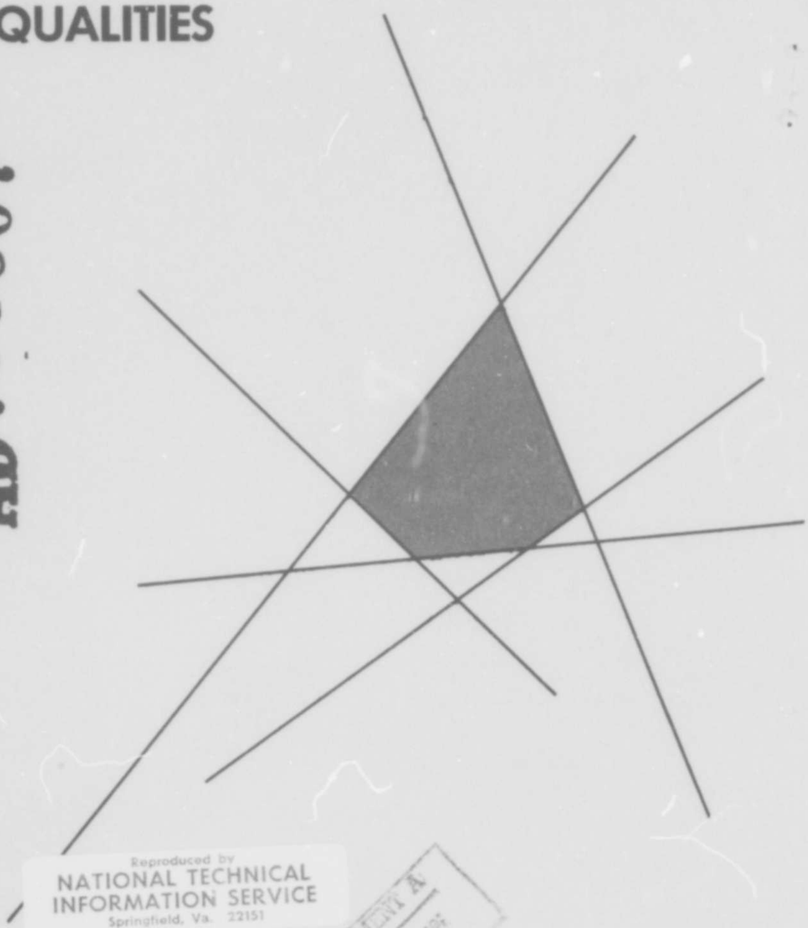
ORC 71-36
DECEMBER 1971

ON THE FINITE CONVERGENCE OF THE RELAXATION METHOD FOR SOLVING SYSTEMS OF INEQUALITIES

by

Jean Louis Goffin



OPERATIONS RESEARCH CENTER

R

COLLEGE OF ENGINEERING

UNIVERSITY OF CALIFORNIA · BERKELEY

156

ERRATA

ON THE FINITE CONVERGENCE OF THE RELAXATION METHOD FOR SOLVING SYSTEMS OF INEQUALITIES

by

Jean Louis Goffin
ORC 71-36
December 1971

Page 103, Lines -1 and -2, replace by:

Case 1. If $C_p(x^*)$ is smooth enough, then the proof given on Page 78 is valid here.

Page 130, Line -7, replace by:

(i) There is one obvious root of equation (*) *squared*:

$$\alpha^q = \frac{\pi}{2} - \alpha .$$

ADDENDUM TO

ON THE FINITE CONVERGENCE OF THE RELAXATION METHOD FOR SOLVING SYSTEMS OF INEQUALITIES

by

Jean Louis Goffin

ORC 71-36
December 1971

1. Using a more delicate method of bounding some parameters, we have improved Theorem (3.3.1).

The value $\frac{2}{1+\nu(P)}$ given for λ^* can be replaced by:

(a) $\lambda^* = 1$ if $\nu(P) \geq \frac{\sqrt{2}}{2}$.

(b) $\lambda^* = \dfrac{2}{1 + 2\nu(P)\sqrt{1 - \nu^2(P)}}$, if $\nu(P) \leq \frac{\sqrt{2}}{2}$.

Comparing to the example of infinite convergence given in (A.1.2.1) this means that this new version of Part (ii) of our main Theorem (3.3.1) *cannot be improved*, at least within the framework of describing a polyhedron by its inner measure $\nu(P)$.

2. The smooth enough property completely characterizes the class of polyhedra for which finite convergence can be *guaranteed* for any value of the relaxation parameter between 1 and 2 : if P is not smooth enough, then an initial point can be chosen such that infinite convergence occurs for $\lambda = 1$.

This means that Part (i) of Theorem (3.3.1) *cannot be improved at all.*

3. In Theorem (3.3.1), Lemma (3.3.2) and Corollary (3.3.3), a crucial assumption was not made explicit, even though it was used in their

proofs: the polyhedron P is defined by the intersection of a *finite* family of halfspaces. These results can fail if this assumption is lifted.

4. Corollary (3.3.3) can be strengthened to:

A necessary and sufficient condition for infinite convergence is: $\sum_{q=0}^{\infty} (\varepsilon^q)^2 = +\infty$. This also ensures finite convergence under the assumptions of Theorem (3.3.1).

5. As a result of this, in (4.1.2), the assumption of Lines -5 and -4, Page 102, can be weakened to: Every index of I is repeated infinitely often.

6. Lemma (3.3.2) and Corollary (3.3.3) are still valid verbatim if the polyhedron P is defined by the intersection of an infinite family of halfspaces, under the following assumption: in choosing a halfspace H at a point x of the relaxation sequence, any positive linear combination of the halfspaces *violated* by x is permitted (the original family still has to be finite).

ON THE FINITE CONVERGENCE OF THE RELAXATION METHOD FOR SOLVING SYSTEMS OF INEQUALITIES

by

Jean Louis Goffin
Operations Research Center
University of California, Berkeley

DECEMBER 1971 ORC 71-36

This research has been partially supported by the Office of Naval Research under Contract N00014-69-A-0200-1010 and the National Science Foundation under Grant GP-30961X with the University of California.

Unclassified

Security Classification

DOCUMENT CONTROL DATA - R & D

(Security classification of title, body of abstract and indexing annotation must be entered when the overall report is classified)

1. ORIGINATING ACTIVITY (Corporate author) University of California, Berkeley	2a. REPORT SECURITY CLASSIFICATION Unclassified 2b. GROUP

3. REPORT TITLE

ON THE FINITE CONVERGENCE OF THE RELAXATION METHOD FOR SOLVING SYSTEMS OF INEQUALITIES

4. DESCRIPTIVE NOTES (Type of report and inclusive dates)

Research Report

5. AUTHOR(S) (First name, middle initial, last name)

Jean Louis Goffin

6. REPORT DATE December 1971	7a. TOTAL NO. OF PAGES 141	7b. NO. OF REFS 69
8a. CONTRACT OR GRANT NO. N00014-69-A-0200-1010 b. PROJECT NO. NR 047 033 c. Research Project No.: RR 003 07 01 d.	9a. ORIGINATOR'S REPORT NUMBER(S) ORC 71-36 9b. OTHER REPORT NO(S) (Any other numbers that may be assigned this report)	

10. DISTRIBUTION STATEMENT

This document has been approved for public release and sale; its distribution is unlimited.

11. SUPPLEMENTARY NOTES Also supported by the National Science Foundation under Grant GP-30961X.	12. SPONSORING MILITARY ACTIVITY Mathematical Sciences Division Office of Naval Research Department of the Navy Arlington, Virginia 22217

13. ABSTRACT

SEE ABSTRACT.

DD FORM 1 NOV 65 1473 (PAGE 1)

S/N 0101-807-6811

Unclassified

Security Classification

Unclassified

Security Classification

14 KEY WORDS	LINK A		LINK B		LINK C	
	ROLE	WT	ROLE	WT	ROLE	WT
Relaxation Method						
Systems of Inequalities						

DD FORM 1 NOV 65 1473 (BACK)

S/N 0101-807-6821

Unclassified

Security Classification

A-11409

ACKNOWLEDGEMENTS

To my parents, my woman, my friends, and my self, whose infinite faith counterpoised my infinite self doubts.

To Professor Richard M. Karp, whose finite faith and unusual dedication was what made the balance positive and this work possible.

To the other members of my committee, Professor David Gale and Professor Pravin Varaiya.

To Ms. Ruth Suzuki, who will be remembered as the woman of the perfect type.

ABSTRACT

The main concern of this work is to rejuvenate the relaxation method for solving linear inequalities, which uses as primitive the notion of hyperplanes, instead of the more derived concept of vertices or bases. The main result is that the method converges finitely for a wide range of values of the relaxation parameter. The smooth enough property is defined, and it delineates a class of problems where the method works particularly well. It is hoped that the relaxation method might become a powerful alternative to the decomposition, or column generation, techniques for large scale programs in which the theoretical finiteness of the simplex method breaks down to a practical transfiniteness.

SYNOPSIS

We investigate the problem of finding a solution to a system of linear inequalities.

The relaxation method constructs a sequence of points $\{x^q\}$ in the following manner: at step q we find the constraint most violated by the point x^q of the sequence, and move on the interior normal of the halfspace H^{i^q} corresponding to this constraint by a distance which is a given multiple λ--the relaxation parameter--of the distance between x^q and H^{i^q}.

In the first chapter, we introduce briefly the results obtained by Agmon, Motzkin, and Schoenberg. They show that if $\lambda \in (0,2)$, and if the solution set P is not empty, then the sequence generated by the relaxation method converges to a point of P; furthermore, if $\lambda = 2$, and if P is full dimensional, then the sequence terminates at a point of P after a finite number of steps.

The concept of Fejér-monotone sequences, due to Motzkin and Schoenberg is introduced, and their main result is reproduced. It will allow us to prove compactly the convergence of the relaxation sequence $\{x^q\}$, as it is, for $\lambda \in (0,2]$, a Fejér-monotone sequence with respect to P.

In order to reduce the proofs of the convergence of the relaxation method to their essence, we gather in the second chapter the tools, classical and new, which this study requires, and we prove various results of a more technical nature.

Some elementary definitions from the theory of lattices are presented, as the semantics involved clarify some arguments.

In paragraphs 2, 3, and 4, we prove some basic results pertaining to the theory of convex sets and polyhedra, by using the properties of the projection map onto a closed convex set K and the family of its inverse images--the normal cones associated to K.

Let P be a polyhedron, then every point x of R^n is shown to have a unique decomposition x = y+z, such that y and z belong to dual elements of the face lattice $\mathcal{F}(P)$ of P, and of the normal cone complex Γ_P associated to P; y is the projection of x on P, and z belongs to $N_P(y)$, the normal cone to P at y.

The smooth enough property for polyhedra is defined. It delineates a class of polyhedra which are as easily tractable by the relaxation method as smooth convex sets.

The basic quantitative concept in this study of the relaxation method is that of the measure of a cone C: we define the outer measure $\mu(C)$ and the inner measure $\nu(C)$. The outer measure is shown to be always greater than the inner measure. We also prove that the inner measure is determined by some maximal linearly independent subset of the halfspaces defining C.

In paragraph 7, the outer measure is used to show that the distance and the residual distance to a polyhedron are equivalent distance functions.

The supporting hyperplane theorem is extended, and we define the smooth enough property for convex sets: smooth compact convex bodies are seen to be smooth enough.

Chapter III is the theoretical core of this work.

We first show the infinite convergence of the relaxation method, and derive a bound on the rate of geometric convergence.

The finiteness of the relaxation method applied to a convex body defined by its complete family of supporting halfspaces holds for any $\lambda \in [1,2]$.

Paragraph 3 contains the main theorems of this work.

Let $P = \bigcap_{i \in I} H^i$ be a full dimensional polyhedron, where I is a finite index set; then the sequence generated by the relaxation method converges finitely to a point of P if:

(i) P is smooth enough, and $\lambda \in [1,2]$

or (ii) if $\lambda \in (\frac{2}{1 + \nu(P)}, 2]$, where $\nu(P)$ is the inner measure of P.

Furthermore, if $\lambda > 2$, then the sequence either converges finitely to a point of P, or it does not converge.

In various specific cases, the finiteness of the relaxation method is proved for some values of the relaxation parameter greater than 2.

The behavior of the relaxation method applied to an inconsistent system of inequalities is examined in paragraph 4: after a finite number of steps, the sequence remains in a neighborhood of the set of Čebyšev points of the system, if $\lambda \in (0,2)$.

The relaxation method is extended to a system of concave inequalities, and the infinite convergence of the sequence is shown.

In the last chapter, we establish the validity of a number of computational procedures: the maximal residual method, scanning of the index set, partitioning, budding, and anti-jamming.

Some intuitive guidelines for the choice of a sequence of relaxation parameters are given: a good strategy will start with values of λ close to 1, and as the sequence approaches P, λ is increased to or above 2.

The application of the relaxation method to the linear programming problem, and to the transportation problem (taken as an example of the column generation techniques) is briefly examined: both these problems are transformed into Čebyšev problems.

We conclude by attempting to justify the hope that the relaxation method will give rise to some good algorithms of solution of these problems.

TABLE OF CONTENTS

PAGE

CHAPTER I

THE RELAXATION METHOD: STATE OF THE NOTION

1.0 Primitive Versus Basic

The main concern of this work is to rejuvenate a rather old method of solving linear inequalities, which uses as primitive the notion of hyperplanes, instead of the more derived concept of vertices or bases.

It is hoped that this approach might grow to become a powerful alternative to the decomposition, or column generation, techniques for large scale programs, in which the theoretical finiteness of the simplex algorithm breaks down to a practical transfiniteness.

The relaxation method has the rather surprising feature that, fixing the dimensionality of the underlying space, the convergence is eased by a larger number of hyperplanes, to the point of becoming trivial if the complete family of supporting hyperplanes to the feasible set is used.

This thesis is dedicated to proving that this claim is more than just a quip.

1.1 Preliminaries

Throughout this work, R represents the real number system, and R^n is the n-dimensional Euclidean space, that is the vector space of n-tuples of reals topologized by the Euclidean norm.

We denote by:

$\langle x,y\rangle = \sum_{i=1}^{n} x_i y_i$ the *inner product* of two vectors

$x = (x_1,\dots,x_n) \in R^n$

$y = (y_1,\dots,y_n) \in R^n$

$\|x\| = \sqrt{\langle x,x\rangle}$ the *Euclidean norm*

$S_r(x) = \{y \in R^n \mid \|x-y\| \leq r\}$ the *closed solid sphere* of center x and radius r

$B = \{x \in R^n \mid \|x\| = 1\}$ the *unit n-1 dimensional spherical surface*

$R_+ = \{x \in R \mid x \geq 0\}$ the *set of non-negative real numbers*

$Z = \{x \in R \mid x \text{ is integer}\}$ the *set of integers*

$Z_+ = R_+ \cap Z$ the *set of non-negative integers*

$S_1 \times S_2 = \{(x,y) \mid x \in S_1,\ y \in S_2\}$ the *cartesian product* of the sets S_1 and S_2

Set inclusion will be denoted by $\subseteq$ and strict inclusion by $\subset$.

Let S be a subset of R^n and we define:

$$\mathscr{L}(S) = \{x \in R^n \mid x = \sum_{i\in I} \mu^i x^i,\ x^i \in S,\ \mu^i \in R,\ I \text{ finite}\}$$

the *linear hull of* S.

$$\mathscr{C}(S) = \{x \in R^n \mid x = \sum_{i\in I} \mu^i x^i,\ x^i \in S,\ \mu^i \in R_+,\ I \text{ finite}\}$$

the *convex cone hull of* S.

$$\mathscr{M}(S) = \{x \in R^n \mid x = \sum_{i\in I} \mu^i x^i,\ x^i \in S,\ \sum_{i\in I} \mu^i = 1,\ I \text{ finite},\ \mu^i \in R\}$$

the *affine (or manifold) hull of* S.

$$\mathcal{H}(S) = \{x \in R^n \mid x = \sum_{i\in I} \mu^i x^i,\ x^i \in S,\ \sum_{i\in I} \mu^i = 1,\ \mu^i \in R_+,\ I \text{ finite}\}$$

the *convex hull of* S.

A set K is *convex* if $K = \mathcal{H}(K)$.

A set C is a *convex cone* if $C = \mathcal{C}(C)$.

A set M is a *manifold* if $M = \mathcal{M}(M)$.

A set L is a *subspace* if $L = \mathcal{L}(L)$.

A set C is a *cone* if $x \in C \Rightarrow \mu x \in C$ for all $\mu \in R_+$.

Those properties are all closed under intersection, which implies that $\mathcal{H}(S)$, $\mathcal{C}(S)$, $\mathcal{M}(S)$, $\mathcal{L}(S)$, are respectively the smallest convex set, convex cone, manifold, subspace containing S. In addition $\mathcal{M}(S)$ and $\mathcal{L}(S)$ are closed sets.

The *dimension of a set* S is the dimension of the linear subspace parallel to $\mathcal{M}(S)$ through the origin:

$$\dim S = \dim(\mathcal{M}(S) - x^o)$$

with $x^o \in \mathcal{M}(S)$. If S is empty, we set $\dim S = -1$.

A convex set K is a *convex body* if $\mathcal{M}(K) = R^n$, or equivalently if it has an interior point.

A point x of a set S is a *relatively interior point* if it is an interior point of S with respect to the relative topology induced in $\mathcal{M}(S)$ by the Euclidean topology, i.e., $x \in$ r.i.(S) if there exists $\varepsilon > 0$ such that $S_\varepsilon(x) \cap \mathcal{M}(S) \subseteq S$.

The *relative interior* r.i.(S) of a set S is the set of its relatively interior points.

Every non-empty convex set has a relatively interior point (see [55], pp. 45-46).

A set S is relatively open if $S = r.i.(S)$.

We denote by $\bar{S}$ the *closure* of S, by ∂S the *boundary* of S, and by Int S the interior of S.

The *addition and subtraction* of two sets S_1 and S_2 are defined by:

$$S_1 + S_2 = \{z = x+y \mid x \in S_1,\ y \in S_2\} \quad ,$$

$$S_1 - S_2 = \{z = x-y \mid x \in S_1,\ y \in S_2\} \quad .$$

If $S_2 = \{x\}$ is a singleton, we will use the notations S_1+x, and S_1-x for $S_1+\{x\}$, and $S_1-\{x\}$.

The *set difference* will be:

$$S_1/S_2 = \{x \in S_1 \mid x \notin S_2\} \quad .$$

The *orthogonal complement of a set* S is:

$$S^{\perp} = \{y \in R^n \mid \langle y,x\rangle = 0, \text{ for all } x \in S\} \quad .$$

It is the largest subspace which is orthogonal to S, and whose intersection with $S/\{0\}$ is void:

$$S^{\perp} = \mathscr{L}(S^{\perp}) = (\mathscr{L}(S))^{\perp} \quad .$$

The sum of two sets S_1 and S_2 which are orthogonal, i.e., $\langle x,y\rangle = 0$ for all $x \in S_1$, $y \in S_2$ is called their *direct sum* and denoted by $S_1 \oplus S_2$.

The *distance from a point* x *to a set* S is

$$d(x,S) = \operatorname*{Inf}_{z \in S} \|x-z\| \quad .$$

If S is closed, we define

$$P_S(x) = \{z \in S |\ \|x-z\| = d(x,S)\}$$

the *set of points closest to* x *in* S.

Let M be a manifold in R^n, x be a point of M, and $r \in R_+$ be a positive real number, then we define $S_r(M,x)$ the *spherical surface with axis* M, *center* x *and radius* r by

$$S_r(M,x) = x + ((M-x)^{\perp} \cap \{z \in R^n |\ \|z\| = r\}) \quad .$$

1.2 The Relaxation Method for Linear Inequalities

In this section we summarize the results obtained in [1] and [51] by Agmon, Motzkin and Schoenberg on the convergence of the relaxation method applied to the solution of a finite system of linear inequalities in R^n.

(1.2.1) Let $\langle a^i,x\rangle + b^i \geq 0$, $i \in I$ be a finite system of linear inequalities where $a^i \in R^n$, $b^i \in R$, $x \in R^n$.

We will assume without loss of generality that $\|a^i\| = 1$ for all $i \in I$.

(1.2.2) The *solution set* is

$$\begin{aligned} P &= \{x \in R^n |\ \langle a^i,x\rangle + b^i \geq 0,\ \forall i \in I\} \\ &= \{x \in R^n |\ Ax + b \geq 0\} \quad . \end{aligned}$$

The problem to be solved is to find a point of P, which we will assume non-void.

Each of the inequalities defines a closed halfspace in R^n:

$$H^i = \{x \in R^n \mid \langle a^i, x\rangle + b^i \geq 0\} \quad .$$

The solution set is a convex polyhedron:

$$P = \bigcap_{i \in I} H^i \quad .$$

The boundary of the halfspace H^i is the hyperplane

$$E^i = \{x \in R^n \mid \langle a^i, x\rangle + b^i = 0\} \quad .$$

(1.2.3) The *relaxation method* constructs a sequence of points in the following manner:

(i) Choose $x^o \in R^n$ arbitrary.

(ii) If $x^q \in P$ the sequence terminates.

If $x^q \notin P$, then some linear inequality is violated.

Let i^q be the index of the most violated constraint:

$$-\langle a^{i^q}, x^q\rangle - b^{i^q} \geq -\langle a^i, x^q\rangle - b^i \quad \text{for all} \quad i \in I$$

and t^q the projection of x^q on E^{i^q}:

$$t^q = x^q + (-\langle a^{i^q}, x^q\rangle - b^{i^q}) \cdot a^{i^q} \quad .$$

(iii) Define

$$x^{q+1} = x^q + \lambda^q(t^q - x^q) \quad , \quad \lambda^q \in R_+ \quad .$$

This procedure is called the relaxation method because at each step we attempt to satisfy only one violated constraint, ignoring or relaxing, the other ones.

We will study the convergence of the relaxation method for a fixed value of the parameter λ. This should be looked upon as a working assumption rather than a restriction. It will follow from the nature of the proofs that if a theorem is valid for some values of λ contained in a closed interval, then it will be valid if the sequence $\{\lambda^q\}$ is contained in this same interval.

According to the values of λ, the *relaxation parameter*, the method has been called:

$\lambda = 1$	projection
$\lambda = 2$	reflexion
$0 < \lambda < 1$	underrelaxation
$1 < \lambda < 2$	overrelaxation
$\lambda > 2$	overreflexion

(1.2.4) <u>THEOREM</u> (Agmon). Let $P = \bigcap_{i \in I} H^i$ be the solution set of a consistent system of linear inequalities, and let $\{x^q\}$ be the sequence defined by the projection method, that is $\lambda = 1$.

Then either the process terminates or the sequence $\{x^q\}$ converges to a solution x^*.

Furthermore,

$$\|x^q - x^*\| \leq 2d(x^o, P)\theta^q$$

where $\theta \in (0,1)$ depends only on the matrix A.

(1.2.5) <u>THEOREM</u> (Motzkin, Schoenberg). Assume $\dim P = n$. Let $\{x^q\}$ be the sequence generated by the relaxation method. Then therefore it converges:

$$r(z) = \lim_{q\to\infty} \|x^q - z\| \quad .$$

Let

$$B(z) = \{x \in R^n \mid \|x-z\| = r(z)\}$$

and

$$X = \bigcap_{z\in S} B(z) \quad .$$

a) Every limit point x^* of $\{x^q\}$ is in X, because $x^* \in B(z)$ for all $z \in S$. So X is not empty as the bounded sequence $\{x^q\}$ has at least one limit point.

b) Assume that X contains more than one point, say x^*, x'^*. Let $E = \{x \in R^n \mid \|x-x^*\| = \|x-x'^*\|\}$ the hyperplane equidistant to x^* and x'^*. So $z \in S$ implies $x^*, x'^* \in B(z)$. Hence $z \in E$, and we conclude that $S \subseteq E$, a contradiction.

c) The bounded sequence $\{x^q\}$ has only one limit point, so it converges.

<u>Case 2</u>: Assume that the sequence does not converge. Let x^* be a limit point of the bounded sequence $\{x^q\}$. Clearly

$$X = \{x \in R^n \mid \|x-z\| = \|x^*-z\|, \forall z \in S\}$$

as $r(z) = \|x^*-z\|$. Hence, by Lemma (2.3.10)

$$X = \{x \in R^n \mid \|x-z\| = \|x^*-z\|, \forall z \in \mathcal{M}(S)\}$$

is a spherical surface with axis $\mathcal{M}(S)$.

It contains all the limit points of the sequence. Q.E.D.

CHAPTER II

TOOLS

2.1 Lattices (see Birkhoff [5])

(2.1.1) A *poset* P is a set on which a binary relation $x \leq y$ is defined, which satisfies for all x, y, z the following conditions:

(1) For all x, $x \leq x$ (reflexive).

(2) If $x \leq y$ and $y \leq z$ then $x \leq z$ (transitivity).

(3) If $x \leq y$ and $y \leq x$ then $x = y$ (antisymmetry).

Example. Let Σ be a class of sets. It is a poset under set inclusion.

$$x \leq y \text{ means } x \subseteq y , \quad x, y \in \Sigma \quad .$$

(2.1.2) If they exist, the elements O(P) and 1(P) such that

$$O(P) \leq x \leq 1(P) \quad \text{for all} \quad x \in P$$

are called the *least* and the *greatest* elements of the poset, or its universal bounds.

(2.1.3) A poset which satisfies (4) is said to be totally ordered and is called a *chain*:

(4) Given $x, y \in P$ either $x \leq y$ or $y \leq x$.

(2.1.4) To every partial order $\leq$, we can associate a *strict partial order* <.

$$x < y \text{ if } x \leq y , \quad x \neq y \quad .$$

It is an antireflexive and transitive binary relation.

(2.1.5) The *converse* $\geq$ of a partial ordering $\leq$ is defined by:

$$x \geq y \text{ if } y \leq x \quad .$$

(2.1.6) The *dual* of a poset P is that poset P^D defined by the converse partial ordering on the same elements.

(2.1.7) THEOREM (Duality Principle). The converse of any partial ordering is itself a partial ordering.

(2.1.8) Let P and Q be two posets. A bijection θ is called an (i) *isomorphism* if

$$\theta(x) \leq \theta(y) \text{ if } x \leq y \quad ;$$

If P = Q then it is an *automorphism*;

(ii) *dual-isomorphism* or *anti-isomorphism* if

$$\theta(y) \geq \theta(x) \text{ if } x \leq y \quad .$$

If P = Q it is called an *anti-automorphism.*

(2.1.9) Many important posets are self-dual (anti-isomorphic with themselves). Such self-dualities are called *involutions*.

Examples. The ordering is the set inclusion:

(i) Let P be the class of subsets of a set S, with $\theta(A) = S/A$, the complement of A in S. It defines an involution.

(ii) Let P be the class of closed convex cones in R^n, with $\theta(C) = C^P$, the polar cone associated to C (see 2.2.2). It defines an involution (see 2.3.6).

(2.1.10) To any poset P we can associate a *graph*:

$$G = (P,A) \quad .$$

That is, to each point in P we associate a node, and the set of arcs is:

$$A = \{(x,y) \in P \times P \mid x < y\} \quad .$$

<u>Property</u>. G does not contain any cycles.

(2.1.11) An element a of a poset P is said to *cover* an element $b \in P$ if

$$b < a \quad \text{and} \quad b < x < a \quad \text{for no} \quad x \in P \quad .$$

(2.1.12) The *diagram* of a poset P is the graph $G' = (P,A')$ where

$$A' = \{(x,y) \in P \times P \mid y \text{ covers } x\} \quad .$$

(2.1.13) A finite poset P is said to be *graded* if there exists a function $g: P \to Z$ such that

(i) $x > y$ implies $g(x) > g(y)$.

(ii) x covers y implies $g(x) = g(y) + 1$.

(2.1.14) In a finite poset P with a 0 (universal lower bound), the *dimension or height* h(x) of an element $x \in P$ is the least upper bound of the length of the chains linking 0 and x.

Clearly in a graded finite poset, the height function is

$$h(x) = g(x) - g(0).$$

(2.1.15) An *upper* (*lower*) *bound* of a subset X of a poset P is an element $a \in P$ containing (contained in) every $x \in X$.

(2.1.16) The *least upper* (*greatest lower*) bound is an upper (lower) bound contained in (containing) every other upper (lower) bound.

(2.1.17) A *lattice* is a poset any two elements of which have a G.L.B. or meet or inf denoted by $x \wedge y$ and a L.U.B. or join or sup denoted by $x \vee y$.

Clearly the lattice property can be visualized on the graph of the poset. The intersection of the subgraphs rooted at x and y is a subgraph *rooted* at $x \vee y$. Similarly for $x \wedge y$, in the "converse" graph.

(2.1.18) A lattice L is *complete* if each of its subsets has a L.U.B. and a G.L.B. in L. A finite lattice is trivially complete.

Example. Let P be the class of closed convex cones in R^n, ordered by set inclusion. Then

$$C_1 \wedge C_2 = C_1 \cap C_2$$

$$C_1 \vee C_2 = \overline{\mathscr{C}}(C_1 \cup C_2) = \overline{C_1 + C_2} \quad .$$

This follows from the fact that $\mathscr{C}(C_1 \cup C_2) = C_1 + C_2$ (see [55], p. 22), but $C_1 + C_2$ might not be closed.

Hence, P is a complete lattice.

On the other hand, the similarly defined lattice of polyhedral cones is not complete.

(2.1.19) A *sublattice* of a lattice L is a subset $X \subseteq L$ such that $a \in X$, $b \in X$ imply $a \wedge b \in X$ and $a \vee b \in X$.

In a lattice L, given any element x, the *set of predecessors* $\mathscr{P}(x) = \{y \in L \mid y \leq x\}$, and the *set of successors* $\mathscr{S}(x) = \{y \in L \mid x \leq y\}$ are both sublattices.

Clearly

$$1(\mathscr{P}(x)) = x$$
$$0(\mathscr{S}(x)) = x$$

and if L has universal bounds

$$0(\mathscr{P}(x)) = 0(L)$$
$$1(\mathscr{S}(x)) = 1(L) \quad .$$

2.2 <u>Cones Associated to a Set</u>

Let x be any point in R^n, and S a subset of R^n.

We define

(2.2.1) (i) The *normal cone* to S at x

$$N_S(x) = \{u \in R^n \mid \langle u, z-x\rangle \leq 0, \ \forall z \in S\} \quad ,$$

that is the set of the exterior normals to those closed half-spaces (possibly degenerate) which contain S, and which contain x in their boundary hyperplanes.

Clearly $N_S(x)$ is a non-empty, closed convex cone.

(2.2.2) $N_S(0)$ is called the *polar cone* associated to S, and is denoted by S^P.

We have: $N_S(x) = (S-x)^p$.

(2.2.3) (ii) The *supporting cone* to S at x

$$C_S(x)$$

is the intersection of all the closed halfspaces which contain S-x and which contain the origin in their boundary hyperplanes.

Clearly $C_S(x)$ is a non-empty closed convex cone.

Trivially $C_S(x) = (N_S(x))^p = C_{S-x}(0)$.

(2.2.4) (iii) The *cone of feasible directions of* S at x

$$D_S(x) = \{z \in R^n | \ x + \eta z \in S \text{ for some } \eta > 0\} \cup \{0\}$$

$$D_S(x) = D_{S-x}(0) \quad .$$

(2.2.5) Following are some immediate properties of the polar cones:

(i) $S_1 \subseteq S_2 \Rightarrow S_1^p \supseteq S_2^p$.

(ii) $S^p = \{u \in R^n | \ \langle u,x\rangle \leq 0, \ \forall x \in S\}$

$$= \bigcap_{x \in S} \{u \in R^n | \ \langle u,x\rangle \leq 0\} .$$

$$S^{pp} = \bigcap_{u \in S^p} \{x \in R^n | \ \langle u,x\rangle \leq 0\}$$

(iii) $S \subseteq S^{pp} = C_S(0)$

That is S^{pp} is the intersection of the homogeneous closed halfspaces which contain S.

(iv) $S^p = S^{ppp}$

(v) $S^p = (\mathscr{C}(S))^p$

(2.2.6) A convex body K is said to be *smooth* at a point x of its boundary if it has only one supporting hyperplane at x.

(2.2.7) Dually it is *rotund* at x, if every supporting hyperplane at x, intersects K at x only.

A convex body is smooth (rotund) if it is smooth (rotund) at every point of its boundary.

(2.2.8) THEOREM (Stoer, Witzgall, p. 104). Suppose $S \subseteq R^n$ and $x \in R^n$. Between the normal cone $N_S(x)$, the supporting cone $C_S(x)$, and the cone of feasible directions $D_S(x)$, the following relations hold:

(i) $D_S^P(x) = N_S(x)$

(ii) $N_S^P(x) = C_S(x)$

(iii) $C_S^P(x) = N_S(x)$.

Proof. (ii) has been mentioned earlier. (iii) follows from (i) and (ii):

$$N_S(x) = D_S^P(x) = D_S^{PPP}(x) = N_S^{PP}(x) = C_S^P(x) \quad .$$

We show first that $N_S(x) \subseteq D_S^P(x)$.

Let $u \in N_S(x)$. For all $z \in D_S(x)$, there exists $\eta > 0$ such that $x + \eta z \in S$. Hence $\langle u, x+\eta z-x \rangle \leq 0$ for all $z \in D_S(x)$ for some $\eta > 0$ or $\langle u,z \rangle \leq 0$ for all $z \in D_S(x)$, and $u \in D_S^P(x)$.

We show then that

$$D_S^P(x) \subseteq N_S(x) \quad .$$

Let $u \in D_S^p(x)$ and $z \in S$. Clearly $z-x \in D_S(x)$. Hence

$$\langle u,z-x\rangle \leq 0 \quad \text{for all} \quad z \in S \quad .$$

That is $u \in N_S(x)$. Q.E.D.

(2.2.9) Let $K \subseteq R^n$ be a convex set. A convex subset W of K is an *extreme subset* of K if $x, y \in K$

$$z = \mu x + (1-\mu)y \in W \quad \text{for some} \quad \mu \in (0,1) \Rightarrow x, y \in W \quad .$$

(2.2.10) A subset W of a convex set K is called *exposed* if some linear form assumes its maximum precisely on W.

(2.2.11) A point y of the boundary of a convex body $K \subseteq R^n$ is called a *vertex of order* d, if $\dim N_K(y) = n-d$. If $\dim N_K(y) = n$, then it is called a *vertex* of K.

(2.2.12) We define $L(C)$ or L the *lineality space* of a closed convex cone C by the equivalent statements:

(i) $L(C) = C \cap (-C)$, the largest subspace contained in C.

(ii) $L(C)$ is the smallest non-empty extreme subset of C.

(iii) $L(C) = (C^P)^\perp$.

(iv) If $C = \{x \in R^n | \langle a^i,x\rangle \geq 0,\ i \in I\}$, then

$$L(C) = \{x \in R^n | \langle a^i,x\rangle = 0,\ i \in I\} \quad .$$

(2.2.13) If $L(C) = \{0\}$, C is called *pointed*. A cone C is *blunt* if $\mathscr{L}(C) = R^n$.

(2.2.14) <u>THEOREM</u> (Stoer, Witzgall, p. 60). Every closed convex cone C is the direct sum of its lineality space and the pointed cone $L^{\perp} \cap C$:

$$C = L \oplus (L^{\perp} \cap C) \quad .$$

(2.2.15) Finally, we define $0^{+}(K)$ the *recession cone* of the convex set K by

$$0^{+}(K) = \{y \in R^{n} | \; K+y \subseteq K\} \quad .$$

It is the set of directions to infinity of K. It can be shown that

$$0^{+}(K) = \{y \in R^{n} | \; \{x+\mu y | \; \mu \geq 0\} \subseteq K, \; x \in K\} \quad ,$$

where the right hand side is independent of x. If K is closed, then so is $0^{+}(K)$ (Rockafellar, pp. 60-71).

(2.2.16) Separation Theorem for General Convex Sets (Stoer, Witzgall, pp. 96-99). In R^{n}, two non-empty convex sets K_1, K_2 can be separated by an hyperplane E such that E does not contain $K_1 \cup K_2$ if and only if K_1 and K_2 have disjoint relative interiors, i.e.,

$$[r.i.(K_1)] \cap [r.i.(K_2)] = \emptyset$$

(2.2.17) THEOREM (Rockafellar, p. 164). Let K be a non-empty convex subset of R^{n}, and let $\mathcal{F}$ be the collection of all the relative interiors of the non-empty extreme subsets of K then $\mathcal{F}$ is a partition of K. Every relatively open convex subset of K is contained in one of the sets of $\mathcal{F}$, and these are the maximal relatively open convex subsets of K.

2.3 The Projection Map

(2.3.1) The *distance from a point* x *to a set* $S \subseteq R^n$ is defined by

$$d(x,S) = \inf_{z \in S} \|x-z\| \quad .$$

If S is closed, then the infimum is attained, and the set of points of S where it is attained is the set of points closest to x in S:

$$P_S(x) = \{y \in S \mid \|x-y\| = d(x,S)\} \quad .$$

(2.3.2) If the set K is closed and convex, then $P_K(x)$ reduces to one point, which will also be denoted by $P_K(x)$, that we call the *projection of* x *into* K. Hence $P_K(x)$ is a function (retraction) from R^n to K.

In fact, the existence of a unique closest point for every point of the underlying space is a characterizing property of closed convex sets in finite dimensional normed spaces, if the unit sphere is smooth and rotund (Motzkin, see in [64]).

The uniqueness follows from:

(2.3.3) THEOREM. Let K be a closed convex subset of R^n. Then

$$\{\{z+N_K(z)\},\ z \in K\}$$

is a partition of R^n.

Furthermore the member of the partition to which x belongs is $P_K(x)$, the unique closest point to x in K.

Proof. First we show that

$$x \in y + N_K(y)$$

where y is a closest point to x in K:

$$\|x-y\| = d(x,K) \ , \quad y \in K \ .$$

a) If $x \in K$, then $y = x$ and

$$x = y \in y + \{0\} \subseteq y + N_K(y) \quad .$$

b) If $x \notin K$, let $z \in K$ arbitrary. By the convexity of K:

$$\mu z + (1-\mu)y \in K \text{ for all } \mu \in [0,1] \ .$$

By definition of y and the convexity of K:

$$\|x-y\|^2 \leq \|x - \mu z - (1-\mu)y\|^2 \text{ for all } z \in K \ , \ \mu \in [0,1]$$

$$\|x-y\|^2 \leq \|x-y\|^2 - 2\mu\langle x-y, z-y\rangle + \mu^2\|z-y\|^2 \text{ for all } z \in K, \ \mu \in [0,1]$$

$$\langle x-y, z-y\rangle \leq \frac{\mu}{2}\|z-y\|^2 \text{ for all } z \in K, \ \mu \in (0,1]$$

$$\langle x-y, z-y\rangle \leq 0 \text{ for all } z \in K$$

i.e.,

$$x - y \in N_K(y) \quad .$$

Now assume that x belongs to two members of the family

$$x \in y_1 + N_K(y_1) \ , \quad y_1 \in K \qquad \text{(i)}$$

$$x \in y_2 + N_K(y_2) \ , \quad y_2 \in K \qquad \text{(ii)}$$

(i) means

$$\langle x-y_1, z-y_1\rangle \le 0 \quad \text{for all} \quad z \in K$$

$$\langle x-y_1, y_2-y_1\rangle \le 0 \quad \text{as} \quad y_2 \in K \quad .$$

(ii) similarly

$$\langle x-y_2, y_1-y_2\rangle \le 0 \quad .$$

Adding

$$\langle y_1-y_2, y_1-y_2\rangle \le 0$$

$$\|y_1-y_2\| = 0 \quad .$$

which implies

$$y_1 = y_2 \quad .$$

Q.E.D.

(2.3.4) COROLLARY. The projection map satisfies the Lipschitz condition:

$$\|P_K(x) - P_K(y)\| \le \|x - y\|$$

(Cheney, Goldstein).

Proof.

$$\langle P_K(x)-P_K(y), P_K(y)-y\rangle \ge 0$$

by Theorem (2.3.3)

$$\langle P_K(y)-P_K(x), P_K(x)-x\rangle \ge 0$$

Adding:

$$\langle P_K(x)-P_K(y), x-y-P_K(x)+P_K(y)\rangle \ge 0$$

or

$$\|P_K(x)-P_K(y)\|^2 \le \langle x-y, P_K(x)-P_K(y)\rangle \le \|x-y\|\cdot\|P_K(x)-P_K(y)\|$$

using Schwarz inequality. Q.E.D.

This implies the continuity of the map P_K. Furthermore, if $y \in K$, then $P_K^{-1}(y) = y + N_K(y)$. If $Y \subseteq K$ is closed, it follows that $\bigcup_{y\in Y} (y + N_K(y)) = P_K^{-1}(Y)$ is closed. It is also true that $\bigcup_{y\in Y} N_K(y)$ is closed.

(2.3.5) A set S is the *orthogonal sum* of S_1 and S_2 if every $x \in S$ admits a unique *orthogonal decomposition* $x = y + z$ with $y \in S_1$, $z \in S_2$ and $\langle y,z\rangle = 0$.

$$S = S_1 \boxplus S_2 \quad .$$

(2.3.6) THEOREM. Let C be a closed convex cone in R^n, then

(i) $C^P = \{u \in R^n | \ d(u,C) = \|u\|\}$

(ii) Any $x \in R^n$ admits a unique orthogonal decomposition into elements $y \in C$ and $z \in C^P$, which are respectively the projections of x into C and C^P, i.e., $R^n = C \boxplus C^P$.

(iii) $C^{PP} = C$.

Proof. (i) by Theorem (2.3.3)

$$u \in 0 + N_C(0) = C^P \quad \text{iff} \quad P_C(u) = 0 \quad .$$

(ii) by Theorem (2.3.3)

$$\langle x - P_C(x), z - P_C(x)\rangle \le 0 \quad \text{for all} \quad z \in C.$$

Hence $$(\mu - 1)\langle x - P_C(x), P_C(x)\rangle \le 0 \quad \text{for all} \quad \mu \ge 0,$$

as $$\mu P_C(x) \in C \quad \text{for all} \quad \mu \ge 0 \quad .$$

It implies that:

$$<x-P_C(x),P_C(x)> = 0$$

and $$<x-P_C(x),z> \leq 0 \quad \text{for all} \quad z \in C.$$

Hence $P_C(x)$ and $x-P_C(x) \in C^P$ is an orthogonal decomposition of x into C and C^P.

Now, let $x = y + z$ be an orthogonal decomposition into C and C^P.

Take $u \in C^P$, $v \in C$.

Then, we have

$$\|x-v\|^2 - \|x-y\|^2 = \|v-y\|^2 - 2<v,z> \geq 0$$
$$\|x-u\|^2 - \|x-z\|^2 = \|u-z\|^2 - 2<u,y> \geq 0$$

which implies:

$$y = P_C(x)$$
$$z = P_{C^P}(x)$$

and the uniqueness is proved.

(iii) Let $x \in C^{PP}$. By (ii),

$$x = P_C(x) + P_{C^P}(x) \quad .$$

But $x \in C^{PP}$ implies

$$<x,P_{C^P}(x)> \leq 0 \quad \text{as} \quad P_{C^P}(x) \in C^P \quad .$$

Hence $$<P_C(x)+P_{C^P}(x),P_{C^P}(x)> \leq 0 \quad .$$

It implies, using (ii), that

$$\|P_{C^P}(x)\|^2 \leq 0$$

and $$P_{C^P}(x) = 0 \quad .$$

But, then $$x = P_C(x) \in C.$$

It follows that $C^{PP} \subseteq C$ and trivially $C \subseteq C^{PP}$. Hence $C = C^{PP}$.

Q.E.D.

(2.3.7) <u>COROLLARY</u>. If C is a convex cone in R^n then

$$C^{PP} = \bar{C} \quad .$$

<u>Proof</u>. $\bar{C} = (\bar{C})^{PP}$ by (2.3.6). $(\bar{C})^{PP} \supseteq C^{PP} \supseteq C$. But C^{PP} is closed, thus $C^{PP} = \bar{C}$. Q.E.D.

<u>Remark</u>. It follows from this that $C = C^{PP}$ is a characterizing property of the closed convex cones in R^n.

(2.3.8) <u>COROLLARY</u>. If in Theorem (2.2.6), we require S to be convex, then we also have

(iv) $D_S^{PP}(x) = \overline{D_S(x)} = C_S(x)$.

<u>Proof</u>. If S is convex, then $D_S(x)$ is convex, and (iv) follows from (2.3.7). Q.E.D.

(2.3.9) <u>COROLLARY</u>: Norm-duality (see [45]). Let K be a closed convex subset of R^n and Σ the class of closed halfspaces which contain K:

$$\Sigma = \{H \subseteq R^n \mid H \text{ is a closed half space, } H \supseteq K\} \quad .$$

Let x be any point in R^n. Then the following duality holds:

$$\min_{z\in K}\|x-z\| = \max_{H\in\Sigma} d(x,H) \quad .$$

Furthermore

(i) If $x \notin K$, the minimum and the maximum are unique and given by

$$z_{opt} = P_K(x)$$

$$H_{opt} = \{z \in R^n \mid \langle x-P_K(x), z-P_K(x)\rangle \leq 0\} \quad .$$

(ii) If $x \in K$, $z_{opt} = x$ is unique and any $H \in \Sigma$ is optimal.

Proof. *Weak duality:*

$H \supseteq K$, $z \in K$ implies

$$d(x,H) = \operatorname*{Inf}_{z \in H} \|x-z\| \leq \operatorname*{Inf}_{z \in K} \|x-z\| \leq \|x-z\| \quad .$$

Strong duality: Using the fact that if

$$H = \{x \in R^n \mid \langle a,x\rangle + b \geq 0\}, \text{ with } \|a\| = 1$$

$$d(x,H) = \text{Max}\{0, -\langle a,x\rangle - b\} \quad .$$

It follows trivially from Theorem (2.3.3) that z_{opt} and H_{opt} give equal values to their respective programs. Q.E.D.

(2.3.10) LEMMA. Let M be a manifold in R^n, and x be a point of R^n, then the set

$$\{y \in R^n \mid \|x-z\| = \|y-z\|, \forall z \in M\}$$

is

$$S_{d(x,M)}(M, P_M(x)) \quad .$$

Proof. By Theorem (2.3.3)

$$\|y - P_M(x)\|^2 = d^2(y,M) + \|P_M(y) - P_M(x)\|^2 = d^2(x,M)$$

$$\|x - P_M(y)\|^2 = d^2(x,M) + \|P_M(x) - P_M(y)\|^2 = d^2(y,M) \quad .$$

Hence

$$P_M(y) = P_M(x)$$

$$d(y,M) = d(x,M) \quad . \qquad \text{Q.E.D.}$$

(2.3.11) LEMMA. Let $C \subseteq R^n$ be a closed convex cone. Then $P_C(-C^P) \subseteq -C^P$.

Proof. Using Theorem (2.3.6), for any $u \in -C^P$ we have:

$$u = P_C(u) + P_{C^P}(u) \quad .$$

Hence: $$P_C(u) = u - P_{C^P}(u) \in -C^P \quad . \qquad \text{Q.E.D.}$$

(2.3.12) COROLLARY. Let C be a closed convex cone in R^n. Then $C \cap (-C^P) = \{0\}$ if and only if C (and hence C^P) is a subspace.

Proof. (i) If C is a subspace, then $C^P = C^\perp = -C^P$ is the orthogonal complement of C. Hence

$$C \cap (-C^P) = C \cap C^P = \{0\} \quad .$$

(ii) Assume that $C \cap (-C^P) = \{0\}$. By Lemma (2.3.11), $P_C(-C^P) \subseteq -C^P$. Thus $P_C(-C^P) \subseteq (-C^P) \cap C = \{0\}$. But, by Theorem (2.3.6)

$$C^P = \{u \in R^n | \; P_C(u) = 0\} \quad .$$

Hence $$-C^P \subseteq C^P$$

and $$C^P \subseteq -C^P \quad .$$

But $C^P = -C^P$ implies that C^P is a subspace. Hence C is a subspace. Q.E.D.

(2.3.13) LEMMA. Let C be a closed convex cone in R^n, then $[r.i.(C)] \cap [r.i.(-C^P)]$ is not empty.

Proof. If the intersection is empty, then C and $-C^P$ can be separated, by Theorem (2.2.11). There exists $a \neq 0$ such that

$$\langle a,C\rangle \geq 0 \geq \langle a,-C^P\rangle .$$

Hence

$$-a \in C^P$$

$$-a \in C$$

but $C \cap C^P = \{0\}$, a contradiction. Q.E.D.

(2.3.14) THEOREM. Let C be a closed convex cone in R^n, then the sets $r.i.(C)$ and $r.i.(-C^P)$ have a non zero vector in common if and only if C is not a subspace.

Proof. (i) If C is a subspace

$$C \cap (-C^P) = C \cap C^P = \{0\}$$

and, using Lemma (2.3.13)

$$\emptyset \subset [r.i.(C)] \cap [r.i.(-C^P)] \subseteq C \cap (-C^P) = \{0\} ,$$

hence

$$[r.i.(C)] \cap [r.i.(-C^P)] = \{0\} .$$

(ii) By Theorem (2.2.14), $C \cap L^{\perp}$ is a pointed cone. If C is not a subspace, then

$$C \cap L^{\perp} \supset \{0\} .$$

Also,

$$C^P \cap L^{\perp} = C^P \text{ and } C \cap L^{\perp}$$

are polar cones in the Euclidean space $L^{\perp}$.

It follows from the definition of relative interiors that

$$r.i.[C \cap L^{\perp}] \quad \text{and} \quad r.i.[C^P]$$

are identically defined within $L^{\perp}$ and R^n. Furthermore

$$0 \notin r.i.[C \cap L^{\perp}] \quad .$$

Hence, applying Lemma (2.3.13) in the space $L^{\perp}$, $r.i.[-C^P]$ and $r.i.[C \cap L^{\perp}]$ have a non zero vector in common. And as

$$r.i.[C \cap L^{\perp}] \subseteq r.i.[(C \cap L^{\perp}) \oplus L] = r.i.[C]$$

the theorem follows. Q.E.D.

2.4 Face Lattices of Polyhedral Cones and Polytopes

(2.4.1) A *polyhedron* is the intersection of a finite family of closed halfspaces.

(2.4.2) A *polytope* is a bounded polyhedron.

(2.4.3) A *polyhedral cone* is the intersection of a finite number of homogeneous closed half spaces.

(2.4.4) In those three cases, the notion of extreme and exposed sets coincide and together with the empty set are called *faces*.

Polyhedral sets are much better behaved than general convex sets because they can be described by finite families of hyperplanes, or dually of points and rays.

Basic to the theory of polyhedral convexity is:

(2.4.5) FARKAS' LEMMA.

$$\{x \in R^n \mid \langle a^i, x\rangle \geq 0,\ i \in I,\ I \text{ finite}\}^P = \mathscr{C}\{-a^i \mid i \in I\} \quad .$$

Note: This follows from Theorem (2.3.6) and the observation that $\mathscr{C}\{-a^i \mid i \in I\}$ is a closed convex cone.

Let $P = \{x \in R^n \mid \langle a^i, x\rangle + b^i \geq 0,\ i \in I\}$ be a polyhedron. Then the cones associated to P at a boundary point y of P are polyhedral cones, and have the simple expressions:

$$N_P(y) = \mathscr{C}\{-a^i \mid i \in I(y)\}$$
$$C_P(y) = D_P(y) = \{z \in R^n \mid \langle a^i, z\rangle \geq 0,\ i \in I(y)\}$$

where $I(y) = \{i \in I \mid \langle a^i, y\rangle + b^i = 0\}$, the set of the indices of the constraints which are tight at y.

This means that $D_P(y)$ is closed for all $y \in \partial P$.

If P is a polytope, it follows from Farkas' Lemma, and the finiteness of the set of vertices (or extreme, or exposed points) that $D_P(y)$ is closed for any $y \in R^n$.

Now, let C be a polyhedral cone. The family of all non-empty faces of C: $\mathscr{F}(C) = \{F \neq \emptyset \mid F \text{ is a face of } C\}$ can be endowed with a lattice structure, if we use set inclusion as partial ordering. Clearly

$$F \wedge G = F \cap G$$
$$F \vee G = \bigcap_{H \in \mathscr{S}(F) \wedge \mathscr{S}(G)} H$$
$$\mathscr{S}(F) = \{H \in \mathscr{F}(C) \mid F \subseteq H\}.$$

That is $F \vee G$ is the smallest face which contains both F and G.

(2.4.6) $\mathscr{F}(C)$ is a finite lattice.

(2.4.7) Define the *polar facial map*

$$\mathscr{P}: \mathscr{F}(C) \to \mathscr{F}(C^P)$$

by $\mathscr{P}(F) = F^{\perp} \cap C^P$.

(2.4.8) <u>THEOREM</u> (Stoer, Witzgall, p.70). The polar facial map is an anti-isomorphism from $\mathscr{F}(C)$ onto $\mathscr{F}(C^P)$, and its inverse $\mathscr{P}^{-1}$ is defined by

$$\mathscr{P}^{-1}(G) = G^{\perp} \cap C \quad .$$

(2.4.9) If a face F covers a face G in $\mathscr{F}(C)$, G is called a *facet* of F.

(2.4.10) <u>THEOREM</u> (Stoer, Witzgall, p.71 and Goldman, Tucker). Each face F of C is

(i) the intersection of facets G of C $(F \neq C)$.

(ii) the conical hull of faces H which cover the lineality space of C $(F \neq L)$.

(2.4.11) <u>THEOREM</u> (Stoer, Witzgall, p.71). Let F and G be faces of a polyhedral cone C; if F covers G, then

$$\dim G = \dim F - 1 \quad .$$

This implies that $\mathscr{F}(C)$ is graded by the dimension of the faces.

The height function is:

$$h(F) = g(F) - g(L(C)) = \dim F - \dim L \quad .$$

(2.4.12) THEOREM (Stoer, Witzgall, pp. 71-73). The face lattice $\mathcal{F}(C)$ of a polyhedral cone C is relatively complemented, i.e., if

$$H,\ F,\ G \in \mathcal{F}(C) \quad \text{and} \quad H \subseteq F \subseteq G \quad .$$

Then there exists an $F' \in \mathcal{F}(C)$

$$F \wedge F' = H\ , \qquad F \vee F' = G \quad .$$

Moreover it can be assumed that $\dim F + \dim F' = \dim G + \dim H$.

To each polyhedron $P = \{x \in R^n \mid Ax + b \geq 0\}$, we can associate a polyhedral cone in R^{n+1}, through the homogenization map $\mathcal{G}$:

$$\mathcal{G}(P) = \overline{\{\binom{x}{1} \mid x \in P\}}^{PP} = \{\binom{x}{z} \mid Ax + bz \geq 0,\ z \geq 0\}$$
$$= \mathcal{C}[P \times \{1\}] \quad .$$

The faces of P form also a lattice $\mathcal{F}(P)$ under set inclusion, provided one includes the empty set.

(2.4.13) THEOREM (Stoer, Witzgall, p.75). If P is a polytope then the face lattices $\mathcal{F}(\mathcal{G}P)$ and $\mathcal{F}(P)$ are isomorphic.

The homogenization technique can be used to show that Theorems (2.4.11) and (2.4.12) are still valid in the case of a polytope; Theorem (2.4.11) is true for a polyhedron. Finally Theorem (2.4.10) can be rewritten as:

Each face F of a polyhedron P is the intersection of some facets of P $(F \neq P)$; each face F of a polytope P is the convex hull of some vertices.

For a polytope:

$$h(P) = \dim(F) - \dim(\emptyset) = \dim F + 1 \quad .$$

(2.4.14) The graph corresponding to $\mathcal{F}(P)$ provides a visualization of the face lattice, and will be called the *facial graph*.

Let $y \in P$ be a point of a polyhedron P. By Theorem (2.2.17), there exists a unique face $F(y)$ of $\mathcal{F}(P)$ such that $y \in \text{r.i.}F(y)$.

Clearly, the set of indices of the tight constraints is the same for all $y \in \text{r.i.}F$; hence we can define

$$I(F) = I(y) \quad \text{for all} \quad y \in \text{r.i.}F \quad ,$$

the set of constraints which define F:

$$F = P \cap \{x \in R^n \mid \langle a^i, x \rangle + b^i = 0,\ i \in I(F)\}$$
$$= P \cap [\bigcap_{i \in I(F)} E^i] \quad .$$

It follows that the supporting (normal) cone is identical for all the relatively interior points of a given face.

Hence, it makes sense to define $C_P(F)$ $(N_P(F))$ to be the supporting (normal) cone at any relatively interior point of the face $F \in \mathcal{F}(P)$:

$$C_P(F) = C_P(y) \quad \text{for some (or any)} \quad y \in \text{r.i.}F$$
$$N_P(F) = N_P(y) \quad \text{for some (or any)} \quad y \in \text{r.i.}F \quad .$$

Clearly

$$C_P(F) = \{z \in R^n \mid \langle a^i, z \rangle \geq 0,\ i \in I(F)\}$$
$$N_P(F) = \mathcal{C}\{-a^i \mid i \in I(F)\} \quad .$$

(2.4.15) THEOREM. Let P be a polyhedron in R^n, and $\mathcal{F}(P)$ the lattice of its faces. Then

$$\{(r.i.F + N_P(F)\}, \ F \in \mathcal{F}(P)\}$$

is a finite partition of R^n.

Proof. Follows from Theorems (2.3.3) and (2.2.17). Q.E.D.

Remark. For a closed convex set K in R^n, the same result will hold if we use the lattice of the extreme subsets of K, except of course that the partition is not finite. Furthermore, the lattice structure can then be useless.

(2.4.16) A *geometric cone complex* $\Gamma = \{C^j \mid j \in J\}$ is a finite family of polyhedral convex cones such that:

(i) Every face of an element of Γ is an element of Γ.

(ii) The intersection of any two elements of Γ is a face of each of them.

Clearly, under set inclusion, Γ is a poset. Provided Γ has a greatest element, Γ can be endowed with a lattice structure:

$$C^i \wedge C^j = C^i \cap C^j$$

$$C^i \vee C^j = \bigcap_{C \in \mathcal{S}(C^i) \cap \mathcal{S}(C^j)} C \ .$$

(2.4.17) The family of the normal cones associated to the faces of a given polyhedron P forms such a complex, which we will call the *normal cone complex associated to* P:

$$\Gamma_P = \{N_P(F) \mid F \in \mathcal{F}(P)\} \ .$$

Clearly $R^n = N_P(\emptyset)$ is the greatest element of Γ_P and hence Γ_P is a lattice.

It follows easily from the definition of $C_P(F)$ and $N_P(F)$ that the sublattice $\mathcal{S}(F)$ of $\mathcal{F}(P)$ is isomorphic to the face lattice $\mathcal{F}(C_P(F))$ of $C_P(F)$, and anti-isomorphic to the face lattice $\mathcal{F}(N_P(F))$ of $N_P(F)$.

Hence, the map which to every face F of $\mathcal{F}(P)$ associates the normal cone $N_P(F)$ is an anti-isomorphism from $\mathcal{F}(P)$ onto Γ_P.

In fact, it is possible to give to Γ_P a rather stronger interpretation.

Assume that P is a polytope in R^n such that $0 \in \text{Int } P$. Then $\mathcal{G}(P) = \mathcal{C}[P \times \{1\}] \subseteq R^{n+1}$ is a polyhedral convex cone.

Let $\sigma: R^{n+1} \to R^{n+1}$ be defined by $\sigma((x,z)) = (x,-z)$, $x \in R^n$, $z \in R$, the symmetry map with respect to $z = 0$. Finally, if $\tilde{P} \subseteq R^{n+1}$, let

$$\mathcal{D}(\tilde{P}) = \{x \in R^n \mid (x,1) \in \tilde{P}\} \subseteq R^n .$$

(2.4.18) The *dual polytope* P^D associated to P can be defined by

$$P^D = \mathcal{D}\sigma[\mathcal{G}(P)]^P = \{y \in R^n \mid \langle y,x\rangle \leq 1, \text{ for all } x \in P\} .$$

(2.4.19) The *dual cone* associated to a set P is defined by (see Valentine, pp. 57-67):

$$\{(x,z) \in R^{n+1} \mid z \geq \delta^*(x|P)\}$$

where $\delta^*(x|P) = \sup_{y \in P} \langle x,y\rangle$ is the support function of P (see 3.2.1).

The dual cone is also: $\sigma[\mathcal{G}(P)]^P$.

The face lattices $\mathcal{F}(P)$ and $\mathcal{F}(P^D)$ are anti-isomorphic and the anti-isomorphism is (Grunbaum [30], p.47):

$$F^D = \mathcal{A}(F) = \{y \in P^D \mid \langle x,y \rangle = 1, \text{ for all } x \in F\} \quad .$$

From this it follows that if $P = \{x \in R^n \mid \langle a^i,x\rangle + b^i \geq 0,$ for all $i \in I\}$ is a polytope such that $0 \in \text{Int } P$, i.e., $b^i > 0$, then

$$P^D = \mathcal{H}\{-\frac{a^i}{b^i} \mid i \in I\} \quad .$$

Furthermore

$$F^D = \mathcal{A}(F) = \mathcal{H}\{-\frac{a^i}{b^i} \mid i \in I(F)\} \quad ,$$

and as

$$N_P(F) = \mathcal{C}\{-a^i \mid i \in I(F)\}$$

this entails the following:

(2.4.20) <u>THEOREM</u>. The normal cone complex Γ_P associated to a polytope P having the origin in its interior is:

(i) The family of the cone hulls of the faces of P^D

or (ii) The family of the projections of the faces of the dual cone $\sigma[\mathcal{G}(P)]^P$ on the hyperplane $z = 0$.

(2.4.21) <u>THEOREM</u>. Let $P \subseteq R^n$ be a polytope with the origin in its interior and let x be any point of R^n/P. Then there is a unique decomposition $x = y + \lambda z$ such that

(i) $y \in P$

(ii) $z \in P^D$, and $\lambda > 0$

(iii) y and z belong to dual faces of P and P^D.

__Proof__. Let $y = P_P(x)$, $\lambda z = x - P_P(x)$. Then by Theorems (2.3.3) and (2.4.20) y and z satisfy (i), (ii) and (iii), for some unique $\lambda > 0$.

Conversely, let $x = y + \lambda z$ be such a decomposition, with

$$y \in F$$
$$z \in F^D , \quad \lambda > 0 \quad .$$

Let $F(y)$ be the least face containing y, i.e., the unique face of P such that $y \in r.i.(F(y))$:

$$F(y) = \{x \in P | <a^i,x> + b^i = 0, \; i \in I(y)\} \quad .$$

Clearly $F(y) \subseteq F$, and $I(y) \supseteq I(F)$. Hence, as $z \in F^D$

$$z \in \mathscr{C}\{-a^i | \; i \in I(F)\} \subseteq \mathscr{C}\{-a^i | \; i \in I(y)\}$$

or $\lambda z \in N_P(y)$. Hence

$$x = y + \lambda z \in y + N_P(y)$$

which implies using Theorem (2.3.3) that $y = P_P(x)$ and the uniqueness follows. Q.E.D.

__Remark__. Using the normal cone complex Γ_P, the theorem becomes: Let P be a polyhedron in R^n, then every $x \in R^n$ admits a unique decomposition $x = y + z$ into elements y and z belonging to dual elements of $\mathscr{F}(P)$ and Γ_P.

(2.4.22) A convex polyhedron P in R^n will be said to be *smooth enough at a point* y of its boundary if:

$$-N_P(y) \subseteq C_P(y) \quad .$$

(2.4.23) A convex polyhedron is *smooth enough* if it is smooth enough at every boundary point, or equivalently if:

$$-N_P(F) \subseteq C_P(F) \quad \text{for all} \quad F \in \mathcal{F}(P)/\{\emptyset\} \quad .$$

From this it follows that a polyhedral cone C is smooth enough if and only if $-C^P \subseteq C$.

For any $F \in \mathcal{F}(C)$

$$-N_C(F) \subseteq -N_C(L) = -C^P \subseteq C \subseteq C_C(F) \quad .$$

Similarly, a polytope P is smooth enough if and only if

$$-N_P(F) \subseteq C_P(F) \quad \text{for all} \quad F \in \text{Vert}(P)$$

where $\text{Vert}(P) = \{F \in \mathcal{F}(P) \mid \dim F = 0\}$, the set of vertices of P.

A polyhedron P is smooth enough if and only if

$$-N_P(F) \subseteq C_P(F)$$

for the minimal elements of the poset $\mathcal{F}(P)/\{\emptyset\}$.

(2.3.24) <u>THEOREM</u>. Let $C = \{x \in R^n \mid \langle a^i, x\rangle \geq 0,\ i \in I\}$ be a polyhedral cone, then C is smooth enough if and only if:

(i) $a^i \in C$ for all $i \in I$

or equivalently

(ii) the Gram matrix of the vectors a^i, $i \in I$ is non-negative:

$$\Gamma\{a^i \mid i \in I\} = (\langle a^i, a^j\rangle) \geq 0 \quad .$$

<u>Proof</u>. (i) follows from (2.4.5) as $-C^P = \mathcal{C}\{a^i \mid i \in I\} \subseteq C$

if and only if $a^i \in C$ for all $i \in I$.

(ii) follows trivially from (i). Q.E.D.

Remark. It is easy to see that a closed convex cone C is smooth enough if and only if

$$\|u+v\| \geq \|u\| \quad \text{for all} \quad u, v \in C^P \quad .$$

A cone C^P satisfying this relation has been called "normal" (see Schaeffer, p.1010), and hence a closed convex cone is smooth enough if and only if its polar cone is "normal".

2.5 Spherical Cones

(2.5.1) Given a unit vector $e \in R^n$ and an angle $\alpha \in [0,\pi]$, the set of vectors which make with e an angle not superior to α, is called the *spherical cone* of axis e and half-aperture angle α, and is denoted by:

$$C_\alpha(e) = \{e \in R^n | \langle x,e\rangle \geq \|x\|\cos\alpha\} \quad .$$

(2.5.2) **Properties**

(i) For all $\alpha \in [0,\pi]$, $C_\alpha(e)$ is a closed cone.

(ii) For all $\alpha \in [0,\frac{\pi}{2}]$, $C_\alpha(e)$ is a convex set.

Proof. Let $x_1, x_2 \in C_\alpha(e)$ and $x_1 + x_2 \neq 0$. Then

$$\frac{\langle x_1+x_2,e\rangle}{\|x_1+x_2\|} \geq \left(\frac{\|x_1\|}{\|x_1+x_2\|} + \frac{\|x_2\|}{\|x_1+x_2\|}\right)\cos\alpha$$

$$\geq \cos\alpha \quad ,$$

that is $x_1 + x_2 \in C_\alpha(e)$. Q.E.D.

(iii) For $\alpha \in [0,\frac{\pi}{2}]$, if we let $\nu = \sin\alpha$, then: $C_\alpha(e) \supseteq S_\nu(e)$ and hence $C_\alpha(e) \supseteq \mathscr{C}[S_\nu(e)]$.

__Proof.__ Let $x \in S_\nu(e)$, $x \neq 0$, then:

$$\frac{<e,x>}{\|x\|} \geq \frac{1}{2}[\frac{1-\nu^2}{\|x\|} + \|x\|] \geq \sqrt{1-\nu^2} = \cos\alpha \quad .$$

If $\alpha = \frac{\pi}{2}$, then $x = 0 \in S_1(e)$ and trivially $0 \in C_{\pi/2}(e)$.

Q.E.D.

(iv) For all $\alpha \in [0,\frac{\pi}{2})$, $\mathscr{C}[S_\nu(e)] \supseteq C_\alpha(e)$ and hence $C_\alpha(e) = \mathscr{C}[S_\nu(e)]$.

__Proof.__ Let $x \in C_\alpha(e)$, $x \neq 0$, then

$$\|\mu x-e\|^2 = \mu^2\|x\|^2 - 2\mu<e,x> + 1$$

is minimum for $\mu^* = \frac{<e,x>}{\|x\|^2} > 0$. Hence

$$\|\mu^* x-e\|^2 = 1 - \frac{(<x,e>)^2}{\|x\|^2} \leq 1 - [1-\nu^2] = \nu^2 \quad .$$

Q.E.D.

__Remark.__ If $\alpha = \frac{\pi}{2}$, then $C_{\pi/2}(e) = \overline{\mathscr{C}[S_1(e)]}$.

(v) The ray $\mathscr{C}(x)$ spanned by $x \neq 0$ is an extreme ray of $C_\alpha(e)$, with $\alpha \in [0,\frac{\pi}{2})$ if and only if $<x,e> = \|x\| \cos\alpha$.

(2.5.3) __THEOREM.__ If $C_\alpha(e)$ is the spherical cone of axis e and half-aperture $\alpha \in [0,\frac{\pi}{2}]$, then

$$[C_\alpha(e)]^P = C_{\frac{\pi}{2}-\alpha}(-e) \quad .$$

__Proof.__ Let $L = \mathscr{S}(e)$, the line through 0 and e. Let $x \in C_\alpha(e) \cap B$, where $B = \{x \in R^n | \|x\| = 1\}$. Clearly,

$$u \in [C_\alpha(e)]^P$$

if and only if

$$\underset{x \in C_\alpha(e) \cap B}{\text{Max}} \langle u,x \rangle \leq 0 \quad .$$

We will assume that $u \in B$. As $e \in C_\alpha(e)$, $\langle u,e \rangle \leq 0$ for all $u \in [C_\alpha(e)]^P$. The vectors x and u have a unique orthogonal decomposition into elements of L and $L^\perp$:

$$x = P_L(x) + P_{L^\perp}(x) = \langle x,e \rangle e + P_{L^\perp}(x) \quad .$$

Hence

$$\langle x,u \rangle = \langle x,e \rangle \langle u,e \rangle + \langle P_{L^\perp}(x), P_{L^\perp}(u) \rangle$$

and

$$\left| P_{L^\perp}(x) \right|^2 = 1 - (\langle x,e \rangle)^2$$

$$\left| P_{L^\perp}(u) \right|^2 = 1 - (\langle u,e \rangle)^2 \quad .$$

For any $u \in B$, such that $\langle u,e \rangle \leq 0$, we have, using Schwarz inequality:

$$\underset{x \in C_\alpha(e) \cap B}{\text{Max}} \langle u,x \rangle = \underset{\delta \geq \sqrt{1-\nu^2}}{\text{Max}} [\delta \langle u,e \rangle + \sqrt{1-\delta^2}\sqrt{1-(\langle u,e \rangle)^2}] \quad .$$

Hence any $u \in B$ such that $\langle u,e \rangle \leq 0$ belongs to $[C_\alpha(e)]^P$ if and only if:

$$\langle u,e \rangle \sqrt{1-\nu^2} + \nu\sqrt{1-(\langle u,e \rangle)^2} \leq 0$$

or

$$(\langle u,e \rangle)^2 \geq \nu^2$$

or

$$\langle u,-e \rangle \geq \nu \quad . \qquad \text{Q.E.D.}$$

2.6 Cone Measures

In the study of the finite convergence of the relaxation method, the crucial quantitative concept is that of the measure of a cone. In this paragraph we attempt to quantify the idea of how large a cone is, in two different ways:

Let $C = \bigcap_{i \in I} H^i$ be a closed convex cone given by the intersection of a family I of halfspaces

$$H^i = \{x \in R^n \mid \langle a^i, x\rangle \geq 0\}$$

where the a^i's are normalized

$$\|a^i\| = 1 \text{ for all } i \in I.$$

If I is a finite set, then C is a polyhedral cone.

We will assume only that $\{a^i \mid i \in I\}$ is a non-empty compact subset of $B = \{x \in R^n \mid \|x\| = 1\}$.

Any subset $\{a^i \mid i \in I\}$ of B such that: $\overline{\mathscr{C}}\{a^i \mid i \in I\} = -C^P$ defines the same cone C. Hence, to every closed convex cone C we will associate:

(i) A maximal family of hyperplanes, that we will call the *complete family of supporting hyperplanes*:

$$\{a^i \mid i \in I^*\} = -C^P \cap B$$

(ii) If C is a blunt closed convex cone then there exists a minimal compact subset $\{a^i \mid i \in I_*\}$ of B which defines C. It clearly is defined by the set of extreme rays of $-C^P$.

If, in addition, C is polyhedral, then:

$a^i \in \{a^i \mid i \in I_*\}$ if and only if

$\{x \in C \mid \langle a^i,x\rangle = 0\}$ is a facet of C.

The distance between a point $x \in R^n$ and the halfspace H^i is:

$$d(x,H^i) = \min_{z \in H^i} \|x-z\| = \max\{0,-\langle a^i,x\rangle\} .$$

(2.6.1) We define the *residue function* by

$$R(x) = \max_{i \in I}\{-\langle a^i,x\rangle\} .$$

It is a convex function on R^n and hence it is continuous. If I is finite, then it is a polyhedral function. By definition of C,

$$C = \{x \in R^n \mid R(x) \leq 0\} .$$

(2.6.2) We define the *residual distance* from x to C as the distance from x to the furthest halfspace:

$$d_C(x) = \max\{0,R(x)\} = \max_{i \in I}\{d(x,H^i)\} .$$

$d_C(x)$ is convex on R^n, continuous, positively homogeneous and polyhedral if I is finite. Also

$$C = \{x \in R^n \mid d_C(x) = 0\}$$
$$R^n/C = \{x \in R^n \mid d_C(x) > 0\} .$$

<u>Remark</u>. Neither $R(x)$ nor $d_C(x)$ are stricto sensu defined by C. They depend on the specific family of halfspaces used to define C.

(2.6.3) <u>The Outer Measure of a Cone</u>: The *outer measure* of a closed convex cone $C = \bigcap_{i\in I} H^i$ is defined by:

$$\mu(C) = \operatorname*{Inf}_{u\in C^P} \frac{d_C(u)}{\|u\|} \quad .$$

(2.6.4) <u>THEOREM</u>. The outer measure $\mu(C)$ is positive:

$$\mu(C) > 0 \quad .$$

<u>Proof</u>.
$$\mu(C) = \operatorname*{Inf}_{u\in C^P\cap B} d_C(u)$$
$$= \operatorname*{Inf}_{u\in C^P\cap B} \operatorname*{Max}_{i\in I}\{-\langle a^i,u\rangle\}$$

but $C \cap C^P = \{0\}$ and $d_C(u) > 0$ for all $u \in C^P \cap B$.

Noting that $C^P \cap B$ is compact and $d_C(u)$ is continuous the theorem follows. Q.E.D.

(2.6.5) <u>Properties</u>

(i) $\mu(C)$ depends on the representation of C: If $C = \bigcap_{i\in I} H^i = \bigcap_{i\in J} H^i$ with $\{a^i \mid i \in I\} \subseteq \{a^i \mid i \in J\}$ then

$$\mu(C = \bigcap_{i\in I} H^i) \le \mu(C = \bigcap_{i\in J} H^i) \quad .$$

(ii) $\mu(C)$ can be interpreted as follows:

$$\mu(C) = \operatorname*{Inf}_{u\in C^P\cap B} \{\cos\alpha \mid C_\alpha(u) \cap \{-a^i \mid i \in I\} = \emptyset\}$$
$$= \operatorname*{Max}_{u\in C^P\cap B} \{\cos\alpha \mid C_\alpha(u) \cap \{-a^i \mid i \in I\} \neq \emptyset\} \quad ,$$

that is, the cosine of the half-aperture angle of the largest spherical cone such that its axis belongs to C^P, and it is contained in the cage defined by the rays (bars)

$$\mathscr{C}\{-a^i\}\ ,\quad i \in I.$$

(iii) $\mu(C) = 1$ if and only if C is defined by the complete family $\{a^i \mid i \in I^*\}$ of supporting halfspaces.

(iv) If C is blunt, then $\mu(C)$ can be made to depend on C only by considering the unique minimal family $\{a^i \mid i \in I_*\}$ of defining halfspaces.

Furthermore, it follows from (i) that:

$$\mu(C = \bigcap_{i \in I_*} H^i) = \underset{J \subseteq I^*}{\mathrm{Min}}\ \mu(C = \bigcap_{i \in J} H^i) \quad .$$

In this case, the cage mentioned in (ii) is simply the set of extreme rays of C^P.

(2.6.6) <u>The Inner Measure of a Cone</u>

The *inner measure* $\nu(C)$ of a closed convex cone C is the sine of the half-aperture angle of the largest spherical cone contained in C.

$$\begin{aligned}
\nu(C) &= \mathrm{Sup}\{\nu = \sin\alpha \mid C_\alpha(e) \subseteq C,\ e \in B\} \\
&= \underset{e \in B}{\mathrm{Sup}}\ \underset{r>0}{\mathrm{Sup}}\ \{r \mid S_r(e) \subseteq C\} \\
&= \underset{e \in B}{\mathrm{Sup}}\ \underset{r>0}{\mathrm{Sup}}\ \{r \mid S_r(e) \subseteq H^i,\ i \in I\} \\
&= \underset{e \in C \cap B}{\mathrm{Sup}}\ \underset{i \in I}{\mathrm{Min}}\ \langle a^i, e\rangle \\
&= \underset{e \in B}{\mathrm{Sup}}\ \underset{i \in I}{\mathrm{Min}}\ \langle a^i, e\rangle \quad .
\end{aligned}$$

(2.6.7) If C is not blunt, it is convenient to define the *relative inner measure* $\nu^{r.i.}(e)$ as the inner measure of C considered as a subset of the subspace $\mathscr{S}(C)$.

We emphasize that I is assumed to be non-empty and hence $C = R^n$ is ruled out.

(2.6.8) <u>THEOREM</u>. There is a unique point $e(C) \in C \cap B$ where $\nu(C) = \underset{i\in I}{\text{Min}} \langle a^i, e\rangle$, provided $\nu(C) > 0$. Furthermore $e(C) \in C \cap L^\perp$.

<u>Proof</u>. The existence of a point $e \in C \cap B$ such that $\nu(C) = \underset{i\in I}{\text{Min}} \langle a^i, e\rangle$ follows from the continuity of $\underset{i\in I}{\text{Min}} \langle a^i, e\rangle$ on the compact set $C \cap B$ (The continuity follows from the concavity and the finiteness of $\underset{i\in I}{\text{Min}} \langle a^i, x\rangle$ on R^n).

First assume that $e \notin C \cap L^\perp$, then $e = P_L(e) + P_{L^\perp}(e)$ with $P_L(e) \neq 0$. Hence

$$\nu(C) = \underset{i\in I}{\text{Min}} \langle a^i, e\rangle = \underset{i\in I}{\text{Min}} \langle a^i, P_{L^\perp}(e)\rangle$$

as $a^i \in -C^P \subseteq L^\perp$. Clearly:

$$P_{L^\perp}(e) \in C \cap L^\perp$$

$$P_{L^\perp}(e) \neq 0$$

as $\nu(C) > 0$ and

$$\frac{P_{L^\perp}(e)}{\|P_{L^\perp}(e)\|} \in C \cap L^\perp \cap B .$$

Hence

$$\nu(C) = \|P_{L^\perp}(e)\| \underset{i\in I}{\text{Min}} \langle a^i, \frac{P_{L^\perp}(e)}{\|P_{L^\perp}(e)\|}\rangle$$

and

$$\underset{i\in I}{\text{Min}} \langle a^i, \frac{P_{L^\perp}(e)}{\|P_{L^\perp}(e)\|}\rangle = \frac{\nu(C)}{\|P_{L^\perp}(e)\|} > \nu(C)$$

as

$$\|P_{L^\perp}(e)\|^2 = 1 - \|P_L(e)\|^2 < 1$$

a contradiction with the definition of $\nu(C)$.

Assume now that $e, e' \in C \cap L^{\perp} \cap B$, and

$$\nu(C) = \underset{i \in I}{\mathrm{Min}} \langle a^i, e\rangle = \underset{i \in I}{\mathrm{Min}} \langle a^i, e'\rangle$$

then $\frac{e+e'}{\|e+e'\|} \in C \cap L^{\perp} \cap B$ as $\|e+e'\| \neq 0$, because $C \cap L^{\perp}$ is pointed. Hence

$$\underset{i \in I}{\mathrm{Min}} \langle a^i, \frac{e+e'}{\|e+e'\|}\rangle \geq \frac{1}{\|e+e'\|}(\underset{i \in I}{\mathrm{Min}} \langle a^i, e\rangle + \underset{i \in I}{\mathrm{Min}} \langle a^i, e'\rangle)$$

$$= \frac{1}{\|e+e'\|} 2\nu(C) \quad .$$

But $$\|e+e'\| < \|e\| + \|e'\| = 2 \quad .$$

Hence $$\underset{i \in I}{\mathrm{Min}} \langle a^i, \frac{e+e'}{\|e+e'\|}\rangle > \nu(C)$$

a contradiction and the theorem is proved. Q.E.D.

(2.6.9) <u>COROLLARY</u>. Similarly, there is a unique point $e^{r.i.}(C) \in C \cap B$ where $\nu^{r.i.}(C) = \underset{i \in I^{r.i.}}{\mathrm{Min}} \langle a_o^i, e\rangle$, with $I^{r.i.} = \{i \in I \mid P_{\mathscr{L}(C)}(a^i) \neq 0\}$, and $a_o^i = P_{\mathscr{L}(C)}(a^i) / \|P_{\mathscr{L}(C)}(a^i)\|$ for $i \in I^{r.i.}$ (provided C is not a subspace, i.e., $I^{r.i.} \neq \emptyset$).

(2.6.10) <u>**Properties**</u>

(i) $\nu(C) > 0$ if and only if $\dim C = n$.

(ii) $C_1 \subseteq C_2$ implies $\nu(C_1) \leq \nu(C_2)$.

(iii) $\nu(C)$ does not depend on the representation of C.

(iv) $\nu(C) = 1$ if and only if C is a halfspace.

(v) $\nu(C) = +\infty$ if and only if $C = R^n$.

(vi) Let $C_1 \Subset C_2$ be the partial ordering of the cones defined by $C_1 \Subset C_2$ if $C_1 \subseteq TC_2$ for some rotation T, then

$$C_1 \Subset C_2 \text{ implies } \begin{bmatrix} \dim C_1 \\ \nu^{r.i.}(C_1) \end{bmatrix} <_{\ell ex} \begin{bmatrix} \dim C_2 \\ \nu^{r.i.}(C_2) \end{bmatrix}$$

(vii) If $C_\alpha(e)$ is a spherical cone,

$$\nu(C_\alpha(e)) = \sin \alpha$$

(viii) $\nu(C_1 \vee C_2) \geq \operatorname{Max}\{\nu(C_1),\nu(C_2)\}$.

(ix) If C is a blunt polyhedral cone and $\mathcal{F}(C)$ its face lattice, then $F_1, F_2 \in \mathcal{F}(C)$, $F_1 \subseteq F_2$ implies

$$C_C(F_1) \subseteq C_C(F_2)$$

and hence
$$\nu(C_C(F_1)) \leq \nu(C_C(F_2)) \quad .$$

It follows that:

$$\inf_{x \in \partial C} \nu(C_C(x)) = \min_{F \in \mathcal{F}(C)} \nu(C_C(F)) = \nu(C_C(L))$$
$$= \nu(C_C(0)) = \nu(C) \quad .$$

(2.6.11) We can define the *inner measure of a polyhedron* P by:

$$\nu(P) = \inf_{x \in \partial P} \nu(C_P(x)) = \min_{F \in \mathcal{F}(P)/\{\emptyset\}} \nu(C_P(F)) \quad .$$

Clearly, the minimum needs to be taken only on the minimal elements of $\mathcal{F}(P)/\{\emptyset\}$.

For a polytope P, we have

$$\nu(P) = \min_{F \in \operatorname{Vert} P} \nu(C_P(F))$$

where $\operatorname{Vert} P = \{F \in \mathcal{F}(P) \mid \dim F = 0\}$, the set of vertices of P.

(2.6.12) <u>LEMMA</u>. A sufficient condition for P to be smooth enough is

$$\nu(P) \geq \frac{\sqrt{2}}{2} \quad .$$

Proof. For all $x \in \partial P$, we have $\nu(C_p(x)) \geq \frac{\sqrt{2}}{2}$. Hence, there exists a spherical cone

$$C_{\frac{\pi}{4}}(e) \subseteq C_p(x) \quad .$$

This implies

$$-N_p(x) = -C_p^P(x) \subseteq -C_{\frac{\pi}{4}}^P(e) = C_{\frac{\pi}{4}}(e) \subseteq C_p(x) \quad .$$

Q.E.D.

(2.6.13) THEOREM. Let $C \neq R^n$ be a blunt closed convex cone and $C_\alpha(e)$ be the largest spherical cone contained in C. Then

(i) $C_{\frac{\pi}{2}-\alpha}(-e)$ is the unique smallest spherical cone containing C^P.

(ii) $-e \in C^P$.

Proof. The polarity map is an involution, i.e., $C_\alpha(e) \subseteq C$ if and only if $C^P \subseteq [C_\alpha(e)]^P = C_{\frac{\pi}{2}-\alpha}(-e)$, which implies (i).

Now assume that $-e \notin C^P$. Let

$$e^* = -P_{C^P}(-e) = P_{-C^P}(e)$$

$$e^{**} = e - e^*$$

Lemma (2.3.11) applied to C^P implies:

$$e^* \in C \text{ and } -C^P \quad .$$

Also, by Theorem (2.3.6):

$$-e^{**} \in C \quad .$$

Hence $\quad \langle a^i, e^{**}\rangle \leq 0$ for all $i \in I$

and $\quad \langle a^i, e^*\rangle = \langle a^i, e\rangle - \langle a^i, e^{**}\rangle \geq \langle a^i, e\rangle$ for all $i \in I$.

As $\|e^*\| < 1$, we have

$$\nu(C) = \underset{i\in I}{\text{Min}} \langle a^i, e\rangle \leq \underset{i\in I}{\text{Min}} \langle a^i, e^*\rangle < \underset{i\in I}{\text{Min}} \frac{\langle a^i, e^*\rangle}{\|e^*\|}$$

a contradiction. Q.E.D.

Remark. This theorem could be used to prove Theorem (2.3.14), and in fact, strengthen it.

(2.6.14) COROLLARY. If C is a blunt closed convex cone, then

$$[\nu(C)]^2 + [\nu^{r.i.}(C^p)]^2 \leq 1 \quad ,$$

with equality if and only if $C \cap L^\perp$ is a spherical cone in the subspace $L^\perp$.

Proof. This follows from the previous theorem and the observation that the smallest spherical cone containing $C \cap L^\perp$ and the largest spherical cone contained in $C \cap L^\perp$ are identical if and only if $C \cap L^\perp$ is a spherical cone (the spherical cones being defined in $L^\perp$). Q.E.D.

(2.6.15) THEOREM. Let $C \neq R^n$ be a blunt closed convex cone. Then

$$\nu(C) = \underset{x\in C}{\text{Max}} \ \underset{u\in C^p}{\text{Min}} \frac{-\langle u, x\rangle}{\|u\|\|x\|} \quad .$$

<u>Proof</u>. For any $x \in C$ and $\beta \geq 0$, $\{u \in R^n \mid \frac{-<u,x>}{\|u\|\|x\|} \geq \beta\}$ is a convex spherical cone. This means that $\frac{-<u,x>}{\|u\|\|x\|}$ is a quasiconcave function of u on

$$\{u \in R^n \mid <u,x> \leq 0\} \supseteq C^P \quad .$$

Furthermore it is homogeneous of degree zero. Hence, as C^P is pointed,

$$\underset{u \in C^P}{\text{Min}} \frac{-<u,x>}{\|u\|\|x\|}$$

is attained on extreme rays of C^P. We conclude that: For all $x \in C$,

$$\underset{u \in C^P}{\text{Min}} \frac{-<u,x>}{\|u\|\|x\|} = \underset{i \in I}{\text{Min}} \frac{<a^i,x>}{\|x\|}$$

Q.E.D.

(2.6.16) <u>COROLLARY</u>. If $C \neq R^n$ is a blunt closed convex cone, we have:

$$v^{r.i.}(C^P) = \underset{u \in C^P}{\text{Max}} \; \underset{x \in \mathcal{O}\mathcal{L}^{\perp}}{\text{Min}} \frac{-<u,x>}{\|u\|\|x\|} \quad .$$

A subset S of the unit $n-1$ dimensional spherical surface B is called Robinson-convex if it is the intersection of closed hemispheres, where $B = \{x \in R^n \mid \|x\| = 1\}$, and a closed hemisphere is the intersection with B of a closed halfspace containing the origin in its boundary (see [14], pp. 155-157).

Clearly, if C is a closed convex cone, $C \cap B$ is Robinson-convex, and if S is a Robinson-convex set, then

$$\mathscr{C}(S) = \{x \in R^n \mid x = \mu y, \mu \in R_+, y \in S\}$$

is a closed convex cone.

Also, if $C_\alpha(e)$ is a convex spherical cone, then

$$C_\alpha(e) \cap B = \{u \in B \mid \langle u,e\rangle \geq \cos \alpha\}$$

could be called a cap.

From this it is clear that most of the results of this paragraph can be considered as results for the theory of the Robinson-convex sets.

Surprisingly enough, some properties of the Robinson-convex sets are quite a bit stronger than similar properties of the closed convex sets in R^{n-1}.

For instance, Theorem (2.6.8) could be interpreted as: If a Robinson-convex set $S \subseteq B$ is full-dimensional, then there is a unique largest cap contained in S.

Clearly the uniqueness of a largest sphere contained in a closed convex set $K \subseteq R^{n-1}$ does not hold in general.

To every Robinson-convex set S, we will associate a polar set:

$$S^P = [\mathscr{C}(S)]^P \cap B = \{v \in B \mid \langle u,v\rangle \leq 0, \text{ for all } u \in S\} \quad .$$

(2.6.17) A Robinson-convex set S (a closed convex cone C) will be called *autopolar* if $S = -S^P$ $(C = -C^P)$. They can be seen as the fixed points of the map $S \to -S^P$ $(C \to -C^P)$ of the set of Robinson-convex sets onto itself (respectively closed convex cones).

In order to prove the following theorem we need a variant of

Helly's theorem for convex subsets of B. If every 2n members of a family of Robinson-convex subsets of B have a non-empty intersection, the whole family has a non-empty intersection (see [64], p.175).

For our purpose here, we need a stronger theorem.

We will define a subset S of B to be *strongly convex* by the equivalent statements:

(i) S is Robinson-convex and contained in an open hemisphere.

(ii) S is the intersection with B of a pointed closed convex cone $C \subseteq R^n$.

(2.6.18) <u>THEOREM</u> (Grunbaum [31]). If the intersection of every n or fewer members of a family of strongly convex subsets of $B = \{x \in R^n \mid \|x\| = 1\}$ is non-empty and if the union of every n+1 members of the family fails to cover B, then the intersection of all the members of the family is not empty.

Now, let $C = \{x \in R^n \mid \langle a^i, x\rangle \geq 0, \text{ for all } i \in I\} \neq R^n$ be a non-empty blunt closed convex cone. Define $\alpha(C)$ by $\sin \alpha(C) = \nu(C) > 0$. Let

$$E_\alpha(C) = \{x \in R^n \mid \min_{i \in I} \langle a^i, x\rangle \geq \|x\| \sin \alpha\}$$
$$= \bigcap_{i \in I} C_{\frac{\pi}{2}-\alpha}(a^i) \quad .$$

It is a closed convex cone. By definition of $\alpha(C)$, and Theorem (2.6.8), we have:

$$E_{\alpha(C)}(C) = \mathscr{C}\{e(C)\} \subseteq C \cap L^{\perp} \quad ,$$

and $$\alpha(C) = \text{Max}\{\alpha \mid E_\alpha(C) \neq \{0\}\}$$
$$= \text{Max}\{\alpha \mid \bigcap_{i\in I} C_{\frac{\pi}{2}-\alpha}(a^i) \neq \{0\}\}$$
$$= \text{Max}\{\alpha \mid \bigcap_{i\in I} [C_{\frac{\pi}{2}-\alpha}(a^i) \cap B \cap L^\perp] \neq \emptyset\}$$

using Theorem (2.6.8). For any non-empty $J \subseteq I$, define $C_J = \bigcap_{i\in J} H^i$. Clearly $J \subseteq J' \subseteq I$ implies $C_J \supseteq C_{J'} \supseteq C_I$, and hence $\alpha(C_J) \geq \alpha(C_{J'}) \geq \alpha(C)$.

(2.6.19) THEOREM. Let $C \neq R^n$ be a blunt closed convex cone. Then

$$\alpha(C) = \text{Min}\{\alpha(C_J) \mid \emptyset \subset J \subseteq I,\ \text{card } J \leq n-d\}$$

where d is the dimension of the lineality space of C, or equivalently $n-d$ is the rank of the matrix $\{a^i \mid i \in I\}$.

Proof. Let $\alpha^* = \text{Inf}\{\alpha(C_J) \mid \emptyset \subset J \subseteq I,\ \text{card } J \leq n-d\}$. Clearly $\alpha^* \geq \alpha(C) > 0$.

First, we show that the infimum is attained.

For $\emptyset \subset J \subseteq I$ such that $\text{card } J \leq n-d$, $\nu(C_J)$ can be viewed as a function defined on $[\{a^i \mid i \in I\}]^{(n-d)} \subseteq B^{(n-d)} \subseteq R^{n(n-d)}$. Both these sets are compact in $R^{n(n-d)}$. Note that if $\text{card } J < n-d$, we will enlarge $\{a^i \mid i \in J\}$ by repeating some a^i, $i \in J$.

Now, let $(a^i,\ i = 1,\dots,n-d) \in [\{a^i \mid i \in I\}]^{n-d}$, then:

(i) $\displaystyle\text{Min}_{i=1,\dots,n-d} \langle a^i, e\rangle$ is a continuous function of $(a^i,\ i = 1,\dots,n-d)$, as the minimum of a finite family of continuous functions.

(ii) $\nu[C_J = \{x \in R^n \mid \langle a^i,x\rangle \geq 0,\ i = 1,\ldots,n-d\}]$

$$= \sup_{e\in B} \min_{i=1,\ldots,n-d} \langle a^i,e\rangle$$

is a lower semicontinuous function of $(a^i,\ i = 1,\ldots,n-d)$, as the supremum of a family of continuous functions.

(iii) Clearly now

$$\nu^* = \sin \alpha^* = \inf_{(a^i,i=1,\ldots,n-d)\in[\{a^i|i\in I\}]^{n-d}} \sup_{e\in B} \min_{i=1,\ldots,n-d} \langle a^i,e\rangle$$

is attained, because it is the infimum of a lower-semicontinuous function on a compact set.

Hence, there exists a set of n-d vectors (possibly not distinct) $a^{*i} \in \{a^i \mid i \in I\}$ such that, if we let

$$C^* = \{x \in R^n \mid \langle a^{*i},x\rangle \geq 0,\ i = 1,\ldots,n-d\} ,$$

then $\nu(C^*) = \nu^*$. Let $e^* = e(C^*)$, then by Theorem (2.6.8)

$$E_{\alpha^*}(C^*) = \bigcap_{i=1}^{n-d} C_{\frac{\pi}{2}-\alpha^*}(a^{*i}) = \mathscr{C}(e^*) \subseteq C^* \cap L^{\perp}(C^*) \subseteq C^* \cap L^{\perp} .$$

Hence, by definition of α^*, we have:

$$\bigcap_{i\in J} [C_{\frac{\pi}{2}-\alpha^*}(a^i) \cap L^{\perp} \cap B] \neq \emptyset .$$

for all $\emptyset \subset J \subseteq I$, such that card $J \leq n-d$. Define $S^i = C_{\frac{\pi}{2}-\alpha^*}(a^i) \cap L^{\perp} \cap B$.

(i) If $\alpha^* = \frac{\pi}{2}$, then $C_0(a^i) = \mathscr{C}(a^i)$. This implies that C_J is a halfspace for any $J \subseteq I$ such that $1 \leq$ card $J \leq n-d$. Hence C is a halfspace and the theorem is trivial, as card $I = 1$.

(ii) Assume then that $\alpha^* < \frac{\pi}{2}$. Clearly $B \cap L^{\perp}$ is the unit n-d-1 dimensional spherical surface with axis L.

Hence $\{S^i \mid i \in I\}$ is a family of strongly convex subsets of $B \cap L^{\perp}$, such that the intersection of n-d or fewer members is not empty.

We still need to show that no $J \subseteq I$, with card J = n-d+1 is such that $\bigcup_{i \in J} S^i = B \cap L^{\perp}$. Assume there is such a J. Clearly $S^i \subset (\text{Int } H^i) \cap B \cap L^{\perp}$ where $\text{Int } H^i = \{x \in R^n \mid \langle a^i, x\rangle > 0\}$. Hence $\bigcap_{i \in J} [(B \cap L^{\perp})/S^i] = \emptyset$ but $(B \cap L^{\perp})/S^i \supset B \cap L^{\perp} \cap (-H^i)$, as $-H^i = R^n/\text{Int } H^i$.

It follows now that:

$$\emptyset = \bigcap_{i \in J} [B \cap L^{\perp} \cap (-H^i)] = B \cap L^{\perp} \cap (-C_J)$$

$$= -(B \cap L^{\perp} \cap C_J) \quad .$$

Hence $C_J \cap (B \cap L^{\perp}) = \emptyset$, a contradiction, as

$$C_J \cap (B \cap L^{\perp}) \supseteq C \cap (B \cap L^{\perp}) \neq \emptyset \quad .$$

Using Theorem (2.6.18) it follows that $\bigcap_{i \in I} S^i$ is not empty and hence $\alpha^* = \alpha(C)$.

Furthermore, as $\bigcap_{i=1}^{n-d} C_{\frac{\pi}{2}-\alpha^*}(a^{*i}) = \mathscr{C}(e^*)$, it follows that $e^* = e(C)$. Q.E.D.

Remark. This result is again stronger than a similar result in R^n, where the largest sphere contained in a compact convex body K has its radius, but not its position, determined by a family of n+1 supporting halfspaces.

(2.6.20) THEOREM. Let $C = \{x \in R^n | \langle a^i,x\rangle \geq 0,\ i \in I\}$ be a blunt closed convex cone. Then $\mu(C) \geq \nu(C)$.

Proof. Define the function:

$$f(x) = \langle h,x\rangle + \frac{1}{\mu(C)} \operatorname*{Min}_{i\in I} \{\langle a^i,x\rangle\}$$

where $h \in C^P \cap B$ satisfies $d_C(h) = \mu(C)$.

We have:

$$f(h) = 0$$

$$f(e(C)) = \langle h,e(C)\rangle + \frac{\nu(C)}{\mu(C)} \geq -1 + \frac{\nu(C)}{\mu(C)} \quad .$$

As $e(C) \in \text{Int } C$ and $h \notin C$, there exists a unique $\alpha \in (0,1)$ such that:

$$y = \alpha e(C) + (1-\alpha)h \in \partial C \quad .$$

We have

$$f(y) = \langle h,y\rangle \leq 0$$

$$f(y) \geq \alpha f(e(C)) + (1-\alpha)f(h) \quad \text{by the concavity of } f.$$

Hence
$$0 \geq \alpha[-1 + \frac{\nu(C)}{\mu(C)}]$$

or
$$\mu(C) \geq \nu(C) \quad .$$
Q.E.D.

2.7 Distances to a Polyhedron

Let $P = \bigcap_{i\in I} H^i$ be a polyhedron defined as the intersection of a finite family of halfspaces:

$$H^i = \{x \in R^n | \langle a^i,x\rangle + b^i \geq 0\}$$

where we assume that $\|a^i\| = 1$ for all $i \in I$.

(2.7.1) We define the *residue function*:

$$R_P(x) = \operatorname*{Max}_{i\in I} \{-\langle a^i,x\rangle - b^i\}$$

and the *residual distance*:

$$d_P(x) = \operatorname*{Max}_{i\in I} d(x,H^i) = \operatorname{Max} \{0,R_P(x)\} \quad .$$

Clearly $R_P(x)$ and $d_P(x)$ are convex, continuous, polyhedral functions, and

$$P = \{x \in R^n |\ R_P(x) \leq 0\} = \{x \in R^n |\ d_P(x) = 0\} \quad .$$

These two functions depend on the specific family of halfspaces which defines P.

(2.7.2) THEOREM. Let $P = \bigcap_{i\in I} H^i$ be a non-empty polyhedron in R^n. Then there exists a real number $\mu^* > 0$ such that for all $x \in R^n$ we have:

$$\mu^* d(x,P) \leq d_P(x) \leq d(x,P) \quad .$$

Proof.

(i) $d(x,P) = \operatorname*{Min}_{z\in \bigcap_{i\in I} H^i} \|x-z\| \geq \operatorname*{Min}_{z\in H^i} \|x-z\| = d(x,H^i)$. Hence

$$d(x,P) \geq \operatorname*{Max}_{i\in I} d(x,H^i) = d_P(x) \quad .$$

(ii) If $x \in P$, then $d(x,P) = d_P(x) = 0$. Let $x \notin P$, and $y = P_P(x)$. Then, using (2.6.3) and (2.3.3) we get:

$$\mu(C_P(y)) = \operatorname*{Inf}_{u\in N_P(y)} \operatorname*{Max}_{i\in I(y)} \frac{-\langle a^i,u\rangle}{\|u\|}$$

$$\leq \underset{i\in I(y)}{\text{Max}} \frac{-\langle a^i,u\rangle}{\|u\|} \quad \text{for all} \quad u \in N_P(y)$$

$$= \underset{i\in I(y)}{\text{Max}} \frac{d(x,H^i)}{d(x,P)} \quad \text{for all} \quad x \in y + N_P(y)$$

$$\leq \underset{i\in I}{\text{Max}} \frac{d(x,H^i)}{d(x,P)} = \frac{d_P(x)}{d(x,P)} \quad \text{for all} \quad x \in y + N_P(y) \quad .$$

Hence

$$d_P(x) \geq \mu[C_P(F)]\cdot d(x,P)$$

for all $x \in \text{r.i.}(F) + N_P(F)$, for all $F \in \mathcal{F}(P)/\{\emptyset\}$. By (2.4.15)

$$d_P(x) \geq \underset{F\in\mathcal{F}(P)/\{\emptyset\}}{\text{Min}} \mu[C_P(F)]\cdot d(x,P)$$

for all $x \in R^n$. Let $\mu^* = \underset{F\in\mathcal{F}(P)/\{\emptyset\}}{\text{Min}} \mu[C_P(F)]$. By (2.6.4):

$$\mu^* > 0 \quad . \qquad \text{Q.E.D.}$$

The next theorem will give an intuitive background for the choice of the relaxation parameter.

Let P be a polytope.

(2.7.3) We define a *parallel set* P_ε to P by:

$$P_\varepsilon = \bigcup_{x\in P} S_\varepsilon(x) = \{y \in R^n \mid d(y,P) \leq \varepsilon\} \quad .$$

If P is not empty, P_ε is a smooth closed ccnvex body for $\varepsilon > 0$ and

$$d(x,P_\varepsilon) = \text{Max}\{d(x,P)-\varepsilon,0\} \quad .$$

The boundary of P_ε is $\partial P_\varepsilon = \{x \in R^n \mid d(x,P) = \varepsilon\}$. For any $F \in \mathcal{F}(P)$ let

$$\partial P_\varepsilon(F) = \{x \in \partial P_\varepsilon \mid P_P(x) \in \text{r.i.}(F)\} \quad ,$$

i.e.,

$$\partial P_\varepsilon(F) = \text{r.i.}(F) + [N_P(F) \cap \{x \in R^n \mid \|x\| = \varepsilon\}] \quad .$$

Let $\mu^{(n-1)}[\partial P_\varepsilon]$ be the n-1 dimensional Lebesgue measure of ∂P_ε.

Clearly, as $\{\partial P_\varepsilon(F) \mid F \in \mathscr{F}(P)\}$ is a finite partition of ∂P_ε, we have:

$$\mu^{(n-1)}[\partial P_\varepsilon] = \sum_{F \in \mathscr{F}(P)} \mu^{(n-1)}[\partial P_\varepsilon(F)] \quad .$$

We also have:

$$\mu^{(n-1)}[\partial P_\varepsilon(F)] = \mu^{(\dim F)}[F] \times \mu^{(n-\dim F-1)}[N_P(F) \cap \{x \mid \|x\| = \varepsilon\}]$$

and

$$\mu^{(n-\dim F-1)}[N_P(F) \cap \{x \mid \|x\| = \varepsilon\}] = \varepsilon^{n-\dim F-1}\mu^{\perp}[F]$$

as we define $\mu^{\perp}[F] = \mu^{(n-\dim F-1)}[N_P(F) \cap B]$.

Clearly $\mu^{(\dim F)}[F] > 0$ and $\mu^{\perp}[F] > 0$.

Let $\mu_d = \sum_{\substack{F \in \mathscr{F}(P) \\ \dim F = d}} \mu^{(\dim F)}[F] \times \mu^{\perp}[F]$. Then

$$\mu^{(n-1)}[\partial P_\varepsilon] = \sum_{d=0}^{n-1} \mu_d \varepsilon^{n-d-1} \quad .$$

Hence, if we define

$$\mu_d(\varepsilon) = \mu^{(n-1)}\{x \in \partial P_\varepsilon \mid P_P(x) \in \text{r.i.}(F),\ F \in \mathscr{F}(P),\ \dim F = d\} \ .$$

The next theorem follows, as $\mu_d(\varepsilon) = \mu_d \varepsilon^{n-d-1}$.

(2.7.4) <u>THEOREM</u>. $\lim_{\varepsilon \downarrow 0} \dfrac{\mu_d(\varepsilon)}{\mu_{d'}(\varepsilon)} = 0$ if $d < d'$

$$\lim_{\varepsilon\uparrow\infty} \frac{\mu_d(\varepsilon)}{\mu_{d'}(\varepsilon)} = 0 \quad \text{if} \quad d > d' \quad .$$

Remark. A similar result holds if we define a parallel set using $d_P(x)$ as distance function.

This theorem can be related to the following result: The set of points where a closed convex body of R^n is not smooth has zero $n-1$ dimensional Lebesgue measure; furthermore, the set of vertices is at most countable, but it can be everywhere dense in the boundary of K (see Appendix 2 and Valentine, pp. 135-136).

2.8 Convex Sets

(2.8.1) THEOREM. Let K be a closed convex body and y a point of its boundary ∂K. Then there exists an interior point s of K such that $s-y$ is the interior normal to a supporting halfspace to K at y.

Proof. Using (2.2.8) and (2.3.8), we get:

$$N_K(y) = [D_K(y)]^P = [\overline{D_K(y)}]^P \quad .$$

By Theorem (2.3.14) $\text{r.i.}[-N_K(y)]$ and $\text{r.i.}[\overline{D_K(y)}] = \text{Int } D_K(y)$ have a ray in common, say $\mathscr{C}(e)$ where $\|e\| = 1$. Clearly, there exists a $\eta^* > 0$ such that

$$s = y + \eta e \in K \quad \text{for all} \quad \eta \in [0,\eta^*] \quad .$$

It remains to show that

$$L = \{s \in R^n \mid s = y + \eta e, \ \eta \in [0,\eta^*]\}$$

intersects the interior of K. Assume it does not; then by

Theorem (2.2.16) K and L can be separated: There exists $a \in R^n$, $\|a\| = 1$ such that $<a,L> \geq <a,K>$; but $y + \eta e \in K \cap L$ for all $\eta \in [0,\eta^*]$. Hence $<a,e> = 0$,

$$<a,L> = <a,y> \geq <a,K> \quad .$$

Now as $e \in \text{Int } D_K(y)$, there exists $\varepsilon > 0$ such that $S_\varepsilon(e) \subseteq D_K(y)$. Let $e' = e + \varepsilon a \in S_\varepsilon(e) \subseteq D_K(y)$. Hence there exists an $\eta' > 0$ such that

$$y + \eta' e' \in K$$

and

$$<a,y+\eta' e'> \leq <a,y>$$

or

$$\eta' \varepsilon \leq 0 \quad ,$$

a contradiction. It follows that some $s \in L$ is in the interior of K. Q.E.D.

(2.8.2) THEOREM. Let y be a point of the boundary ∂K of a closed convex body $K \subseteq R^n$. Then there exists a spherical neighborhood $S_\varepsilon(y)$, an interior point s of K and an angle $\alpha \in (0,\frac{\pi}{2})$ such that

(i) $<s-y,K-y> \geq 0$

(ii) $<\frac{s-y}{\|s-y\|},\frac{u}{\|u\|}> < -\cos\alpha$ for all $u \in \bigcup_{x \in S_\varepsilon(y) \cap \partial K} N_K(x)$.

In other words, s-y "strictly" separates K-y from the union of the normal cones at the points of $S_\varepsilon(y) \cap \partial K$.

Proof. By Theorem (2.8.1) there exists a point s such that $s \in \text{Int } K$, i.e., there exists $r > 0$ such that $S_r(s) \subseteq K$, and

$s-y$ is the interior normal to a supporting hyperplane of K at y. Let $u \in B$, $\varepsilon > 0$. Then

$$u \in \bigcup_{x \in S_\varepsilon(y) \cap \partial K} N_K(x)$$

if and only if there exists $x \in S_\varepsilon(y) \cap \partial K$ such that for all $z \in K$

$$\langle u, z-x\rangle \leq 0 \quad .$$

It implies that there exists $x \in S_\varepsilon(y)$ such that for all $z \in S_r(s)$

$$\langle u,z\rangle \leq \langle u,x\rangle$$

or

$$\operatorname*{Max}_{z \in S_r(s)} \langle u,z\rangle \leq \operatorname*{Max}_{x \in S_\varepsilon(y)} \langle u,x\rangle$$

or

$$\langle u, s+ru\rangle \leq \langle u, y+\varepsilon u\rangle$$

or

$$\langle u, \frac{y-s}{\|y-s\|}\rangle \geq \frac{r-\varepsilon}{\|y-s\|} \quad .$$

Choose any $\varepsilon \in (0,r)$ and define $0 < \cos\alpha = \frac{r-\varepsilon}{\|y-s\|} < 1$.

Q.E.D.

Let $y \in \partial K$ be a point of the boundary of a closed convex body $K \subseteq R^n$. We define the function

$$\eta: N_K(y) \cap B \to R_+$$

by

$$\eta(u) = \operatorname{Sup}\{\eta \in R_+ \mid y - \eta u \in K\} \quad .$$

The supremum is attained unless it is infinite.

The *gauge function* of a convex set K is ([55], p.28):

$$\delta(x|K) = \operatorname{Inf}\{\lambda \geq 0 \mid x \in \lambda C\} \quad .$$

Clearly $\delta(x|K-y)$ restricted to $-N_K(y)$ is simply

$$\frac{\|x\|}{\eta(\frac{x}{\|x\|})} \quad .$$

We also define the function:

$$\eta^*: \partial K \to R_+$$

by
$$\eta^*(y) = \operatorname*{Inf}_{u \in N_K(y) \cap B} \eta(u) \quad .$$

(2.8.3) The closed convex body K will be said to be *smooth enough at* $y \in \partial K$ if $\eta^*(y) > 0$.

A necessary condition for this is clearly $-N_K(y) \subseteq D_K(y)$.

(2.8.4) The closed convex body K will be said to be *smooth enough* if:

$$\eta^* = \operatorname*{Inf}_{y \in \partial K} \eta^*(y) > 0$$

and η^* will be called the *modulus of smoothness* of K.

A necessary condition for this is that K be smooth enough at every point of its boundary.

It follows that a closed convex cone is smooth enough if and only if $-C^P \subseteq C$:

(i) $\eta^*(0) > 0$ implies that $-\eta^*(0)u \in C$ for all $u \in N_C(0) \cap B = C^P \cap B$. Hence $-u \in C$ for all $u \in C^P$.

(ii) If $-C^P \subseteq C$, then for any $y \in \partial C$, we have:

$$-N_C(y) \subseteq -C^P \subseteq C \subseteq D_C(y) \quad .$$

Hence, for any $u \in N_C(y) \cap B$, $\eta(u) = +\infty$. That is $\eta^*(y) = +\infty$

for all $y \in \partial C$. Hence $\eta^* = +\infty$.

The smooth enough closed convex cones can then be characterized, using (2.6.17) by the following theorem.

(2.8.5) THEOREM. A closed convex cone is smooth enough if and only if it contains an autopolar cone.

Proof. Clearly all autopolar cones are closed, convex, pointed and blunt.

(i) We first show that if $-C^P \subset C$, then C contains a smaller smooth enough cone.

There exists an $e \in B \cap C$ such that $e \notin -C^P$. Using the separation theorem there exists $a \in B$ such that

$$\langle a,e\rangle < 0$$
$$\langle a,-C^P\rangle \geq 0 \quad .$$

This implies that $a \in C$, $a \notin -C^P$. Let $D = C \cap H$ where $H = \{x \in R^n | \langle a,x\rangle \geq 0\}$. It follows that (using [55], pp.150, 22 and 75)

$$D^P = (C \cap H)^P = \overline{\mathscr{C}}(C^P \cup -\mathscr{C}(a)) = \overline{[C^P + \mathscr{C}(-a)]} = C^P + \mathscr{C}(-a) \ .$$

Now we have

$$\left.\begin{array}{l} -C^P \subseteq H \Rightarrow \quad -C^P \subseteq D \\ \mathscr{C}(a) \subseteq H \Rightarrow \mathscr{C}(a) \subseteq D \end{array}\right\} \Rightarrow -D^P \subseteq D \quad .$$

(ii) Let P be the lattice of closed convex cones in R^n. Then $P(C) = \{D \in P | -D^P \subseteq D, D \subseteq C\}$ is a poset, and by Hausdorff Maximality Principle, it contains a maximal chain $P' \subseteq P(C)$.

Clearly $C \in P'$.

Now P is a complete lattice and hence $P' \subseteq P$ has a greatest lower bound in P. Clearly it is $C^* = \bigcap_{D \in P'} D$. Also, $-D^P \subseteq C$ for any $D, C \in P'$: Clearly as P' is a chain, either $D \subseteq C$ or $C \subseteq D$; in the first case we get $-D^P \subseteq D \subseteq C$, in the second $-D^P \subseteq -C^P \subseteq C \subseteq D$. Hence $-D^P \subseteq C^*$ for any $D \in P'$ and also

$$- \bigcup_{D \in P'} D^P \subseteq C^* \quad .$$

The right hand side is a closed convex cone. Hence

$$-\overline{\mathscr{C}}\left(\bigcup_{D \in P'} D^P \right) \subseteq C^* \quad .$$

But we also have ([55], p.150):

$$-(C^*)^P = -\overline{\mathscr{C}}\left(\bigcup_{D \in P'} D^P \right) \quad .$$

Hence $C^* \in P'$.

Now if $-(C^*)^P \subset C^*$, by (i) there is a $D' \in P$ which is smooth enough such that $D' \subset C^*$. But then $P' \cup \{D'\}$ is a chain in $P(C)$ which contradicts the maximality of P'. The sufficiency of the theorem is then established.

(iii) Finally if C contains an autopolar cone D then $C \supseteq D = -D^P \supseteq -C^P$ and the necessity follows. Q.E.D.

(2.8.6) <u>THEOREM</u>. A smooth compact convex body $K \subseteq R^n$ is smooth enough.

<u>Proof</u>. For any $y \in \partial K$, let $n(y) \in B$ be the interior normal to the supporting halfspace of K at y: $\{n(y)\} = -N_K(y) \cap B$.

(i) There exists a point $s(y) = y + \eta_y n(y) \in \text{Int } K$ where $\eta_y > 0$. Hence $S_{r(y)}(s(y)) \subseteq K$ for some $r(y) > 0$.

(ii) $n(x)$ is a continuous function of x on ∂K (see [10], p.6). The function of η defined by $\|x+\eta n(x)-s(y)\|^2$ is minimum for $\eta = -\langle n(x),x-s(y)\rangle > 0$. Its minimal value is

$$\sigma(x) = \|x-s(y)\|^2 - (\langle n(x),x-s(y)\rangle)^2 \geq 0 \quad .$$

We can interpret $\sigma(x)$ as the square of the distance from $s(y)$ to the half-line $\{x+\eta n(x) \mid \eta \geq 0\}$.

Clearly $\sigma(y) = 0$ and $\sigma(x)$ is a continuous function of $x \in \partial K$.

Hence there exists a neighborhood N_y of y, open in the relative topology induced on ∂K by the Euclidean topology of R^n such that $x \in N_y \subseteq \partial K$ implies $\sigma(x) \leq r^2(y)$.

Let $\eta^*(x) = \operatorname{Sup}\{\eta \geq 0 \mid x+\eta n(x) \in K\}$. It follows easily that for any $x \in N_y$, we have $\eta^*(x) \geq r(y)$, as

$$S_{r(y)}(s(y)) \subseteq K \subseteq \{z \in R^n \mid \langle n(x),z-x\rangle \geq 0\} \quad .$$

But $\bigcup_{y \in \partial K} N_y$ is an open covering of the compact space ∂K; it has a finite subcovering: $\partial K = \bigcup_{j=1}^{m} N_{y_j}$. Clearly then

$$\begin{aligned} \operatorname*{Inf}_{x \in \partial K} \eta^*(x) &\geq \operatorname{Inf}\{r(y_j) \mid j = 1,\ldots,m\} \\ &= \operatorname{Min}\{r(y_j) \mid j = 1,\ldots,m\} > 0 \quad . \end{aligned}$$

Q.E.D.

(2.8.7) <u>COROLLARY</u>. Let K be a smooth closed convex body in R^n and Y be any compact subset of ∂K. Then

$$\operatorname*{Inf}_{y \in Y} \eta^*(y) > 0 \quad .$$

CHAPTER III

THE RELAXATION METHOD

3.1 Infinite Convergence of the Relaxation Method

The following lemma is basic to the study of the relaxation method:

(3.1.1) LEMMA. Let x, z, t be points of R^n and let

$$x' = x + \lambda(t-x)$$

where $\lambda \in R$. Then

$$\|x'-z\|^2 = \|x-z\|^2 - \lambda(2-\lambda)\|t-x\|^2 + 2\lambda\langle t-z,t-x\rangle \quad .$$

Proof.
$$\begin{aligned} \|x'-z\|^2 &= \|x+\lambda(t-x)-z\|^2 \\ &= \|x-z\|^2 + \langle 2(x-z)+\lambda(t-x),\lambda(t-x)\rangle \\ &= \|x-z\|^2 - \lambda(2-\lambda)\|t-x\|^2 + \lambda\langle 2(x-z)+2(t-x),t-x\rangle \\ &= \|x-z\|^2 - \lambda(2-\lambda)\|t-x\|^2 + 2\lambda\langle t-z,t-x\rangle \quad . \end{aligned}$$

Q.E.D.

Let $P = \bigcap_{i\in I} H^i$ be a closed convex set given by the intersection of a family of halfspaces:

$$H^i = \{x \in R^n \mid \langle a^i,x\rangle + b^i \geq 0\}$$

where we assume that $\|a^i\| = 1$. Define

$$d_P(x) = \sup_{i\in I} d(x,H^i) \quad ,$$

the residual distance from x to P.

The relaxation method chooses at the iteration q, the index

i^q of the most violated halfspace, i.e.,

$$d(x^q, H^{i^q}) = d_P(x^q) = -\langle a^{i^q}, x^q\rangle - b^{i^q} .$$

It then moves on the direction a^{i^q} of the interior normal to H^{i^q}.

Let

$$t^q = P_{H^{i^q}}(x^q) = x^q + d_P(x^q)a^{i^q} .$$

Then the next point of the sequence is defined by

$$x^{q+1} = x^q + \lambda[t^q - x^q] = x^q + \lambda d_P(x^q)a^{i^q} .$$

Hence Lemma (3.1.1) can be written as:

(3.1.2) <u>LEMMA</u>.

$$\|x^{q+1}-z\|^2 = \|x^q-z\|^2 - \lambda d_P(x^q)[(2-\lambda)d_P(x^q) - 2\langle a^{i^q}, t^q-z\rangle] .$$

<u>Remark</u>. Some regularity assumptions need to be made on the defining family of halfspaces in order that the supremum in the definition of $d_P(x)$ be attained (and finite).

We need to rule out families like:

$$P = \{x \in \mathbb{R} \mid x \geq -\tfrac{1}{k},\ k \in Z_+/\{0\}\}$$

or

$$P = \{x \in \mathbb{R} \mid x \geq k,\ k \in Z_+\} .$$

(3.1.3) We will assume either one of the two conditions:

(i) I is finite.

(ii) P is not empty and $\{(a^i, b^i) \in B \times \mathbb{R} \mid i \in I\}$ is closed in $\mathbb{R}^{n+1}$ where $B = \{x \in \mathbb{R}^n \mid \|x\| = 1\}$.

Trivially, if (i) is satisfied, then

$$d_P(x) = \max_{i\in I} d(x,H^i) < +\infty \quad .$$

If (ii) is satisfied, then

$$0 \leq d(x,H^i) \leq d(x,P) < +\infty \quad \text{for all} \quad i \in I, \quad x \in R^n.$$

Hence, as $0 \leq d_P(x) = \sup_{i\in I} d(x,H^i) \leq d(x,P)$ for all $x \in R^n$, $d_P(x)$ is not only convex, but also finite on R^n and it follows that it is also continuous.

If $x \in P$, $d(x,H^i) = 0$ for all $i \in I$ and clearly $d_P(x) = 0$.

If $x \notin P$, then $d_P(x) = \varepsilon > 0$. Hence let

$$\begin{aligned}\tilde{D}_\varepsilon &= \{(a^i,b^i) \mid i \in I,\ d(x,H^i) \geq \tfrac{\varepsilon}{2}\} \\ &= \{(a^i,b^i) \mid i \in I,\ \tfrac{\varepsilon}{2} \leq -\langle a^i,x\rangle - b^i \leq \varepsilon\} \\ &= \{(a^i,b^i) \mid i \in I\} \\ &\quad \cap \{(a,b) \in B \times R \mid \tfrac{\varepsilon}{2} \leq -\langle a,x\rangle - b \leq \varepsilon\} \ .\end{aligned}$$

Clearly $\tilde{D}_\varepsilon$ is compact in R^{n+1}. Also

$$d_P(x) = \sup_{i\in I} d(x,H^i) = \sup_{(a,b)\in\tilde{D}_\varepsilon} \{-\langle a,x\rangle - b\}$$

and the supremum is attained, as

$-\langle (a,b),(x,1)\rangle$ is a continuous (linear) function of (a,b).

To summarize, under the condition (3.1.3), $d_P(x)$ is convex, continuous and $d_P(x) = \max_{i\in I} d(x,H^i)$. Also

$$P = \{x \in R^n \mid d_P(x) \leq 0\} \quad .$$

We will assume from now on that (3.1.3) is satisfied.

(3.1.4) <u>THEOREM</u>. Let $P = \bigcap_{i\in I} H^i$ be a non-empty convex subset of R^n. Then the sequence of points generated by the relaxation method is a Fejér-monotone sequence with respect to P, if $\lambda \in (0,2]$.

Proof. If $z \in P$, then $z \in H^{i^q}$. Hence $<a^{i^q}, t^q - z> \leq 0$ for all $z \in P$. By Lemma (3.1.2) it implies

$$\|x^{q+1} - z\| \leq \|x^q - z\| \quad \text{for all } z \in P.$$

Q.E.D.

(3.1.5) Remarks. (i) In fact x^{q+1} is pointwise closer than x^q from the set:

$$\begin{aligned} H &= \{z \in R^n \mid (2-\lambda)d_p(x^q) \geq 2<a^{i^q}, t^q - z>\} \\ &= \{z \in R^n \mid <a^{i^q}, z> + b^{i^q} + \frac{2-\lambda}{2} d_p(x^q) \geq 0\} \\ &= \{z \in R^n \mid <a^{i^q}, z - \frac{x^q + x^{q+1}}{2}> \geq 0\} \end{aligned}$$

which is the halfspace, parallel to H^{i^q}, such that its boundary hyperplane passes through the center of the segment $[x^q, x^{q+1}]$. Clearly

$$\begin{aligned} \lambda < 2 &\Rightarrow H \supset H^{i^q} \\ \lambda = 2 &\Rightarrow H = H^{i^q} \\ \lambda > 2 &\Rightarrow H \subset H^{i^q} \end{aligned} .$$

(ii) Theorem (3.1.4) still holds if i^q is chosen in such a way that H^{i^q} is any violated halfspace.

(3.1.6) THEOREM. Let $P = \bigcap_{i \in I} H^i$ be a full dimensional closed convex set defined by the intersection of a family of closed half-spaces. Then the sequence of points generated by the relaxation method converges finitely or infinitely to a point of P, for any given value of the relaxation parameter λ such that $0 < \lambda \leq 2$.

Furthermore, if the convergence is infinite, then the limit of the sequence is a point of the boundary of P.

Proof. Assume that the sequence $\{x^q\}$ does not terminate. By Theorem 3.1.4, $\{x^q\}$ is Fejér-monotone with respect to P. By Theorem 1.3.2, $\{x^q\}$ converges to a point, say x^*, But

$$d_P(x^q) = \|t^q - x^q\| = \frac{1}{\lambda}\|x^{q+1} - x^q\| \quad .$$

Hence
$$\lim_{q\to\infty} d_P(x^q) = \frac{1}{\lambda} \lim_{q\to\infty} \|x^{q+1} - x^q\| = 0 \quad .$$

But $d_P(x)$ is continuous. Hence $d_P(x^*) = 0$, which implies that $x^* \in P$. Clearly $x^* \in \partial P$ as $x^q \notin P$ for all q. Q.E.D.

(3.1.7) THEOREM. Let $P = \bigcap_{i\in I} H^i$, where I is a finite set, be a non-empty polyhedron; then the relaxation method generates a sequence which converges finitely or infinitely to a point $x^* \in P$, for any given value of the relaxation parameter λ such that $\lambda \in (0,2)$.

Furthermore, we have:

$$\|x^q - x^*\| \leq 2d(x^o, P)\theta^q$$

where $\theta \in [0,1)$ depends only on the matrix $A = (a^i, i \in I)$ and on λ.

Proof. Assume that $x^q \notin P$. Lemma (3.1.2) implies that:

$$d^2(x^{q+1}, P) \leq \left|x^{q+1} - P_P(x^q)\right|^2$$

$$\leq d^2(x^q,P) - \lambda(2-\lambda)d_p^2(x^q)$$
$$\leq d^2(x^q,P)[1-\lambda(2-\lambda)\mu^{*2}]$$

using Theorem (2.7.2).

Let $\theta = \sqrt{1-\lambda(2-\lambda)\mu^{*2}}$. Clearly $\theta \in [0,1)$. Hence $d(x^q,P) \leq \theta^q d(x^o,P)$ and if the sequences does not terminate, it implies that:

$$\lim_{q\to\infty} d(x^q,P) = 0 \quad .$$

Using Theorem (1.3.2), it implies that $\{x^q\}$ converges to a point $x^* \in \partial P$ (if the sequence is infinite).

We also have:

$$\|x^*-P_p(x^q)\| \leq \|x^q-P_p(x^q)\| = d(x^q,P)$$

by the Fejér-monotonicity of $\{x^q\}$.

Hence x^* and x^q both belong to the sphere

$$S_{d(x^q,P)}(P_p(x^q))$$

and
$$\|x^*-x^q\| \leq 2d(x^q,P) \leq 2\theta^q d(x^o,P) \quad .$$

If the sequence terminates, this holds if we take x^* to be the last point of the sequence. Q.E.D.

<u>Remark</u>. The assumption that i^q is chosen so that H^{i^q} is the most violated halfspace can be almost ignored. See (3.3.3).

3.2 <u>Relaxation With Respect to a Convex Body Defined by Its Complete Family of Supporting Halfspaces</u>

In this paragraph, we present a theorem which is an extension, as well as a minor correction, of a theorem due to Motzkin and Schoenberg ([51], pp. 402-404).

Let K be a closed convex subset of R^n.

(3.2.1) Its *support function* is defined by:

$$\delta^*(a|K) = \text{Sup}\ \{\langle a,x\rangle |\ x \in K\} \quad .$$

Clearly, it is a positively homogeneous function, hence it is enough to define it on $B = \{a \in R^n |\ \|a\| = 1\}$. Define

$$b(a) = \delta^*(-a|K) \quad .$$

Then, classically (see Rockafellar [55], pp. 112 et seq.):

$$K = \{x \in R^n |\ \langle a,x\rangle + b(a) \geq 0,\ \text{for all } a \in B\}$$
$$= \{x \in R^n |\ \langle a,x\rangle + b(a) \geq 0,\ \text{for all } a \in B^*\}$$

where $B^* = \{a \in B|\ b(a) < +\infty\}$. It can be shown that $\overline{\mathscr{C}}(-B^*) = [0^+(K)]^P$, where the closure operation, contrary to a claim by Stoer-Witzgall ([62], p.133), is necessary. A parabola in R^2 shows that B^* need not be closed.

$\mathscr{C}(-B^*)$ is the barrier cone of K.

(3.2.2) The *complete family of supporting halfspaces* associated to a closed convex set K is:

$$\{H = \{x \in R^n |\ \langle a,x\rangle + b(a) \geq 0\} |\ a \in B^*\} \quad .$$

But in any case $\{(a,b)|\ b = b(a),\ a \in B^*\}$ is closed.

(3.2.3) <u>THEOREM</u>. Let K be a closed convex body in R^n, defined by its complete family of supporting halfspaces. Then the relaxation method converges finitely to a point of K if $\lambda \in [1,2]$. Furthermore, if $\lambda > 2$, the process either converges finitely to a point of K, or it does not converge.

Proof. Using Theorem (3.1.6), it is enough to show that infinite convergence cannot occur. Assume it does. Then $\{x^q\}$ converges to $x^* \in \partial K$. Also, using Corollary (2.3.9):

$$t^q = P_{H^{i^q}}(x^q) = P_K(x^q) \in \partial K \quad .$$

Hence, if $\lambda = 1$, there is convergence in one step, as $x^{q+1} = t^q \in K$. If $\lambda > 1$, as $t^q = x^q + \frac{1}{\lambda}(x^{q+1}-x^q)$

$$\lim_{q\to\infty} t^q = x^* \quad .$$

Using Theorem (2.8.2), there exists an $\varepsilon > 0$, an angle $\alpha < \frac{\pi}{2}$, and a ray e with $\|e\| = 1$ such that:

(i) $\langle e,K-x^*\rangle \geq 0$

(ii) $\langle e,u\rangle < -\cos\alpha$

for all $$u \in \bigcup_{x\in S_\varepsilon(x^*)\cap\partial K} [N_K(x) \cap B] \quad .$$

Clearly, there exists an integer Q such that for all $q \geq Q$, we have:

$$t^q \in S_\varepsilon(x^*) \cap \partial K \quad .$$

By Theorem (2.3.3), $x^q-t^q \in N_K(t^q)$. Hence

$$\langle e,x^q-t^q\rangle < -\|x^q-t^q\|\cos\alpha \quad \text{for all} \quad q \geq Q$$

$$\langle e,t^q-x^*\rangle \geq 0 \quad .$$

As $x^{q+1} = t^q + (\lambda-1)(t^q-x^q)$, we have

$$\langle e,x^{q+1}\rangle = \langle e,t^q\rangle + (\lambda-1)\langle e,t^q-x^q\rangle$$
$$> \langle e,x^*\rangle + (\lambda-1)d(x^q,K)\cos\alpha$$

for all $q \geq Q$, which contradicts

$$\lim_{q\to\infty} x^{q+1} = x^* \quad .$$

Q.E.D.

Remarks. (i) By the norm duality theorem, if a family of supporting hyperplanes contains the complete family, then only hyperplanes from the complete family will be considered by the relaxation method.

(ii) If a closed convex body is smooth, then any family of supporting hyperplanes contains the complete family.

3.3 Finite Convergence of the Relaxation Method Applied to a Polyhedron

(3.3.1) THEOREM. Let $P = \bigcap_{i\in I} H^i$ be a full dimensional polyhedron in R^n. Then the sequence generated by the relaxation method converges finitely to a point of P if:

(i) P is smooth enough and $\lambda \in [1,2]$

or (ii) $\lambda \in (\lambda^*,2]$

where $\lambda^* = \frac{2}{1+\nu(P)}$ and $\nu(P)$ is the inner measure of P. Furthermore, if $\lambda > 2$, then the sequence either converges finitely to a point of P or it does not converge.

We first prove:

(3.3.2) LEMMA. If the sequence $\{x^q\}$ generated by the relaxation method converges infinitely to a point x^*, where $x^* \in \partial P$, then there exists an integer Q such that $x^q \in x^* + N_P(x^*)$ for all $q \geq Q$ and this for any $\lambda > 0$.

Proof. Let $I(x^*) = \{i \in I | \langle a^i, x^* \rangle + b^i = 0\}$ as x^q converges to x^* and as I is finite, there exists an integer Q such that

$$\langle a^i, x^q \rangle + b^i > 0 \text{ for all } q \geq Q, \quad i \in I/I(x^*) \quad .$$

This means that for $q \geq Q$, the relaxation method ignores the halfspaces H^i, $i \in I/I(x^*)$ and hence it is equivalent to studying the convergence of the sequence $\{u^q = x^q - x^*\}_{q \geq Q}$ which can be viewed as generated by the relaxation method as applied to the polyhedral cone

$$C_P(x^*) = \{x \in R^n | \langle a^i, x \rangle \geq 0, \ i \in I(x^*)\} \quad .$$

Clearly $\lim_{q \to +\infty} u^q = 0$. Assume that for some $q^o \geq Q$, we have

$$x^{q^o} \notin x^* + N_P(x^*)$$

or equivalently

$$u^{q^o} \notin N_P(x^*) = [C_P(x^*)]^P \quad .$$

Hence there exists a vector $z \in C_P(x^*)$ such that $\langle z, u^{q^o} \rangle > 0$. It follows that:

$$\langle z, u^{q+1} \rangle = \langle z, u^q \rangle + \lambda \langle z, a^{i^q} \rangle d_P(x^q)$$

where $d_P(x^q) = \|t^q - x^q\| = d_{C_P(x^*)}(u^q)$. But $z \in C_P(x^*)$ implies $\langle a^{i^q}, z \rangle \geq 0$ for all $q \geq q^o$. Hence $\{\langle z, u^q \rangle\}_{q \geq q^o}$ is a monotone increasing sequence of positive numbers and this contradicts

$$\lim_{q \to +\infty} u^q = 0$$

Q.E.D.

Proof of Theorem (3.3.1). Using Theorem (3.1.6) and Lemma (3.3.2), we conclude that it suffices to show that the assumption of an infinite sequence $\{u^q = x^q - x^*\}_{q \geq Q}$ within $N_p(x^*)$ is untenable.

Case 1. $C_p(x^*)$ is smooth enough, i.e.,

$$-N_p(x^*) \subseteq C_p(x^*) \quad .$$

It follows that:

$$u^{q+1} = u^q + \lambda \langle -a^{i^q}, u^q \rangle a^{i^q}$$

as

$$d_{C_p(x^*)}(u^q) = \langle -a^{i^q}, u^q \rangle = d_p(x^q) \; .$$

Hence we first have

$$\langle a^{i^q}, u^{q+1} \rangle = (\lambda - 1) d_p(x^q) \quad .$$

But if $\lambda > 1$, this means that

$$\langle a^{i^q}, u^{q+1} \rangle > 0$$

and by Theorem (2.3.24), $a^{i^q} \in C_p(x^*)$. Hence $u^{q+1} \notin N_p(x^*)$ $= [C_p(x^*)]^P$ and the contradiction.

We also have:

$$\begin{aligned} \langle a^i, u^{q+1} \rangle &= \langle a^i, u^q \rangle + \lambda \langle a^{i^q}, a^i \rangle d_p(x^q) \\ &\geq \langle a^i, u^q \rangle \quad . \end{aligned}$$

This means that the quantities $\langle a^i, u^q \rangle$ do not decrease. Furthermore, for $\lambda = 1$ at least one of those quantities becomes zero at each iteration. Hence after at most card$[I(x^*)]$ steps, we have $\langle a^i, u^q \rangle \geq 0$ for all $i \in I(x^*)$, and the contradiction follows for the case where $\lambda = 1$.

Remark. If the set to which we apply the relaxation method is a smooth enough polyhedral cone: $C = \{x \in R^n | \langle a^i,x\rangle \geq 0,\ i \in I\}$ then the relaxation method will terminate after at most card I steps for any $\lambda \geq 1$.

Furthermore the residual distance is monotone decreasing:

$$\begin{aligned} d_C(u^{q+1}) &= \operatorname*{Max}_{i\in I} [\langle -a^i,u^q\rangle + \lambda d_C(u^q)\langle -a^i,a^{i^q}\rangle] \\ &\leq d_C(u^q) + \lambda d_C(u^q) \operatorname*{Max}_{i\in I} [\langle -a^i,a^{i^q}\rangle] \\ &\leq d_C(u^q) \quad . \end{aligned}$$

Case 2. Let $C_\alpha(e)$ be the largest spherical cone contained in $C_p(x^*)$: $\nu = \sin\alpha$ is the inner measure of both these cones.

We define:

$$D^q = \frac{\langle u^q,-e\rangle}{\|u^q\|} = \cos\alpha^q \quad \text{for} \quad q \geq Q \quad ,$$

the cosine of the angle between $-e$ and u^q. By Theorem (2.6.3), $u^q \in N_p(x^*)$ implies that $u^q \in C_{\frac{\pi}{2}-\alpha}(-e)$; and

$$\frac{\langle u^q,-e\rangle}{\|u^q\|} \geq \sin\alpha = \nu \quad .$$

But

$$u^{q+1} = u^q + \lambda d_p(x^q)a^{i^q}$$

where $d_p(x^q) = d_{C_p(x^*)}(u^q) = -\langle a^{i^q},u^q\rangle$. Hence

$$\langle u^{q+1},-e\rangle = \langle u^q,-e\rangle - \lambda d_p(x^q)\langle a^{i^q},e\rangle$$

and we get the following recurrence relation on D^q:

$$D^{q+1} = D^q \frac{\|u^q\|}{\|u^{q+1}\|} - \lambda d_p(x^q)\langle a^{i^q},e\rangle \frac{\|u^q\|}{\|u^{q+1}\|} \quad .$$

We define

$$\mu^q = \frac{d_p(x^q)}{d(x^q,P)} = \frac{<-a^{i^q},u^q>}{\|u^q\|} .$$

Letting $\mu = \mu[C_p(x^*) = \bigcap_{i\in I(x^*)} (H^i-x^*)]$, it follows from (2.7.2) that:

$$0 < \mu \leq \mu^q \leq 1 .$$

Lemma (3.1.2) applied to $z = x^*$ becomes:

$$\|u^{q+1}\|^2 = \|u^q\|^2 - \lambda(2-\lambda)d_p^2(x^q) + 2\lambda d_p(x^q)<a^{i^q},t^q-x^*> .$$

As $<a^{i^q},t^q-x^*> = 0$, we have:

$$\|u^{q+1}\|^2 = \|u^q\|^2 - \lambda(2-\lambda)d_p^2(x^q)$$

$$= \|u^q\|^2[1-\lambda(2-\lambda)(\mu^q)^2] .$$

Define

$$\beta^q = [1-\lambda(2-\lambda)(\mu^q)^2]^{-1/2} .$$

Then

$$\|u^q\| = \beta^q\|u^{q+1}\| .$$

Hence the recurrence relation on D^q becomes:

$$D^q - D^{q+1} = D^q[1-\beta^q] + \lambda\mu^q\beta^q<a^{i^q},e> .$$

As $-a^{i^q} \in N_p(x^*)$, $<a^{i^q},e> \geq \nu$

$$D^q - D^{q+1} \geq D^q[1-\beta^q] + \lambda\mu^q\beta^q\nu .$$

The assumption that $\{u^q\}_{q\geq Q} \subseteq N_p(x^*)$ implies that

$$\nu \leq D^q \leq 1 \quad \text{for all} \quad q \geq Q \ .$$

Observing that $\lambda \geq 2 \Rightarrow \beta^q \leq 1$, $\lambda < 2 \Rightarrow \beta^q > 1$, it follows that

$$D^q - D^{q+1} \geq F_\lambda(\mu^q) \quad \text{for all} \quad q \geq Q$$

where

$$F_\lambda(\mu^q) = \begin{cases} (1-\beta^q) + \lambda\mu^q\beta^q\nu & \text{for} \quad \lambda < 2 \\ \nu(1-\beta^q) + \lambda\mu^q\beta^q\nu & \text{for} \quad \lambda \geq 2. \end{cases}$$

Clearly, $F_\lambda(\mu^q)$ is a lower bound on the decrease of the cosine of the angle between u^q and $-e$ at each iteration.

It is a continuous function of μ^q on $[\mu,1]$.

Hence, if
$$\min_{\mu^q \in [\mu,1]} F_\lambda(\mu^q) = \varepsilon > 0 \quad ,$$
then

$$D^q - D^{q+1} \geq \varepsilon \quad \text{for all} \quad q \geq Q$$

and

$$D^{q+1} \leq D^q - \varepsilon \leq D^Q - (q-Q)\varepsilon \ .$$

Hence, if
$$q - Q > \frac{D^Q - \nu}{\varepsilon}$$

we have $D^{q+1} < \nu$, and the contradiction follows as this implies that

$$u^{q+1} \notin N_p(x^*) \quad .$$

1. If $\lambda \geq 2$, then $F_\lambda(\mu^q) > 0$ for all $\mu^q \in [\mu,1]$. Hence $\min_{\mu^q \in [\mu,1]} F_\lambda(\mu^q) > 0$ and for this case, the theorem follows.

2. If $1 < \lambda < 2$, then we have:

$$\frac{\partial F_\lambda(\mu^q)}{\partial \mu^q} = \lambda(\beta^q)^3[\nu - (2-\lambda)\mu^q] \quad .$$

Hence $\frac{\partial F_\lambda(\mu^q)}{\partial \mu^q} \gtreqless 0$ if $\mu^q \lesseqgtr \frac{\nu}{2-\lambda}$. This implies that $F_\lambda(\mu^q)$ is a quasiconcave function of μ^q and attains its minimum either for $\mu^q = \mu$ or $\mu^q = 1$, i.e.,

$$\underset{\mu^q \in [\mu,1]}{\text{Min}} \quad F_\lambda(\mu^q) = \text{Min}\,\{F_\lambda(\mu), F_\lambda(1)\}$$

and

$$\underset{\mu^q \in [\mu,1]}{\text{Min}} \quad F_\lambda(\mu^q) > 0$$

if and only if $F_\lambda(\mu) > 0$ and $F_\lambda(1) > 0$.

First we show that $F_\lambda(\mu^q)$ is an increasing function of λ for all $\mu^q \in [0,1]$:

$$\frac{\partial F_\lambda(\mu^q)}{\partial \lambda} = \mu^q(\beta^q)^3[\nu + (\lambda-1)\mu^q - \nu\lambda(\mu^q)^2] \quad .$$

But $G(\mu^q) = \nu + (\lambda-1)\mu^q - \nu\lambda(\mu^q)^2$ is a concave function of μ^q and

$$G(0) = \nu > 0$$

$$G(1) = (\lambda-1)(1-\nu) > 0 \ ,$$

hence

$$G(\mu^q) > 0 \quad \text{for all} \quad \mu^q \in [0,1],$$

and

$$\frac{\partial F_\lambda(\mu^q)}{\partial \lambda} > 0 \quad \text{for all} \quad \mu^q \in [0,1] \quad .$$

a) $F_\lambda(1) > 0$.

$$F_\lambda(1) = 1 - \frac{1}{\sqrt{1-\lambda(2-\lambda)}}(1-\lambda\nu) = 1 - \frac{1-\lambda\nu}{\lambda - 1} \quad .$$

Hence

$$F_\lambda(1) > 0 \ \text{if and only if} \ \lambda-1 > 1-\lambda\nu$$

or $$\lambda > \frac{2}{1+\nu} \quad .$$

b) $F_\lambda(\mu) > 0$.

$$F_\lambda(\mu) = 1 - \frac{1}{\sqrt{1-\lambda(2-\lambda)\mu^2}}(1-\lambda\mu\nu) \quad .$$

If $\lambda\mu\nu > 1$, then trivially $F_\lambda(\mu) > 0$. If $\lambda\mu\nu \leq 1$, then $F_\lambda(\mu) > 0$ if and only if

$$1 - \lambda(2-\lambda)\mu^2 > 1 - 2\lambda\mu\nu + \lambda^2\mu^2\nu^2$$

or $$\lambda > \frac{2(\mu-\nu)}{\mu(1-\nu^2)} = \frac{2}{1-\nu^2} - \frac{2\nu}{\mu(1-\nu^2)} \quad .$$

As we have

$$\frac{2}{1+\nu} - \frac{2(\mu-\nu)}{\mu(1-\nu^2)} = \frac{2\nu[1-\mu]}{\mu[1-\nu^2]} \geq 0 \quad ,$$

it follows clearly that

$$\inf_{\mu^q \in [\mu,1]} F_\lambda(\mu^q) > 0 \text{ if and only if } F_\lambda(1) > 0$$

or $$\lambda > \frac{2}{1+\nu} \quad ,$$

although $\inf_{\mu^q \in [\mu,1]} F_\lambda(\mu^q)$ does not need to be equal to $F_\lambda(1)$.

Hence from this if

$$2 \geq \lambda > \lambda^* = \frac{2}{1+\nu(P)} \quad ,$$

an infinite sequence is impossible in any of the normal cones to points of the boundary of P; the convergence is then finite.

Q.E.D.

We have assumed all along that the halfspace H^{i^q} was chosen to be the most distant halfspace from x^q or equivalently in view of the normalization of the $\{a^i\}$, i^q was also the index of the most violated inequality.

These assumptions can be almost entirely ignored.

We choose i^q to be the index of a violated inequality and let

$$\varepsilon^q = \frac{d(x^q, H^{i^q})}{d_P(x^q)} \quad .$$

Clearly $0 < \varepsilon^q \leq 1$.

(3.3.3) COROLLARY. Theorem (3.3.1) is valid provided:

$$\limsup \varepsilon^q = \varepsilon > 0 \quad .$$

Proof. (i) The proof of the infinite convergence follows easily: The sequence $\{x^q\}$ is still Fejér-monotone

$$d(x^q,P) \leq \frac{1}{\mu} d_P(x^q) = \frac{1}{\lambda\mu\varepsilon^q} \| x^{q+1} - x^q \| \ .$$

Also $d(x^q,P)$ is a non-negative, monotone decreasing sequence and $\{x^q\}$ converges to a point. So

$$\lim_{q\to\infty} \| x^q - x^{q+1} \| = 0$$

$$\lim_{q\to\infty} d(x^q,P) \geq 0$$

and if $\limsup \varepsilon^q = \varepsilon > 0$, a contradiction with $\lim d(x^q,P) > 0$ follows. Hence

$$\lim_{q\to\infty} d(x^q,P) = 0$$

and by continuity, $x^q \to x^* \in \partial P$.

(ii) In the proof of the finiteness of the convergence, we have

$$\mu^q = \frac{\|t^q - x^q\|}{d(x^q,P)} = \frac{d(x^q,H^{i^q})}{d_P(x^q)} \frac{d_P(x^q)}{d(x^q,P)} \geq \varepsilon^q \mu \quad .$$

But the condition on λ which ensures finite convergence is:

$$F_\lambda(1) > 0 \quad .$$

Hence $\lambda > \frac{2}{1+\nu(P)}$ still holds, provided of course that $\mu^q = \varepsilon^q \mu$ is bounded away from zero infinitely often, or

$$\limsup \varepsilon^q > 0 \quad . \qquad \text{Q.E.D.}$$

In particular, if the $\{a^i\}$ are not normalized, we can choose i^q to be the index of the most violated constraint, and Theorem (3.3.1) is still valid. It follows from the finiteness of I.

(3.3.4) LEMMA. Let $K \subseteq R^n$ be a closed convex body which is either defined by its complete family of supporting halfspaces or by a finite family (and hence K is a polyhedron).

Let $S_r(s)$ be any sphere contained in K. Then if at some step in the sequence $\{x^q\}$ generated by the relaxation method we have

$$2 \leq \lambda < 2 + \frac{r}{\|x^q - s\| - r} \quad .$$

The sequence $\{x^q\}$ converges finitely to a point of K.

Proof. (Figure 1) Lemma (3.1.2) gives

$$\|x^{q+1} - z\|^2 = \|x^q - z\|^2 - \lambda d_K(x^q)[(2-\lambda)d_K(x^q) - \langle a^{i^q}, t^q - z\rangle] \quad .$$

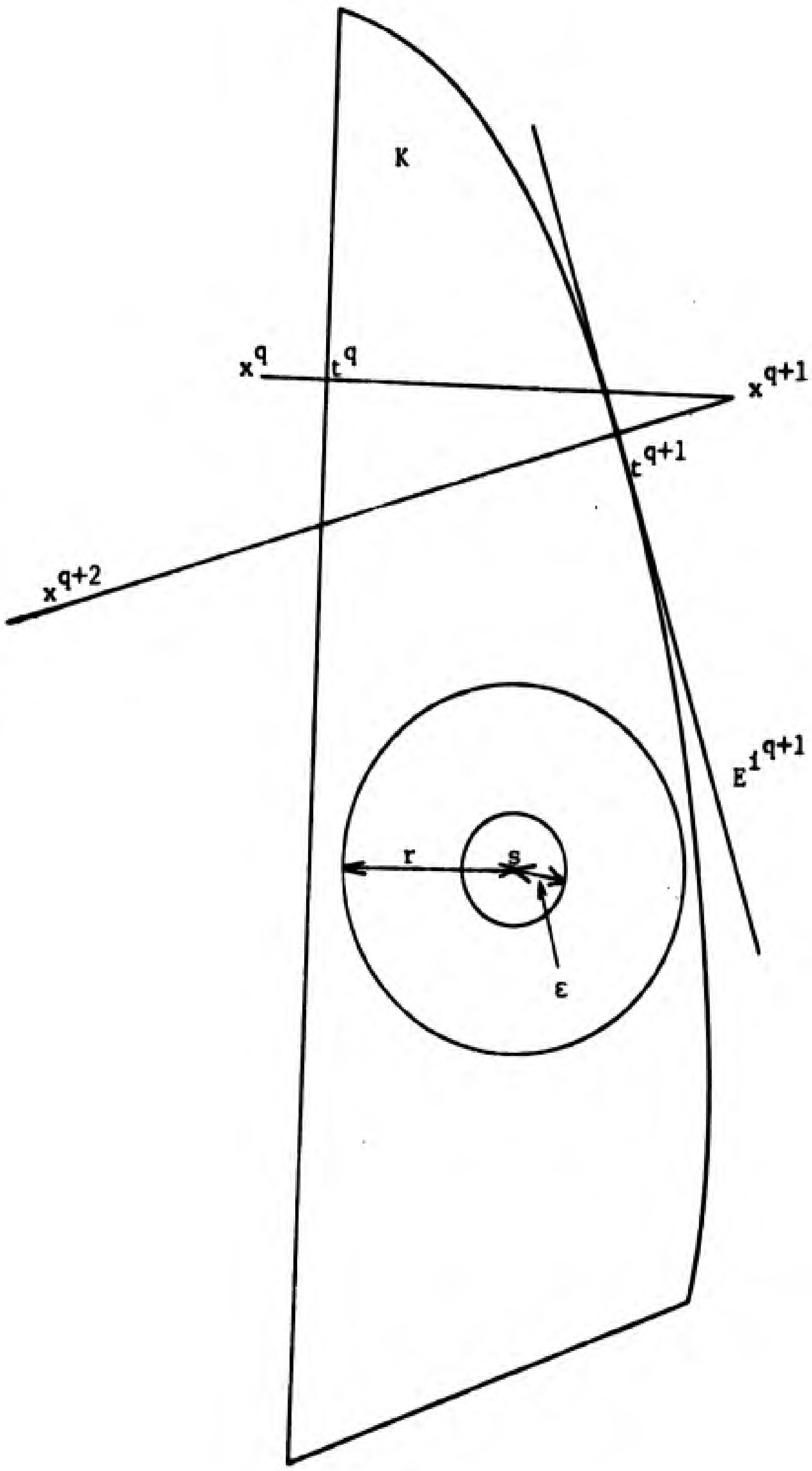


Figure 1

Hence
$$\|x^{q+1}-z\|^2 \leq \|x^q-z\|^2$$

if and only if
$$(\lambda-2)d_K(x^q) \leq \langle a^{i^q},t^q-z\rangle \quad .$$

But, as $S_r(s) \subseteq K$

$$\langle a^{i^q},t^q-z\rangle \geq r-\varepsilon \quad \text{for all} \quad z \in S_\varepsilon(s)$$

for $\varepsilon \in (0,r]$. So that

$$(\lambda-2)d_K(x^q) \leq r-\varepsilon$$

implies that x^{q+1} is closer than x^q from every point of $S_\varepsilon(s)$.

This, however, does not imply that the sequence $\{x^q\}$ is Fejér-monotone with respect to $S_\varepsilon(s)$ because there is no guarantee that $(\lambda-2)d_K(x^{q+1}) \leq r-\varepsilon$.

Now we have:

$$d_K(x) \leq d(x,K) \leq \|x-s\| - r \quad .$$

Hence
$$\|x^q-s\| - r \leq \frac{r-\varepsilon}{\lambda-2}$$

implies that
$$d_K(x^q) \leq \frac{r-\varepsilon}{\lambda-2} \quad .$$

It follows from this that

$$\lambda \leq 2 + \frac{r-\varepsilon}{\|x^q-s\|-r}$$

implies that x^{q+1} is closer than x^q from the points of $S_\varepsilon(s)$ and furthermore the same equation is satisfied at x^{q+1} as $\|x^{q+1}-s\| \leq \|x^q-s\|$.

Hence $\{x^q\}$ is a Fejér-monotone sequence with respect to $S_\varepsilon(s)$. By Theorem (1.3.2), the sequence $\{x^q\}$ either terminates at a point of K or converges to a point x^*.

But $\{x^q\} \to x^*$ implies that $d_K(x^*) = 0$ and hence $x^* \in \partial K$. This is ruled out by Theorems (3.3.1) and (3.2.3).

Therefore the sequence $\{x^q\}$ converges finitely, if for some $\varepsilon > 0$

$$\lambda \leq 2 + \frac{r-\varepsilon}{\|x^q-s\|-r}$$

and hence if

$$\lambda < 2 + \frac{r}{\|x^q-s\|-r} \quad .$$

(3.3.5) THEOREM. Let K be a closed convex body in R^n which is either defined by its complete family of supporting halfspaces, or by a finite family. If the recession cone $0^+(K)$ of K is full dimensional and has inner measure ν, then the relaxation method generates a finitely convergent sequence for any initial point $x^o \in R^n$ if

$$2 \leq \lambda < 2 + \frac{\nu}{1-\nu} \quad .$$

Proof. Let $C_\alpha(e)$ with $\|e\| = 1$ and $\sin \alpha = \nu$ be the largest spherical cone contained in $0^+(K)$. It follows that

$$z + C_\alpha(e) \subseteq K \quad \text{for all} \quad z \in K \quad .$$

Let $s = z + \eta e$ for $\eta \geq 0$. Then

$$S_{\eta\nu}(s) \subseteq z + C_\alpha(e) \subseteq K \quad \text{for all} \quad \eta \geq 0 \quad .$$

Also $x-s = x-z-\eta e$. Hence by Lemma (3.3.4) the sequence $\{x^q\}$ converges finitely if for some $\eta > 0$ we have

$$2 \leq \lambda < 2 + \frac{\eta\nu}{\|x-z-\eta e\| - \eta\nu} \quad .$$

Letting $\eta \to \infty$, the theorem follows. Q.E.D.

(3.3.6) THEOREM. Let K be a smooth enough convex body defined by its complete family of supporting halfspaces and η^* its modulus of smoothness. Then for any given $\lambda > 2$, the sequence generated by the relaxation method converges finitely for any initial point such that

$$d(x^o,K) < \frac{\eta^*}{\lambda-2} \quad .$$

Furthermore, if $d(x^o,K) \leq \frac{\eta^*}{\lambda-1}$, then one step convergence occurs.

Proof. Clearly, we have

$$x^{q+1} = x^q + \lambda(t^q-x^q) = x^q + \lambda d(x^q,K)a^{i^q}$$

where $t^q \in \partial K$, $a^{i^q} = \dfrac{t^q-x^q}{\|t^q-x^q\|}$. By definition of η^* as

$$a^{i^q} \in -N_K(t^q)$$

we have $$t^q + \eta a^{i^q} \in K \quad \text{for all} \quad \eta \in [0,\eta^*] \quad .$$

Hence (Figure 2)

$$d(x^{q+1},K) \leq \text{Max } \{<x^{q+1}-t^q-\eta^* a^{i^q},a^{i^q}>,0\}$$

$$= \text{Max } \{(\lambda-1)d(x^q,K)-\eta^*,0\} \quad .$$

Case 1. $d(x^q,K) \leq \frac{\eta^*}{\lambda-1}$. It implies that $d(x^{q+1},K) = 0$ and hence $x^{q+1} \in K$.

Case 2. $d(x^q,K) < \frac{\eta^*}{\lambda-2}$. Let $\varepsilon = \eta^* - (\lambda-2)d(x^q,K) > 0$. It follows that

$$d(x^{q+1},K) \leq d(x^q,K) - \varepsilon(\lambda-1)$$

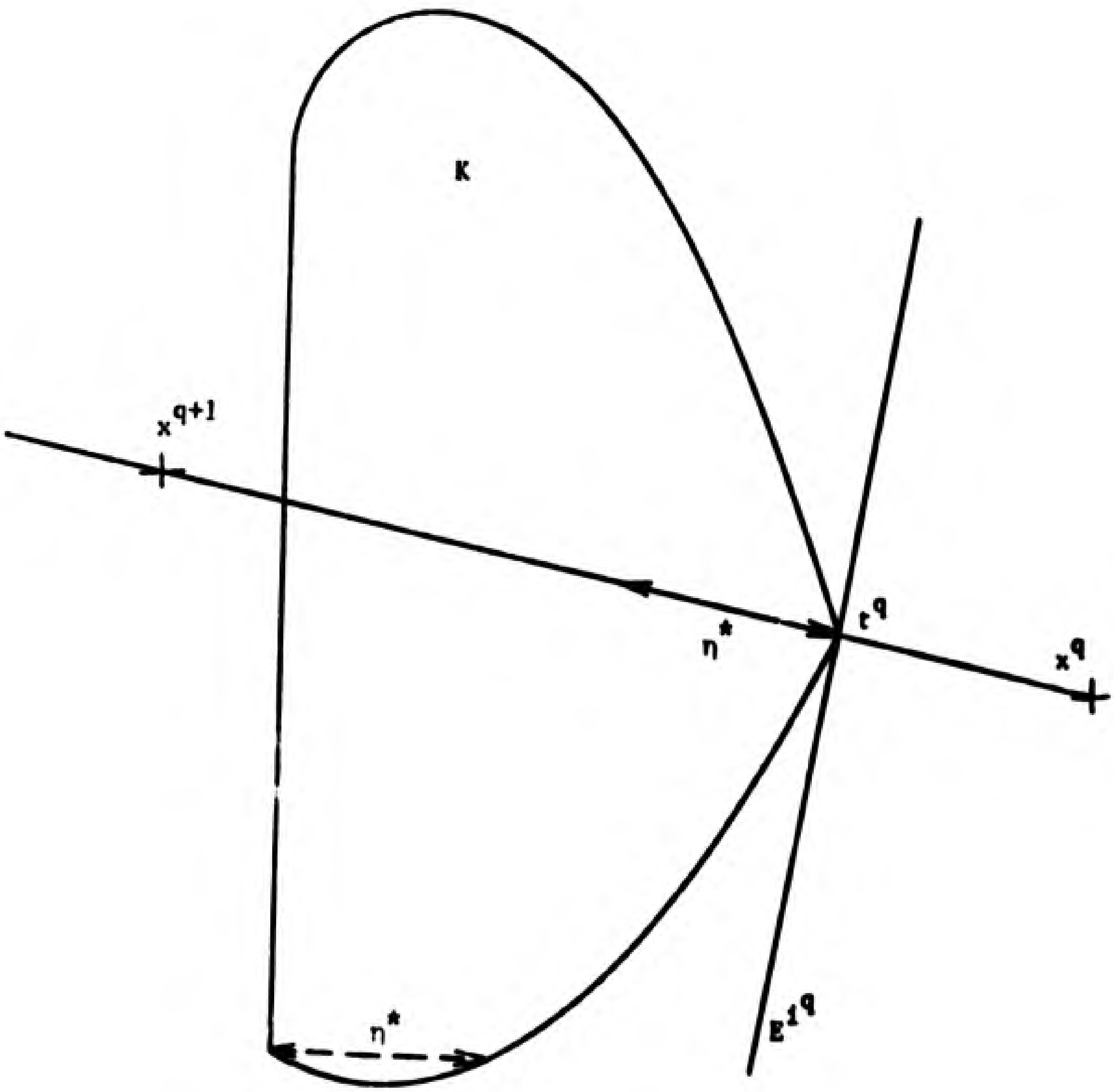


Figure 2

which clearly implies that

$$d(x^q,K) \leq d(x^o,K) - q\varepsilon(\lambda-1) \leq 0$$

for $q \geq \frac{d(x^o,K)}{\varepsilon(\lambda-1)}$. And hence there is finite convergence.

Q.E.D.

(3.3.7) <u>COROLLARY</u>. If C is a smooth enough closed convex cone defined by its complete family of supporting halfspaces, then the relaxation method converges after one iteration for any $\lambda > 1$.

3.4 <u>Relaxation With Respect to an Empty Set</u>

Let $P = \{x \in R^n | \langle a^i,x\rangle + b^i \geq 0,\ i \in I\}$, where I is finite, be an "empty polyhedron", given by an incompatible system of inequalities. We define a family of "parallel" sets by:

$$P_v = \{x \in R^n | \langle a^i,x\rangle + b^i + v \geq 0,\ i \in I\}$$

for $v \in R_+$. The residue function associated to P_v is

$$R_{P_v}(x) = \underset{i \in I}{\text{Max}} \{-\langle a^i,x\rangle - b^i - v\}$$

$$= R_P(x) - v$$

and the residual distance is

$$d_{P_v}(x) = \text{Max}\{R_P(x) - v, 0\} \quad .$$

We define

$$v^* = \underset{x \in R^n}{\text{Inf}}\ d_P(x) = \underset{x \in R^n}{\text{Inf}}\ R_P(x) > 0 \quad ,$$

the Čebyšev deviation of the system of inequalities.

Clearly, it follows that

$$w \in [0,w^*) \Rightarrow P_w = \emptyset$$

$$w = w^* \quad \Rightarrow P_w \neq \emptyset \text{ and } \dim P_w < n$$

$$w \in (w^*,\infty) \Rightarrow P_w \neq \emptyset \text{ and } \dim P_w = n \quad .$$

The set $P_{w^*} = \{x \in R^n | \; R_p(x) = w^*\}$ is the set of points which minimize the function $R_p(x)$, and is known as the set of Čebyšev solutions to the incompatible system of inequalities.

Generally $P_w = \{x \in R^n | \; R_p(x) \leq w\}$.

Let x^q be a point of the sequence generated by the relaxation method applied to P.

Clearly,

$$w^q = d_p(x^q) = R_p(x^q) \geq w^*$$

and if $w \leq w^q$, then $d_{P_w}(x^q) = w^q - w$; $w \geq w^q$ then $d_{P_w}(x^q) = 0$.

At the point x^q, the relaxation method applied to any P_w such that $w < w^q$ would choose the same direction of movement a^{i^q}.

Let

$$H_w^{i^q} = \{x \in R^n | \; \langle a^{i^q},x\rangle + b^{i^q} + w \geq 0\}$$

$$t_w^q = P_{H_w^{i^q}}(x^q) = x^q + d_{P_w}(x^q)a^{i^q}$$

$$= x^q + (d_p(x^q) - w)a^{i^q} \quad .$$

We will assume that $\lambda \in (0,2)$. Hence, by Lemma (3.1.2), x^{q+1} is closer from a point $z \in R^n$ if and only if

$$(2-\lambda)d_p(x^q) - 2\langle a^{i^q}, t^q - z\rangle \geq 0$$

or $$\frac{2-\lambda}{2} d_P(x^q) \geq \langle a^{i^q}, t_w^q + w a^{i^q} - z \rangle$$

or $$\frac{2-\lambda}{2} d_P(x^q) - w \geq \langle a^{i^q}, t_w^q - z \rangle \quad .$$

But we have $P_w \subseteq H_w^{i^q}$. Hence

$$\langle a^{i^q}, t_w^q - z \rangle \leq 0 \quad \text{for all} \quad z \in P_w \quad .$$

From this it follows that x^{q+1} is pointwise closer than x^q from P_w provided w satisfies

$$w \leq \frac{2-\lambda}{2} d_P(x^q)$$

or $$d_{P_w}(x^q) \geq \frac{\lambda}{2-\lambda} w \quad .$$

We also have

$$\frac{1}{\lambda} \| x^q - x^{q+1} \| = d_P(x^q) = [w + d_{P_w}(x^q)] \quad ,$$

if $w \leq d_P(x^q)$.

(i) We first show that for any $w > w^*$ the equation $d_{P_w}(x^q) \geq \frac{\lambda}{2-\lambda} w$ cannot hold for all q.

<u>Proof</u>. If it did, then $\{x^q\}$ would be a Fejér-monotone sequence with respect to the full dimensional polyhedron P_w.

By Theorem (1.3.2) it converges to a point, if the sequence does not terminate at a point of P_w. Hence

$$\lim_{q \to \infty} [w + d_{P_w}(x^q)] = \frac{1}{\lambda} \lim_{q \to \infty} \| x^q - x^{q+1} \| = 0$$

and $$\lim_{q \to \infty} d_{P_w}(x^q) = -w < 0 \quad ,$$

a contradiction. Q.E.D.

Note that

$$d_{P_w}(x^q) \geq \frac{\lambda}{2-\lambda} w$$

if and only if $\quad d_{P_{w^*}}(x^q) \geq \frac{\lambda}{2-\lambda} w^* + \frac{2}{2-\lambda}(w-w^*)$.

(ii) The equation $d_{P_{w^*}}(x^q) \geq \frac{\lambda}{2-\lambda} w^*$ can hold for all q, but then $\lim_{q\to\infty} d_{P_{w^*}}(x^q) = \frac{\lambda}{2-\lambda} w^*$ and this implies that the family of hyperplanes has some properties of symmetry.

Proof. For $w = w^*$, $\dim P_{w^*} < n$ and infinite convergence is ruled out as in (i).

By Theorem (1.3.2) it is possible that the sequence $\{x^q\}$ might have more than one limit point, but then the set of limit points belongs to a spherical surface having $\mathcal{M}(P_{w^*})$ as its axis.

Letting $\mu^* = \mu(P_{w^*})$ (see Theorem (2.7.2)), it follows that any limit point x^* will satisfy the following relations:

$$d_{P_{w^*}}(x^*) = \frac{\lambda}{2-\lambda} w^*$$

$$\frac{\lambda}{2-\lambda} w^* \leq d(x^*, P_{w^*}) \leq \frac{1}{\mu^*} \frac{\lambda}{2-\lambda} w^*$$

and hence $d(x^*, P_{w^*})$ is bounded. Also $d(x^*, P_{w^*})$ has the same value for all the limit points and

$$\lim_{q\to\infty} \|x^q - x^{q+1}\| = \lambda[w^* + \frac{\lambda}{2-\lambda} w^*] = \frac{2\lambda}{2-\lambda} w^* \quad .$$

All of this implies that P_{w^*} is rather symmetrical.

As exemplified in Appendix (1.1.2), if there are two limit points x^* and y^*, then it follows that x^* and y^* are symmetrical points with respect to a hyperplane E^* containing

P_{w^*} and in fact after a finite number of steps the relaxation process only uses two hyperplanes E_1 and E_2 which are parallel to E^* at a distance w^* of E^*. If there are more limit points, then similar results ensue. Q.E.D.

Now if we assume that (ii) does not occur, then

$$d_{P_{w^*}}(x^q) \geq \frac{\lambda}{2-\lambda} w^* \ , \quad \text{or equivalently} \quad d_p(x^q) \geq \frac{2}{2-\lambda} w^*$$

cannot hold for all q and hence the Fejér-monotonicity of $\{x^q\}$ with respect to P_{w^*} does not necessarily hold for all q. We have:

$$\|x^{q+1}-z\|^2 = \|x^q-z\|^2 + \lambda d_p(x^q)[-(2-\lambda)d_p(x^q) + 2w^* + 2<a^{i^q}, t^q_{w^*}-z>] \ .$$

But as

$$<a^{i^q}, t^q_{w^*}-z> \leq 0 \quad \text{for all} \quad z \in P_{w^*} \quad ,$$

it follows that

$$\|x^{q+1}-z\|^2 \leq \|x^q-z\|^2 + \lambda d_p(x^q)[2w^* - (2-\lambda)d_p(x^q)]$$

for all $z \in P_{w^*}$. Theorem (2.7.2) applied to P_{w^*} gives:

$$\mu^* d(x^q, P_{w^*}) \leq d_p(x^q) - w^* \leq d(x^q, P_{w^*}) \quad .$$

<u>Case 1</u>. $d_p(x^q) \geq \frac{2}{2-\lambda} w^*$ implies that

$$d^2(x^{q+1}, P_{w^*}) \leq d^2(x^q, P_{w^*}) \quad .$$

<u>Case 2</u>. $w^* \leq d_p(x^q) \leq \frac{2}{2-\lambda} w^*$. We have

$$d^2(x^{q+1}, P_{w^*}) \leq \frac{1}{\mu^{*2}}[d_p(x^q) - w^*]^2 + \lambda d_p(x^q)[2w^* - (2-\lambda)d_p(x^q)] \ .$$

This is a convex function of $d_p(x^q)$, as the coefficient of $d_p^2(x^q)$ is positive for any $\lambda \in (0,2)$. Hence

$$d^2(x^{q+1},P_{w*}) \leq \text{Max } \{\lambda^2 w^{*2}, \frac{\lambda^2 w^{*2}}{(2-\lambda)^2 \mu^{*2}}\}$$

and

$$d(x^{q+1},P_{w*}) \leq \frac{\lambda w^*}{(2-\lambda)\mu^*} \text{Max } \{1,(2-\lambda)\mu^*\}$$

All of this can be roughly summarized in:

(3.4.1) THEOREM. Let $\{\langle a^i,x\rangle + b^i \geq 0,\ i \in I\}$ be an incompatible finite system of linear inequalities. Then the sequence $\{x^q\}$ generated by the relaxation method, for any given $\lambda \in (0,2)$, has a bounded distance from P_{w*}, the set of Čebyšev points. Furthermore after a finite number Q of steps we have:

$$d(x^q,P_{w*}) \leq \frac{2\lambda w^*}{(2-\lambda)\mu^*} \quad \text{for all} \quad q \geq Q \quad .$$

This theorem can be compared to the following result of Eremin ([19], [20]), where $\{\lambda^q |\ q \in Z_+\} \subseteq R_+$ is a sequence of relaxation parameters: Under the conditions of Theorem (3.4.1), a necessary condition for the convergence of the sequence $\{x^q\}$ to a point of the set P_{w*} is

$$\sum_{q=1}^{\infty} \lambda^q = \infty, \qquad \lambda^q \to 0 \quad .$$

3.5 Relaxation With Respect to a Convex Set Defined by Nonlinear Inequalities

Let $K = \{x \in R^n |\ g^i(x) \geq 0,\ i \in I\}$ where

I is finite

$g^i: R^n \to R$ is concave and differentiable.

We also assume that there exists a point $x \in K$ such that $g^i(x) > 0$ for all $i \in I$. It follows from this that K is a closed convex body and also that $g^i(x) > 0$ for all $i \in I$, $x \in \text{Int } K$.

Let $x \in R^n/K$. It implies that the set of indices of the constraints violated by x is not empty:

$$J(x) = \{i \in I \mid g^i(x) < 0\} \neq \emptyset \quad \text{for all} \quad x \notin K \ .$$

From the concavity of g^i, it follows that, if we denote by $\nabla g^i(x)$ the gradient of g^i

$$<\nabla g^i(x), y-x> \geq g^i(y) - g^i(x)$$

for all $y \in R^n$. Clearly, it follows that

$$<\nabla g^i(x), y-x> + g^i(x) \geq 0 \quad \text{for all} \quad i \in I, \quad y \in K.$$

Define

$$H_x^i = \{z \in R^n \mid <z-x, \nabla g^i(x)> + g^i(x) \geq 0\}$$

for $i \in I$, $x \in R^n$.

We have

(i) H_x^i is a closed halfspace which contains K for all $i \in I$, $x \in R^n$.

(ii) $x \notin H_x^i$ for all $i \in J(x)$.

(3.5.1) The boundary hyperplane E_x^i has been called the *cutting plane* associated to the point x and the constraint g^i.

To every point x, we can associate a polyhedron $P_x = \bigcap_{i \in I} H_x^i$ which contains K.

From the definition of H_x^i, it follows that

$$K = \bigcap_{x \in R^n} P_x = \bigcap_{x \notin K} P_x = \bigcap_{x \notin K} [\bigcap_{i \in J(x)} H_x^i] \quad .$$

At every point x^q of the relaxation sequence, we generate a finite subfamily $\{H^i_{x^q} \mid i \in J(x^q)\}$ of the family $\{H_x^i \mid i \in J(x),$ $x \in R^n\}$.

The point x^{q+1} will be generated from x^q through one iteration of the relaxation method applied to the polyhedron P_{x^q}.

(3.5.2) The *residual distance* to K can be defined by

$$d_K(x) = d_{P_x}(x) = \text{Max}\ \{0, \underset{i \in I}{\text{Max}}\ \{\frac{-g^i(x)}{\|\nabla g^i(x)\|}\}\}$$
$$= \text{Max}\ \{0, \underset{i \in J(x)}{\text{Max}}\ \{\frac{-g^i(x)}{\|\nabla g^i(x)\|}\}\} \quad .$$

It might be simpler, as far as choosing the constraint on which we project, to use

$$\ell_K(x) = \text{Max}\ \{0, \underset{i \in J(x)}{\text{Max}}\ \{-g^i(x)\}\} \quad .$$

(3.5.3) LEMMA. Let $s \in \text{Int}\ K$ and r be an arbitrary positive number. Then there are finite positive numbers $a, b, c, d,$ depending on s and r only, such that

(i) $b\ell_K(x) \leq d_K(x) \leq a\ell_K(x)$ for all $x \in S_r(s)$

(ii) $c \leq \|\nabla g^i(x)\| \leq d$ for all $i \in J(x)$, $x \in S_r(s)/\text{Int}\ K$.

Proof. If $x \in K$, it is trivial. If $x \notin K$, then

$$\langle \nabla g^i(x), s-x \rangle \geq g^i(s) - g^i(x) \quad .$$

Now for all $i \in J(x) \cup I(x)$, where $I(x) = \{i \in I \mid g^i(x) = 0\}$,

$g^i(x) \leq 0$. Also for all $i \in I$, $g^i(s) > 0$.

Hence using Schwarz inequality

$$\|\nabla g^i(x)\| \geq \frac{-g^i(x)+g^i(s)}{\|x-s\|} \quad \text{for all} \quad i \in J(x) \cup I(x) \quad .$$

Let $g = \underset{i \in I}{\text{Min}}\ g^i(s)$. Then for any $x \notin \text{Int } K$ and any $i \in J(x) \cup I(x)$ we have

$$\|\nabla g^i(x)\| \geq \frac{g^i(s)}{\|x-s\|} \geq \frac{g}{\|x-s\|} \geq \frac{g}{r} \quad .$$

The existence of an upper bound d on $\|\nabla g^i(x)\|$ follows from the continuity of $\nabla g^i(x)$.

We also have for $i \in J(x)$

$$-\frac{g^i(x)}{d} \leq -\frac{g^i(x)}{\|\nabla g^i(x)\|} \leq -\frac{g^i(x)}{g}\|x-s\|$$

and

$$\frac{\ell_K(x)}{d} \leq d_K(x) \leq \frac{\|x-s\|}{g}\ell_K(x) \quad .$$

Q.E.D.

A reasoning similar to the one used in Corollary (3.3.3) showsthat the convergence properties of the relaxation method are not affected by which function of ℓ_K or d_K is used to choose the index i^q of the constraint towards which we move. In any case, the direction of movement is

$$a^{i^q} = \frac{\nabla g^{i^q}(x^q)}{\|\nabla g^{i^q}(x^q)\|} \quad .$$

Let $H^{i^q} = H^{i^q}_{x^q} = \{y \in R^n \mid \langle \nabla g^{i^q}(x^q), y-x^q\rangle + g^{i^q}(x^q) \geq 0\}$.

We have

$$t^q = x^q + \frac{-g^{i^q}(x^q)}{\|\nabla g^{i^q}(x^q)\|^2} \nabla g^{i^q}(x^q) \quad .$$

Hence x^{q+1} will be

$$x^{q+1} = x^q + \lambda d_K(x^q) a^{i^q}$$

or

$$x^{q+1} = x^q + \lambda \frac{\ell_K(x^q)}{\|\nabla g^{i^q}(x^q)\|} a^{i^q} \quad ,$$

depending on whether we use d_K or ℓ_K to choose i^q.

It also follows that both ℓ_K and d_K are continuous on R^n, take the value zero on K, and positive values on R^n/K, but only ℓ_K is convex.

(3.5.4) LEMMA. The sequence $\{x^q\}$ generated by the relaxation method is Fejér-monotone with respect to K, for any $\lambda \in (0,2]$.

Proof. At any step in the sequence x^{q+1} is pointwise closer than x^q from P_{x^q} (using Theorem (3.1.4)). But $K \subseteq P_{x^q}$ and the Lemma follows. Q.E.D.

(3.5.5) THEOREM. The sequence $\{x^q\}$ generated by the relaxation method applied to the closed convex body $K = \{x \in R^n | \ g^i(x) \geq 0, i \in I\}$ converges finitely or infinitely to a point of K for any $\lambda \in (0,2]$.

If the convergence is infinite, then the limit of the sequence is on the boundary of K.

Proof. Similar to Theorem (3.1.6) Q.E.D.

(3.5.6) CONJECTURE. Theorem (3.3.1) is valid in this case.

The proof of this will be left for a later work.

The conditions under which Theorem (3.5.5) is valid can be weakened in various directions: See Eremin ([22], [23]), Germanov and Spiridonov ([28]), Polyak ([53]), Schlechter ([56]), Spiridnov ([58]) and the references given in those papers.

CHAPTER IV

CUI BONO?

4.1 Procedures

The conditions which ensure the finite convergence of the relaxation method applied to a full dimensional polyhedron defined by a finite system of linear inequalities are very lax, if we assume that λ satisfies the conditions of Theorem (3.3.1).

Corollary (3.3.3) validates a vast number of computational schemes.

Let $P = \bigcap_{i \in I} H^i$ be a full dimensional polyhedron where

$H^i = \{x \in R^n \mid \langle a^i, x \rangle + b^i \geq 0\}$

$a^i \neq 0$

I is a finite set.

(4.1.1) The Maximal Residual Method. The direction of movement is the interior normal of the halfspace corresponding to a most violated constraint. Theorem (3.3.1) holds without change.

(4.1.2) Systematic Scanning of the Index Set. Assume for convenience that $\|a^i\| = 1$ for all $i \in I$. Let $i(.): Z_+ \to I$ be a function from the set of non-negative integers onto the index set I. We assume that there is an integer p such that

$$i([q, q+p] \cap Z_+) = I \qquad \text{for all } q \in Z_+ .$$

This means that every index of I is repeated after at most $p+1$ iterations.

At step q, we choose the index of the halfspace on which we project to be $i(q)$

$$x^{q+1} = x^q + \lambda d(x^q, H^{i(q)}) a^{i(q)} \quad .$$

Then Theorem (3.3.1) holds without change.

Proof. The only case in which (3.3.1) could fail is if $\lim_{q\to\infty} \varepsilon^q = 0$ where

$$1 \geq \varepsilon^q = \frac{d(x^q, H^{i(q)})}{d_P(x^q)} \geq 0 \quad .$$

(i) Infinite Convergence

$\{x^q\}$ is Fejér-monotone with respect to P and $\dim P = n$. Hence by (1.3.2) $\{x^q\}$ converges to a point, say x^*.

Assume that $x^* \notin P$. Then clearly $d_P(x^*) > 0$ and $d_P(x^*) = d(x^*, H^{i^*})$ for some $i^* \in I$.

We have $\lim_{q\to\infty} \|x^q - x^{q+1}\| = 0$ which implies that $\lim_{q\to\infty} d(x^q, H^{i(q)}) = 0$

But

$$d(x^*, H^{i(q)}) = \text{Max } \{0, -<a^{i(q)}, x^q> - b^{i(q)} + <a^{i(q)}, x^q - x^*>\} \quad .$$

It follows that

$$\lim_{q\to\infty} d(x^*, H^{i(q)}) = 0 \quad .$$

This contradicts the fact that i^* is repeated infinitely often.

(ii) Finite Convergence

We use the notations of the proof of Theorem (3.3.1).

Case 1. If $C_P(x^*)$ is smooth enough, then the proof used for $\lambda = 1$ will be valid here for any $\lambda \geq 1$.

Case 2. For the same values of λ, we still have $D^q - D^{q+1} \geq 0$. Assume that $\lim_{q\to\infty} \varepsilon^q = 0$. Using Theorem (2.7.2)

$$\mu^* \|u^q\| \leq d_{C_p(x^*)}(u^q) \leq \|u^q\| \quad .$$

It implies that $\mu^* \varepsilon^q \|u^q\| \leq d(x^q, H^{i(q)}) \leq \varepsilon^q \|u^q\|$ and

$$\lim_{q\to\infty} \text{Max}\ \{0, <-a^{i(q)}, \frac{u^q}{\|u^q\|}>\} = 0 \quad .$$

Hence, given any $\delta > 0$, there exists an integer Q such that for all $q \geq Q$, we have

$$<a^{i(q)}, \frac{u^q}{\|u^q\|}> + \delta \geq 0 \quad .$$

As we assume that $\{u^q\} \subseteq N_p(x^*)$, clearly $\{\frac{u^q}{\|u^q\|}\} \in N_p(x^*) \cap B$. Using the fact that $\lim_{q\to\infty} \varepsilon^q = 0$

$$u^{q+1} = u^q + \lambda d(x^q, H^{i(q)}) a^{i(q)}$$

leads to

$$\lim_{q\to\infty} \frac{\|u^{q+1}\|}{\|u^q\|} = 1 \quad .$$

Also

$$\frac{u^{q+1}}{\|u^{q+1}\|} = \frac{u^q}{\|u^q\|} \frac{\|u^q\|}{\|u^{q+1}\|} + \lambda \frac{d(x^q, H^{i(q)})}{\|u^q\|} \frac{\|u^q\|}{\|u^{q+1}\|} a^{i(q)} \quad .$$

It follows that if for some subsequence $Z^* \subseteq Z_+$, $\frac{u^q}{\|u^q\|}$ converges to u^*, then we also have

$$\lim_{\substack{q\to\infty \\ q\in Z^*}} \frac{u^{q+1}}{\|u^{q+1}\|} = u^* \quad .$$

Hence for any integer $t > 0$

$$\lim_{\substack{q\to\infty \\ q\in Z^*+t}} \frac{u^q}{\|u^q\|} = u^* \quad ,$$

though not necessarily uniformly in t. As $\{\frac{u^q}{\|u^q\|}\} \in N_p(x^*) \cap B$, a compact set, the sequence has at least one limit point $u^* \in N_p(x^*) \cap B$.

Let Z^* be a subsequence of Z_+ such that

$$\lim_{\substack{q\to\infty \\ q\in Z^*}} \frac{u^q}{\|u^q\|} = u^* \quad .$$

We can then find an integer Q_1 such that

$$<a^{i(q)},u^*> + 2\delta \geq 0$$

for all $q \geq Q_1$ and such that $q \in Z^* + \{0,1,\ldots,p\}$. This implies that

$$<a^i,u^*> + 2\delta \geq 0 \quad \text{for all} \quad i \in I(x^*)$$

or equivalently

$$d_{C_p(x^*)}(u^*) = \max_{i\in I(x^*)} <-a^i,u^*> \leq 2\delta \quad .$$

But this implies in turn that

$$\mu^* \leq 2\delta \quad ,$$

a contradiction as $\mu^* > 0$ and $\delta > 0$ was arbitrary.

Q.E.D.

(4.1.3) Partitioning

Let $I = I_1 \cup I_2$ be a partition of I and let $P_j = \bigcap_{i\in I_j} H^i$ for $j \in \{1,2\}$. Then $P = P_1 \cap P_2$.

Define a function $j(.): Z_+ \to \{1,2\}$ satisfying the condition given in (4.1.2). Then, at step q, we go through one iteration of the relaxation method as applied to $P_{j(q)}$.

Clearly I can be partitioned into more than two sets. Theorem (3.3.1) holds in this case, but the values of λ which guarantee finite convergence depend on the properties of the set P and not on the separate properties of P_1 and P_2.

Example. Let $P = \{x \in R^n | \; Ax + b \geq 0, \; x \geq 0\}$ where A is an $m \times n$ matrix and b an m vector.

Let $P_1 = \{x \in R^n | \; Ax + b \geq 0\}$

$P_2 = R^n_+ = \{x \in R^n | \; x \geq 0\}$.

At step $2q$, we will apply one iteration of the relaxation method to P_1. At step $2q+1$, we apply the relaxation method to P_2, a number of times equal to the number of negative components of x^{2q+1}.

Clearly those steps can be summarized by

$$x_k^{2q+2} = (\lambda-1)\left|x_k^{2q+1}\right| \quad \text{if} \quad x_k^{2q+1} < 0$$

$$x_k^{2q+2} = x_k^{2q+1} \quad \text{if} \quad x_k^{2q+1} \geq 0$$

where the index k represents the components of x in R^n.

Clearly Theorem (3.3.1) is valid and this generalizes a result obtained by Eremin in [21].

(4.1.4) Budding

The comparison of Theorems (3.2.3) and (3.3.1) suggests the idea that a larger family of supporting halfspaces could improve the convergence properties of the relaxation method.

Any halfspace

$$H = \{x \in R^n | \; <\sum_{i \in I} k^i a^i, x> + \sum_{i \in I} k^i b^i \geq 0\}$$

where $k^i \in R_+$ contains P and hence any finite family of such halfspaces can be used.

At every point $x \in R^n$ we define

$$J(x) = \{i \in I \mid \langle a^i, x\rangle + b^i < 0\} \quad ,$$

the set of the indices of the inequalities violated by x.

For any index set $J \subseteq I$ we define

$$P_J = \{x \in R^n \mid J(x) = J\} \quad ,$$

the *cavity* associated to the index set J (see [47]).

Clearly $\{P_J \mid J \subseteq I\}$ is a finite partition of R^n and every P_J is either empty or full dimensional.

Merzlyakov goes one step further and partitions the cavities into *subcavities* according to the index of the most violated constraint (if there is a tie, it is broken by a preassigned ordering on I). For any $j \in J$, $J \subseteq I$

$$P_J^j = \{x \in P_J \mid d(x,H^j) \geq d(x,H^i) \text{ for all } i \in J, \text{ and } d(x,H^j) > d(x,H^i) \text{ for all } i < j,\ i \in J\}.$$

Clearly $\{P_J^j \mid j \in J,\ J \subseteq I\}$ is a finite partition of R^n.

To every non-empty subcavity P_J^j, we associate non-negative numbers $k^i(J,j)$ for $i \in J$. We assume that

$$k^j(J,j) > 0$$

and

$$\sum_{i \in J} k^i(J,j) a^i \neq 0 \quad .$$

This last assumption is not necessary if P is not empty. This follows from Gordan's theorem (see [27], p.48):

If $\sum_{i\in J} k^i a^i = 0,\ \ k^i \geq 0,\ \ k^j > 0,\ \ j \in J$ has a solution, then

$$\langle a^i, u\rangle > 0,\ \ i \in J$$

has no solution. But if $x \in P$ and $y \in P^j_J$, then $u = x-y$ solves this last system.

The halfspace associated to the subcavity P^j_J is by definition

$$H^j_J = \{x \in R^n \mid \langle \sum_{i\in J} k^i(J,j)a^i, x\rangle + \sum_{i\in J} k^i(J,j)b^i \geq 0\} \quad .$$

We have the following properties:

(i) $H^j_J \supseteq P$

(ii) $H^j_J \cap P^j_J = 0$

(iii) The family $\{H^j_J \mid j \in J,\ J \subseteq I\}$ contains the halfspaces H^i whose boundary hyperplanes intersect P in a facet of P.

(iv) Hence $P = \bigcap_{J\subseteq I} \ \bigcap_{j\in J} H^j_J$.

This family of hyperplanes is finite and can be used to replace, at least conceptually, the original family $\{H^i \mid i \in I\}$.

To every $x^q \in R^n$ corresponds a unique subcavity P^j_J containing x^q. We will use the halfspace H^j_J to define the next iteration of the relaxation sequence

$$x^{q+1} = x^q + \lambda d(x^q, H^j_J) a^j_J$$

where

$$a^j_J = \frac{\sum_{i\in J} k^i(J,j)a^i}{\| \sum_{i\in J} k^i(J,j)a^i \|} \quad ,$$

$$d(x^q, H^j_J) = \frac{\sum_{i\in J} k^i(J,j)d(x^q, H^i)}{\| \sum_{i\in J} k^i(J,j)a^i \|} \quad .$$

It follows that

$$d(x^q,H^j_J) \geq \frac{k^j(J,j)}{\sum_{i\in J} k^i(J,j)} d_P(x^q) \quad .$$

As Corollary (3.3.3) is satisfied, Theorem (3.3.1) holds without change in this case. This generalizes a theorem of Merzlyakov [47]. This process of generating halfspaces, which we describe as budding, does not need to be carried through in advance. We only need to generate the halfspace associated to the subcavity in which x^q lies, and the choice of the coefficients can be done then. We could, of course, associate to every subcavity a finite number of halfspaces without changing Theorem (3.3.1). This idea could give rise to numerous specific algorithms.

(4.1.5) Anti-Jamming

Lemma (3.3.2) states that the relaxation sequence $\{x^q\}$ converges infinitely to a point $x^* \in \partial P$ only if the whole sequence $\{x^q\}_{q\geq Q}$ "jams" into $x^* + N_P(x^*)$.

A similar phenomenon occurs in feasible direction methods (see [69], pp. 274-280), and the procedure devised to alleviate jamming can be adapted here.

Let $x^q \notin P$, H^{i^q} be the most distant halfspace, i.e., $d_P(x^q) = d(x^q,H^{i^q})$, $J_1 \subseteq I$ be a set of indices of violated constraints, and $J_2 \subseteq I$ be a set of indices of mildly satisfied constraints.

The anti-jamming procedure consists in using at step q the halfspace

$$H = \{x \in R^n \mid \langle a,x\rangle + b \geq 0\}$$

where

(i) $a = \dfrac{\sum_{i\in J_1\cup J_2} k^i a^i}{\left\| \sum_{i\in J_1\cup J_2} k^i a^i \right\|}$ $\qquad b = \dfrac{\sum_{i\in J_1\cup J_2} k^i b^i}{\left\| \sum_{i\in J_1\cup J_2} k^i a^i \right\|}$

(ii) $k^i \geq 0, \quad i \in J_1 \cup J_2$

(iii) $d(x^q,H) = \dfrac{\sum_{i\in J_1} k^i d(x^q,H^i) - \sum_{i\in J_2} k^i d(x^q,E^i)}{\left\| \sum_{i\in J_1\cup J_2} k^i a^i \right\|} \geq \epsilon d_p(x^q)$

for some $\epsilon > 0$ and where $E^i = \partial H^i$. Then $x^{q+1} = x^q + \lambda d(x^q,H)a$. Under the condition that we generate only a finite number of such hyperplanes, Theorem (3.3.1) holds without change.

This method attempts to generate the halfspace

$$H^* = \{x \in R^n \mid \langle -x^q + P_p(x^q), x - P_p(x^q)\rangle \geq 0\}$$

which for $\lambda = 1$ would give convergence in one step to $P_p(x^q)$. In fact, if we assume that we know the set of indices of the constraints satisfied with equality by $P_p(x^q)$, say I^*, then we can find numbers $k_*^i \geq 0$, $i \in I^*$ which will generate the halfspace H^*. We number the constraints so that $I^* = \{1,\ldots,m\}$. Let $A^* = (a^1,\ldots,a^m)'$ be the $m \times n$ matrix composed of the interior normals of the halfspaces H^i for $i \in I^*$, where $'$ denote transposition. Let $b^* = (b^1,\ldots,b^m)'$.

We assume for simplicity that the rank of A^* is m.

Define the manifold $M = \{x \in R^n \mid A^*x + b^* = 0\}$. It is the affine hull of the least face of P containing $P_p(x^q)$.

We claim that

$$P_M(x^q) = [I_n - A^{*\prime}(A^*A^{*\prime})^{-1}A^*]x^q - A^{*\prime}(A^*A^{*\prime})^{-1}b^*$$

where I_n is the $n \times n$ unit matrix. This follows from

(i) $A^*P_M(x^q) + b^* = 0$ and hence $P_M(x^q) \in M$

(ii) Defining $(k_*^1,\ldots,k_*^m)' = -(A^*A^{*\prime})^{-1}(A^*x^q + b^*)$, we get $x^q = P_M(x^q) - \sum_{i=1}^{m} k_*^i a^i$, or $x^q \in P_M(x^q) + N_M(P_M(x^q))$ and the claim follows from Theorem (2.3.3).

If our guess I^* of the index set $I(P_p(x^q))$ was right, then in addition we have

(iii) $P_p(x^q) = P_M(x^q) \subset$

(iv) $k_*^i \geq 0$, $i = 1,\ldots,m$

(v) $H^* = \{x \in R^n | <\sum_{i=1}^{m} k_*^i a^i, x> + \sum_{i=1}^{m} k_*^i b^i \geq 0\}$, and $x^{q+1} = P_{H^*}(x^q) = P_p(x^q)$.

This idea could be used to alleviate the infinite convergence of the relaxation method. If $\{x^q\}$ converges infinitely to $x^* \in \partial P$, then

$$\lim_{q\to\infty} (<a^i,x^q> + b^i) \begin{cases} = 0 & \text{if } i \in I(x^*) \\ > 0 & \text{if } i \in I/I(x^*) \end{cases}$$

and this allows the identification of $I(x^*)$, at least in a limiting sense.

It could also lead to the development of more simplex-like methods, in which the coefficients k^i are computed. Let $J \subseteq I$ and $M_J = \{x \in R^n | <a^i,x> + b^i = 0, i \in J\}$ the manifold defined by the constraints $i \in J$. Then, as before, coefficients k^i, $i \in J$ can be computed and if $k^i \geq 0$, for all $i \in J$, then

$$H = \{x \in R^n \mid \langle \sum_{i \in J} k^i a^i, x \rangle + \sum_{i \in J} k^i b^i \geq 0\}$$

can be used to define a "Fejér feasible" iteration

$$x^{q+1} = x^q + \lambda d(x^q, H) \frac{\sum_{i \in J} k^i a^i}{\|\sum_{i \in J} k^i a^i\|} .$$

4.2 Solution Strategies

The choice of a sequence $\{\lambda^q\}$ of relaxation parameters can be guided by the results obtained in this work.

It should be clear, though, that the development of reasonably tight bounds on the number of steps needed to converge and an extensive experimentation will lead to more precise guidelines.

An interesting line of investigation would be the idea of using the information gained through the sequence $\{d_p(x^q)\}$ to choose the relaxation parameters.

We will list below some facts which are relevant to the choice of $\{\lambda^q\}$.

We assume that $P = \bigcap_{i \in I} H^i$ is a full dimensional polyhedron $P = \{x \in R^n \mid Ax + b \geq 0\}$.

Fact 1. Every vector $x^q \in R^n$ can be decomposed uniquely in two vectors $y^q = P_p(x^q)$ and z^q belonging to dual elements of $\mathcal{F}(P)$ and Γ_p (see Theorem (2.4.21)).

Hence if there is infinite convergence of the sequence $\{x^q\}$ to a point $x^* \in \partial P$, then after a finite number of iterations Q, we have (see Lemma (3.3.2))

$$y^q = x^* \quad \text{for all } q \geq Q$$
$$z^q \in N_P(x^*) \quad \text{for all } q \geq Q \ .$$

We could say that the sequence $\{x^q\}_{q \geq Q}$ falls into a normal cone trap $x^* + N_P(x^*)$. This happens if λ is not large enough to steer the sequence out of the trap.

A normal cone trap can be escaped by increasing λ above $\frac{2}{1+\nu(C_P(x^*))}$, or by the use of the anti-jamming and budding procedures.

Fact 2. Theorem (3.1.7) shows that for $\lambda \in [1,2)$ the convergence rate is geometric

$$\|x^q - x^*\| \leq 2d(x^0,P)\theta^q$$
$$d(x^q,P) \leq d(x^0,P)\theta^q$$

where $\theta = \sqrt{1-\lambda(2-\lambda)\mu^{*2}}$.

Clearly θ is minimum for $\lambda = 1$, and $\theta = 1$ for $\lambda = 2$. This indicates that the guaranteed reduction in distance at each step is greater if $\lambda = 1$ and that for $\lambda = 2$, we could have $d(x^{q+1},P) = d(x^q,P)$.

Fact 3. Theorem (2.7.4) indicates that if the initial point x^0 were chosen at random "close" to the set P, the probability that it lies in the normal cone of a facet of P is high. If x^0 were chosen at random "far" from P (assumed to be bounded), it would be likely to lie in the normal cone of a vertex of P.

We could say that the cone traps vanish when we get close enough to P, but that if the initial point is far enough from

P, and if λ is close to 1, the relaxation method manages to steer the sequence into them. The smooth enough property is very pleasant, in the sense that it rules out the existence of any cone traps.

<u>Fact 4</u>. The finiteness of the convergence of the relaxation method was proved by contradiction in Theorem (3.3.1). This technique of proof has the drawback that even if it shows that finite convergence cannot not happen, it does not explain why it happens. The following considerations throw some light on this and in fact they sketch a constructive proof.

To every point x^q of the sequence, we associate the point $y^q = P_P(x^q)$ and the least face F^q of $\mathcal{F}(P)$ which contains y^q. $\{F^q\}$ is a sequence of nodes on the facial graph representing the face lattice $\mathcal{F}(P)$. As λ is large enough to escape any cone trap, this means that at some iteration q, we have $y^{q+1} \neq y^q$. If x^q is close enough to P, then $\|y^{q+1}-y^q\| \leq \|x^{q+1}-x^q\| = \lambda d_P(x^q)$ is small. Given any $y \in \partial P$, there exists a neighborhood in which there are no vertices of lower order (see (2.2.11)) than y, and the vertices of equal order belong to $F(y)$, the least face containing y. This means that there is a tendency for dim (F^q) to increase when there is a transition to another node of the facial graph.

On the other hand, if $\dim(F^{q+1}) < \dim F^q$, this means that y^{q+1} is "far" from y^q, which implies, most of the time, that $d(x^{q+1},P)$ is sizably less than $d(x^q,P)$.

Roughly said, normally when x^q escapes a cone trap, either

$\dim(F^q)$ increases, or $d(x^q, P)$ diminishes sizably.

Noticing that if $\dim(F^q) = n-1$, one step convergence is likely, if we are close enough to P, this explains the finiteness of the convergence.

A constructive proof can be obtained if all the adverbs were quantified in mathematical formulae.

<u>Conclusion</u>. All of this indicates that a good strategy will start with a relaxation parameter close to 1. At this point the budding procedure should be used, as it has the effect of increasing μ^* and hence of decreasing θ.

When the sequence approaches P, the relaxation parameter should be gradually increased to or above 2.

4.3 <u>Linear Programming and Decomposition</u>

Let $P = \{x \in R^n | \; Ax + b \geq 0, \; x \geq 0\}$ be a non-empty polyhedron, where A is an $m \times n$ matrix and $b \in R^m$.

The linear programming problem consists of finding the points of P where $\langle c,x \rangle$ is maximum.

Classically, either of the following alternatives can occur:

(i) c does not belong to the barrier cone of P and hence $\langle c,x \rangle$ is unbounded on P.

(ii) c belongs to the relative interior of a unique element $N_P(F)$ of $\Gamma_P/(R^n)$. Then the set of optimal points is the face F of P

$$F = \{x \in P | \; \langle c,x \rangle = \underset{z \in P}{\mathrm{Max}} \langle c,z \rangle = \delta^*(c|P)\} \quad .$$

(4.3.1) Primal-Dual Method (Polyak [53])

In $R^n \times R^m = R^{n+m}$, we define the two polyhedra

$$\tilde{P} = \{(x,k) \in R^n \times R^m \mid x \in P\}$$

$$\tilde{D} = \{(x,k) \in R^n \times R^m \mid k \in D\}$$

with $$D = \{k \in R^m \mid A'k + c \leq 0,\ k \geq 0\} \quad .$$

Classically the dual programs

$$\text{Max } \{\langle c,x\rangle \mid x \in P\}$$

$$\text{Min } \{\langle b,k\rangle \mid k \in D\}$$

have equal values, provided D and P are not empty. Furthermore the solutions to the primal and dual programs are the solutions of the following inequalities in R^{n+m}

(i) $(x,k) \in \tilde{P}$

(ii) $(x,k) \in \tilde{D}$

(iii) $-\langle b,k\rangle + \langle c,x\rangle \geq 0$

The polyhedron of solutions is clearly not full dimensional, as (iii) can be written as an equality.

(4.3.2) Direct Approach Penalty Functions

Let $P = \{x \in R^n \mid \langle a^i,x\rangle + b^i \geq 0,\ \|a^i\| = 1,\ i \in I\}$ be a full dimensional polyhedron.

The relaxation method can be used to find a point of P, say x^*.

Define $P_1 = \{x \in P \mid \langle c,x\rangle \geq \langle c,x^*\rangle + \varepsilon\}$, where $\varepsilon > 0$ is small enough so that $\dim P_1 = n$ (this is possible if x^* is not optimal). The relaxation method can then be used to find a point of P_1 and so on.

Clearly we should choose a^{i^q} so that $\langle c, a^{i^q}\rangle > 0$, if it is possible. This can be clarified by the following theorem, which is similar to a theorem due to Zangwill [67].

Assume that $\|c\| = 1$. Let $\mu(P) = \min_{F \in \mathcal{F}(P)} \mu[C_P(F)]$, as defined in Theorem (2.7.2).

<u>THEOREM</u>. The set of optimal solutions of the linear programming problem $\text{Max}\ \{\langle c,x\rangle \mid x \in P\}$ and the set of optimal solutions of the unconstrained maximization problem $\text{Max}\ \{\langle c,x\rangle - rd_P(x) \mid x \in R^n\}$ are equal for all $c \in B = \{u \in R^n \mid \|u\| = 1\}$ if and only if $r > \frac{1}{\mu(P)}$.

<u>Proof</u>. Clearly $\langle c,x\rangle - rd_P(x)$ is a concave function of x (we assume $r > 0$) which equals $\langle c,x\rangle$ only on P.

(i) Let F be a face of P such that $\mu[C_P(F)] = \mu(P)$ and $x^* \in \text{r.i.}[F]$. By definition (2.6.3)

$$\mu[C_P(F)] = \mu[C_P(x^*)] = \min_{u \in N_P(x^*) \cap B} d_{C_P(x^*)}(u) \quad .$$

Let $h \in N_P(x^*) \cap B$ be a point where the minimum is achieved. Choose $c = h$, then there is some $\varepsilon^* > 0$ such that for all $\varepsilon \in (0,\varepsilon^*]$ we have

$$d_P(x^*+\varepsilon c) = d_{C_P(x^*)}(\varepsilon c) = \varepsilon\mu(P) \quad .$$

Clearly x^* solves $\text{Max}\ \{\langle c,x\rangle \mid x \subset P\}$ and

$$\langle c,x^*+\varepsilon c\rangle - rd_P(x^*+\varepsilon c) = \langle c,x^*\rangle + \varepsilon - r\varepsilon\mu(P) \quad .$$

Hence if $r \le \frac{1}{\mu(P)}$, $x^* + \varepsilon c \notin P$ gives to the unconstrained maximization problem a value greater than or equal to the value of the linear program.

(ii) Let x^* be a solution of Max $\{\langle c,x\rangle \mid x \in P\}$, that is $c \in N_P(x^*)$.

We only need to show that

$$\langle c,x^*+u\rangle - rd_P(x^*+u) < \langle c,x^*\rangle$$

for all u such that $x^* + u \notin P$.

a) If $u \notin C_P(x^*)$, then

$$\begin{aligned}\langle c,x^*+u\rangle - rd_P(x^*+u) - \langle c,x^*\rangle \\ \le \langle c,u\rangle - rd_{C_P(x^*)}(u) \quad .\end{aligned}$$

By Theorem (2.3.6) u has a unique orthogonal decomposition into elements z and v of $N_P(x^*)$ and $C_P(x^*)$ with $\|z\| \neq 0$.

Hence

$$\begin{aligned}\langle c,u\rangle - rd_{C_P(x^*)}(u) &\le \langle c,v\rangle + \langle c,z\rangle - r\mu(P)\|z\| \\ &\le \|z\|(1-r\mu(P)) < 0\end{aligned}$$

if $r > \frac{1}{\mu(P)}$.

b) $u \in C_P(x^*)$, but $u \notin C_{F^*}(x^*)$ where $F^* = \{x \in P \mid \langle c,x\rangle = \langle c,x^*\rangle\}$, the set of optimal points. Then $x^* + \eta u \in P/F^*$ fc all $\eta \in (0,\eta^*]$ and some $\eta^* > 0$.

Clearly

$$\langle c,x^*+\eta u\rangle - rd_P(x^*+\eta u) = \langle c,x^*+\eta u\rangle < \langle c,x^*\rangle$$

for all $\eta \in (0,\eta^*]$.

For $\eta > \eta^*$, the same equation follows from the concavity of $\langle c,x\rangle - rd_P(x)$.

c) $u \in C_{F^*}(x^*)$. Let $\eta^* = \mathrm{Sup}\,\{\eta > 0 \mid x^*+\eta u \in P\}$ and $x^{**} = x^* + \eta^* u$. Then the points $x^* + \eta u$ for $\eta \in [0,\eta^*]$ are optimal for the linear program.

For the points corresponding to $\eta > \eta^*$, we simply compare them to x^{**}, using the proof of a). Q.E.D.

This theorem shows that the linear programming problem can be solved by finding a solution of the Čebyšev problem

$$\underset{x\in\mathbb{R}^n}{\mathrm{Max}}\ \mathrm{Min}\,\{\langle\varepsilon c,x\rangle,\ \underset{i\in I}{\mathrm{Min}}\,\{\langle\varepsilon c+a^i,x\rangle + b^i\}\}$$

where $0 < \varepsilon < \mu(P)$. This can be handled if we solve the linear inequalities

$$\varepsilon\langle c,x\rangle \geq w$$
$$\langle\varepsilon c+a^i,x\rangle + b^i \geq w$$

where w is a parameter.

Various strategies can be devised to choose w, but they should be screened by an extensive experimentation.

(4.3.3) <u>Large Scale Programming Decomposition</u>

This is, as announced in [35] by Held and Karp, and in our introduction, what appears to be the most interesting field of application of the relaxation method. It could be applied to the traveling salesman problem ([34], [35]), the multicommodity flow problem ([63]), the transportation problem ([66]), or in fact to any problem within the reach of the general decomposition principle ([41]).

We will treat briefly the transportation problem. A standard formulation goes as follows: Given n origins with stocks $a_k > 0$ of a single commodity and m destinations with demands $d_j > 0$, find a minimum cost shipping schedule which meets all the constraints.

Let y_{kj} = amount shipped from k to j

b_{kj} = cost of shipping one unit from k to j

The problem is

$$\text{Min} \sum_{k=1}^{n} \sum_{j=1}^{m} b_{kj} y_{kj}$$

subject to

$$\sum_{j=1}^{m} y_{kj} = a_k , \quad k = 1,\dots,n$$

$$\sum_{k=1}^{n} y_{kj} = d_j , \quad j = 1,\dots,m$$

$$y_{kj} \geq 0$$

where we assume that $\sum_{k=1}^{n} a_k = \sum_{j=1}^{m} d_j$.

If m is much greater than n and too large to be handled by network methods, then the decomposition principle can be used.

Let $\{(y_{kj}^{i} |\ k = 1,\dots,n;\ j = 1,\dots,m) |\ i \in I\}$ be a list of the extreme points of

$$\sum_{k=1}^{n} y_{kj} = d_j , \quad j = 1,\dots,m$$

$$y_{kj} \geq 0 \quad .$$

Then the transportation problem is equivalent to the master program:

$$\text{Min} \sum_{i\in I} b^i \mu^i$$

subject to

$$\sum_{i\in I} \mu^i = 1$$

$$-\sum_{i\in I} \mu^i \overline{a^i_k} = 0 \ , \quad k = 1,\ldots,n$$

$$\mu^i \geq 0$$

where

$$b^i = \sum_{k=1}^{n} \sum_{j=1}^{m} b_{kj} y^i_{kj}$$

$$\overline{a^i_k} = a^i_k - \sum_{j=1}^{m} y^i_{kj} \quad .$$

The dual program is

$$\text{Max } w$$

subject to

$$w \leq b^i + \langle a^i, x\rangle$$

where $x \in R^n$ is the vector of dual variables and

$$a^i = (\overline{a^i_k},\ k=1,\ldots,n)' \in R^n \quad .$$

This is a Čebyšev problem, i.e.,

$$\underset{x\in R^n}{\text{Max}} \ \underset{i\in I}{\text{Min}} \ \{\langle a^i,x\rangle + b^i\}$$

which can be solved parametrically using the relaxation method, as in (4.3.2).

We find a point $x \in P_w = \{z \in R^n \mid \langle a^i,z\rangle + b^i \geq w,\ i \in I\}$ for increasing values of w.

The hyperplane H^{i^q} that we will use will be chosen to maximize the residue and is generated by the following subproblem

$$\text{Min} \sum_{j=1}^{m} \sum_{k=1}^{n} (b_{kj}-x_k)y_{kj}$$

subject to $\qquad \sum_{k=1}^{n} y_{kj} = d_j \ , \quad j = 1,\ldots,m$

$$y_{kj} \geq 0$$

where x_k is the k^{th} component of x, the actual point in the relaxation sequence.

The subproblem separates into m subproblems whose solutions are trivially: Let k_j be an index which minimizes $(b_{kj}-x_k)$ over k. Then

$$y_{kj}^{i^q} \begin{cases} = 0 & \text{if } k \neq k_j \\ = d_j & \text{if } k = k_j \end{cases} .$$

This leads to

$$a^{i^q} = (a_k^{i^q} - \sum_{j=1}^{m} y_{kj}^{i^q}, \ k = 1,\ldots,n)'$$

$$b^{i^q} = \sum_{k=1}^{n} \sum_{j=1}^{m} b_{kj} y_{kj}^{i^q} \quad .$$

The simplex method fails to converge fast because the family of halfspaces defining the set $\tilde{P} = \{(w,x) \in R \times R^n | \ w \leq \langle a^i,x\rangle + b^i, \ i \in I\}$ is very large and hence the number of vertices is extremely large. The abundance of faces of $\tilde{P}$ makes such a set easier to handle by techniques designed for smooth sets. A smooth enough set is loosely a set which behaves, as far as the relaxation method is concerned, as well as a smooth set. It is hoped that for some classes of problems, the smooth enough property could be established theoretically. Those problems would be easy to solve using the relaxation method.

It should also be clear that from time to time we could have a vertex check. Assume that at step q we have reached a feasible point x^q of $P_{\bar{w}} \subseteq R^n$. If $w^q = \min_{i \in I} \{\langle a^i, x^q \rangle + b^i\}$, then (w^q, x^q) belongs to the boundary $\partial\tilde{P}$ of $\tilde{P}$, which is a convex hypersurface in R^{n+1}: $w(x) = \min_{i \in I} \{\langle a^i, x \rangle + b^i\}$. Using an ascent algorithm on $\partial\tilde{P}$ (see Cheney [11], pp. 51-54), we can arrive in at most n steps to a vertex x' of $\tilde{P}$.

After checking for optimality at x', a new target value $\bar{w}' > w(x')$ is chosen and the relaxation method will be applied to $P_{\bar{w}'}$.

4.4 Conclusions

The relaxation method as described in this work can be used to solve any problem which can be put in the form of a system of inequalities. The range of applications is thus vast, from the solution of systems of equations to the general mathematical programming problem.

The main result of this work is that the finiteness of the convergence is not only a theoretical curiosity, but also a computationally useful fact. The range of values of the relaxation parameter which guarantees finite convergence is far wider than was previously known.

The field of application where the relaxation method is most promising is in our opinion the class of problems amenable to a treatment by the general decomposition principle. One main reason for this hope is that no method has conquered in a definitive manner this class of problems.

The claim that we made in the introduction is more than just a quip, but the question, how much more, is still open, and can be answered only through experimentation.

To simplex, to relax: This thesis' question
Whether 'tis faster on P to iterate
On the narrowing edge slung between vertices
Or to take the normal against a sea of planes
And by opposing it, to leap to end today.

APPENDIX 1

In the development of this work, two examples played a key role, providing us not only with an invaluable intuitive background, but also with counterexamples to some tempting conjectures.

(A.1.1) Relaxation With Respect to a Polytope P in R

(A.1.1.1) P is Full Dimensional

Clearly, it is enough to study

$$P = \{x \in R \mid x+1 \geq 0,\ -x+1 \geq 0\} = [-1,+1] \quad .$$

If $x^0 > 1$ is the initial point, then $a^{i^0} = -1$, and $x^1 = x^0(1-\lambda) + \lambda$.

If $x^0 < -1$ is the initial point, then $a^{i^0} = 1$, and $x^1 = x^0(1-\lambda) - \lambda$.

We will assume that $x^0 > 1$.

Case 1. $\lambda \in (0,1]$

If $\lambda = 1$ there is convergence in one step to $x^* = 1$.

If $\lambda < 1$, then

$$d(x^q,P) = x^q - 1 = (x^0-1)(1-\lambda)^q$$

and there is infinite convergence to $x^* = 1$.

Case 2. $\lambda = 2$

$$d(x^q,P) = \text{Max}\ \{0, d(x^0,P) - 2q\} \quad .$$

Hence, there is finite convergence after the completion of $\left\lceil \frac{d(x^0,P)}{2} \right\rceil$ iterations, where $\lceil x \rceil$ is the smallest integer greater than or equal to x.

<u>Case 3</u>. $\lambda \in (1,2)$

We have similarly

$$d(x^{q+1},P) = \text{Max}\,\{0,(\lambda-1)d(x^q,P) - 2\}$$

and hence

$$d(x^q,P) = \text{Max}\,\{0,(\lambda-1)^q d(x^o,P) + \frac{2}{2-\lambda}[(\lambda-1)^q-1]\}$$

from which it follows that there is finite convergence after

$$\left\lceil \frac{\log_e[1+\frac{2-\lambda}{2}d(x^o,P)]}{\log_e \frac{1}{\lambda-1}} \right\rceil \quad \text{iterations.}$$

<u>Case 4</u>. $\lambda > 2$

(i) If $d(x^o,P) < \frac{2}{\lambda-2}$, then there is finite convergence and the number of iterations is given by

$$\left\lceil \frac{\log_e[1-\frac{\lambda-2}{2}d(x^o,P)]^{-1}}{\log_e(\lambda-1)} \right\rceil \; .$$

(ii) If $d(x^o,P) = \frac{2}{\lambda-2}$, then $d(x^q,P) = d(x^o,P)$ for all q and the sequence oscillates with $x^{2q} = x^o = \frac{\lambda}{\lambda-2}$ and $x^{2q+1} = -x^o = -\frac{\lambda}{\lambda-2}$.

(iii) If $d(x^o,P) > \frac{2}{\lambda-2}$, then the sequence diverges and is unbounded

$$d(x^q,P) = (\lambda-1)^q[d(x^o,P) - \frac{2}{\lambda-2}] + \frac{2}{\lambda-2} \; .$$

<u>Remark</u>. The results of this paragraph are valid for the relaxation method applied to

(i) A sphere in R^n

(ii) Any convex body, which has two parallel supporting hyperplanes E_1 and E_2 such that some point of $E_1 \cap K$ and some point $E_2 \cap K$ lie on a common normal to E_1 and E_2.

(A.1.1.2) P is Empty

Let $P = \{x \in R \mid x-1 \geq 0, -x-1 \geq 0\}$.

If $x^o > 0$, then $x^1 = x^o(1-\lambda) - \lambda$.

If $x^o < 0$, then $x^1 = x^o(1-\lambda) + \lambda$.

We have $P_{w^*} = \{0\}$ and we assume that $x^o > 0$.

Case 1. $\lambda = 2$

$d(x^q, P_{w^*}) = |x^q| = x^o + 2q$ and the sequence is unbounded.

Case 2. $\lambda > 2$

We have similarly

$$d(x^{q+1}, P_{w^*}) - 1 = [d(x^q, P_{w^*}) - 1 + 2][\lambda-1]$$

and hence

$$d(x^q, P_{w^*}) - 1 = [d(x^o, P_{w^*}) - 1](\lambda-1)^q + 2\frac{\lambda-1}{\lambda-2}[(\lambda-1)^q - 1]$$

$$|x^q| = d(x^q, P_{w^*}) = (\lambda-1)^q[d(x^o, P_{w^*}) + \frac{\lambda}{\lambda-2}] - \frac{\lambda}{\lambda-2} .$$

The sequence is unbounded.

Case 3. $\lambda \in (0,2)$

Define $x^*(\lambda) = \frac{\lambda}{2-\lambda}$, then we have

$$|x^q| - x^*(\lambda) = (\lambda-1)^q[|x^o| - x^*(\lambda)]$$

and hence the sequence has two limit points: $x^*(\lambda)$ and $-x^*(\lambda)$.

(A.1.2) Relaxation With Respect to a Closed Convex Cone in R^2

Let $C_\alpha(e)$, $\alpha \in [0, \frac{\pi}{2}]$ be a closed convex cone in R^2

(clearly it is spherical).

If $\alpha \geq \frac{\pi}{4}$, then $C_\alpha(e)$ is smooth enough and convergence takes at most two steps.

We assume then that $\alpha \in (0,\frac{\pi}{4})$.

(A.1.2.1) Example of Infinite Convergence [51] (See Figure 3)

Let $e = (-1,0)$, then define

$$C_\alpha(e) = \{(x,y) \in R^n \mid \langle(-\sin\alpha,-\cos\alpha),(x,y)\rangle \geq 0$$
$$\langle(-\sin\alpha,\cos\alpha),(x,y)\rangle \geq 0\} .$$

We choose the initial point to be

$$u^o = (\cos(\alpha+\gamma),\sin(\alpha+\gamma))$$

where $0 \leq \gamma < \frac{\pi}{2} - 2\alpha$ so that $u^o \in [C_\alpha(e)]^P$.

Let the relaxation parameter be

$$\lambda = 1 + \frac{tg\gamma}{tg(\gamma+2\alpha)} .$$

It follows, after some reductions, that

$$u^1 = \frac{\cos(2\alpha+\gamma)}{\cos\gamma}(\cos(\alpha+\gamma),-\sin(\alpha+\gamma))$$

$$u^2 = (\frac{\cos(2\alpha+\gamma)}{\cos\gamma})^2(\cos(\alpha+\gamma),\sin(\alpha+\gamma)) .$$

It is then clear that

$$\{u^q\} \subseteq [C_\alpha(e)]^P = C_{\frac{\pi}{2}-\alpha}(-e) \quad \text{for all } q .$$

Furthermore

$$d(u^q,C_\alpha(e)) = \|u^q\| = [\frac{\cos(2\alpha+\gamma)}{\cos\gamma}]^q .$$

The convergence is clearly infinite. For a given cone $C_\alpha(e)$,

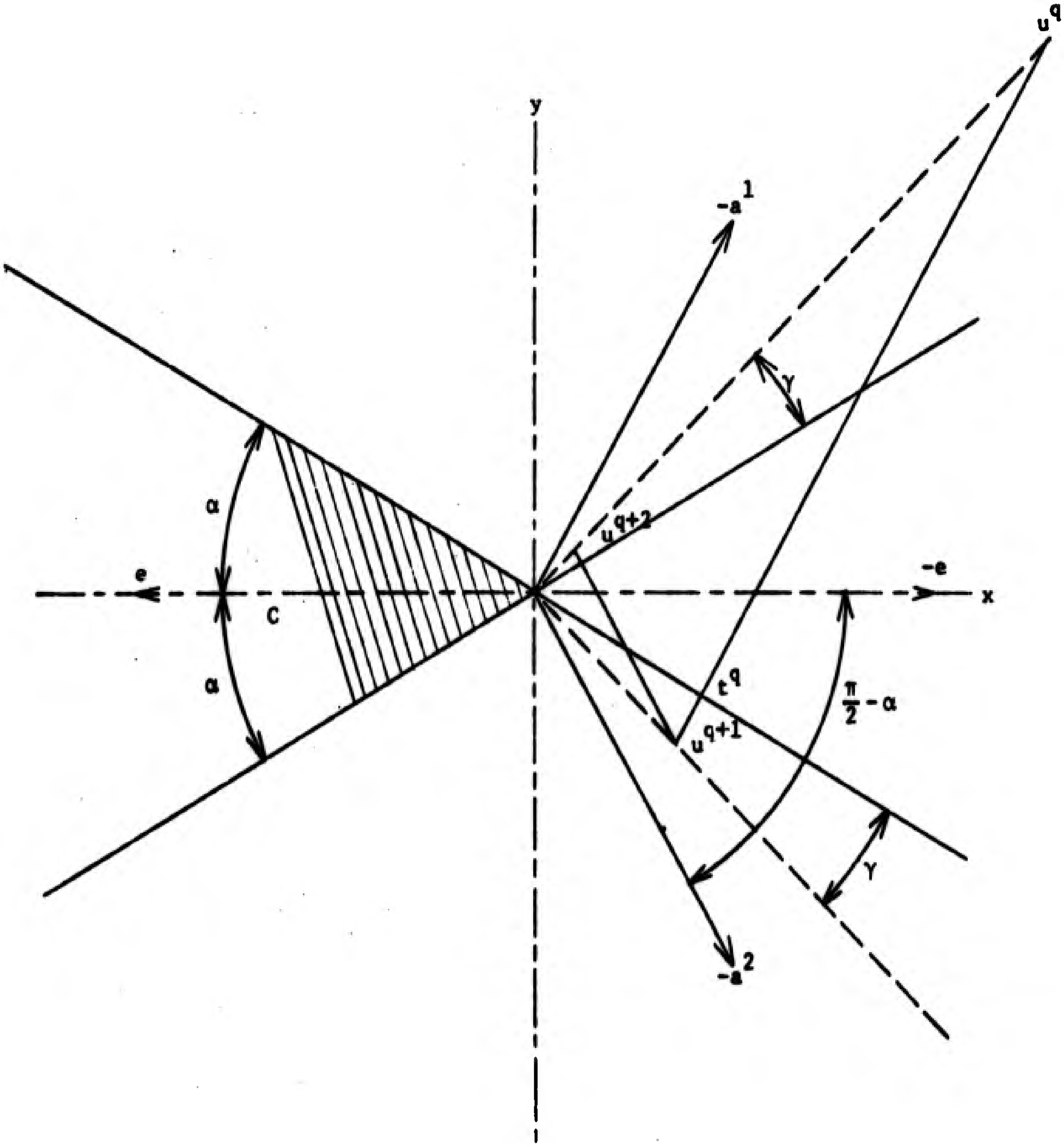


Figure 3

the largest λ for which this example gives infinite convergence is

$$\lambda^* = \max_{\gamma\in[0,\frac{\pi}{2}-2\alpha]} \left[1+\frac{tg\gamma}{tg(\gamma+2\alpha)}\right] .$$

The maximum is attained for $\gamma^* = \frac{\pi}{4}-\alpha$ and $\lambda^* = \frac{2}{1+\sin 2\alpha} = \frac{2}{1+2\nu\sqrt{1-\nu^2}}$, where $\nu = \sin\alpha$.

(A.1.2.2) Study of the Finite Convergence

As in (3.3) let $D^q = \cos\alpha^q = \frac{\langle u^q,-e\rangle}{\|u^q\|}$. Here the recurrence relation on D^q is exact

$$\cos\alpha^q - \cos\alpha^{q+1} = \cos\alpha^q[1-\beta^q] + \lambda\mu^q\beta^q\nu$$

where $\mu^q = \sin(\alpha+\alpha^q)$. It follows that finite convergence occurs if and only if

$$\min_{\alpha^q\in[0,\frac{\pi}{2}-\alpha]} \{\cos\alpha^q(1-\beta^q) + \lambda\mu^q\beta^q\nu\} > 0 .$$

We will study the roots of

$$\beta^q[\cos\alpha^q - \lambda\mu^q\nu] = \cos\alpha^q . \qquad (*)$$

(i) There is one obvious root: $\alpha^q = \frac{\pi}{2}-\alpha$

(ii) Clearly, if $\cos\alpha^q \leq \lambda\mu^q\nu$, the equation (*) has no other roots.

(iii) If $\cos\alpha^q \geq \lambda\mu^q\nu$, then the roots of the equation (*) are the roots of

$$(\beta^q)^2[\cos\alpha^q - \lambda\mu^q\nu]^2 = \cos^2\alpha^q .$$

Letting $z = \cos^2\alpha^q$, $\delta = \nu^2$ and using the fact that

$\alpha^q = \frac{\pi}{2} - \alpha$ solves this equation, we get

$$(z - \frac{1}{2})^2 = \frac{1}{4} - \delta(1-\delta)\frac{\lambda^2}{(2-\lambda)^2} \quad . \qquad (**)$$

Hence, depending on whether

$$\frac{1}{4} \begin{matrix} < \\ = \\ > \end{matrix} \delta(1-\delta)\frac{\lambda^2}{(2-\lambda)^2}$$

the equation (**) has $\begin{matrix} \text{no} \\ \text{one} \\ \text{two} \end{matrix}$ solutions in real numbers.

Hence (**) has no real solutions if and only if $\lambda > \frac{2}{1+\sin 2\alpha}$.

This will imply finite convergence as

(i) For $\alpha^q = 0$

$$\cos \alpha^q(1-\beta^q) + \lambda\mu^q\beta^q\nu = 1 - \beta^q(1-\lambda \sin^2\alpha) ,$$

which is positive for any $\lambda > 0$.

(ii) The root $\alpha^q = \frac{\pi}{2} - \alpha$ can be dismissed, as the recurrence equation on α^q is valid only if $u^{q+1} \in [C_\alpha(e)]^p$, but if $\alpha^q = \frac{\pi}{2} - \alpha$, then $\alpha^{q+1} = \frac{\pi}{2} + \alpha$ and $u^{q+1} \notin [C_\alpha(e)]^p$.

We conclude that $\lambda > \frac{2}{1+\sin 2\alpha}$ implies finite convergence for any initial point.

(A.1.2.3) Study of the Overprojection (Figure 4)

We use the same example, but we choose

$$u^o = (-\cos(\alpha+\delta), \sin(\alpha+\delta))\varepsilon$$

as the initial point, where $0 \leq \delta < \frac{\pi}{2} - 2\alpha$. If we choose

$$\lambda = 1 + \frac{tg(\delta+2\alpha)}{tg\delta} \quad ,$$

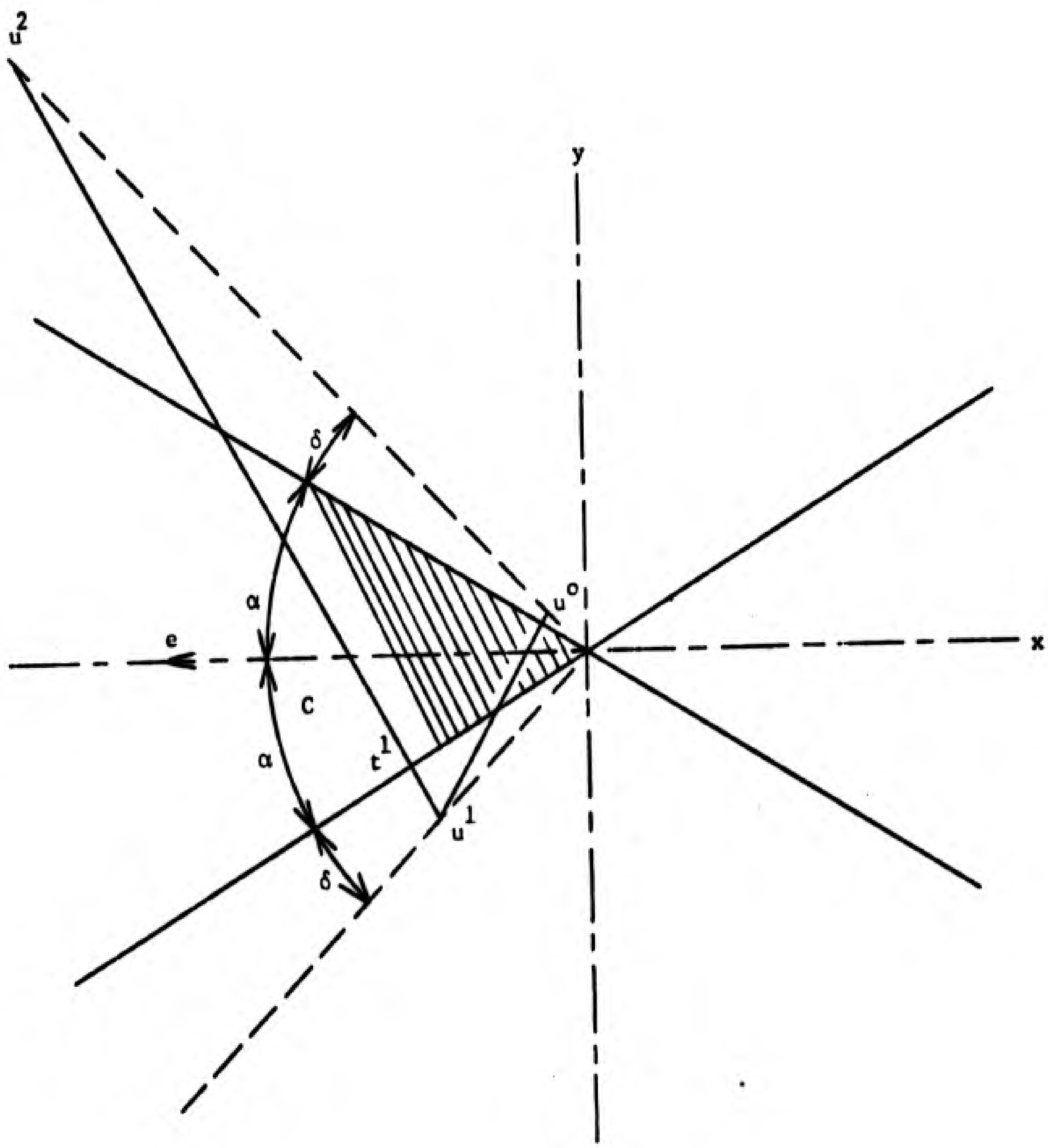


Figure 4

it follows that

$$u^1 = \varepsilon\frac{\cos\delta}{\cos(2\alpha+\delta)}(-\cos(\alpha+\delta),-\sin(\alpha+\delta))$$

$$u^2 = \varepsilon\left(\frac{\cos\delta}{\cos(2\alpha+\delta)}\right)^2(-\cos(\alpha+\delta),\sin(\alpha+\delta)) \quad .$$

The sequence diverges. The minimum such λ will be given as in (A.1.2.1) if

$$\delta^* = \frac{\pi}{4} - \alpha \quad \text{and hence} \quad \lambda^* = \frac{2}{1-\sin 2\alpha} \; .$$

This example shows that the smooth enough assumption is required for Theorem (3.3.6) to hold, as $\varepsilon > 0$ can be chosen arbitrarily small.

Remark. The results of (A.1.2) can be applied to a spherical cone in R^n, or to the direct sum of a subspace L and a spherical cone in $L^{\perp}$.

APPENDIX 2

Example of a Convex Set for Which the Vertex Set is Dense in its Boundary

Let $\{a_n\}_{n \in Z}$ be a summable sequence of real positive numbers, i.e.,

$$\sum_{n=0}^{\infty} a_n < \infty, \quad a_n > 0 \text{ for all } n \in Z \ .$$

Let R_a be the set of rationals of $[0,1]$. Every $r \in R_a$ has a unique representation as an irreducible fraction. R_a can be totally ordered by the lexicographic ordering of the irreducible fractions, according to their denominatorsfirst, and then their numerators.

Let $h(r)$ be the height of r in this ordering. It is a one-to-one function onto Z.

Now, define $\qquad F: [0,1] \to R_+$

by
$$F(x) = \sum_{r \in [0,x]} (x-r) a_{h(r)} \ .$$

We can interpret $a_{h(r)}$ as the increase in slope at r.

Clearly

(i) F is convex and continuous.

(ii) F is differentiable at every irrational point of $[0,1]$ and

$$\frac{dF}{dx} = \sum_{r \in [0,x]} a_{h(r)} = \sum_{r \in [0,x)} a_{h(r)} \ .$$

(iii) F is not differentiable at any rational point and

$$\left.\frac{dF^+}{dx}\right|_{x=r^*} = \sum_{r \in [0,r^*]} a_{h(r)}$$

$$\left.\frac{dF^-}{dx}\right|_{x=r^*} = \sum_{r\in[0,r^*)} a_{h(r)} = \left.\frac{dF^+}{dx}\right|_{x=r^*} - a_{h(r^*)} \quad .$$

Now let

$$K = \{(x,y) \in R^2 |\ F(x_1) \le x_2 \le 2F(1) - F(x_1) \text{ for } x_1 \in [0,1]$$
$$\text{or } F(-x_1) \le x_2 \le 2F(1) - F(-x_1) \text{ for } x_1 \in [-1,0]\} \ .$$

It is a compact convex body and every boundary point with rational first coordinate is a vertex.

BIBLIOGRAPHY

[1] Agmon, Shmuel, "The Relaxation Method for Linear Inequalities," Canadian J. of Math., Vol. 6, pp. 382-392, (1954).

[2] Aleksandrov, P.S., COMBINATORIAL TOPOLOGY, Graylock Press, Rochester, New York, (1956).

[3] Allen, D.N. de G., RELAXATION METHODS IN ENGINEERING AND SCIENCE, McGraw-Hill, New York, New York, (1954).

[4] Bateman, P.T., et al., SEMINAR ON CONVEX SETS, Institute for Advanced Study, Princeton, New Jersey, (1949-1950).

[5] Birkhoff, Garrett, LATTICE THEORY, American Mathematical Society, Providence, Rhode Island, (1940).

[6] Blaschke, Wilhelm, KREIS UND KUGEL, Walter de Gruyter & Co., Berlin, Germany, (1956).

[7] Block, H.D. and S.A. Levin, "On the Boundedness of an Iterative Procedure for Solving a System of Linear Inequalities," Proc. Amer. Math. Soc., Vol. 26, pp. 229-235, (1970).

[8] Bonnesen, T., and W. Fenchel, THEORIE DER KONVEXEN KORPER, Verlag von Julius Springer, Berlin, Germany, (1934).

[9] Bregman, L.M., "The Relaxation Method of Finding the Common Point of Convex Sets and Its Application to the Solution of Problems in Convex Programming," Zh. vychisl. Mat. mat. Fiz. 7, 3, 620-631, (1967) = USSR Comp. Math. and Math. Phys., 3, pp. 200-217, (1967).

[10] Buseman, Herbert, CONVEX SURFACES, Interscience Pub., New York, New York, (1958).

[11] Cheney, E.W., APPROXIMATION THEORY, McGraw-Hill, New York, New York, (1966).

[12] Cheney, Ward and Allen A. Goldstein, "Proximity Maps for Convex Sets," Proc. Amer. Math. Soc., 6, pp. 448-450, (1959).

[13] Daniel, James W., THE APPROXIMATE MINIMIZATION OF FUNCTIONALS, Prentice-Hall, Englewood Cliffs, New Jersey, (1971).

[14] Danzer, Ludwig, Branko Grunbaum, and Victor Klee, "Helly's Theorem and Its Relatives," in [39], pp. 101-180.

[15] Demyanov, Vladimir F. and Aleksandr M. Rubinov, APPROXIMATE METHODS IN OPTIMIZATION PROBLEMS, American Elsevier Pub. Co., Inc., New York, New York, (1970).

[16] Dunford, Nelson and Jacob T. Schwartz, LINEAR OPERATORS: PART I: GENERAL THEORY, (with the assistance of William G. Bade and Robert G. Bartle), Interscience Pub., New York, New York, (1967).

[17] Eggleston, H.G., CONVEXITY, Cambridge University Press, Cambridge, England, (1963).

[18] Eggleston, H.G., THE PROBLEMS IN EUCLIDEAN SPACE: APPLICATION OF CONVEXITY, Pergamon Press, New York, New York, (1957).

[19] Eremin, I.I., "Incompatible Systems of Linear Inequalities," Dokl. Akad. Nauk SSSR 138, 1280-1283, (1961) = Soviet Math. Dokl. pp. 821-824, (1961). MR. 24 #A735.

[20] Eremin, I.I., "An Iterative Method for Čebyšev Approximations of Incompatible Systems of Linear Inequalities," Dokl. Akad. Nauk SSSR 143 1254-1256, (1962) = Soviet Math. Dokl.,3, pp. 570-572, (1962). MR 24 #A3166.

[21] Eremin, I.I., "A Generalization of the Motzkin-Agmon Relaxational Method," USP. Mat. Nauk, 20, 2, pp. 183-187, (1965), in Russian.

[22] Eremin, I.I., "On Systems of Inequalities with Convex Functions in the Left Sides," Amer. Math. Soc. Transl. (2), Vol. 88, pp. 67-83, (1970), translation of Izv. Akad. Nauk SSSR Ser. Mat 30 pp. 265-278, (1966).

[23] Eremin, I.I. "Methods of Fejér's Approximations in Convex Programming," Math. Notes of Acad. Sci. USSR, Vol. 3, No. 2, pp. 139-149, (Feb. 1968).

[24] Faddeev, D.K. and V.N. Faddeva, COMPUTATIONAL METHODS OF LINEAR ALGEBRA, trans. by Robert C. Williams, W.H. Freeman & Co., San Francisco, California, (1963).

[25] Favard, J., "Sur les Corps Convexes," J. de Math. Pures et Appliquées, 12, pp. 219-282, (1933).

[26] Fejér, L., "Ueber die Lage der Nulstellen von Polynomen die aus Minimumforderungen gewisser Arten entspringen," Math. Annalen, 85, pp. 41-48, (1922).

[27] Gale, David, THEORY OF LINEAR ECONOMIC MODELS, McGraw-Hill, New York, New York, (1960).

[28] Germanov, M.A. and V.S. Spiridonov, "On a Method of Solving Systems of Non-linear Inequalities," Zh. vychisl. Mat. mat. Fiz. 6, 2, pp. 335-336, (1966) = USSR Comp. Math. and Math. Phys., 2, pp. 194-196, (1966).

[29] A.J. Goldman and A.W. Tucker, "Polyhedral Convex Cones", pp. 19-40, in [40].

[30] Grunbaum, Branko, CONVEX POLYTOPES, (with the cooperation of Victor Klee, M.A. Perles, and G.C. Shephard), Interscience Pub., New York, New York, (1967).

[31] Grunbaum, Branko, "Subsets of S^n Contained in a Hemisphere", Anais Acad. Brasil. Cienc., Vol. 32, pp. 323-328, (1960).

[32] Grunbaum, Branko, "Polytopes, Graphs, and Complexes," Bul. Amer. Math. Soc., 76, pp. 1131-1201, (1970).

[33] Grunbaum, Branko and G.C. Shephard, "Convex Polytopes," Bul. of the London Math. Soc., No. 3, pp. 257-300, (Nov. 1969).

[34] Gubin, L.G., B.T. Polyak, and E.V. Raik, "The Method of Projections for Finding the Common Point of Convex Sets," Zh. vychisl. Mat. mat. Fiz. 7, 6, pp. 1211-1228, (1967) = USSR Comp. Math. and Math. Phys., 6, pp. 1-24, (1967).

[35] Held, M. and R.M. Karp, "The Traveling Salesman Problem and Minimum Spanning Trees," Operations Research, Vol. 18, pp. 1138-1162, (1970).

[36] Held, M. and R.M. Karp, "The Traveling Salesman Problem and Minimum Spanning Trees--Part II," ORC 70-33, Operations Research Center, Univ. Calif., Berkeley, California, (Oct. 1970).

[37] Kantorovich, L.V. and G.P. Akilov, FUNCTIONAL ANALYSIS IN NORMED SPACES, ed. by Dr. A.P. Robertson, trans. from the Russian by D.E. Brown, Pergamon Press, New York, New York, (1964).

[38] Klee, Victor, "Some Characterizations of Convex Polyhedra," ACTA MATHEMATICA, Almquist Wiksells Boktryckeri, Uppsala, (1959).

[39] Klee, Victor, ed., CONVEXITY, PROCEEDINGS OF SYMPOSIA IN PURE MATHEMATICS, Vol. VII, Amer. Math. Soc., Providence, Rhode Island, (1963).

[40] Kuhn, H.W. and A.W. Tucker, LINEAR INEQUALITIES AND RELATED SYSTEMS, Princeton University Press, Princeton, New Jersey, (1956).

[41] Lasdon, L.S., OPTIMIZATION THEORY FOR LARGE SYSTEMS, The Macmillan Co., New York, New York, (1970).

[42] Ljubich, Yu. I.,and G.D. Maistrovskii, "The General Theory of Relaxation Processes for Convex Functionals," Russian Math. Surveys, Vol. 25, No. 1, pp. 57-117, (Jan.-Feb. 1970).

[43] Ljubich, Yu. I. and G.D. Maistrovskii, "On the Stability of Relaxational Processes," Dokl. Akad. Nauk SSSR, Tom 191, No. 1, (1970) = Soviet Math. Dokl., Vol. II, No. 2, pp. 311-313, (1970).

[44] Luenberger, David G., "The Conjugate Residual Method for Constrained Minimization Problems," SIAM J. Numer. Anal., Vol. 7, No. 3, (Sept. 1970).

[45] Luenberger, David G., OPTIMIZATION BY VECTOR SPACE METHODS, John Wiley & Sons, New York, New York, (1969).

[46] Mangasarian, Olvi L., NONLINEAR PROGRAMMING, McGraw-Hill, New York, New York, (1969).

[47] Merzlyakov, Yu. I., "On a Relaxation Method of Solving Systems of Linear Inequalities," Zh. vychisl. Mat. mat. Fiz., 2, 3, pp. 482-487, (1962) = USSR Comp. Math. and Math. Phys., pp. 504-510, (1964).

[48] Minsky, M. and S. Papert, PERCEPTRONS, Massachusetts Institute of Technology Press, Cambridge, Mass., (1969).

[49] Moreau, Jean-Jacques, "Proximité et Dualité dans un Espace Hilbertien," Bul. Soc. Math. France, 93, pp. 273-299, (1965).

[50] Motzkin, Th., BETRÄGE ZUR THEORIE DER LINEAREN UNGLEICHUNGEN, Buchdruckerei Azriel, Jerusalem, Israel, (1936).

[51] Motzkin, Th. and I.J. Schoenberg, "The Relaxation Method for Linear Inequalities," Canadian J. of Math., Vol. 6, pp. 393-404, (1954).

[52] Niethammer, W., "Über- und Unter-relaxation bei Linearen Gleichungssystemen," Computing, Vol. 5, Fasc 1, pp. 303-311, (Springer-Verlag 1970).

[53] Polyak, B.T., "Gradient Methods for Solving Equations and Inequalities," Zh. vychisl. Math. math. Fiz., 4, pp. 995-1005, (1964) = USSR Comp. Math. and Math. Phys., 4, 6, pp. 17-32, (1964).

[54] Polyak, B.T., "Some Methods of Speeding up the Convergence of Iteration Methods," Zh. vychisl. Math. math. Fiz., 4, 5, pp. 791-803, (1964) = USSR Comp. Math and Math. Phys., 5, pp. 1-17, (1964).

[55] Rockafellar, R. Tyrrell, CONVEX ANALYSIS, Princeton University Press, Princeton, New Jersey, (1970).

[56] Schlechter, Samuel, "Relaxation Methods for Convex Problems," SIAM J. Numer. Anal., Vol. 5, No. 3, pp. 601-612, (1968).

[57] Schaefer, Helmut,"Some Spectral Properties of Positive Linear Operators," Pacific J. Math., Vol. 10, No. 3, pp. 1009-1019, (Fall 1960).

[58] Spiridonov, V.S., "Application of the Gradient Relaxation Method for the Solution of Systems of Non-linear Equations," Zh. vychisl. Mat. mat. Fiz., 8, 4, pp. 872-873, (1968) = USSR Comp. Math and Math. Phys., 4, pp. 231-233, (1968).

[59] Southwell, R.V., RELAXATION METHODS IN ENGINEERING SCIENCE, Oxford University Press, Oxford, England, (1940).

[60] Stiefel, E., "Recent Developments in Relaxation Techniques," Proc. Int. Congress of Math., Vol. 1, pp. 384-391, (1957).

[61] Stoelinga, Th. G.D., CONVEXE PUNTVERZAMELINGEN, H.J. Paris, Amsterdam, Holland, (1932).

[62] Stoer, J. and C. Witzgall, CONVEXITY AND OPTIMIZATION IN FINITE DIMENSIONS I, Springer-Verlag, Berlin-Heidelberg-New York, (1970).

[63] Tomlin, J.S., "Minimum-Cost Multi-Commodity Network Flows," J. ORSA, 14 (1), pp. 45 - 51, (Feb. 1966).

[64] Valentine, Frederick A., CONVEX SETS, McGraw-Hill, New York, New York, (1964).

[65] Weyl, H., "The Elementary Theory of Convex Polyhedra," CONTRIBUTIONS TO THE THEORY OF GAMES, ed. by H.W. Kuhn and A.W. Tucker, Princeton University Press, Princeton, New Jersey, (1950).

[66] Williams, A.C., "A Treatment of Transportation Problems by Decomposition," J. Soc. Ind. Appl. Math., 10, No. 1, pp. 35-48, (1962).

[67] Zangwill, W.I., "Non-linear Programming via Penalty Functions," Management Science, Vol. 13, No. 5, (Jan. 1967).

[68] Zangwill, W.I., "An Algorithm for the Chebyshev Problem--with an Application to Concave Programming," Management Science, Vol. 14, No. 1, (Sept. 1967).

[69] Zangwill, W.I., NONLINEAR PROGRAMMING: A UNIFIED APPROACH, Prentice-Hall, Inc., Englewood Cliffs, New Jersey, (1969).