$

Now, by (3) and (4)

$$g(x)-f(x) \leqq g(x^0)-f(x^0) = \varphi(\xi^0)-\psi(\xi^0) \leqq \varphi(\xi)-\psi(\xi)$$

for $x \in C \cap D$, $\xi \in \Gamma \cap \Delta$, which shows that x^0 and ξ^0 are solutions of I and II respectively.

By interchanging the roles of f and φ and of ψ and g it is immediately seen that Problem III' also is equivalent with the pair of Problems I and II.

The main theorem of the theory of the zero-sum two-person game is a particular case of 49.

Let A be a given m by n matrix. Let C denote the set of all points x for which $x \geqq 0$, $\sum_{j=1}^{n} x_j = 1$ and define $f(x) = 0$ in C. Let Δ be the set of all points $\xi = A'u$, $u \geqq 0$, $\sum_{i=1}^{m} u_i = 1$, and define $\psi(\xi) = 0$ in Δ. Then

$$x'\xi - f(x) - \psi(\xi) = u'Ax$$

for $x \geqq 0$, $\sum_{j=1}^{n} x_j = 1$, $u \geqq 0$, $\sum_{i=1}^{m} u_i = 1$, and both Γ and D are the whole n-space since C and Δ are bounded. Hence, 49 yields

$$\min_{\xi \in \Delta} \max_{x \in C} u'Ax = \max_{x \in C} \min_{\xi \in \Delta} u'Ax.$$

The existence of the extreme values is obvious in this case.

Let A be an m by n matrix, b an m-dimensional vector, and c an n-dimensional vector. A pair of basic, mutually dual, linear programming problems is:

1) to find the maximum of $c'x$ subject to the conditions $x \geqq 0$, $Ax \leqq b$;

2) to find the minimum of $b'u$ subject to the conditions $u \geqq 0$, $A'u \geqq c$.

If $Ax \leqq b$ for some $x \geqq 0$ and $A'u \geqq c$ for some $u \geqq 0$, both problems have solutions and $\max c'x = \min b'u$.

To show that this is a particular case of the preceding results suppose first that $m = n$ and that A is non-singular. Define C to be the set of all x satisfying $Ax \leqq b$ and put $f(x) = 0$ in C. Define D to be the positive orthant $x \geqq 0$ and put $g(x) = c'x$ in D. Then Problem I reduces to the linear programming Problem 1. To determine the conjugate functions $\Gamma, \varphi(\xi)$ and $\Delta, \psi(\xi)$ introduce a parameter vector u by $u = A'^{-1}\xi$. Then

$$\varphi(\xi) = \sup_{Ax \leqq b} \xi' x = \sup_{Ax \leqq b} u'Ax.$$

Since Ax assumes any value less than or equal to b as x varies in C, $u'Ax$ is bounded above if and only if $u \geqq 0$. Thus, Γ is the set of all $\xi = A'u$, $u \geqq 0$ and $\varphi(\xi) = u'b$.

Further

$$\psi(\xi) = \inf_{x \geqq 0} (\xi - c)'x$$

where the right side is finite (then having the value zero) if and only if $\xi \geqq c$. Thus, Δ is the set of all $\xi = A'u \geqq c$ and $\psi(\xi) = 0$ in Δ. This shows that Problem II reduces to the linear Problem 2.

The general case where A is arbitrary rectangular may be reduced to the case just considered in the following way. Denote by E_i the i by i identity matrix. Instead of A consider the non-singular $m + n$ by $m + n$ matrix $\begin{pmatrix} A & E_m \\ -E_n & 0^m \end{pmatrix}$. Complete the vectors b, c, x, ξ, u to $(m+n)$-dimensional vectors $\binom{b}{0}, \binom{c}{0}, \binom{x}{y}, \binom{\xi}{\eta}, \binom{u}{v}$.

Then the two linear problems take the forms: 1) to maximize $c'x$ subject to the conditions

$$\begin{pmatrix} A & E_m \\ -E_n & 0 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} \leqq \begin{pmatrix} b \\ 0 \end{pmatrix}, \begin{pmatrix} x \\ y \end{pmatrix} \geqq 0,$$

which can be written $Ax + y \leqq b$, $x \geqq 0$, $y \geqq 0$; 2) to minimize $b'u$ subject to the conditions

$$\begin{pmatrix} A' & -E_n \\ E_m & 0 \end{pmatrix} \begin{pmatrix} u \\ v \end{pmatrix} \geqq \begin{pmatrix} c \\ 0 \end{pmatrix}, \quad \begin{pmatrix} u \\ v \end{pmatrix} \geqq 0,$$

or $A'u - v \geqq c$, $u \geqq 0$, $v \geqq 0$. Since $c'x$ and $b'u$ do not depend on y and v respectively, these problems are equivalent with the original Problems 1 and 2. For, if x^0, y^0 and u^0, v^0 are solutions of the new problems, x^0 and u^0 solve

1 and 2; and if x^o and u^o are solutions of the latter problems, x^o, y^o and u^o, v^o solve the new ones for arbitrary y^o and v^o satisfying $0 \leqq y^o \leqq b - Ax^o$, $0 \leqq v^o \leqq A'u^o - c$.

Since the functions occurring here are piecewise linear, the assumptions of 47 guarantee the existence of the two extreme values. These assumptions take here the following form: there exist $x \geqq 0$, $y \geqq 0$ satisfying $Ax + y \leqq b$ and there exist $u \geqq 0$, $v \geqq 0$ satisfying $A'u - v \geqq c$. Obviously, it is sufficient to require the existence of at least one $x \geqq 0$ such that $Ax \leqq b$ and of at least one $u \geqq 0$ such that $A'u \geqq c$, for this x and this u together with $y = 0$ and $v = 0$ satisfy the stated conditions. Herewith the statement concerning the linear programming Problems 1 and 2 is completely proved.

7. THE LEVEL SETS OF A CONVEX FUNCTION

Consider an arbitrary real function $\varphi(x)$ defined over a set D in A^n. For a given real number τ the subset L_τ of D consisting of those points x of D for which $\varphi(x) \leqq \tau$ will be called the <u>level set</u> of $\varphi(x)$ for the level τ. Clearly, L_τ is empty for $\tau < \inf \varphi$, and $L_\tau = D$ for $\tau > \sup \varphi$. Therefore τ will be restricted to the smallest interval Ω containing the whole range of φ. This interval may be finite or infinite, open, half open, or closed. To exclude the trivial case when $\varphi(x)$ is a constant, it will be assumed that Ω has interior points. In the following all numbers $\tau, \tau_o, \ldots$ are supposed to belong to Ω. On observing that $\varphi(x) \leqq \tau_o$ is equivalent to $\varphi(x) \leqq \tau$ for all $\tau > \tau_o$, it is immediately seen that the family of level sets L_τ has the following properties:

I. $\bigcup_{\tau \in \Omega} L_\tau = D.$

II. $L_{\tau_1} \subset L_{\tau_2}$ if $\tau_1 < \tau_2$.

III. $\bigcap_{\tau > \tau_0} L_\tau = L_{\tau_0}$, and $\bigcap_{\tau \in \Omega} L_\tau$ is empty if Ω is open to the left.

Conversely, given a set D in A^n and a family of subsets L_τ indexed by the real numbers of some interval and satisfying Conditions I-III, there is a function $\varphi(x)$ defined over D for which the sets L_τ are the level sets. To exhibit such a function define $\varphi(x) = \inf_{L_\tau \supset x} \tau$. Then, $\varphi(x)$ is finite for all $x \in D$ because for every $x \in D$ I ensures that some L_τ contains x while III ensures that if Ω is unbounded below, there is some $L_{\tau'}$ which does not contain x. The level set corresponding to τ_0 of this function consists of all x such that $\inf_{L_\tau \supset x} \tau \leqq \tau_0$. Thus, x is in this level set if and only if, to every $\varepsilon > 0$, there is a $\tau < \tau_0 + \varepsilon$ such that $x \in L_\tau$. Because of II this means $x \in L_\tau$ for all $\tau > \tau_0$ and hence, by III, $x \in L_{\tau_0}$. A further consequence of III is that $\varphi(x) = \min_{L_\tau \supset x} \tau$. This equation establishes a one-to-one correspondence between the functions $\varphi(x)$ defined over D and the indexed families of subsets of D satisfying I - III.

It is well known that a function $\varphi(x)$ with level sets L_τ is lower semicontinuous if and only if for all $\tau \in \Omega$:

IV. L_τ is closed relative to D.

The condition for upper semicontinuity:

statement holds:

> 50. The level sets of a function $\varphi(x)$, $x \in D$, are convex if and only if $\varphi(x)$ is quasi-convex.

To prove the necessity let x and y be arbitrary points of D and define $\tau = \max(\varphi(x), \varphi(y))$. Then $x \in L_\tau$, $y \in L_\tau$ and, since L_τ is convex, $(1-\theta)x + \theta y \in L_\tau$. Hence $\varphi((1-\theta)x + \theta y) \leqq \tau$. To prove the sufficiency let L_τ be an arbitrary level set of $\varphi(x)$. If $x \in L_\tau$, $y \in L_\tau$, it follows that $\varphi(x) \leqq \tau$, $\varphi(y) \leqq \tau$. Because of the quasi-convexity of $\varphi(x)$, $\varphi((1-\theta)x + \theta y) \leqq \tau$, that is $(1-\theta)x + \theta y \in L_\tau$.

A family of subsets L_τ of D satisfying I - V, that is the family of level sets of a lower semicontinuous, quasi-convex function $\varphi(x)$ defined over D with range Ω, is briefly called a <u>quasi-convex family</u>. Suppose now L_τ is transformable into the family of level sets K_t, $t \in W$, of a convex function $f(x) = F(\varphi(x))$, briefly called a <u>convex family</u>. Then both $f(x)$ and $\varphi(x)$ are continuous. The interval W, the image of Ω by $t = F(\tau)$, is open to the right since a convex function in an open domain has no maximum. Hence Ω must have the same property. This implies in particular that all sets $L_\tau = K_t$ are proper subsets of D. If W is closed to the left, Ω is closed to the left, and conversely, and we have $a = F(\alpha)$. Thus, with the notations $\alpha = \inf \varphi(x)$, $\beta = \sup \varphi(x)$, $a = \inf f(x)$, and $b = \sup f(x)$ where $-\infty \leqq \alpha < \beta \leqq \infty$ and $-\infty \leqq a < b \leqq \infty$, W is $a \underset{(=)}{\leqq} t < b$, and Ω is $\alpha \underset{(=)}{\leqq} \tau < \beta$, where the equalities can only occur simultaneously (and, of course, only if a and α are finite). The open intervals $a < t < b$ and $\alpha < \tau < \beta$ are denoted by W_0 and Ω_0 respectively.

A rather obvious necessary condition which a quasi-

$\bigcup_{\tau < \tau_0} L_\tau$ is open relative to D, will not be used explicitly.

Let $t = F(\tau)$ be a strictly increasing continuous function defined for $\tau \in \Omega$. Denote by W the range of $F(\tau)$, $\tau \in \Omega$, and let $\tau = \Phi(t)$, $t \in W$, be the inverse of F. Then the family of sets $K_t = L_{\Phi(t)}$, $t \in W$, is the family of level sets of the function $f(x) = F(\varphi(x))$ and satisfies Conditions I - IV if L_τ, $\tau \in \Omega$, does. For the sake of brevity two families like L_τ and K_t obtained from each other by a strictly increasing and continuous index transformation $t = F(\tau)$ will be said to be <u>transformable</u> into each other.

The problem to be discussed in the following may now be formulated:

> Under what conditions is a family of sets L_τ satisfying I - IV transformable into the family of level sets of a convex function. To avoid inessential difficulties the domain D will henceforth be assumed to be convex and open.

An obvious necessary condition is:

V. L_τ is convex for $\tau \in \Omega$.

However, this condition is not sufficient. Call a function $\varphi(x)$ defined over D <u>quasi-convex</u> if

$$\varphi((1-\theta)x+\theta y) \leqq \max\ (\varphi(x), \varphi(y))$$

for $0 \leqq \theta \leqq 1$ and all x and y in D. The following convex family L_τ must satisfy in order that it be transformable into a convex family is

$$\overline{\bigcup_{\tau < \tau_o} L_\tau} = L_{\tau_o} \quad \text{for} \quad \tau_o \in \Omega_o.$$

This expresses the fact that a convex function cannot assume a constant value except possibly its minimum on a relatively open subset of its domain. This condition will not, however, be used explicitly. The further discussion of the problem stated above will be based on the following characterization of a convex family:

51. A quasi-convex family K_t, $t \in W$, is a convex family if and only if

$$(*) \qquad (1-\theta)K_{t_o} + \theta K_{t_1} \subset K_{t_\theta}$$

where $0 \leqq \theta \leqq 1$, $t_o \in W$, $t_1 \in W$, $t_\theta = (1-\theta)t_o + \theta t_1$.

To prove this, suppose K_t are the level sets of the convex function $f(x)$, $x \in D$. Let $x^\theta = (1-\theta)x^o + \theta x^1$, where $x^o \in K_{t_o}$, $x^1 \in K_{t_1}$, be an arbitrary point of $(1-\theta)K_{t_o} + \theta K_{t_1}$. Then

$$f(x^\theta) \leqq (1-\theta)f(x^o) + \theta f(x^1) \leqq (1-\theta)t_o + \theta t_1 = t_\theta.$$

Hence $x^\theta \in K_{t_\theta}$. Conversely, let $(*)$ be satisfied and define $f(x) = \min_{K_t \supset x} t$. As mentioned above, this function has the level sets K_t. Let x^o and x^1 be arbitrary points of D and put $f(x^o) = t_o$, $f(x^1) = t_1$, and $x^\theta = (1-\theta)x^o + \theta x^1$. Then $x^o \in K_{t_o}$, $x^1 \in K_{t_1}$, and $x^\theta \in K_{t_\theta}$ because of (*). Hence

$$f(x^\theta) = \min_{K_t \supset x^\theta} t \leqq t_\theta = (1-\theta)f(x^o) + \theta f(x^1).$$

This proves the statement.

Let M be a point set. As in Chapter II, Section 5, the cone with vertex at the origin consisting of all directions in which M is bounded will be denoted by $B(M)$. The following rather obvious properties of cones B will be used: For any two point sets M, N

$$\begin{aligned} &B(M) \supset B(N) && \text{if} \quad M \subset N, \\ &B(\lambda M) = B(M) && \text{for} \quad \lambda > 0, \\ &B(M+N) = B(M) \cap B(N). \end{aligned}$$

For a quasi-convex family L_τ, $\tau \in \Omega$, transformable into a convex family:

VI. All sets L_τ, $\tau \in \Omega_o$, are bounded in the same directions, that is $B = B(L_\tau)$, $\tau \in \Omega_o$, is independent of τ. If L_α exists, $B \subset B(L_\alpha) \subset \overline{B}$.

Since this statement is invariant under index transformations, it suffices to prove it for a family K_t satisfying (*). Let $t \in W_o$, $t_1 \in W_o$, $t_1 > t$, be given and choose $t_o < t$ in W. With $\theta = (t-t_o)/(t_1-t_o)$ the relation (*) yields

$$(1-\theta)K_{t_o} + \theta K_{t_1} \subset K_t.$$

Hence, because $K_{t_o} \subset K_t \subset K_{t_1}$,

$$B(K_{t_1}) \subset B(K_t) \subset [B(K_{t_o}) \cap B(K_{t_1})] = B(K_{t_1}).$$

Thus $B(K_t) = B(K_{t_1})$ which proves the statement.

If L_α exists, $B \subset B(L_\alpha)$ because $L_\alpha \subset L_\tau$, $\tau \in \Omega_o$. It only remains to prove that $B(L_\alpha) \subset \overline{B}$ when L_α exists. Let $\xi \neq 0$ be in $B(L_\alpha)$ and let H be the supporting hyperplane of L_α with normal direction ξ. In L_α there is some point p whose distance from H is less that a given $\varepsilon > 0$. Denote by H_ε that hyperplane parallel to H at distance ε which is separated from p by H. In H_ε consider the $(n-1)$-dimensional closed (solid) unit sphere $\overline{U}$ whose center is the orthogonal projection of p on H_ε. The compact set $\overline{U}$ having a positive distance from L_α, there is by III some $t > a$ such that K_t and $\overline{U}$ are disjoint. By the Separation Theorem 28, Section 6, Chapter II, there is a hyperplane H' separating K_t and $\overline{U}$. The normal vector ξ' of H' which is directed towards $\overline{U}$ belongs to B because K_t is bounded in this direction. The tangent of the angle formed by ξ and ξ' is less than 2ε since H' separates p from $\overline{U}$. Hence the ray (ξ) is a limit ray of rays $(\xi') \in B$. This proves $B(L_\alpha) \subset \overline{B}$.

Since the asymptotic cone $A(M)$ of a convex set M is the polar cone, $(B(M))^* = \overline{B(M)}^*$ of $B(M)$ (Proposition 26, Section 5, Chapter II), the preceding result yields:

52. All level sets of a convex function have the same asymptotic cone.

Now let L_τ, $\tau \in \Omega$, be a family of subsets of D satisfying conditions I - VI. Denote by

$$h(\tau, \xi) = h_{L_\tau}(\xi)$$

the support function of L_τ. From VI it follows that for fixed $\tau \in \Omega_0$, $h(\tau, \xi)$ is defined over the cone B and nowhere else. If α is finite and $\alpha \in \Omega$, $h(\alpha, \xi)$ is defined not only over B, but possibly on certain boundary rays of B which do not belong to B. However, in the sequel it will be sufficient to consider $h(\alpha, \xi)$ for $\xi \in$ B. Furthermore, it suffices to consider unit vectors ξ. By II, $h(\tau, \xi)$ for fixed ξ is an increasing function of $\tau \in \Omega$ which may be interpreted as follows. Let $\tau = \varphi(x)$ be the function with the level sets L_τ. In the (n+1)-dimensional space x, τ consider the set $[D, \varphi]$. Its orthogonal projection upon the 2-flat A^2 spanned by the τ-axis and the vector $(\xi, 0)$, $\xi \in$ B, laid off from the origin is called the ξ-<u>profile</u> of φ. If $-\xi$ is also in B, the $(-\xi)$-profile is identical with the ξ-profile. In A^2 introduce the τ,y-coordinate system consisting of the τ-axis and the oriented line determined by $(\xi, 0)$ as y-axis. Every line $\tau = \tau_0$, $\tau_0 \in \Omega$, in A^2 parallel with the y-axis intersects the ξ-profile in a segment or a ray (in the direction $-\xi$) whose end-point in the direction ξ has the y-coordinate $h(\tau, \xi)$. This follows because for $\|\xi\| = 1$, $h(\tau_0, \xi)$ is the distance from the origin to the supporting (n-1)-flat with normal direction ξ of L_{τ_0}. Thus $y = h(\tau, \xi)$ or, in case $-\xi \in$ B, $y = h(\tau, \xi)$ and $y = -h(\tau, -\xi)$ are the equations of the boundary of the ξ-profile.

Suppose now there is a strictly increasing continuous function $t = F(\tau)$ such that $f(x) = F(\varphi(x))$ is convex in

D. Then the sets $K_t = L_{\Phi(t)}$, $\tau = \Phi(t)$ the inverse of $t = F(\tau)$, satisfy (*) and, hence, by the properties of support functions stated at the end of Section 1

$$(**) \quad h(\Phi(t_\theta), \xi) \geqq (1-\theta)h(\Phi(t_o), \xi) + \theta h(\Phi(t_1), \xi)$$

where $t_\theta = (1-\theta)t_o + \theta t_1$. This means that $h(\Phi(t), \xi)$ is a concave function of t for fixed $\xi \in B$ in accordance with the fact that the ξ-profiles of $F(\varphi(x))$ are convex sets.

Conversely, suppose there exists a strictly increasing continuous function $t = F(\tau)$, $\tau \in \Omega$, $\tau = \Phi(t)$, $t \in W$, such that for a family L_τ, $\tau \in \Omega$, the function $h(\Phi(t), \xi)$ is a concave function of t for every fixed $\xi \in B$, that is the ξ-profiles of $F(\varphi(x))$ are all convex. It follows from this hypothesis that $F(\varphi(x))$ is a convex function in D. To prove this it is sufficient to prove (*). Now (**) is valid and for two point sets M and N, $h_M(\xi) \leqq h_N(\xi)$ implies $\overline{\{M\}} \subset \overline{\{N\}}$. Hence

$$\overline{K_{t_\theta}} \supset \overline{(1-\theta)K_{t_o} + \theta K_{t_1}} \supset (1-\theta)K_{t_o} + \theta K_{t_1}.$$

Condition IV implies $\overline{K}_t \cap D = K_t$. Consequently

$$K_{t_\theta} \supset D \cap ((1-\theta)K_{t_o} + \theta K_{t_1}) = (1-\theta)K_{t_o} + \theta K_{t_1}.$$

The latter equality follows from the inclusions $K_{t_o} \subset D$, $K_{t_1} \subset D$, and the convexity of D. This completes the proof of the following theorem:

53. Let L_τ, $\tau \in \Omega$, be the family of level sets of a lower semicontinuous, quasi-convex function $\varphi(x)$

such that the cone $B(L_\tau) = B$ is independent of τ for $\tau \in \Omega_o$. Let $h(\tau, \xi)$, $\xi \in B$, be the support function of L_τ. Further let $t = F(\tau)$ be a strictly increasing continuous function and $\tau = \Phi(t)$, $t \in W$, its inverse. Then $F(\varphi(x))$ is convex for $x \in D$ if and only if $h(\Phi(t), \xi)$ for every fixed $\xi \in B$ is a concave function of $t \in W$, that is

$$\frac{h(\tau_2, \xi) - h(\tau_1, \xi)}{F(\tau_2)-F(\tau_1)} \geqq \frac{h(\tau_3, \xi) - h(\tau_2, \xi)}{F(\tau_3)-F(\tau_2)}$$

for any three numbers $\tau_1 < \tau_2 < \tau_3$ in Ω.

This condition may be given a different form. If $h(\tau_2, \xi) = h(\tau_1, \xi)$, the inequality implies $h(\tau_3, \xi) = h(\tau_2, \xi)$ since $h(\tau, \xi)$ increases with τ. The inequality being trivially satisfied in this particular case, it is equivalent to

$$\frac{F(\tau_3) - F(\tau_2)}{F(\tau_2) - F(\tau_1)} \geqq \frac{h(\tau_3, \xi) - h(\tau_2, \xi)}{h(\tau_2, \xi) - h(\tau_1, \xi)}$$

the right-hand side being interpreted as 0 whenever the denominator vanishes. The quantity

$$\mathcal{H}(\tau_1, \tau_2, \tau_3) = \sup_{\xi \in B} \frac{h(\tau_3, \xi) - h(\tau_2, \xi)}{h(\tau_2, \xi) - h(\tau_1, \xi)}$$

which only depends on the family L_τ, is used the state the necessary condition:

VII. There is a strictly increasing

continuous function $F(\tau)$, $\tau \in \Omega$,
such that

(***) $$F(\tau_3)-F(\tau_2) \geqq (F(\tau_2)-F(\tau_1))\,\mathcal{H}(\tau_1, \tau_2, \tau_3)$$

for any three numbers $\tau_1 < \tau_2 < \tau_3$ in Ω.

From the preceding it is clear that I - VII are necessary and sufficient conditions in order that a family of subsets of an open convex set D suitably indexed by real numbers forms the family of level sets of a convex function defined over D. While I - VI are simple and intuitive, VII is rather complicated. There is no simple test to decide whether the function $\mathcal{H}(\tau_1, \tau_2, \tau_3)$ is such as to admit a strictly increasing continuous solution of the functional inequality above. Both local and global properties of $\mathcal{H}(\tau_1, \tau_2, \tau_3)$ enter decisively. Compared with the original problem there seems to be no progress. However, VII has the advantage of leading to a kind of construction of the required function $F(\tau)$. To indicate the procedure the following remarks may be added.

Let $\tau_0 < \tau_1 < \tau$ be fixed in Ω. Select numbers τ_i, $i = 1,\ldots,p + 1$ such that

$$\tau_1 < \tau_2 < \ldots < \tau_p < \tau_{p+1} = \tau.$$

Then (***) yields

$$F(\tau_{i+1})-F(\tau_i) \geqq (F(\tau_i)-F(\tau_{i-1}))\,\mathcal{H}(\tau_{i-1}, \tau_i, \tau_{i+1})$$

for $i = 1,\ldots,p$. Multiplication of these inequalities for $i = 1,\ldots,j \leqq p$, gives

$$F(\tau_{j+1})-F(\tau_j) \geqq (F(\tau_1)-F(\tau_o)) \prod_{i=1}^{j} \mathcal{H}(\tau_{i-1}, \tau_i, \tau_{i+1}).$$

Summation over j gives

$$F(\tau)-F(\tau_1) \geqq (F(\tau_1)-F(\tau_o)) \sum_{j=1}^{p} \prod_{i=1}^{j} \mathcal{H}(\tau_{i-1}, \tau_i, \tau_{i+1}).$$

With the notation

$$k(\tau_o, \tau_1, \tau) = \sup \sum_{j=1}^{p} \prod_{i=1}^{j} \mathcal{H}(\tau_{i-1}, \tau_i, \tau_{i+1})$$

where the sup is taken over all subdivisions $\tau_1 < \tau_2 < \dots < \tau_p < \tau$ of the interval $\tau_1, \tau,$

$$F(\tau)-F(\tau_1) \geqq (F(\tau_1)-F(\tau_o))\, k(\tau_o, \tau_1, \tau).$$

Thus $k(\tau_o, \tau_1, \tau)$ has to be finite for all $\tau_o < \tau_1 < \tau$ in Ω. This involves a mixture of local and global conditions on $\mathcal{H}$. If k is finite, a function $F(\tau)$ which has the desired properties for $\tau > \tau_1$ may be obtained as follows. It is easily seen that the values of $F(\tau)$ at two points, τ_o and τ_1 say, may be prescribed arbitrarily. Then any strictly increasing continuous function $F(\tau)$, $\tau > \tau_1$, satisfying

$$F(\tau) \geqq F(\tau_1) + (F(\tau_1)-F(\tau_o))\, k(\tau_o, \tau_1, \tau)$$

can be shown to have the required properties. Such functions exist since $k(\tau_o, \tau_1, \tau)$ is increasing in τ. In similar ways the function can be constructed for τ between τ_o and τ_1 and for τ less that τ_o.

In the next section the construction is carried through in the case of smooth functions.

8. SMOOTH CONVEX FUNCTIONS WITH PRESCRIBED LEVEL SETS

Let D be an open convex set in A^n. The problem discussed in the preceding section will now be solved under the assumption that the prescribed subsets L_τ of D are the level sets of a twice differentiable function $\tau = \varphi(x)$. As in Section 7 we set $\alpha = \inf \varphi(x)$, $\beta = \sup \varphi(x)$. We ask for a twice differentiable strictly increasing function $F(\tau)$, $\alpha \underset{(=)}{\leq} \tau < \beta$, such that $f(x) = F(\varphi(x))$ is convex in D. We start by deriving necessary conditions, which will turn out to be sufficient. The results of Section 7 will not be used.

We introduce the notations

$$\frac{\partial \varphi}{\partial x_i} = \varphi_i, \qquad \frac{\partial f}{\partial x_i} = f_i,$$

$$i, j = 1, \ldots, n.$$

$$\frac{\partial^2 \varphi}{\partial x_i \partial x_j} = \varphi_{ij}, \quad \frac{\partial^2 f}{\partial x_i \partial x_j} = f_{ij},$$

The derivatives of $f(x) = F(\varphi(x))$ may then be written

(1) $$f_i = F' \varphi_i,$$

(2) $$f_{ij} = F'' \varphi_i \varphi_j + F' \varphi_{ij}.$$

Suppose now $f(x) = F(\varphi(x))$ is convex. Then $f(x)$ has no critical points except possibly those at which it attains its absolute minimum. Obviously, $\varphi(x)$ must have the same property. We formulate this as the first necessary condition:

A. $\varphi(x)$ has no critical points

except those where it attains its absolute minimum, if such a minimum exists.

From (1) and $F'(\tau) \geq 0$ for $\alpha \underset{(\leqq)}{\leq} \tau < \beta$ it therefore follows that $F'(\tau) > 0$ for $\tau > \alpha$. Now $f(x)$ is convex if and only if for every fixed $x \in D$ the quadratic form

$$\sum_{i,j} f_{ij}(x) y_i y_j = F''(\varphi(x))(\sum_i \varphi_i(x) y_i)^2 + F'(\varphi(x)) \sum_{i,j} \varphi_{ij}(x) y_i y_j$$

in the variables y_i, $i = 1, \ldots, n$, is positive semidefinite.

If $\varphi(x)$, and hence $f(x)$, has a minimum, this condition is obviously satisfied at all points where the minimum is assumed, that is at all $x \in L_\alpha$. This is because $\varphi_i = 0$ and $\sum_{i,j} \varphi_{ij} y_i y_j$ is positive semidefinite at these points. Hence it is sufficient to consider those x for which $\varphi(x) > \alpha$. For such x, $F' > 0$ so that the notations

$$\sigma = \sigma(x) = \frac{F''(\varphi(x))}{F'(\varphi(x))},$$

(3) $$Q(y,y) = \sum_{i,j} \varphi_{ij} y_i y_j + \sigma(\sum_i \varphi_i y_i)^2,$$

may be used to replace the previous condition by: $Q(y,y)$ is positive semidefinite for every x in D but not in L_α.

Let such an x be fixed. The characteristic determinant of $Q(y,y)$ is

$$c_Q(\lambda) = \left| \varphi_{ij} - \lambda \delta_{ij} + \sigma \varphi_i \varphi_j \right|$$

$$= \begin{vmatrix} \varphi_{ij} - \lambda \delta_{ij} + \sigma \varphi_i \varphi_j & \varphi_i \\ 0 & 1 \end{vmatrix}.$$

Subtraction of suitable multiples of the added column from the other columns leads to

$$c_Q(\lambda) = \begin{vmatrix} \varphi_{ij} - \lambda\,\delta_{ij} & \varphi_i \\ -\sigma\varphi_j & 1 \end{vmatrix} .$$

This determinant equals the minor of its lower right hand corner, plus the value of the determinant when 1 is replaced by zero. Thus the characteristic determinant of $Q(y,y)$ takes the form

$$(4) \qquad c_Q(\lambda) = \left|\varphi_{ij} - \lambda\,\delta_{ij}\right| - \sigma \begin{vmatrix} \varphi_{ij} - \lambda\,\delta_{ij} & \varphi_i \\ \varphi_j & 0 \end{vmatrix} .$$

If it is written as a polynomial in λ,

$$c_Q(\lambda) = T_n - T_{n-1}\lambda + \dots + (-1)^n T_o\,\lambda^n,$$

we have $T_o = 1$, and T_ρ, $\rho = 1,\dots,n$, is the ρth elementary symmetric function of the characteristic roots.

The first term on the right side of (4) is the characteristic determinant

$$c_p(\lambda) = S_n - S_{n-1}\lambda + \dots + (-1)^n S_o\,\lambda^n$$

of the quadratic form

$$P(y,y) = \sum_{i,j} \varphi_{ij} y_i y_j .$$

Here $S_0 = 1$, and S_ρ, $\rho = 1,\ldots,n$, is the ρth elementary symmetric function of the characteristic roots of $P(y,y)$. We are going to show that the second term of (4) is essentially the characteristic determinant $C_p^*(\lambda)$ of the quadratic form $P^*(y,y)$ in $n-1$ variables derived by specializing $P(y,y)$ to the hyperplane $\sum_i \varphi_i y_i = 0$. The characteristic roots of $P^*(y,y)$ are the stationary values of $P(y,y)$ subject to the constraints $\sum_i \varphi_i y_i = 0$ and $\sum_i y_i^2 = 1$. Hence, by the multiplier rule, they are the stationary values of the function

$$\sum_{i,j} \varphi_{ij} y_i y_j + 2z \sum_i \varphi_i y_i - \lambda (\sum_i y_i^2 - 1)$$

with y_i unrestricted, $2z$ and λ denoting the multipliers. For the critical points y_i this gives the condition

$$\sum_j \varphi_{ij} y_j + z\, \varphi_i - \lambda y_i = 0 \tag{5}$$

$$\sum_j \varphi_j y_j = 0, \tag{6}$$

$$\sum_i y_i^2 = 1. \tag{7}$$

The existence of a solution y_i, z of this system implies

$$\begin{vmatrix} \varphi_{ij} - \lambda \delta_{ij} & \varphi_i \\ \varphi_j & 0 \end{vmatrix} = 0. \tag{8}$$

Suppose that λ satisfies this equation and that y_i, z solve the system (5), (6), (7). Multiplying (5) by y_i and summing over i, we see that $\sum_{i,j} \varphi_{ij} y_i y_j = \lambda$ so that λ is the stationary value in question. Hence (8) is the characteristic equation of $P^*(y,y)$. Formally the left side of (8) is a polynomial of degree n in λ. However, the coefficient of λ^n vanishes. The coefficient of λ^{n-1}, which is needed for normalization, is obtained by dividing the determinant (8) by λ^{n-1} and letting $\lambda \longrightarrow \infty$. If this is done by dividing each of the first n rows by λ and multiplying thereafter the last column by λ, the coefficient is easily found to be

$$\begin{vmatrix} -1 & 0 & \dots & 0 & \varphi_1 \\ 0 & -1 & \dots & 0 & \varphi_2 \\ \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & \dots & -1 & \varphi_n \\ \varphi_1 & \varphi_2 & \dots & \varphi_n & 0 \end{vmatrix} = (-1)^n \sum_i \varphi_i^2 .$$

With the notation

$$k^2 = \sum_i \varphi_i^2$$

we therefore have

$$C_P^*(\lambda) = -\frac{1}{k^2} \begin{vmatrix} \varphi_{ij} - \lambda \delta_{ij} & \varphi_i \\ \varphi_j & 0 \end{vmatrix} .$$

If this is written as a polynomial

$$C_P^*(\lambda) = S_{n-1}^* - S_{n-2}^* \lambda + \dots + (-1)^{n-1} S_o^* \lambda^{n-1},$$

then $S_o^* = 1$ and S_ρ^* is the ρth elementary symmetric function of the characteristic roots of $P^*(y,y)$. Hence, (4) may be written

$$C_Q = C_p + \sigma k^2 C_p^* \quad .$$

Therefore

$$T_\rho = S_\rho + \sigma k^2 S_{\rho-1}^*, \qquad \rho = 1, \ldots, n. \tag{9}$$

Now $Q(y,y)$ is positive semidefinite if and only if all characteristic values are non-negative, that is

$$T_\rho \geqq 0, \qquad \rho = 1, \ldots, n. \tag{10}$$

As is well known, this implies that if one $T_\rho = 0$, all the following T_ρ vanish too.

Looking for necessary conditions that there may exist an $F(\tau)$ such that $F(\varphi(x))$ is convex, we assume (10) to be valid. The expression (3) shows that $P^*(y,y)$ agrees with $Q(y,y)$ for y_i satisfying $\sum\limits_i \varphi_i y_i = 0$. Hence, $P^*(y,y)$ is positive semidefinite and thus

$$S_{\rho-1}^* \geqq 0, \qquad \rho = 1, \ldots, n.$$

Let

$$\mu_1 \geqq \mu_2 \geqq \cdots \geqq \mu_n$$

and

$$\mu_1^* \geqq \mu_2^* \geqq \cdots \geqq \mu_{n-1}^*$$

be the characteristic values of $P(y,y)$ and $P^*(y,y)$

respectively. By the maximum-minimum properties of the characteristic values of a quadratic form,

$$\mu_1 \geqq \mu_1^* \geqq \mu_2 \geqq \cdots \geqq \mu_\rho \geqq \mu_\rho^* \geqq \cdots \geqq \mu_{n-1}^* \geqq \mu_n.$$

If $r - 1$ denotes the rank of $P^*(y,y)$ (which, of course, may depend on x), then

$$\mu_1^* > 0, \ \ldots, \ \mu_{r-1}^* > 0, \ \mu_r^* = \cdots = \mu_{n-1}^* = 0.$$

Hence

$$\mu_1 > 0, \ \ldots, \ \mu_{r-1} > 0,$$

and if $r < n$,

$$\mu_r \geqq 0, \ \mu_{r+1} = \cdots = \mu_{n-1} = 0, \quad \mu_n \leqq 0.$$

This shows that the rank of $P(y,y)$ is at most $r + 1$, and that

$$S_{r+1} = \mu_1 \cdots \mu_r \, \mu_n \leqq 0$$

if $r < n$. On the other hand, because $S_r^* = 0$, (9) and (10) for $\rho = r + 1$ yield $S_{r+1} \geqq 0$. Hence $S_{r+1} = 0$, that is $\mu_r = 0$ or $\mu_n = 0$. Thus the rank of $P(y,y)$ is actually at most r.

B. In order that there may exist a twice differentiable strictly increasing function $F(\tau)$ such that $F(\varphi(x))$ is convex, it is necessary that for each fixed $x \in D$ the quadratic form $\sum_{i,j} \varphi_{ij}(x) y_i y_j$ restricted to the hyperplane $\sum_i \varphi_i(x) y_i = 0$ be positive semidefinite, and if $r - 1$ denotes its rank, the rank of the same form without the restriction be at most r.

This has only been proved for x not in L_α. However, for $x \in L_\alpha$ we have $\varphi_i(x) = 0$ and the statement is obviously true.

The first part of the condition, $P^*(y,y)$ positive semidefinite, expresses the convexity of the level sets of $\varphi(x)$. The second part is trivially satisfied when $P^*(y,y)$ has the maximal rank $n - 1$. At points x where $r < n$ it restricts the local behaviour of $\varphi(x)$ in a way indicated by the following example:

Let $n = 2$ and assume that for each τ_0 of a certain subinterval of $\alpha \underset{(=)}{\leq} \tau < \beta$ the curve $\varphi(x) = \tau_0$ (τ_0 a constant) contains a segment depending smoothly on τ_0. Then the rank of $P^*(y,y)$ is zero at the points of the segments. The surfaces $\tau = \varphi(x)$ and, hence, $t = f(x)$ then contain pieces of ruled surfaces whose generators are parallel to the x_1x_2-plane. Such a ruled surface can only be convex if it is a cylinder, that is if the generators and, thus, the segments are mutually parallel. This is just what the condition, rank of $P(y,y)$ at most one, requires in this case.

Even if $\varphi(x)$ is an analytic function, the rank condition may restrict its local behaviour. Take again $n = 2$ and assume that the curvature of a curve where $\varphi(x)$ is a constant vanishes at some point. Then the rank condition requires that the Gaussian curvature of the surface $\tau = \varphi(x)$ also vanish at that point.

Consider again a fixed x not in L_α. In view of (9) and because of $S_\rho = S^*_{\rho-1} = 0$ for $\rho > r$ the condition

(10) reduces to $\sigma \geqq \bar{\sigma}$ where

$$\bar{\sigma} = \bar{\sigma}(x) = \max_{1 \leqq \rho \leqq r} \left(- \frac{S_\rho}{k^2 S^*_{\rho-1}} \right) .$$

Let this maximum be attained for $\rho = \rho_o$. For the coefficients of the characteristic equation of $Q(y,y)$ with σ replaced by $\bar{\sigma}$ we then have

$$\bar{T}_\rho = S_\rho + \bar{\sigma} k^2 S^*_{\rho-1} \geqq 0, \qquad \rho = 1, \ldots, r,$$

the equality sign being valid for $\rho = \rho_o$. As mentioned above, this implies $\bar{T}_i = 0$ for $\rho > \rho_o$; hence, in particular, for $\rho = r \geqq \rho_o$. This gives

$$\bar{\sigma} = - \frac{S_r}{k^2 S^*_{r-1}} . \tag{11}$$

For each fixed x not in L_α and $\tau = \varphi(x)$ we therefore have

$$\frac{F''(\tau)}{F'(\tau)} = \sigma(\tau) \geqq \sup_{\varphi(x)=\tau} \bar{\sigma}(x) = \sup_{\varphi(x)=\tau} \left(- \frac{S_r}{k^2 S^*_{r-1}} \right) . \tag{12}$$

where the sup has to be taken over all $x \in D$ for which $\varphi(x) = \tau$. Thus, we have the further necessary condition:

C. If for a twice differentiable, strictly increasing function $F(\tau)$, $\alpha \underset{(=)}{\leq} \tau < \beta$, the function $F(\varphi(x))$ is convex, then

$$\frac{F''(\tau)}{F'(\tau)} \geqq \sup_{\varphi(x)=\tau} \left(-\frac{S_r}{k^2 S^*_{r-1}} \right) .$$

Conversely, let there be given a twice differentiable function $\tau = \varphi(x)$, $x \in D$, and a twice differentiable, strictly increasing function $F(\tau)$, $\alpha \underset{(=)}{\leq} \tau < \beta$, where $\alpha = \inf \varphi$ and $\beta = \sup \varphi$, such that conditions A, B, and C are satisfied. Then $f(x) = F(\varphi(x))$ is convex in D. We have to show that the quadratic form $\sum_{i,j} f_{ij} y_i y_j$ is positive semidefinite for each $x \in D$. For the points $x \in L_\alpha$, if any, this is obviously the case as mentioned at the beginning of this section. For x not in L_α we have to show that $Q(y,y)$ is positive semidefinite. Because of C,

$$Q(y,y) = \sum_{i,j} \varphi_{ij} y_i y_j + \frac{F''}{F'} \left(\sum_i \varphi_i y_i\right)^2$$

$$\geqq \sum_{i,j} \varphi_{ij} y_i y_j - \frac{S_r}{k^2 S^*_{r-1}} \left(\sum_i \varphi_i y_i\right)^2 .$$

It therefore suffices to prove that the latter form, call it $Q'(y,y)$, is positive semidefinite. From (3) and (9) it is seen that the coefficients of its characteristic equation are

$$T_\rho' = S_\rho - \frac{S_r}{S^*_{r-1}} S^*_{\rho-1}, \qquad \rho = 1,\ldots,n.$$

Now, $S_\rho = S^*_{\rho-1} = 0$ for $\rho = r+1,\ldots,n$, because of B. Hence

$$T_\rho' = 0, \qquad \rho = r, r+1, \ldots, n,$$

which shows that the rank of $Q'(y,y)$ is at most $r - 1$. On the other hand, $Q'(y,y)$ restricted to the hyperplane $\sum_i \varphi_i y_i = 0$ agrees with $P^*(y,y)$. Because of B, $P^*(y,y)$ has $r - 1$ positive characteristic roots. Hence $Q'(y,y)$ must have the same property. This proves the statement.

HISTORICAL NOTES

CHAPTER I

CONVEX CONES

Sections 1 - 6. Important contributions to the theory of convex cones are contained (more or less explicitly) in Minkowski's posthumous paper [48]. The basic paper on the subject is, however, Part II of Steinitz's paper [57]. Practically all the concepts and results of Sections 1 - 6 are to be found in this paper. Also many of the proofs given here are based on ideas due to Steinitz. Polyhedral convex cones have been the subject of several more recent expositions, namely Weyl [66] (with purely algebraic methods), Gale [21], Gerstenhaber [24].

Section 7. As mentioned in the text, the theory of (polyhedral) convex cones is closely related to the theory of (finite) systems of linear inequalities. For the latter theory and its various geometrical interpretations the reader is referred to Dines and McCoy [16] and especially to the dissertation of Motzkin [49]. Included in the latter is a very complete bibliography up to 1934. Of more recent papers Dines [14], Blumenthal [5], [6], Levi [42], La Menza [40], Nagy [50] may be mentioned. Further references may be found in Contributions to the Theory of Games (Annals of Mathematics Study 24, Princeton, 1950).

For the second interpretation used in Section 7 see also Gale [21]. Theorem 17 for polyhedral cones has been announced by Tucker [63]; the corollaries III - VI are likewise due to Tucker.

CHAPTER II

CONVEX SETS

For the literature up to 1934 concerning basic properties of convex sets the reader is referred to the report [8] by Bonnesen and the author. Attention is called to the dissertation of Straszewicz [60] which gives a comprehensive account for compact sets and to Part I of Steinitz [57] which deals with arbitrary convex sets. For convex polyhedra see also Kirchberger [36] and especially Weyl [66]. More recent, mainly expository articles are Dines [14], Botts [9], Bateman [3], Macbeath [45]. For a generalization of the concept of convex sets see Green and Gustin [25].

Section 2. In Proposition 6 (stating that every point of the convex hull of a point set M is a centroid of at most $n + 1$ points of M) the maximal number $n + 1$ can be replaced by n if the set M has certain properties of connectedness. See [8] p. 9 for references to the first papers on this subject. Further references are Bunt [11], Hanner [28], and especially Hanner and Rådström [29]. The following question is likewise connected with Proposition 6: What is the smallest positive integer p with the property that every point z relative interior to the convex hull of a set M of linear dimension $d > 0$ is relative interior to the convex hull of a subset of M with linear dimension d consisting of at most p points? The answer is $p = 2d$ as is easily seen by applying the Corollary to Theorem 8 (Chapter I) to the cone with vertex z consisting of the rays which join z with the points of M. This result (essentially due to Steinitz) occurs implicitly in the discussion of systems of linear inequalities of the form $Ax \geqq 0$. (Cf. Chapter I, Section 7 and e.g. Dines and McCoy [16], Dines [14].) A direct proof has recently been given by Gustin [26].

Section 4. Projecting cones and normal cones were introduced by Minkowski [48], the cones of directions of boundedness and asymptotic cones by Steinitz [57]. For the theory of asymptotic cones and various applications see Stoker [58].

The concept of s-convexity (under the name of even convexity) is introduced in the author's paper [19].

Section 6. The Separation Theorem 27 is due to Minkowski [48]. The useful statement 28 is slightly more general. Theorem I of Klee's paper [37] may be considered as a generalization of Proposition 27 to an arbitrary finite number of compact convex sets.

Section 7. For literature concerning extreme points and supports see [8], p. 16, further, for polyhedra, Weyl [66]. Straszewicz [61] has shown that in Proposition 33 it is sufficient to consider "exponed points" instead of extreme points. An exponed point of a closed convex set is by definition a point of the set through which there is a (supporting) hyperplane having no other points in common with the set.

Section 8. Convex sets in projective spaces have been considered by Steinitz [57], Part III. (For a problem in connection with the definition see also Kneser [38].) The polarity with respect to the unit sphere has been introduced by Minkowski [48], p. 146-7; cf. also Haar [27], Helly [31], von Neumann [65], Young [67], Bateman [3]. For generalizations to certain unbounded sets see Rådström [54], Lorch [44]. Arbitrary polarities have been considered by Steinitz [57], Part III, and, as in Section 8, for sets which are not necessarily closed or open, by the author [19].

CHAPTER III

CONVEX FUNCTIONS

For the history of the theory of convex functions, various applications, and generalizations as well as extensive bibliographies the reader is referred to Popoviciu [51] and Beckenbach [4]. Apart from some references to basic papers, only more recent papers dealing or connected with the topics of this report are quoted in the sequel. A modern, detailed exposition of many basic properties of convex functions is given in Haupt, Aumann, Pauc [30], I, Section 4.8, Section 5.4.2.1, Section 5.5, II, Section 2.2.5.

Section 1. Proposition 4 which comprises many of the classical inequalities of analysis seems to have been the start of the theory (Hölder [32], Brunn [10], and the basic paper Jensen [34].) Convex functions defined over arbitrary point sets have been considered by Galvani [23], Tortorici [62], and especially Popoviciu [51]. Homogeneous convex functions (gauge functions, supports functions) were introduced by Minkowski [47], [48]. For further references see [8] Section 4. A recent paper is Rédei [55]. See also the exposition by Bateman [3].

It should be pointed out that Propositions 5, 10, 11, 14, which for systematic reasons are deduced directly from the definition of convex functions, are immediate consequences of the existence of a support through every point $x, f(x)$ (proved in Section 4).

Sections 2 - 4. For references concerning the well-known continuity properties of convex functions see Popoviciu [51]. The question whether a convex function is necessarily absolutely continuous has been discussed by Friedman [20], the answer being affirmative for $n = 1$ only. For the behaviour of a convex function at the boundary of its domain (Propositions 24-26) see the author's paper [18].

The first proofs of the existence of the one-sided derivatives of a convex function of one variable and of the directional derivative of a convex function of several variables seem to have been given by Stolz [59], p. 35-36 and Galvani [23]. The latter concept has been applied to the study of homogeneous convex functions by Bonnesen and the author [8], Section 4. The discussion of the directional derivatives of arbitrary convex functions as given in the present Section 4 probably has not been published elsewhere. A new approach to the study of certain smoothness properties of convex functions has been made by Anderson and Klee [2]. Busemann and Feller [12] and Alexandroff [1] have proved the almost everywhere existence of a second differential of a convex function of several variables. A new definition of smooth homogeneous convex functions based on the definiteness of the quadratic form occurring in Proposition 35 has been proposed by Lorch [44].

Section 5. The conjugate of a convex function of one variable has been defined by Mandelbrojt [46]. For the general concept and some of its properties see the author's paper [18]. The inequality stated in Proposition 38 has a well-known analogue for homogeneous functions: Let $F(x)$ and $H(\xi)$ be the gauge function and the support function, (respectively), of a convex body C containing the origin in its interior. Then

$$x'\xi \leqq F(x)H(\xi)$$

for all x and ξ. (Cf. Helly [31], von Neumann [65], Young [67], Lorch [44].) This may be considered as a special case of Proposition 38. For, put $f(x) = 0$ for $x \in C$, that is for $F(x) \leqq 1$. Then $\varphi(\xi) = H(\xi)$ and hence

$$x'\xi \leqq H(\xi) \quad \text{for} \quad F(x) \leqq 1.$$

Because of the homogeneity of F this is equivalent to the above inequality.

The rest of Section 5 is unpublished. The corollary, Proposition 43, is a slight generalization of a theorem due to Bohnenblust, Karlin, Shapley [7]. Helly's Theorem, which appears here as a corollary (Proposition 45) and various generalizations have been the subject of many recent papers: Vincensini [64], Robinson [56], Lannér [41], Dukor [17], Rado

[53], Horn [33], Rademacher and Schoenberg [52], Karlin and Shapley [35], Levi [43], Klee [37]. For references to the older papers see [8] p. 3. Proposition 46 generalizes Minkowski's well-known characterization of the support functions of (compact) convex bodies. See [8], p. 28 for the older literature. Further references are Rédei [55], Bateman [3]. The determination of the support function of the intersection of convex sets following Proposition 46 seems to be noted for the first time by F. Riesz (who communicated it to Lannér, see [41].)

Section 6. Unpublished. The results generalize the duality property of linear programming problems proved by Gale, Kuhn, Tucker [22] to non-linear problems of the type considered by Kuhn and Tucker [39]. The consideration of completely arbitrary closed convex functions is essential for the formulation and the validity of a simple duality theorem. For the theory of programming problems in general the reader is referred to Activity Analysis of Production and Allocation (Cowles Commission Monograph 13, New York 1951).

Section 7. The problem of the existence and the determination of a convex function with prescribed level sets was raised and studied by de Finetti [13] under the assumption that the domain D and, thus, all level sets are compact and convex. In this case the Conditions I - VI are trivially satisfied. The part of Section 7 dealing with these conditions in the general case is not published. Condition VII is a generalization to the case considered here of a result of de Finetti. For details of the construction of a convex function the reader is referred to de Finetti's paper.

Section 8. Unpublished. In a footnote de Finetti [13] states that in his case of a compact D the smoothness of the function $\varphi(x)$ implies the existence of an $F(\tau)$ such that $F(\varphi(x))$ is convex. This contradicts the results of Section 8 of the present report. Apparently de Finetti had overlooked the fact that the smoothness of φ does not imply the smoothness of the support function $h(\xi, \tau)$. This is only the case if the rank $r - 1$ introduced in Section 8 has its maximal value $n - 1$ everywhere in D. Then the quantity $\overline{\sigma}$ (see equation (11)) is easily found to be

$$\overline{\sigma} = \frac{\partial^2 h}{\partial \tau^2} \Big/ \frac{\partial h}{\partial \tau} .$$

At points where $r < n$, the second derivative may not exist even if φ is analytic.

BIBLIOGRAPHY

1. Alexandroff, A. D., Almost everywhere existence of the second differential of a convex function and some properties of convex surfaces connected with it. [Leningrad State Univ. Annals, Math. Ser., 6 (1939), 3-35.]

2. Anderson, R. D., and Klee, V. L. Jr., Convex functions and upper semi-continuous collections. [Duke Math. J. 19 (1952), 349-357.]

Aumann, George, see Haupt, Otto; Aumann, Georg; Pauc, Christian.

3. Bateman, P. T., Introductory material on convex sets in euclidean space. [Seminar on convex sets (The Institute for Advanced Study) Princeton 1949-1950, 1-26.]

4. Beckenbach, E. F., Convex functions. [Bull. Amer. Math. Soc. 54 (1948), 439-460.]

5. Blumenthal, Leonard M., Metric methods in linear inequalities. [Duke Math. J. 15 (1948), 955-966.]

6. ----------, Two existence theorems for systems of linear inequalities. [Pacific J. Math. 2 (1952), 523-530.]

7. Bohnenblust, H. F., Karlin, S., and Shapley, L. S., Games with continuous, convex pay-off. [Contributions to the theory of games (Annals of Mathematics Study 24) Princeton 1950, 181-192.]

8. Bonnesen, T., and Fenchel, W., Theorie der konvexen Körper. [Ergebnisse der Mathematik 3, 1, Berlin 1934, New York 1948, 164 p.]

9. Botts, Truman, Convex sets. [Amer. Math. Monthly, 49 (1942), 527-535.]

10. Brunn, Hermann, Über Curven ohne Wendepunkte. [München 1889, 74 p.]

11. Bunt, L. N. H., Bijdrage tot de theorie der convexe puntverzamelingen. [Dissertation Amsterdam 1934, 108 p.]

12. Busemann, Herbert, and Feller, Willy, Krümmungseigenschaften konvexer Flächen. [Acta Math. 66 (1935), 1-47.]

13. de Finetti, Bruno, Sulle stratificazioni convesse. [Ann. Mat. Pura Appl. (4) 30 (1949), 173-183.]

14. Dines, L. L., Convex extensions and linear inequalities. [Bull. Amer. Math. Soc. 42 (1936), 353-365.]

15. -----, On convexity. [Amer. Math. Monthly 45 (1938), 199-209.]

16. Dines, L. L., and McCoy, N. H., On linear inequalities. [Trans. Roy. Soc. Canada, Sect. III (3) 27 (1933), 37-70.]

17. Dukor, I. G., On a theorem of Helly on collections of convex bodies with common points. [Uspekhi Matem. Nauk. 10 (1944), 60-61."

Feller, Willy, see Busemann, Herbert, and Feller, Willy.

18. Fenchel, W., On conjugate convex functions. [Canadian J. Math. 1 (1949), 73-77.]

19. -------, A remark on convex sets and polarity. [Medd. Lunds Univ. Mat. Sem. (Supplementband tillägnat Marcel Riesz) 1952, 82-89.]

-------, see Bonnesen, T., and Fenchel, W.

20. Friedman, Bernhard, A note on convex functions. [Bull. Amer. Math. Soc. 46 (1940), 473-474.]

21. Gale, David, Convex polyhedral cones and linear inequalities. [Activity Analysis of Production and Allocation (Cowles Commission Monograph 13) New York 1951, 287-297.]

22. Gale, David; Kuhn, Harold W.; Tucker, Albert W., Linear programming and the theory of games. [Activity Analysix of Production and Allocation (Cowles Commission Monograph 13) New York 1951, 317-329.]

23. Galvani, L., Sulle funzioni convesse di una o due variabili definite in un aggregato qualunque. [Rend. Circ. Mat. Palermo, 41 (1916), 103-134.]

24. Gerstenhaber, Murray, Theory of convex polyhedral cones. [Activity Analysis of Production and Allocation (Cowles Commission Monograph 13) New York 1951, 298-316.]

25. Green, J. W., and Gustin, W., Quasiconvex sets. [Canadian J. Math. 2 (1950), 489-507.]

26. Gustin, William, On the interior of the convex hull of a euclidean set. [Bull. Amer. Math. Soc. 53 (1947), 299-301.]

------, see Green, J. W., and Gustin, W.

27. Haar, A., Die Minkowskische Geometrie und die Annäherung an stetige Funktionen. [Math. Ann. 78 (1918), 294-311.]

28. Hanner, Olof, Connectedness and convex hull. [Seminar on convex sets (The Institute for Advanced Study) Princeton 1949-1950, 35-40.]

29. Hanner, Olof, and Rådström, Hans, A generalization of a theorem of Fenchel. [Proc. Amer. Math. Soc. 2 (1951), 589-593.]

30. Haupt, Otto; Aumann, Georg; Pauc, Christian, Differential - und Integralrechnung. I. Band, zweite Auflage, Berlin 1948, 218 p., II. Band, zweite Auflage, Berlin 1950, 210 p.

31. Helly, E., Über Systeme linearer Gleichungen mit unendlich vielen Unbekannten. [Monatsh. Math. Phys. 31 (1921), 60-91.]

32. Hölder, O., Über einen Mittelwertsatz. [Nachr. Ges. Wiss. Göttingen 1889, 38-47.]

33. Horn, Alfred, Some generalizations of Helly's theorem on convex sets. [Bull. Amer. Math. Soc. 55 (1949), 923-929.]

34. Jensen, J. L. W. V., Sur les fonctions convexes et les inégalités entre les valeurs moyennes. [Acta Math. 30 (1906), 175-193.]

. Karlin, S., see Bohnenblust, H. F., Karlin, S., and Shapley, L. S.

35. Karlin, S., and Shapley, L. S., Some applications of a theorem on convex functions. [Ann. of Math. 52 (1950), 148-153.]

36. Kirchberger, P., Über Tschebyschefsche Annäherungsmethoden. [Math. Ann. 57 (1903), 509-540.]

37. Klee, V. L., Jr., On certain intersection properties of convex sets. [Canadian J. Math. 3 (1951), 272-275.]

----, see Anderson, R. D., and Klee, V. L. Jr.

38. Kneser, H., Eine Erweiterung des Begriffes "konvexer Körper". [Math. Ann. 82 (1921), 287-296.]

. Kuhn, H. W., see Gale, David; Kuhn, Harold W.; Tucker, Albert W.

39. Kuhn, H. W., and Tucker, A. W., Nonlinear programming. [Proceedings of the Second Berkeley Symposium on Mathematical Statistics and Probability, Berkeley 1951, 481-492.]

40. La Menza, Francisco, Los sistemas de inecuaciones lineales y sus aplicaciones al estudio de los cuerpos convexos. [An. Soc. Ci. Argent. 121 (1936), 209-248; 122 (1936), 86-122, 297-310, 381-394; 124 (1937), 157-175, 248-274.]

41. Lannér, Folke, On convex bodies with at least one point in common. [Kungl. Fysiografiska Sällskapets i Lund Förhandlinger 13 no. 5, 41-50 (1943), also Medd. Lunds Univ. Mat. Sem. 5 (1943).]

42. Levi, F. W., Ein Reduktionsverfahren für lineare Vektorungleichungen. [Arch. Math. 2 (1949), 24-26.]

43. ----, On Helly's theorem and the axioms of convexity. [J. Indian Math. Soc. (N. S.) Part A, 15 (1951), 65-76.]

44. Lorch, E. R., Differentiable inequalities and the theory of convex bodies. [Trans. Amer. Math. Soc. 71 (1951), 243-266.]

45. Macbeath, A. M., Compactness theorems. [Seminar on convex sets (The Institute for Advanced Study) Princeton 1949-1950, 41-51.]

46. Mandelbrojt, Szolem, Sur les fonctions convexes. [C. R. Acad. Sci. Paris 209 (1939), 977-978.]

McCoy, N. H., see Dines, L. L., and McCoy, N. H.

47. Minkowski, Hermann, Geometrie der Zahlen. Leipzig and Berlin 1896, 1910, 256 p.

48. ---------, Theorie der konvexen Körper, insbesondere Begründung ihres Oberflächenbegriffs. [Gesammelte Abhandlungen, Zweiter Band, Leipzig and Berlin 1911, 131-229.]

49. Motzkin, Th., Beiträge zur Theorie der linearen Ungleichungen. [Inaugural-Dissertation Basel, Jerusalem 1936, 71 p.]

50. Nagy, Béla de Sz., Sur les lattis linéaires de dimension finie. [Comment. Math. Helv. 17 (1945), 209-213.]

Pauc, Christian, see Haupt, Otto; Aumann, Georg; Pauc, Christian.

51. Popoviciu, Tibere, Les fonctions convexes. [Actualités scientifiques et industrielles 992, Paris 1945, 76 p.]

52. Rademacher, Hans; and Schoenberg, I. J., Convex domains and Tchebycheff's approximation problem. [Canadian J. Math. 2 (1950), 245-256.]

53. Rado, R., A theorem on Abelian groups. [J. London Math. Soc. 22 (1947), 219-226.]

54. Rådström, Hans, Polar reciprocity. [Seminar on convex sets (The Institute for Advanced Study) Princeton 1949-1950, 27-29.]

--------, see Hanner, Olof and Rådström, Hans.

55. Rédei, L., Über die Stützebenenfunktion konvexer Körper. [Math. Naturwiss. Anz. Ungar. Akad. Wiss. 60 (1941), 64-69.] (Hungarian, German summary.)

56. Robinson, Charles V., Spherical theorems of Helly type and congruence indices of spherical caps. [Amer. J. Math. 64 (1942), 260-272.]

Schoenberg, I. J., see Rademacher, Hans; and Schoenberg, I. J.

Shapley, L. S., see Bohnenblust, H. F., Karlin, S., and Shapley, L. S.

-------, see Karlin, S., and Shapley, L. S.

57. Steinitz, E., Bedingt konvergente Reihen und konvexe Systeme. I, II, III. [J. Reine Angew. Math. 143 (1913), 128-175; 144 (1914), 1-40; 146 (1916), 1-52.]

58. Stoker, J. J., Unbounded convex sets. [Amer. J. Math. 62 (1940), 165-179.]

59. Stolz, Otto, Grundzüge der Differential- und Integralrechnung. Erster Theil. Leipzig 1893, 460 p.

60. Straszewicz, Stefan, Beiträge zur Theorie der konvexen Punktmengen. [Dissertation Zürich 1914, 57 p.]

61. -----------, Über exponierte Punkte abgeschlossener Punktmengen. [Fund. Math. 24 (1935), 139-143.]

62. Tortorici, P., Sui massimi e minimi delle funzioni convesse. [Atti Accad. Naz. Lincei. Rend. (4) 14 (1931), 472-474.]

63. Tucker, A. W., Extensions of theorems of Farkas and Stiemke. [Bull. Amer. Math. Soc. 56 (1950), 57.]

------, see Gale, David; Kuhn, Harold W.; Tucker, Albert W.

------, see Kuhn, H. W., and Tucker, A. W.

64. Vincensini, Paul, Sur une extension d'un théorème de M. J. Radon sur les ensembles convexes. [Bull. Soc. Math. France 67 (1939), 115-119.]

65. von Neumann, J., Some matrix-inequalities and metrization of matric-space. [Mitt. Forsch. - Inst. Math. Mech. Univ. Tomsk, 1 (1937), 286-299.]

66. Weyl, H., Elementare Theorie der konvexen Polyeder. [Comment. Math. Helv. 7 (1935), 290-306.] (English translation, Contributions to the theory of games, Annals of Mathematics Study no. 24, Princeton 1950, 3-18.)

67. Young, L. C., On an inequality of Marcel Riesz. [Ann. of Math. 40 (1939), 567-574.]

INDEX OF DEFINITIONS