UNCLASSIFIED

THE TECHNOLOGICAL INSTITUTE

NORTHWESTERN UNIVERSITY

EVANSTON ILLINOIS

O. N. R. RESEARCH MEMORANDUM NO. 55

The Technological Institute
The College of Liberal Arts
Northwestern University

DUALITY IN SEMI-INFINITE PROGRAMS AND SOME WORKS OF HAAR AND CARATHÉODORY

by

A. Charnes, W. W. Cooper* and K. Kortanek

*Carnegie Institute of Technology

March, 1962

Part of the research underlying this paper was undertaken for the project Temporal Planning and Management Decision under Risk and Uncertainty at Northwestern University and part for the project Planning and Control of Industrial Operations at Carnegie Institute of Technology. Both projects are under contract with the U. S. Office of Naval Research. Contract Nonr-1228(10), Project NR 047-021, and Contract Nonr-760(01), Project NR 047011.

SYSTEMS RESEARCH GROUP
A. Charnes, Director

DUALITY IN SEMI-INFINITE PROGRAMS
AND SOME WORKS OF HAAR AND CARATHEODORY

by

A. Charnes, W. W. Cooper and K. Kortanck*

Foreword

The following paper was stimulated by a paper of the Hungarian mathematician, A. Haar (possibly one of the all-time great mathematicians). Because it was published in a relatively obscure journal, it has only recently been made generally available through the posthumous publication of his collected works. Since the theorems to be established rest heavily on this work, and Haar's paper is published in German, we present it for ease of access in free translation in the appendix.

Introduction

Conjectured by vonNeumann and proved by Gale, Kuhn and Tucker [1], the dual theorem of linear programming has been unique among dual extremal (or variational) principles (see, for example, K. Friedrichs [2] for classical mathematical physics principles, and J. Dennis [3] and W. S. Dorn [4] for more recent use of Legendre transformations to establish dual "quadratic" programming principles) applying to general systems of constraints involving a finite number of variables in that neither principle contains the variables associated with the other. The theorem has also been shown to be as fundamental for the theory of linear inequalities (see particularly Charnes and Cooper [5] for this approach) as the classic Farkas-Minkowski lemma.

---

* The research of A. Charnes and K. Kortanck at Northwestern University has been supported by O. N. R. contract Nonr-1228(10); that of W. W. Cooper at Carnegie Institute of Technology has been supported by O. N. R. contract Nonr-760(01). Reproduction in whole or in part is permitted for any purpose of the United States Government.

Generalizations to linear mappings between linear topological spaces were forthcoming from S. Karlin and H. F. Bohnenblust [6] (also L. Hurwicz [7]) for the Farkas-Minkowski lemma, and from D. Bratton (also recently K. Kretschmer [8]) for the dual theorem in a brilliant, unpublished but well-known paper [9]. As expected, these generalizations are not nearly as strong or precise as the original finite dimensional theorems.

The Farkas-Minkowski lemma has been extended in another direction in finite dimensional spaces by the Kuhn-Tucker theorem [10] which gives a necessary and sufficient condition for the existence of a minimum to a convex differentiable function over a convex set defined by a finite number of differentiable inequalities subject to certain additional differential-geometric constraint qualifications. This extension is made in terms of an equivalent saddle-point formulation involving additional Lagrangean variables. In general, however, these are no longer related to any dual problem.

Dual theorems for nonlinear functionals and constraints in finite dimensional spaces have been established by W. Fenchel [11] in terms of contact transformations, but the related problems and domains of definition are presented only in highly implicit forms. The most explicit result to date involving non-linearity and with separation of the variables of the dual problems has been achieved by E. Eisenberg [12] in the form of maximization of a concave function homogeneous of the first degree subject to a finite system of inequalities related to the convex homogeneous function to be minimized in the dual problem. The triple is subject to additional qualifications and somewhat implicitly defined dual constraint sets.

Starting from a little-known work of A. Haar [13], we define a notion of dual "Haar" (or "semi-infinite") programs which associate minimization of a linear function of finitely many variables over a convex set defined by an infinite (arbitrary cardinal) system of linear inequalities with maximization of a linear function of infinitely many variables subject to a finite system of linear inequalities. We introduce the notion of "general finite sequence" space for the latter problem and establish that Charnes' theorem (associating linear independence and extreme points) [5] continues to hold and that the

Opposite Sign Property of Charnes and Cooper [14] characterizes, algebraically, convex solution sets spanned by extreme points. (Note: These sets need no longer be bounded.)

We study further, applying the notions of "regularization" to these systems, the straightforward duality relations between these programs, obtaining (1) an extended dual theorem precisely paralleling that of the finite system case, (2) a general dual theorem for the most general case of (finite) convex programs by expressing them in our (infinite) form, and (3) the study of any (real) semi-infinite program is reduced to that of a "Haar" program.

It should be further noted that our dual structure appears to be particularly adapted to probing the borderline between properties which are purely algebraic and those which require topology. Also it appears to offer new possibilities for numerical analysis and effective solution of problems of optimization over convex sets with an infinite number of extreme points since it substitutes direct algebraic manipulation and minimal topology for differential-geometric requirements or qualifications.

Generalized Finite Sequence Spaces as Solutions Spaces

By a generalized finite sequence space, S, with respect to an index set I, we mean the vector space of all (possibly infinite) vectors $[\lambda_i : i \in I]$ over an ordered field F with only finitely many non-zero entries. Let V be a vector space over F and consider a collection of vectors: P_o, $\{P_i : i \in I\}$ in V. We call the subspace R spanned by these vectors the "requirements space, and we call S the "solutions space" because it is in S that the solution set Λ appears, where

$$\Lambda = \{\lambda \in S : \sum_{i \in I} \lambda_i P_i = P_o,\ \lambda \geq 0\} .$$

Clearly Λ is a convex set in S.

<u>Theorem 1</u> (Linear Independence by Association with Extreme Points)

$\lambda \neq 0$ is an extreme point of Λ in S if and only if the non-zero coordinates of λ correspond to coefficients of linearly independent vectors in R.

<u>Proof</u>: Assume that λ is an extreme point of Λ, and let $J = \{i \in I : \lambda_i > 0\}$. Assume on the contrary that the set $\{P_i : i \in J\}$ is linearly dependent. Then there exist β_i for $i \in J$, not all zero, such that

$$\sum_{i \in J} \beta_i P_i = 0 .$$

Define $\beta \in S$ by placing zeros in other coordinate positions. Since J is finite and $\lambda_j > 0$ for all $j \in J$, there exists $k > 0$ such that $\lambda_j + k\beta_j, \lambda_j - k\beta_j > 0$ for all $j \in J$. Set $\lambda^{(1)} = \lambda - k\beta$ and $\lambda^{(2)} = \lambda + k\beta$. Then $\lambda^{(1)} \neq \lambda^{(2)}$ since some $\beta_j \neq 0$. Moreover $\lambda^{(1)}, \lambda^{(2)} \in \Lambda$ and $\lambda = \frac{1}{2}\lambda^{(1)} + \frac{1}{2}\lambda^{(2)}$ implies λ is not an extreme point, which is a contradiction. Hence the set $\{P_j : j \in J\}$ is linearly independent.

On the other hand if $\lambda \in \Lambda$ is such that the set $\{P_j : j \in J\}$ is linearly independent, then λ is an extreme point. For if it were not we could write λ as a convex combination of two distinct points of Λ,

$$\lambda = \mu\lambda^{(1)} + (1-\mu)\lambda^{(2)} \text{ with } \lambda \neq \lambda^{(1)}, \lambda^{(2)}.$$

Now $\sum_{i \in I} \lambda_i^{(1)} P_i = \sum_{i \in I} \lambda_i^{(2)} P_i = \sum_{i \in I} \lambda_i P_i = P_o$; but the non-zero coordinate positions of $\lambda^{(1)}$ appear among those of λ, and therefore the only P_i vectors with non-aero coefficients in $\sum_{i \in I} \lambda_i^{(1)} P_i$ are among $\{P_j : j \in J\}$, a linearly independent set which means that the expression for P_o is unique. Therefore $\lambda_j^{(1)} = \lambda_j$ for all $j \in J$

and therefore $\lambda^{(1)} = \lambda$. Similarly $\lambda^{(2)} = \lambda$; This is a contradiction and we conclude that λ is an extreme point.

Definition: A set K in S is bounded if there exists $M \in F$ such that for any $\lambda \in K$, $|\lambda_i| \leq M$ for all i I (or alternatively, if $\sum_{i \in P} |\lambda_i| \leq M$ all $\lambda \in K$).

Thus if λ, $\alpha \in \Lambda$, the ray $K = \{\lambda + \mu\alpha : \mu \geq 0\}$ is not a bounded set.

Theorem 2 (Opposite Sign Theorem)!

Λ is generated by its extreme points if and only if for any $\alpha \in S$, $\alpha \neq 0$, $\sum_{i \in I} \alpha_i P_i = 0$ implies some α_r and some α_s are of opposite signs.

Proof: Suppose that Λ is generated by its extreme points, and that there exists $\alpha \neq 0$, $\alpha \geq 0$ such that $\sum_{i \in I} \alpha_i P_i = 0$. We will show a contradiction. For any $\mu \geq 0$, $\lambda + \mu\alpha \in \Lambda$ and therefore is a convex combination of extreme points of Λ. The only possible extreme points that could occur in such an expression are those with non-zero coordinate positions among those of $\lambda + \mu\alpha$ for all $\mu \geq 0$. For every such positioning of non-zero entries there corresponds at most one extreme point; otherwise we could express P_o in two different ways with respect to the same linearly independent set. Now if, say, $\lambda + \mu\alpha$ has N non-zero entries, then there are at most $\sum_{m=1}^{n} \binom{N}{m}$ such extreme points, i.e., finite in number. But we saw above that $\lambda + \mu\alpha$, for all $\mu \geq 0$, is in the convex hull of these extreme points, which is impossible since the convex hull of a finite number of points is bounded. Hence there exists no such $\alpha \in S$ as above, and the opposite sign property must hold.

Suppose now that the opposite sign property holds. Given $\lambda \epsilon \Lambda$, we must show that it is a convex combination of a finite number of extreme points in Λ. Suppose λ has N non-zero coordinates and is not an extreme point. We will show that λ is a strictly convex combination of $\lambda^{(1)}, \lambda^{(2)} \epsilon \Lambda$, each of which has at least one fewer positive component than λ. The same construction can be applied to $\lambda^{(1)}$ and $\lambda^{(2)}$ and so on, until an extreme point is encountered. This is a finite process because we will at most encounter 2^{N-1} points with only one non-zero coordinate, each of which is extreme because its associated vector is surely linearly independent. Thus after all extreme points necessary to stop the process are met, we can reverse the steps to express the original λ as a convex combination of these.

It suffices to carry out the first construction. Thus $\sum_{j \epsilon J} \lambda_j P_j = P_o$ and $\sum_{j \epsilon J} \alpha_j P_j = 0$, since the P_j's are linearly dependent. Therefore by the opposite sign assumption some $\alpha_r > 0$ and some $\alpha_s < 0$.

$$\text{Let } \rho_1 = \min_{\alpha_j > 0} \frac{\lambda_j}{\alpha_j} \quad \text{and} \quad \rho_2 = \min_{\alpha_j < 0} \frac{\lambda_j}{|\alpha_j|}, \quad \text{so that } \rho_1, \rho_2 > 0.$$

Set $\lambda^{(1)} = \lambda_j - \rho_1 \alpha_j$ for $j \epsilon J$ and 0 in other components. Similarly

$\lambda^{(2)} = \lambda_j + \rho_2 \alpha_j$. Then $\lambda^{(1)}$ and $\lambda^{(2)} \epsilon \Lambda$ and

$$\lambda = \frac{\rho_2}{\rho_1 + \rho_2} \lambda^{(1)} + \frac{\rho_1}{\rho_1 + \rho_2} \lambda^{(2)}.$$

Now the above minimums will be assumed for some $j_1, j_2 \epsilon J$, and therefore $\lambda^{(1)}$ and $\lambda^{(2)}$ have at least one more zero than does λ.

Q. E. D.

The following example shows that Λ although generated by its extreme points need no longer be bounded.

<u>Example:</u> $\Lambda = \{\lambda : \sum_{k=1}^{\infty} 2^{-k} \lambda_k = 1 \, , \, \lambda \geq 0\}$ evidently has the opposite sign property and therefore is spanned by its extreme points. Since the extreme points are of the form $\lambda^{(k)} = (0, \ldots, 0, 2^k, 0, \ldots)$, where 2^k occurs in the k^{th} position, Λ is unbounded.

<u>The Extended Dual Theorem</u>

We call the following pair of problems formed from the same data "dual semi-infinite" programs.

$$
\begin{array}{llcll}
 & \text{I} & & & \text{II} \\
\min & u^T P_o & & \max & \sum_{i \in I} c_i \lambda_i \\
\text{subject to} & u^T P_i \geq c_i & & \text{subject to} & \sum_{i \in I} P_i \lambda_i = P_o \\
 & & & & \lambda \in S, \ \lambda \geq 0 \, .
\end{array}
$$

When F is the real field, then $u^T \epsilon R_m$ for some m. If in addition the set $M = \{(P_i^T, c_i) : i \epsilon I\}$ is a "canonically closed" set in R_{m+1}, we call these problems dual "Haar" programs. By "canonically closed" is meant that in an equivalent inequality system (for minimization) in which the (P_i^T, c_i) form a bounded set, this set is closed.

In most of his theorems on linear inequalities Haar only includes closure of M as a basic assumption. However, we can see readily (by counter-example) that he really meant a little more than this, that is, that the set M be canonically closed. Haar's theorem on inhomogeneous inequalities is stated as follows.

Theorem 3 (Haar's Theorem on Inhomogeneous Inequalities)

Let $D_k(u_1, u_2, \ldots, u_n) = a_{k1} u_1 + \ldots + a_{kn} u_n + a_{kn+1}$ for all $k \in I$, with $(u_1, u_2, \ldots, u_n)$ viewed as in R_n. If $D(u_1, u_2, \ldots, u_n) = a_1 u_1 + \ldots + a_n u_n + a_{n+1} \geq 0$ is a consequence of the canonically closed system $D_k \geq 0$ $(k \in I)$, then there exist $\lambda_k \geq 0$, $\lambda_o \geq 0$, with at most $n+1$ non-zero such that

$$D(u_1, \ldots, u_n) = \sum_k \lambda_k D_k(u_1, \ldots, u_n) + \lambda_o .$$

The proof of this theorem along with other results of Haar appear in the appendix.

Theorem 4 (Extended Dual Theorem)

For any pair of dual Haar programs precisely one of the following occurs:

(i) $\sup \sum_{i \in I} c_i \lambda_i = \infty$ and I is inconsistent

(ii) $\inf u^T P_o = -\infty$ and II is inconsistent

(iii) I and II are both inconsistent

(iv) $\inf u^T P_o = \sup \sum_{i \in I} c_i \lambda_i = \sum_{i \in I} c_i \lambda_i^*$ for some $\lambda^* \in \Lambda$.

Before we turn to the proof, let us consider some examples of the above situations.

Examples: Let $I = \{1, 2, \ldots\}$

	I	II
(i)	Min u	Max $\sum_k (-1)^k \lambda_k$
	subject to $2^{-k} u \geq (-1)^k$ $k=1,2,\ldots$	subject to $\sum_k 2^{-k} \lambda_k = 1$
	$0 \cdot u \geq 0$	$\lambda_k \geq 0$.

Clearly I is inconsistent. However, for II take k large and even to see that $\lambda = (0, \ldots, 0, 2^k, 0, \ldots)$ is a feasible point which means that the maximum is larger than 2^k. Hence since k can be arbitrarily large Max = $+\infty$.

(ii)

	I	II
	Min u	Max $\sum_k (-\lambda_k)$
	subject to $u(-2^{-k}) \geq -1$	subject to $\sum_{k \in I} (-2^{-k})\lambda_k = 1$
	$0 \cdot u \geq 0$	$\lambda_k \geq 0$

Clearly II is inconsistent, and feasible u consist of $u \leq 0$, i.e., Min = $-\infty$.

(iii)

	I	II
	Min u	Max $\sum_k \lambda_k$
	subject to $u(-2^{-k}) \geq 1$	subject to $\sum_k (-2^{-k})\lambda_k = 1$
	$0 \cdot u \geq 0$	$\lambda_k \geq 0$

Both I and II are inconsistent.

(iv) Consider as the constraint set the points under the curve $y = \tan^{-1}(x)$, with $x \geq 0$ and above the x-axis.

We observe that the equation of the tangent line at the point $(x, \tan^{-1}x)$ is given by $\dfrac{u_2 - \tan^{-1}x}{u_1 - x} = \dfrac{1}{1+x^2}$ and therefore our constraint set is given by the following system of inequalities

$$u_1(1+x^2)^{-1} - u_2 \geq -\tan^{-1}x + x(1+x^2)^{-1} = c_x$$

$$u_2 \geq 0$$

$$-u_2 \geq -\frac{\pi}{2}$$

Let the direct problem, I, be $\min(-u_2)$ subject to the above constraints. In this example the set of coefficient points is

$\{ ((1+x^2)^{-1}, -1, -\tan^{-1}x + x(1+x^2)^{-1}) : x \geq 0 \}$ in addition to

the points $(0,1,0)$ and $(0,-1,-\frac{\pi}{2})$. Clearly this set is compact, since the limit point $(0,-1,-\frac{\pi}{2})$ as $x \to \infty$ is added. Note that $\min -u_2 = -\max u_2 = -\frac{\pi}{2}$ is never attained because feasible points must lie on or under the curve.

For this example, the dual II is as follows.

If we let $P_x = \begin{pmatrix} (1+x^2)^{-1} \\ -1 \end{pmatrix}$, $P_\alpha = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$, $P_\beta = \begin{pmatrix} 0 \\ -1 \end{pmatrix}$, then we have

$$\max \quad \Sigma\, c_x \lambda_x + 0 \cdot \lambda_\alpha + (-\tfrac{\pi}{2})\lambda_\beta$$

$$\text{subject to} \quad \Sigma\, P_x \lambda_x + P_\alpha \lambda_\alpha + P_\beta \lambda_\beta = P_o = \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \quad \text{all } \lambda\text{'s} \geq 0\,.$$

We see that if $\lambda_\beta = 1$ and all other λ's $= 0$, then the maximum $-\frac{\pi}{2}$ is attained.

<u>Lemma 1</u> If both I and II are consistent, then

$$\inf u^T P_o = \sup \sum_{i \in I} u^T P_i \lambda_i = \sum_i c_i \lambda_i^* \qquad \text{for some } \lambda^* \in \Lambda\,.$$

Proof: We observe that

$$u^T P_o = \sum_{i\epsilon I} u^T P_i \lambda_i \geq \sum_{i\epsilon I} c_i \lambda_i . \quad \text{Thus } \inf u^T P_o \geq \sup \Sigma c_i \lambda_i .$$

Let $z^* \equiv \inf u^T P_o$. Then $u^T P_o \geq z^*$ whenever $u^T P_i \geq c_i$ all i.

By Haar's inhomogeneous extension of the Farkas-Minkowski lemma it follows that for all u_j, $u^T P_o - z^* = \sum_{i\epsilon I} (u^T P_i - c_i) \nu_i^* + \nu_o^*$, where $\nu_i^*, \nu_o \geq 0$ and ν^* is in S.

Thus $\quad P_o = \sum_{i\epsilon I} P_i \nu_i^*$, so $\nu^* \epsilon \Lambda$

and $\quad z^* = \sum_{i\epsilon I} c_i \nu_i^* - \nu_o^*$, or $z^* \leq \sum_i c_i \nu_i^*$.

Hence $\quad \sum_i c_i \nu_i^* \geq z^* = \inf u^T P_o \geq \sup \sum_i c_i \lambda_i$,

so that $\quad \sum_i c_i \nu_i^* = \sup \sum_i c_i \lambda_i = \inf u^T P_o$.

Q. E. D.

This may be strengthened to

Theorem 5

The results of the lemma hold if $\{(P_i^T, c_i) : i \epsilon I\}$ has the Farkas-Minkowski property, that is, if for any inequality $D \geq 0$ which is a consequence of the system $D_k \geq 0$, there exist $\lambda_k \geq 0$, $\lambda_o \geq 0$ (only finitely many non-zero) such that $D = \sum_k \lambda_k D_k + \lambda_o$. In the appendix it is shown that canonically closed systems have the Farkas-Minkowski property.

Proof of the Extended Dual Theorem

Adjoining artificial variables and bounding constraints to the given problems, we obtain the following regularized version.[1]

1/ Compare this formulation with the finite problem regularization given in Management Models and Industrial Applications of Linear Programming, A. Charnes and W. W. Cooper, pp. 189-190, Vol. I.

$$\begin{array}{lll} & \mathrm{I}_R & \\ & \min \quad u_o M + u^T P_o & \\ \text{subject to} & u_o + u^T P_i \geq c_i & \\ & u^T I_m \geq -U e_m^T & \\ & -u^T I_m \geq -U e_m^T & \\ & u_o \quad \geq 0 & \end{array}$$

$$\begin{array}{ll} & \mathrm{II}_R \\ & \min \sum_{i\in I} c_i \lambda_i - U e_m^T(\nu^+ + \nu^-) \\ \text{subject to} & \lambda_o + \sum_{i\in I} \lambda_i \leq M \\ & \sum_{i\in I} P_i \lambda_i + I_m(\nu^+ - \nu^-) = P_o \\ & \text{where } \lambda_o, \lambda_i \geq 0,\ i\in I, \text{ and } \nu_j^+, \nu_j^- \geq 0 \end{array}$$

Clearly I_R is consistent, for take $u^T = 0$ and $u_o \geq \sup\{c_i\}$, since $\{c_i : i \in I\}$ is a compact set. As for II_R, take $\lambda_i = 0$ $(i\in I)$ and $\nu^+ - \nu^- = P_o$ with $\lambda_o = M$. Note in addition that the system is still a Haar system because addition of a finite number of coefficient points cannot destroy compactness of the coefficient set. Hence we can apply the lemma to the regularized version to conclude that

$$\inf\{u_o M + u^T P_o\} = \sum_{i\in I} c_i \lambda_i^* - U \sum_{j=1}^{m} \nu_j^*, \quad \text{where } \nu_j = \nu_j^+ - \nu_j^- .$$

Further, the relevant set[1] u_o, u^T for I_R is non-empty and compact since it is the intersection of closed sets (half-spaces) and by regularization is non-empty and bounded. Thus by compactness the "inf" is actually assumed for some u_o^*, u^{*T}. The following possibilities are therefore mutually exclusive and collectively exhaustice for u_o and $\nu^T = (\nu^{+T}, \nu^{-T})$. We tabulate them with the correspondingly numbered conclusion of our theorem.

(i) $u_o^* \neq 0$, $\nu^* = 0$; I no solution, $\sup \sum_i c_i \lambda_i = \infty$

(ii) $u_o^* = 0$, $\nu^* \neq 0$, II no solution, $\inf\ u^T P_o = -\infty$

(iii) $u_o^* \neq 0$, $\nu^* \neq 0$, neither I nor II has a solution

(iv) $u_o^* = 0$, $\nu^* = 0$, both feasible and $\inf u^T P_o = \sup \sum_i c_i \lambda_i = \sum_i c_i \lambda_i^*$ for some λ^* in Λ.

1/ only bounded u_o are relevant .

Although for some purposes it may be vital to deal with a semi-infinite program as presented, for most cases the object of primary interest will be the nature of the solutions to the system $u^T P_i \geq c_i$, $i \in I$. We have already pointed out that any such system is equivalent to one in which the $\{P_i, c_i\}$ are bounded (for each inequality can be separately divided through by the maximum of the absolute value of its P_i, c_i entries). With regularization the most general case is brought under the foregoing by the observation that

<u>Lemma</u>: The canonical closure of a system of linear inequalities has the same solution set as the original system.

<u>Proof</u>: It suffices to show that if $(a^T, \alpha) = \lim_n (a^{(n)^T}, \alpha_n)$ and $a^{(n)^T} + \alpha_n \geq 0$, $n = 1, 2, \ldots$ then $a^T u + \alpha \geq 0$.

Suppose one could have $a^T u + \alpha = -\delta < 0$ with $a^{(n)^T} + \alpha_n \geq 0$ for all n.

But then $$0 \leq a^{(n)^T} + \alpha_n = -\delta - [\,(a^T - a^{(n)^T})\, u + (\alpha - \alpha_n)] \leq -\frac{\delta}{2} \text{ for } n \geq n_o(\delta).$$

This is a contradiction and therefore we conclude $a^T u + \alpha \geq 0$.

Q. E. D.

Thus by reducing them to equivalent Haar programs we have achieved a duality theory for semi-infinite programs as complete as that in the finite situation.

To obtain the general convex programming dual theorem, we move the functional into the constraints and replace it with a linear function as follows. Suppose the direct problem is: min $C(u)$ subject to $G(u) \geq 0$, where $G^T = (\ldots, G_i(u), \ldots)$ is a finite vector of concave functions which defines the convex set W of the u's. Let $u^T P_i \geq c_i$, $i \in I$ be a system of supports for W, and $z - u^T Q_\alpha \geq d_\alpha$, $\alpha \in A$ be a system of supports for $z - C(u) \geq 0$.

Then the direct problem may be rewritten as:

$$\begin{aligned} &\min\ z \\ \text{subject to}\quad & z - u^T Q_\alpha \geq d_\alpha\ , \\ & u^T P_i \geq c_i\ ,\ \alpha \in A\ ,\ i \in I\ . \end{aligned}$$

Thus we have

Theorem 5

Assuming the Farkas-Milkowski property for this system, the extended dual theorem applies to the following dual programs:

$$\begin{aligned} &\min\ z && \max\ \sum_\alpha d_\alpha \mu_\alpha + \sum_i c_i \lambda_i \\ \text{subject to}\quad & z - u^T Q_\alpha \geq d_\alpha && \text{subject to}\quad \sum_\alpha \mu_\alpha = 1 \\ & u^T P_i \geq c_i && -\sum_\alpha Q_\alpha \mu_\alpha + \sum_i P_i \lambda_i = 0 \\ & && \mu_\alpha,\ \lambda_i \geq 0\ . \end{aligned}$$

Since the work of A. Haar utilized above is not available in English, we have prepared a free translation of it which is contained in the following appendix, together with, first, a rendition of the pertinent remarks and theorem of Caratheodory's which Haar refers to and employs.

APPENDIX

A. PRELIMINARY REMARKS BASED ON CARATHEORORY'S (1911) PAPER

In addition to the notion of convexity, Caratheodory's convex sets, K, are closed. He discusses outer points and boundary points of a convex set, agreeing with the usual intuitive meaning. A bounding-hyperplane, h, is a hyperplane which does not meet K, and does not separate points of K, i.e., there can't be points in K for which $h > 0$ and points for which $h < 0$. This concept is extended to cases where K is a closed and bounded set, i.e., K is <u>compact</u>.

Now given any compact set $M \subset R_n$, the space of n-tuples of real numbers, let K be the collection of all points through which no bounding-hyperplane can be drawn for M.

<u>Lemma</u>: K is the smallest convex set containing M.

<u>Proof</u>: K is bounded, for let A = maximum distance of M to the origin O. Then each plane whose distance from O is $> A$ is a bounding-hyperplane for M, and hence each point greater than Adistance from O is not in K. K is closed: for let a be an accumulation point of K. If $a \notin K$, then we can pass a bounding-hyperplane for M through a, say, s, with distance p from M. But this means we can pass bounding-hyperplanes through points of distance less than p from a. Hence a cannot be arbitrarily close to points of K which is a contradiction. Hence $a \in K$, and K is closed in R_n.

Now if $a, b \in K$, $a \neq b$, and c is any point on $\overline{ab}$, then $c \in K$. For if not, pass a bounding-hyperplane s for M through c, then a and b are not on s, and therefore lie on different sides of s. But depending on which side of s M is on we can shift s either through a or b to get a new bounding-hyperplane for M. This is impossible since $a, b \in K$. Hence $c \in K$, and K is convex.

If L is any convex set containing M and $p \notin L$, then we can pass a bounding-hyperplane, s, for L through p. But s is also a bounding-hyperplane for M and therefore $p \notin K$ by definition. Hence $K \subset L$, and K if the smallest convex set containing M.

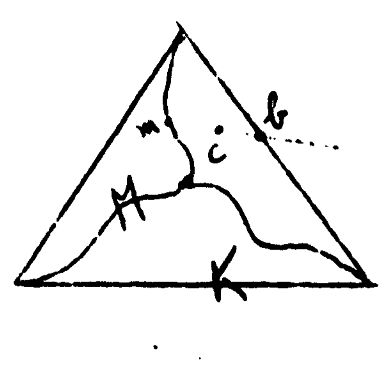
<u>Main Theorem</u>: Let M and K be as above. For any $c \in K$, there exist a finite number of points of m, P_i and masses m_i with $m_i \geq 0$ and

$$\sum_{i=1}^{n} m_i = 1 \qquad \text{such that} \quad c = \sum_{i=1}^{n} m_i p_i .$$

<u>Proof</u>: Dimension $M = 1$: let a and b be the extreme points of M. Then every point of K is of the form $c = ta + (1-t)b$, $0 \leq t \leq 1$.

Assume that the statement is true for (n-1)-dimensional space (or equivalently for (n-1)-dimensional subspaces in R_n).

Let K and M be in R_n. Given $c \in K$, let m be any point in M, and form $\overline{mc}$ and b, the intersection with the boundary of K. If c is interior



to K, then we can find $b \neq m$. Now pass a "supporting-plane," s, through b having properties (1) $b \in S$; (2) all points of K lie on only one side of s; s necessarily intersects M, otherwise it would be a bounding-hyperplane for M. Let M' be this intersection. It has at least 1 lower dimension than M and by assumption there are points $p_i \in M'$ with masses m_i such that $b = \sum_{i=1}^{n} m_i p_i$. Now we know that $c = tm + (1-t)b$ for some t, $0 < t < 1$. Hence $c = tm + \sum_{i=1}^{n} m_i(1-t)p_i$ and $t + \sum_{i=1}^{n} m_i(1-t) = 1$. Hence we have proved the assertion for dimension n, and by induction the assertion holds for all $n \geq 1$.

B. SYSTEMS OF LINEAR INEQUALITIES (THE HAAR PAPER)

1. Homogeneous Inequalities

Let $D_k(u_1, u_2, \ldots, u_n) = a_{k1}u_1 + a_{k2}u_2 + \ldots + a_{kn}u_n$ where k ranges over some indexing set I and a_{ki}, u_i are real numbers. Let $D(u_1, u_2, \ldots, u_n) = a_1u_1 + a_2u_2 + \ldots + a_nu_n$ and view the D_k's and D as linear functionals on R_n. Consider $p_k = (a_{k1}, a_{k2}, \ldots, a_{kn})$, for all $k \in I$, as points in R_n and let $M = \{p_k : k \in I\}$. We say that the system of inequalities $D_k \geq 0$ is closed if M is a closed set in R_n. In most of the theorems Haar includes closure of M as a basic assumption. However, we can see that he really meant a little more than this, that is, that the set M be "canonically closed" in the sense that there exist positive constants $\{c_k : k \in I\}$ such that the set $\overline{M} = \{p_k/c_k : k \in I\}$ is not only closed but bounded (compact). In this case we call the system $D_k \geq 0$ canonically closed. In addition we say that the inequality $D \geq 0$ is a consequence of the system $D_k \geq 0$ if every solution of the system fulfills $D \geq 0$.

<u>Theorem 1</u>: If $D \geq 0$ is a consequence of the canonically closed system $D_k \geq 0$, then there exist non-negative numbers λ_k with at most n of them non-zero such that

$$D = \sum_{k \in I} \lambda_k D_k .$$

<u>Proof</u>: Let $\overline{D} = D_k/c_k$ where $c_k > 0$ are such that $p_k/c_k : k \in I$ is compact. If the theorem is true for $\overline{D}_k$, then it is also true for D_k; for in that case we have

$$D = \sum_{k \in I} \overline{\lambda}_k \overline{D}_k = \sum_{k \in I} (\overline{\lambda}_k/c_k) D_k \text{ so we can set } \lambda_k = \overline{\lambda}_k/c_k .$$

Thus we can assume that $M = \{p_k : k \in I\}$ is compact. Without loss of generality we can assume that there is a vector $\overline{u} = (\overline{u}_1, \overline{u}_2, \ldots, \overline{u}_n)$ in R_n such that $D_k(\overline{u}) > 0$ for all $k \in I$. Otherwise we have $\bigcap_{k \in I} D_k^{-1}\{x : x \geq 0\} = \bigcap_{k \in I} D_k^{-1}(0)$, but the subspace on the right hand side has dimension $\leq n-1$ so that we don't need n variables.

$$\text{Let } P_k = (A_{k1}, A_{k2}, \ldots, A_{kn}) = \frac{1}{D_k(\overline{u})}(a_{k1}, a_{k2}, \ldots, a_{kn}) \text{ for } k \in I .$$

Then P_k all lie in the hyperplane $\overline{E}$ given by $\sum_{i=1}^{n} \overline{u}_i x_i = 1$, since

$\sum_{i=1}^{n} \overline{u}_i A_{ki} = \left(\sum_{i=1}^{n} a_{ki} \overline{u}_i \right) 1/D_k(\overline{u}) = 1$. Consider now the equivalent system $D_k/D_k(\overline{u}) = A_{k1} u_1 + A_{k2} u_2 + \ldots + A_{kn} u_n \geq 0$ where now the coefficient points P_k lie on $\overline{E}$.

__Lemma 1:__ The set $\hat{M} = \{ P_k : k \in I \}$ is compact in R_n.

__Proof:__ For every $p_k = (a_{k1}, a_{k2}, \ldots, a_{kn}) \in M$ define $f_{\bar{u}}(p_k) = D_k(\bar{u})$. Since $D_k(\bar{u}) > 0$ for all $k \in I$, $f_{\bar{u}} > 0$ on M. Clearly $f_{\bar{u}}$ is continuous on M, and therefore assumes its absolute minimum and maximum on M, i.e. we can write $0 < m_1 \leq f_{\bar{u}} \leq m_2$ on M. Now we observe that $\hat{M} = \left\{ \frac{x}{f_{\bar{u}}(x)} : x \in M \right\}$. Immediately we can see that $\hat{M}$ is bounded since M is. $\hat{M}$ is closed; for let $\frac{x_m}{f_{\bar{u}}(x_m)} \longrightarrow y$ as $m \longrightarrow \infty$. Then $\{ x_m \}$, as an infinite sequence in M, has a limit point x in M, and hence a subsequence $x_{mi} \longrightarrow x$ as $i \longrightarrow \infty$. Hence $\frac{x_{mi}}{f_{\bar{u}}(x_{mi})} \longrightarrow \frac{x}{f_{\bar{u}}(x)}$ since $f_{\bar{u}}(x) \neq 0$. Since every subsequence of a convergent sequence converges to that limit of the sequence, we conclude $\frac{x}{f_{\bar{u}}(x)} = y$, i.e. $y \in \hat{M}$, and $\hat{M}$ is closed.

To give this new system a geometric interpretation, let $u_1 = U_1 + \bar{u}_1$, $u_2 = U_2 + \bar{u}_2, \ldots, u_n = U_n + \bar{u}_n$ be a solution of this system and let

$$\theta_k(U_1, U_2, \ldots, U_n) = D_k(U_1 + \bar{u}_1, \ldots, U_n + \bar{u}_n) / D_k(\bar{u}) =$$

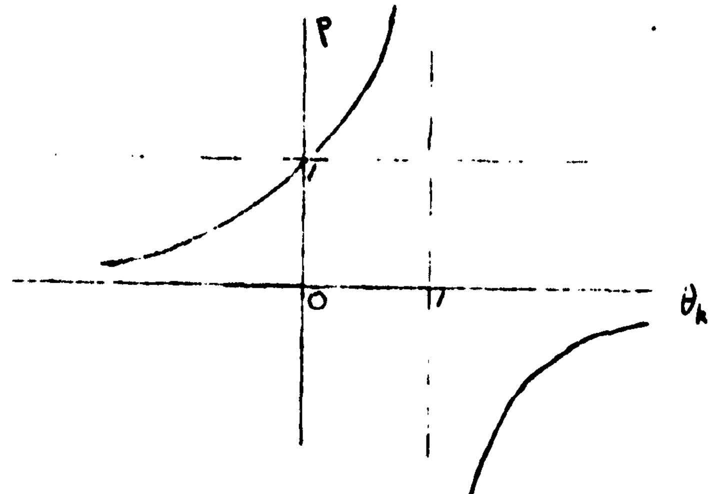
$$= \frac{\sum_{i=1}^{n} a_{ki}(U_i + \bar{u}_i)}{D_k(\bar{u})} = \sum_{i=1}^{n} A_{ki} U_i + 1 \quad \text{for all } k \in I .$$

The system $U_1, U_2, \ldots, U_n$ is then a solution of the inequalities

$$\theta_k(U_1, U_2, \ldots, U_n) \geq 0 \quad (k \in I) .$$

Let us look for intersection points of the hyperplane $U_1 x_1 + U_2 x_2 + \ldots + U_n x_n + 1 = 0$ with the vectors $\overrightarrow{OP_k}$, where O is the origin.

Either there is no intersection with the line $\overrightarrow{OP_k}$ at all in which case $A_{k1}U_1+\ldots+A_{kn}U_n = 0$, or the intersection point has the form.
$x = (pA_{k1},\ldots,pA_{kn})$ where $p = \dfrac{-1}{\theta_k(U_1,\ldots,U_n)-1}$, graphically:



Now $\overrightarrow{OP_k}$ is not cut in its interior if and only if p is not in the interval $\{0 < x \leq 1\}$ which is equivalent (from the graph) to $\theta_k \geq 0$. Hence our hyperplane has no interior intersection with the segment $\overrightarrow{OP_k}$ if and only if $\theta_k(U_1,U_2,\ldots,U_n) \geq 0$.

These hyperplanes can be characterized in another way. We have seen that the points of $\hat{M}$, i.e. $\{P_k : k \in I\}$ all lie in the hyperplane $\bar{E}$ and that $\hat{M}$ is compact. In Caratheodory's sense let $\bar{K}$ be the smallest convex set in $\bar{E}$ containing $\hat{M}$. Connect all points of $\bar{K}$ with the origin O to get an n-dimensional convex set K, i.e. $K = \{(e\bar{x}_1,\ldots,e\bar{x}_n) : (\bar{x}_1,\ldots,\bar{x}_n) \in \bar{K},\ 0 \leq e \leq 1\}$.

Lemma 2: The hyperplane $U_1x_1+\ldots+U_nx_n+1 = 0$ does not have an interior intersection with $\overrightarrow{OP_k}$ for all $k \in I$, if and only if it does not separate K.

Proof: <= : Assume the hyperplane does not separate K, i.e.,

$U_1x_1+\dots+U_nx_n+1$ has the same sign on K; but since $0 \in K$, we have

$U_1x_1+\dots+U_nx_n+1 \geq 0$ on K. Since $P_k \in K$ $(k \in I)$, we have

$0 \leq A_{k1}U_1+\dots+A_{kn}U_n+1 = \theta_k(U_1,\dots,U_n)$ $(k \in I)$ ==>
the hyperplane has no interior intersection with any of the $\overline{OP}_k$.

==> : Assume the hyperplane has no interior intersection with any $\overline{OP}_k(k \in I)$, i.e., $\theta_k(U_1,\dots,U_n) = \sum_{i=1}^{n} A_{ki}U_i + 1 \geq 0$. By the main theorem of Caratheodory, if $(\bar{x}_1,\dots,\bar{x}_n) \in \bar{K}$, we can write $\bar{x}_i = \sum_{k \in I} \mu_k A_{ki}$ with $\mu_k \geq 0$ and $\Sigma \mu_k = 1$ and at most n μ_k's being non-zero. Hence for each k we have

$\sum_{i=1}^{n} \mu_k A_{ki} U_i + \mu_k \geq 0$ and therefore $0 \leq \sum_{i=1}^{n} \left(\sum_k \mu_k A_{ki} \right) U_i + \sum_k \mu_k = \sum_{i=1}^{n} \bar{x}_i U_i + 1$.

Hence for any e, $0 \leq e \leq 1$, we have $0 \leq \sum_{i=1}^{n} (e\bar{x}_i) U_i + 1 = \sum_{i=1}^{n} x_i U_i + 1$; hence for any $(x_1,\dots,x_n) \in K$, $\sum_{i=1}^{n} x_i U_i + 1 \geq 0$, i.e., the hyperplane does not separate K.

Thus far we have been building machinery for the proof of Theorem 1.

Lemma 3 $D_k(\bar{u}) > 0$ for all ==> $D(\bar{u}) > 0$

Proof: We have seen in the proof of lemma 1, that for $f_{\bar{u}}$ defined on M by

$f_{\bar{u}}(p_k) = f_{\bar{u}}(a_{k1},\dots,a_{kn}) = D_k(\bar{u})$, $0 < m_1 \leq f_{\bar{u}} \leq m_2$ on M. It is not hard to show that given $\epsilon > 0$, if u' is in some δ-neighborhood of $\bar{u}$, $N_\delta(\bar{u})$, then $\left| f_{u'}(p_k) - f_{\bar{u}}(p_k) \right| < \epsilon$ for all $p_k \in M$. Choose $\epsilon = \frac{m_1}{2}$ so that for $u' \in N_\delta(\bar{u})$,

$\frac{-m_1}{2} < f_{u'}(p_k) - f_{\bar{u}}(p_k) < \frac{m_1}{2}$ for all $p_k \in M$.

Hence for any $u' \in N_\delta(\bar{u})$,

$$\frac{m_1}{2} < f_{u'}(p_k) < \frac{m_1}{2} + m_2 \implies f_{u'} > 0 \text{ on } M.$$

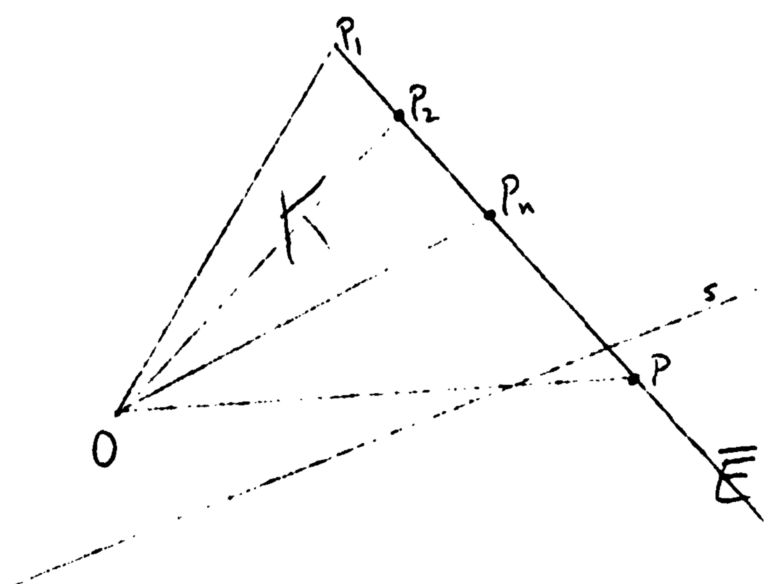
Suppose now $D(\bar{u}) = 0$. Then since D is a linear functional on R_n not identically zero, there exists a $u^* \in R_n$ such that $D(u^*) = 1$.

Set $u' = \bar{u} - \delta \dfrac{u^*}{\|u^*\|}$.

Then $u' \in N_\delta(\bar{u})$ so that $D_k(u') > 0$ for all $k \in I$. Hence since D is a consequence of the system $D_k \geq 0$, we have $D(u') \geq 0$. But $D(u') = 0 - \dfrac{\delta}{\|u^*\|} < 0$ which is a contradiction. Hence $D(\bar{u}) > 0$.

Thus set $A_1 = \dfrac{a_1}{D(\bar{u})}, \ldots, A_n = \dfrac{a_n}{D(\bar{u})}$, where $D(\bar{u}) = \sum_{i=1}^{n} a_i u_i$,

so that $\sum_{i=1}^{n} A_i \bar{u}_i = \left(\sum_{i=1}^{n} a_i \bar{u}_i\right) \dfrac{1}{D(\bar{u})} = 1 \implies$ the point $P = (A_1, \ldots, A_n)$ lies in $\bar{E}$.



<u>Lemma 4</u> Let $(U_1, \ldots, U_n)^{\perp}$ be a hyperplane not separating K, then

$$\theta(U_1, \ldots, U_n) \equiv D(U_1 + \bar{u}_1, \ldots, U_n + \bar{u}_n) / D(\bar{u}) = \sum_{i=1}^{n} A_i U_i + 1 \geq 0 .$$

Proof: For such a hyperplane we know $\theta_k(U_1,\ldots,U_n) \geq 0$ $(k \in I)$ because by lemma 2, this plane has no interior intersection with $\overrightarrow{OP_k}$, i.e., $\theta_k(U_1,\ldots,U_n) \geq 0$. But this means $D_k(U_1+\bar{u}_1,\ldots U_n+\bar{u}_n) \geq 0$, for all k, hence $D(U_1+\bar{u}_1,\ldots,U_n+\bar{u}_n) \geq 0$, hence $\theta(U_1,\ldots,U_n) = \dfrac{D(\bar{U}+\bar{u})}{D(\bar{u})} \geq 0$, which was to be proved.

Consequence: For such hyperplanes, $\theta(U_1,\ldots,U_n) \geq 0$, which means by our previous geometric interpretation that the hyperplane has no interior intersection with $\overrightarrow{OP}$.

Lemma 5: $P = (A_1,\ldots,A_n)$ is in $\bar{K}$.

Proof: We know that $P \in \bar{E}$. Thus it suffices to know $P \in K$. Assume on the contrary that $P \notin K$. Then since K is closed there exist interior points of $\overrightarrow{OP}$ also not in K. We can pass a bounding-hyperplane, s, through one of these since a convex set is the set of all points for which this can't be done. But this contradicts the above consequence of lemma 4. Hence $P \in K$ and therefore $P \in \bar{K}$.

Hence we can write $A_1 = \sum_{k \in I} \mu_k A_{k1},\ldots,A_n = \sum_{k \in I} \mu_k A_{kn}$, $\sum_{k \in I} \mu_k = 1$ with at most n non-zero μ_k. Hence

$$\theta(U) = \sum_{i=1}^{n} A_i U_i + 1 = \sum_k \sum_{i=1}^{n} \mu_k A_{ki} U_i + 1 = \sum_k \mu_k \left(\sum_{i=1}^{n} A_{ki} U_i + 1\right) = \sum_k \mu_k \theta_k(U).$$

Hence
$$D(u) = D(U+\bar{u}) = \frac{\theta(U)}{D(\bar{u})} = \frac{\sum_k \mu_k \theta_k(U)}{D(\bar{u})} = \sum_k \frac{\mu_k \theta_k(U)\, D_k(\bar{u})}{D(\bar{u})\, D_k(\bar{u})} =$$

$$= \sum_{k \in I} \left[\frac{\mu_k D_k(\bar{u})}{D(\bar{u})}\right] D_k(u) = \sum_{k\, I} \lambda_k D_k(u)$$ and theorem 1 is proved,

$\lambda_k \geq 0$, and at most n of them are non-zero.

2. Inhomogeneous Inequalities

Theorem 1 above dealt with the homogeneous case. For the inhomogeneous case let

$$D_k(u_1,\ldots,u_n) = a_{k1}u_1+\ldots+a_{kn}u_n+a_{kn+1},$$

$k\in I$ with $(u_1,\ldots,u_n)$ viewed as in E^n. Let $M=\{p_k=(a_{k1},\ldots,a_{kn}):k\in I\}$ as before.

Theorem 2

If $D(u_1,\ldots,u_n)=a_1u_1+\ldots+a_nu_n+a_{n+1}\geq 0$ is a consequence of the canonically closed system $D_k\geq 0$ $(k\in I)$, then there exist $\lambda_k\geq 0\,(k\in I)$ with at most $n+1$ non-zero such that

$$D(u_1,\ldots,u_n)=\sum_{(k)}\lambda_k D_k(u_1,u_2,\ldots,u_n)+\lambda_o$$

Proof: We can assume existence of $(\bar u_1,\ldots,\bar u_n)=\bar u$ such that $D_k(\bar u)>0$, otherwise $\bigcap_k D_k^{-1}(0)=\bigcap_k D_k^{-1}(\geq 0)$ and $\bigcap_k D_k^{-1}(\geq 0)$ will be a translate of an (n-1)-dimensional subspace so we can reduce the number of variables.

Again, set $\theta_k(U_1,\ldots,U_n)=D_k(U_1+\bar u_1,\ldots,U_n+\bar u_n)/D_k(\bar u_1,\ldots,\bar u_n)$

$$=\Big[\sum_{i=1}^{n}a_{ki}U_i+\Big(\sum_{i=1}^{n}a_{ki}\bar u_i+a_{kn+1}\Big)\Big]/D_k(\bar u_1,\ldots,\bar u_n)$$

$$=\sum_{i=1}^{n}\frac{a_{ki}}{D_k(\bar u)}U_i+1=\sum_{i=1}^{n}A_{ki}U_i+1,\quad \text{where } A_{ki}=\frac{a_{ki}}{D_k(\bar u)}$$

Let $(A_{k1},\ldots,A_{kn})=P_k$.

Then consider the hyperplane, h, given by $U_1x_1+\ldots+U_nx_n+1=0$. Then just as in the homogeneous case we can prove the property:

A) $\overrightarrow{OP_k}$ has no interior intersection with $h \Longleftrightarrow \theta_k(U_1,\ldots,U_n)\geq 0$.

Let K be smallest convex set containing O and P_k. This is the same K as before, and since M is compact the set $\{O, P_k, k\epsilon I\}$ is again compact inside K. As before we see that $D_k(\bar{u}_1,\ldots,\bar{u}_n) > 0 \;\forall k \epsilon I \implies$

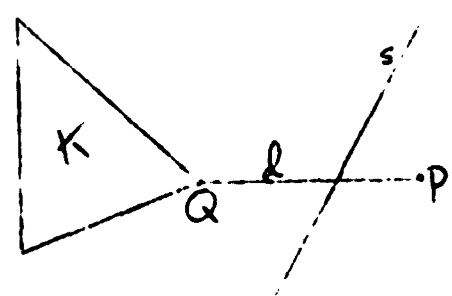
$$D(\bar{u}_1,\ldots,\bar{u}_n) > 0, \text{ and therefore}$$

$$\theta(U_1,\ldots,U_n) = \frac{D(U+\bar{u})}{D(\bar{u})} = \sum_{i=1}^{n} \frac{a_i}{D(\bar{u})} U_i = \sum_{i=1}^{n} A_i U_i + 1 \geq 0$$

Hence we have the following restatement of lemma 2:

B) $D \geq 0$ is a consequence of $D_k \geq 0 \iff$ for any hyperplane $(U_1,\ldots,U_n)^T$ not separating K we have $\theta(U_1,\ldots,U_n) \geq 0$.

Let $P = (A_1,\ldots,A_n)$.



Claim $P \epsilon K$. If not, let d = distance from P to K realized from some point $Q \epsilon K$, since K is compact. Pass a hyperplane, s, through an interior point of $\overrightarrow{QP}$ which does not separate K, say $U_1x_1+\ldots+U_nx_n+1 = 0$. Since $0 \epsilon K$ we know $U_1x_1+\ldots+U_nx_n+1 \geq 0$ on the whole half-space in which K lies. Since P is not this half-space, $\sum_{i=1}^{n} U_iA_i + 1 < 0$ for P, which is impossible by property B). Hence $P \epsilon K$. By Cartheodory's theorem there exist $\mu_0, \mu_k \geqslant 0$ with at most $n+1$ non zero (since the dimension of K is n), such that

$$\sum_k \mu_k + \mu_o = 1 \quad \text{and} \quad A_i = \mu_o \cdot 0 + \sum_{k\epsilon I} \mu_k A_{ki}\,, \quad \text{for } i=1,2,\ldots,n$$

and
$$\sum_{k\epsilon I} \mu_k \theta_k(U_1,\ldots,U_n) = \sum_{k\epsilon I} \mu_k \{\sum_{i=1}^{n} A_{ki} U_i + 1\} =$$

$$= \sum_{i=1}^{n} \Big\{\sum_{k\epsilon I} \mu_k A_{ki}\Big\} U_i + \sum_k \mu_k = \sum_{i=i}^{n} A_i U_i + (1-\mu_o) = \theta(U_1,\ldots,U_n) - \mu_o\,.$$

Hence $\theta(U_1,\ldots,U_n) = \sum_{k\epsilon I} \mu_k \theta_k(U_1,\ldots,U_n) + \mu_o$. Since $D(\bar{u})$ and $D_k(\bar{u}) > 0$ we can make the necessary adjustments to the θ's as before to get

$D = \sum_{k\epsilon I} \lambda_k D_k + \lambda_o$ with $\lambda_k \geq 0$, $k \epsilon I$, and at most $n+1$ of the λ_k's non-zero.

3. Parameter Representation of Linear Inequalities

Since the convex set K is compact, K is bounded by a finite number of hyperplanes, i.e., K is a polyhedron. The hyperplane $\bar{E}$ with equation $\bar{u}_1 x_1 + \ldots + \bar{u}_n x_n = 1$ is one of these and we know that for all $(x_1,\ldots,x_n) \epsilon K$,

$$-\bar{u}_1 x_1 - \bar{u}_2 x_2 - \ldots - \bar{u}_n x_n + 1 \geq 0\,.$$

Since all of the remaining hyperplanes go through the origin 0, they have equations $u_1^{(q)} x_1 + \ldots + u_n^{(q)} x_n = 0$ $(q=1,2,\ldots,N)$, where the coefficients $u_1^{(q)},\ldots,u_n^{(q)}$ can be chosen so that for all points of K we have

$$u_1^{(q)} x_1 + \ldots + u_n^{(q)} x_n \geq 0 \qquad (q=1,2,\ldots,N)\,.$$

Hence the points $(x_1,\ldots,x_n)$ of K can be characterized as those (and only those) satisfying the following $N+1$ inequalities

$$* \begin{cases} u_1^{(q)} x_1 + \ldots + u_n^{(q)} x_n \geq 0 & q=1,2,\ldots,N \\ -u_1 x_1 - \ldots - u_n x_n + 1 \geq 0 & \end{cases}$$

Theorem 3:

$U_1+\bar{u}_1,\ U_2+\bar{u}_2,\dots,U_n+\bar{u}_n$ is a solution of the inequality system $D_k \geq 0 \iff U_1x_1+\dots+U_nx_n+1 \geq 0$ is a consequence of the N+1 inequalities(*).

Proof $\Longrightarrow$ if $D_k(U_1+\bar{u}_1,\dots,U_n+\bar{u}_n) \geq 0, \forall k$, then we have seen that a $(U_1,\dots,U_n)^T$-plane does not separate K, i.e., $U_1x_1+\dots+U_nx_n+1 \geq 0$ if $(x_1,\dots,x_n) \in k$, i.e., if the system (*) holds for $(x_1,\dots,x_n)$.

$\Longleftarrow$: From this assumption it follows that the hyperplane $(U_1,\dots,U_n)^T$ does not separate K $\Longrightarrow$ it does not have an interior intersection with the segments $\overline{OP}_k$ $\Longrightarrow$ $D_k(U_1+\bar{u}_1,\dots,U_n+\bar{u}_n) \geq 0$.

Hence $(U_1+\bar{u}_1,\dots,U_n+\bar{u}_n)$ satisfies $D_k \geq 0$ for all k if and only if there exist non-negative numbers $\lambda^{(q)}$ $(q=1,2,\dots,N)$, $\bar{\lambda}$, and λ_o with at most n+1 different from zero such that

$$U_1x_1+\dots+U_nx_n+1 \equiv \sum_{q=1}^{N} \lambda^{(q)}(u_1^{(q)}x_1+\dots+u_n^{(q)}x_n)+\bar{\lambda}(-\bar{u}_1x_1-\dots-\bar{u}_nx_n+1)+\lambda_o$$

$$\text{or } U_1 = \sum_{q=1}^{N} \lambda^{(q)}u_1^{(q)}-\bar{\lambda}\bar{u}_1,\dots,U_n = \sum_{q=1}^{N} \lambda^{(q)}u_n^{(q)}-\bar{\lambda}\bar{u}_n, \qquad 1=\bar{\lambda}+\lambda_o .$$

This is a consequence of the inhomogeneous case, §2.

Therefore every solution $u_1,u_2,\dots,u_n$ of the systeo of inequalities $D_k \geq 0$ $(k \in I)$ can be put in the following form

$$u_1 = \sum_{q=1}^{N} \lambda^{(q)}u_1^{(q)}+\lambda_o\bar{u}_1, \qquad u_2 = \sum_{q=1}^{N} \lambda^{(q)}u_2^{(q)}+\lambda_o\bar{u}_2,\dots,$$

$$u_n = \sum_{q=1}^{N} \lambda^{(q)}u_n^{(q)}+\lambda_o\bar{u}_n \qquad \text{where } \lambda^{(q)}, \lambda_o \geq 0 .$$

This follows from observing that

$$U_1 + \bar{\lambda}\bar{u}_1 = \sum_{q=1}^{N} \lambda^{(q)} u_1^{(q)} \quad \text{and} \quad U_1 + \bar{\lambda}\bar{u}_1 = U_1 + \bar{u}_1 - \lambda_o \bar{u}_1 \text{ , etc.}$$

The converse of the statement is also true simply by reversing the steps to put things in terms of $U_1, \ldots, U_N$ again and appealing to Theorem 3.

4. Integral Inequalities

Let $a_1(x), a_2(x), \ldots, a_\nu(x)$ and $a(x)$ be continuous functions in the interval $\alpha \leq x \leq \beta$. Let $u(x)$ range over all continuous functions defined in this interval. We say that the inequality

$\int_\alpha^\beta a_1(x)\, u(x)\, dx \geq 0$ is the consequence of the following inequalities

$$\int_\alpha^\beta a_1(x)\, u(x)\, dx \geq 0 \text{ ,} \quad \int_\alpha^\beta a_2(x)\, u(x)\, dx \geq 0 \text{ ,} \ \ldots \ \int_\alpha^\beta a_\nu(x)\, u(x)\, dx \geq 0$$

if it is satisfied by all continuous $u(x)$ which satisfy the above system. We assume that the functions $a_1(x), a_2(x), \ldots, a_\nu(x)$ are all linearly independent. We then assert:

Theorem: In this case we can write $a(x)$ in the following form:

$$a(x) \equiv \lambda_1 a_1(x) + \lambda_2 a_2(x) + \ldots + \lambda_\nu a_\nu(x), \text{ where } \lambda_1, \lambda_2, \ldots, \lambda$$

constants.

Proof: We can solve for $\lambda_1, \lambda_2, \ldots, \lambda_\nu$ from the following system of linear equations:

$$\int_\alpha^\beta a_1(x)\, a(x)\, dx = \lambda_1 \int_\alpha^\beta a_1(x)^2 dx + \lambda_2 \int_\alpha^\beta a_1(x)\, a_2(x)\, dx + \ldots + \lambda_\nu \int_\alpha^\beta a_1(x)\, a_\nu(x)\, dx$$

$$\int_\alpha^\beta a_2(x)\, a(x)\, dx = \lambda_1 \int_\alpha^\beta a_2(x) a_1(x) dx + \lambda_2 \int_\alpha^\beta a_2(x)^2\, dx + \ldots + \lambda_\nu \int_\alpha^\beta a_2(x)\, a_\nu(x)\, dx$$

$$\vdots$$

$$\int_\alpha^\beta a_\nu(x)\, a(x)\, dx = \lambda_1 \int_\alpha^\beta a_\nu(x)\, a_1(x) dx + \lambda_2 \int_\alpha^\beta a_\nu(x)\, a_2(x)\, dx + \ldots + \lambda_\nu \int_\alpha^\beta a_\nu^2(x)\, dx$$

The matrix of this system is positive definite, for let

$$A = (a_{pq}) = \left(\int_{\alpha}^{\beta} a_p(x)\, a_q(x)\, dx \right) \quad \text{be the linear transformation.}$$

Then we have to show $(Ax|x) = 0 \Longrightarrow x = 0$. But

$$(Ax|x) = \sum_{p,q=1}^{n} x_p x_q a_{pq} = \sum_{p,q=1}^{n} x_p x_q \int_{\alpha}^{\beta} a_p(x)\, a_q(x)\, dx =$$

$$= \int_{\alpha}^{\beta} [x_1 a_1(x) + x_2 a_2(x) + \dots + x_\nu a_\nu(x)]^2\, dx .$$

Hence $\quad (Ax|x) = 0$ implies $x_1 a_1(x) + x_2 a_2(x) + \dots + x_\nu a_\nu(x) = 0$ for $x \in [a, \beta]$

$\Longrightarrow x_1 = x_2 = \dots = x_\nu = 0$ since $a_p(x)$'s are linearly independent.

But this means A is non-singular; for if $Ax = 0$, then from $|(Ax,x)| \leq \|Ax\| \|x\|$ we have that $(Ax, x) = 0$ and hence $x = 0$.

Hence there exists a unique solution $\lambda_1, \lambda_2, \dots, \lambda_\nu$.

Let $w(x) \equiv a(x) - \lambda_1 a_1(x) - \lambda_2 a_2(x) - \dots - \lambda_\nu a_\nu(x)$. Then

$$\int_{\alpha}^{\beta} a_1(x)\, w(x)\, dx = 0 , \quad \int_{\alpha}^{\beta} a_2(x)\, w(x)\, dx = 0 , \quad \int_{\alpha}^{\beta} a_\nu(x)\, w(x)\, dx = 0 .$$

Similarly for $-w(x)$. Hence as a consequence we have

$$\int_{\alpha}^{\beta} a(x)\, w(x) \geq 0 \quad \text{and} \quad - \int_{\alpha}^{\beta} a(x)\, w(x) \geq 0$$

$\Longrightarrow \int_{\alpha}^{\beta} a(x)\, w(x)\, dx = 0$. Now replace $a(x)$ by $w(x) + \lambda_1 a_1(x) + \dots + \lambda_\nu a_\nu(x)$

to get $\quad \int_{\alpha}^{\beta} w(x)^2\, dx = 0$ which implies $w(x) \equiv 0$, i.e.,

$$a(x) = \lambda_1 a_1(x) + \lambda_2 a_2(x) + \dots + \lambda_\nu a_\nu(x) .$$

We have to show now that $\lambda_i \geq 0$ $(i=1, 2, \ldots, \nu)$

Consider the system of equations:

$$\int_\alpha^\beta a_2(x)\, a_1(x)\, dx = \mu_2 \int_\alpha^\beta a_2(x)^2\, dx + \mu_3 \int_\alpha^\beta a_2(x)\, a_3(x)\, dx + \ldots + \mu_\nu \int_\alpha^\beta a_2(x) a_\nu(x)\, dx$$

$$\int_\alpha^\beta a_3(x)\, a_1(x)\, dx = \mu_2 \int_\alpha^\beta a_3(x)\, a_2(x)\, dx + \mu_3 \int_\alpha^\beta a_3(x)^2\, dx + \ldots + \mu_\nu \int_\alpha^\beta a_3(x)\, a_\nu(x)\, dx$$

$$\int_\alpha^\beta a_\nu(x)\, a_1(x)\, dx = \mu_2 \int_\alpha^\beta a_\nu(x)\, a_2(x)\, dx + \mu_3 \int_\alpha^\beta a_3(x)^2\, dx + \ldots + \mu_\nu \int_\alpha^\beta a_\nu(x)^2\, dx$$

Since $a_2(x), a_3(x), \ldots, a_\nu(x)$ are linearly independent we see that the matrix of this system is non-singular. Hence there is a unique solution $\mu_2, \mu_3, \ldots, \mu_\nu$.

Let $\bar{w}(x) = a_1(x) - \mu_2 a_2(x) - \mu_3 a_3(x) - \ldots - \mu_\nu a_\nu(x)$

so that we can write the above equalities in the form

$$\int_\alpha^\beta a_2(x)\, \bar{w}(x)\, dx = 0\ , \quad \int_\alpha^\beta a_3(x)\, \bar{w}(x)\, dx = 0, \ldots, \quad \int_\alpha^\beta a_1(x)\, \bar{w}(x)\, dx = 0\ .$$

Furthermore $\int_\alpha^\beta a_1(x)\, \bar{w}(x)\, dx = \int_\alpha^\beta \bar{w}(x)^2\, dx > 0$ otherwise $a_1(x)$ would be a linear combination of the other $a_i(x)$'s .

Hence $\bar{w}(x)$ is a solution of our integral inequalities and as a consequence we have $\int_\alpha^\beta a(x)\, \bar{w}(x)\, dx \geq 0$. But

$$\int_\alpha^\beta a(x)\, \bar{w}(x) dx = \int_\alpha^\beta [\lambda_1 a_1(x) + \lambda_2 a_2(x) + \ldots + \lambda_\nu a_\nu(x)]\ \bar{w}(x)\, dx = \lambda_1 \int_\alpha^\beta \bar{w}(x)^2\, dx \geq 0\ .$$

Hence $\lambda_1 \geq 0$. Similarly we argue for $\lambda_2, \ldots, \lambda_\nu$ by removing the appropriate $a_i(x)$ from $\{a_1(x), a_2(x), \ldots, a_\nu(x)\}$, and considering the smaller system of linear equalities. Hence the assertion is proved.

BIBLIOGRAPHY

[1] Gale, D., H. W. Kuhn and A. W. Tucker, Activity Analysis of Production and Allocation, T. C. Koopmans, ed., New York, J. Wiley and Sons, 1951.

[2] Friedrichs, K., "Ein Verfahren der Variationsrechnung das Minimum eines Integral als das Maximum eines anderen Ausdruckes Darzustellen," Math. Ziet., 1929, pp. 13-20.

[3] Dennis, J. B., Mathematical Programming and Electrical Networks, New York, J. Wiley and Sons, 1959.

[4] Dorn, W. S., "Duality in Quadratic Programming," Quart. Appl. Math., 1960, pp. 155-162.

[5] Charnes, A., and W. W. Cooper, Management Models and Industrial Applications of Linear Programming, Vols. I and II, New York, J. Wiley and Sons, 1961.

[6] Bohnenblust, H. F., and S. Karlin, "On a Theorem of Ville," Contribs. Th. Games, Vol. I, Princeton U. Press, 1950.

[7] Hurwicz, L., Studies in Linear and Non-Linear Programming, K. J. Arrow, ed., Stanford U. Press, 1958.

[8] Kretschmer, K., "Programmes in Paired Spaces," Can. Jour. Math., 1961, pp. 221-238.

[9] Bratton, D., "The Duality Theorem in Linear Programming," Cowles Comm. Disc. Paper, Math. No. 427, Yale U., 1955.

[10] Kuhn, H. W., and A. W. Tucker, "Non-Linear Programming," Proc. 2nd Berkeley Symp. Math. Stat. and Prob., J. Neyman, ed., U. Calif. Press, 1950.

[11] Fenchel, W., "Convex Cones, Sets and Functions," ONR Project Rep., Math. Dept., Princeton U., 1956.

[12] Eisenberg, E., "Duality in Homogeneous Programming," Proc. Am. Math. Soc. 12, No. 5, Oct. 1961, pp. 783-787.

[13] Haar, A., "Uber Lineare Ungleichungen," Acta. Math. (Szeged) 2, 1924, pp. 1-14.

[14] Charnes, A., and W. W. Cooper, "The Strong Minkowski-Farkas-Weyl Theorem for Vector Spaces over Ordered Fields," Proc. Nat. Acad. Sci. U.S.A., 44, No. 9, Sept. 1958, pp. 914-916.