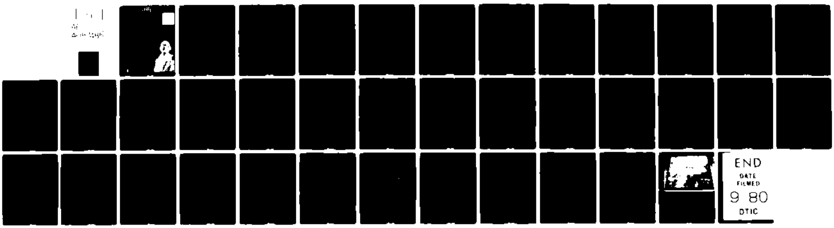
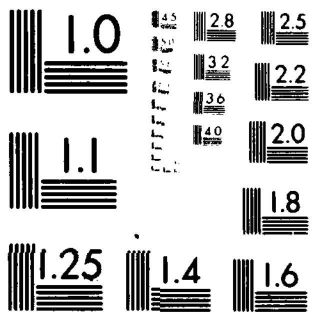
AD-A087 485 GEORGE WASHINGTON UNIV WASHINGTON DC INST FOR MANAGE--ETC F/G 12/1
THE METHOD OF CONTRACTING ELLIPSOIDS,(U)
APR 80 J E FALK DAAG29-79-C-0062
UNCLASSIFIED SERIAL-T-418 ARO-16229.3-M NL





MICROCOPY RESOLUTION TEST CHART
NATIONAL BUREAU OF STANDARDS 1963 A

ARO 16229.3-M

LEVEL II

ADA087485

THE
GEORGE
WASHINGTON
UNIVERSITY

STUDENTS FACULTY STUDY R
ESEARCH DEVELOPMENT FUT
URE CAREER CREATIVITY CC
MMUNITY LEADERSHIP TECH
NOLOGY FRONTI ESIGN
ENGINEERING APP ENC
GEORGE WASHIN NIV

DTIC
ELECTE
AUG 4 1980
C

DDC FILE COPY

SCHOOL OF ENGINEERING
AND APPLIED SCIENCE

THIS DOCUMENT HAS BEEN APPROVED FOR PUBLIC RELEASE AND SALE; ITS DISTRIBUTION IS UNLIMITED

(12)

THE METHOD OF CONTRACTING ELLIPSOIDS

by

James E. Falk

DTIC
ELECTE
AUG 4 1980
S

Serial T-418
18 April 1980

The George Washington University
School of Engineering and Applied Science
Institute for Management Science and Engineering

Contract DAAG 29-79-C-0062
US Army Research Office

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

80 8 1 040

NONE

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE — READ INSTRUCTIONS BEFORE COMPLETING FORM

1. REPORT NUMBER T-418	2. GOVT ACCESSION NO. AD-A087	3. RECIPIENT'S CATALOG NUMBER 485
4. TITLE (and Subtitle) (6) THE METHOD OF CONTRACTING ELLIPSOIDS		5. TYPE OF REPORT & PERIOD COVERED SCIENTIFIC
		6. PERFORMING ORG. REPORT NUMBER T-418
7. AUTHOR(s) (10) JAMES E. FALK	(15)	8. CONTRACT OR GRANT NUMBER(s) DAAG 29-79-C-0062
9. PERFORMING ORGANIZATION NAME AND ADDRESS THE GEORGE WASHINGTON UNIVERSITY INSTITUTE FOR MANAGEMENT SCIENCE & ENGINEERING WASHINGTON, DC 29952		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS (17) 1 [illegible] 9.3-M
11. CONTROLLING OFFICE NAME AND ADDRESS U. S. ARMY RESEARCH OFFICE P.O. BOX 12211 RESEARCH TRIANGLE PARK, NC 27709	(12)	12. REPORT DATE (11) 18 APR 1980
		13. NUMBER OF PAGES 35
14. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office) (14) SERIAL-T-417		15. SECURITY CLASS. (of this report) NONE
		15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

DISTRIBUTION OF THIS REPORT IS UNLIMITED;
APPROVED FOR PUBLIC SALE AND RELEASE.

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

18. SUPPLEMENTARY NOTES

THE VIEW, OPINIONS, AND/OR FINDINGS CONTAINED IN THIS REPORT ARE THOSE OF THE AUTHOR(S) AND SHOULD NOT BE CONSTRUED AS AN OFFICIAL DEPARTMENT OF THE ARMY POSITION, POLICY, OR DECISION, UNLESS SO DESIGNATED BY OTHER DOCUMENTATION.

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

LINEAR PROGRAMMING

THE RUSSIAN ALGORITHM

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

The details of "Khachian's Algorithm," or the "Russian Algorithm," are derived. Previously reported results based on "deep cuts" are verified, and a proof of the polynomial convergence properties is sketched.

DD FORM 1 JAN 73 1473 EDITION OF 1 NOV 65 IS OBSOLETE
S/N 0102-014-6601

NONE 406743

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

THE GEORGE WASHINGTON UNIVERSITY
School of Engineering and Applied Science
Institute for Management Science and Engineering

Abstract
of
Serial T-418
18 April 1980

THE METHOD OF CONTRACTING ELLIPSOIDS

by

James E. Falk

The details of "Khachian's Algorithm," or the "Russian Algorithm," are derived. Previously reported results based on "deep cuts" are verified, and a proof of the polynomial convergence properties is sketched.

Research Supported by

Contract DAAG 29-79-C-0062
US Army Research Office

THE GEORGE WASHINGTON UNIVERSITY
School of Engineering and Applied Science
Institute for Management Science and Engineering

THE METHOD OF CONTRACTING ELLIPSOIDS

by

James E. Falk

1. Introduction

The October 6, 1979 issue of *Science News* [11] featured a cover story entitled "Linear Programming: Solid New Algorithm," which announced the discovery of a "totally new algorithm for linear programming" by the Soviet mathematician L. G. Khachian. *Science* [6] devoted a page and a half of their November 2, 1979 edition to the event, hinting that even the difficult traveling salesman problem might now be rendered tractable.

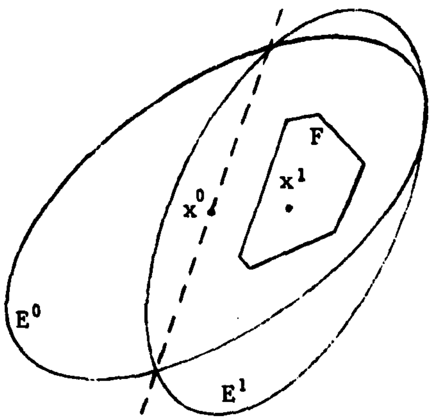
The *New York Times* took over, in articles on November 7, 11, and 27, 1979, amplifying the potential applications of this "surprise discovery by an obscure Soviet mathematician."

The genesis of all this publicity was an article by L. G. Khachian appearing in the January-February 1979 issue of *Doklady* entitled "A Polynomial Algorithm in Linear Programming" [4]. English translations of the Russian version appeared later. Hungarian mathematicians Gâcs and Lovâsz, then at Stanford, distributed a set of their notes on Khachian's paper, which included some variations, comments, and proofs of Khachian's results [2].

The purpose of this paper is to set down a detailed description of the algorithm, including proofs not available in the aforementioned sources. The title, "The Method of Contracting Ellipsoids," is chosen to accurately reflect the nature of the algorithm. It is used in preference to the currently common description, "Khachian's Method," since the algorithm itself has roots in literature predating Khachian's paper (as he himself points out) in articles published by N. Z. Shor [9], [10], D. B. Yudin and A. S. Nemirovskii [13], and A. Yu. Levin [7]. As we will point out, Khachian's (very significant) contribution was to establish the polynomial property of the method when applied to a system of strict linear inequalities.

2. Geometrical Description of the Method

Consider a system of linear inequalities (see Figure 1) describing a set $F \subset R^n$ enclosed in an ellipsoid E^0 with center x^0. It is desired to locate a point $\bar{x} \in F$. If $x^0 \in F$, we are done; otherwise there is some linear inequality $ax \leq \beta$ which is violated by x^0, i.e., $ax^0 > \beta$.



Accession For
NTIS GRA&I
DTIC TAB
Unannounced
Justification
By
Distribution/
Availability Codes
Avail and/or
Dist special
A

Figure 1.--Contracting ellipsoids.

Then the set $F \subset E^1$, where E^1 is an ellipsoid of minimal volume containing the set $E^0 \cap \{x: ax \leq ax^0\}$. The new ellipsoid E^1

has center x^1 , which is then tested for feasibility to F , and the method continues by replacing E^0 and x^0 by E^1 and x^1 .

Thus a sequence $E^0, E^1, \ldots$ is generated. We will show that the volumes $v(E^k)$ decrease geometrically by a factor which is a function of the dimensionality n of the space R^n . Assuming F is not empty, the ellipsoids will tighten up in volume around S , so that eventually an ellipsoid will be produced with center in F .

3. Derivation of the Method

In this section we will derive the updating formulas for the Method of Contracting Ellipsoids. The algorithm is summarized below.

3.1 The Method of Contracting Ellipsoids

We wish to solve (or establish the insolvability of) the system $Ax \leq b$, where A is $m \times n$ and b is $m \times 1$.

Start: Locate $\bar{x} \in R^n$ and a radius $r > 0$ such that the sphere $\{x: ||x - \bar{x}|| \leq r\}$ contains a feasible point (if any). Set $J = rI$.

Step: Test the point $\bar{x}$ for feasibility. If $\bar{x}$ is not feasible, identify a constraint $ax \leq \beta$ such that $a\bar{x} > \beta$.

(i) Set

$$\delta = \frac{a\bar{x} - \beta}{||Ja||} . \tag{1}$$

If $\delta > 1$, there is no solution. If $\delta = 1$, x^+ below is the only solution, provided it is feasible.

(ii)
$$x^+ = \bar{x} - \left(\frac{n\delta+1}{n+1}\right) \frac{J^T Ja}{||Ja||} . \tag{2}$$

$$\text{(iii)} \quad J^+ = n\left(\frac{1-\delta^2}{n^2-1}\right)^{\frac{1}{2}} \left[I - \left(1 - \left(\frac{(n-1)(1-\delta)}{(n+1)(1+\delta)}\right)^{\frac{1}{2}}\right) \frac{Ja(Ja)^T}{||Ja||^2}\right] J \, . \quad (3)$$

These formulas differ from those presented by Shor [10] and Khachian, primarily by the inclusion of the factor δ above. This factor is related to a deeper cut than the aforementioned authors report (they take $\delta = 0$), and leads to faster convergence. The formulas above were presented by Wolfe [12], and will be derived herein.

In the material which follows, B denotes a symmetric, positive definite matrix with eigenvalues $\lambda_1, \ldots, \lambda_n$. Thus,

$$B = P^T D P \, , \quad (4)$$

where P is an orthogonal transformation, and $D = \text{diag}\{\lambda_1, \ldots, \lambda_n\}$. Setting

$$J = D^{\frac{1}{2}} P \, , \quad (5)$$

where $D^{\frac{1}{2}} = \text{diag}\{\lambda_1^{\frac{1}{2}}, \ldots, \lambda_n^{\frac{1}{2}}\}$, we have an equivalent factorization of B as

$$B = J^T J \, . \quad (6)$$

Given B and a vector $\bar{x}$, $E(B,\bar{x})$, or, more simply E , will denote the ellipsoid

$$E = E(B,\bar{x}) = \{x: (x-\bar{x})^T B^{-1}(x-\bar{x}) \leq 1\} \, .$$

<u>Lemma 1</u>: The ellipsoid E is an affine transformation T of the unit sphere $S = \{y: ||y||^2 \leq 1\}$. Specifically,

$$T(S) = E \, ,$$

where $T(y) = \bar{x} + J^T y$.

<u>Proof:</u> $(x-\bar{x})^T B^{-1}(x-\bar{x}) \leq 1$ is equivalent to

$$(x-\bar{x})^T J^{-1}(J^{-1})^T (x-\bar{x}) \leq 1 \, .$$

Setting $y = (J^{-1})^T(x-\bar{x})$ establishes the result.

Figure 2 illustrates this lemma. Note that J is nonsingular from (5). Lemma 1 will be used extensively, since it is easier to work with spheres than ellipsoids.

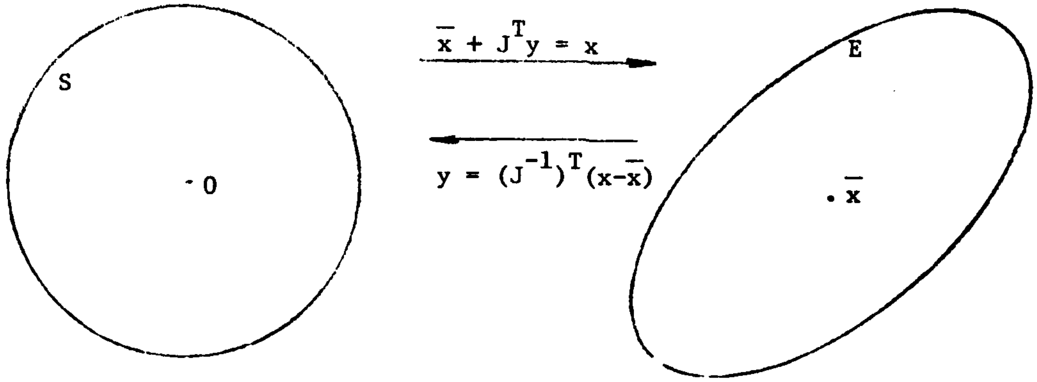


Figure 2.--Ellipsoids are affine transformations of spheres.

Lemma 2: The volume of E is

$$v(E) = \left(\prod_{j=1}^{n} \lambda_j^{\frac{1}{2}}\right) v(s) \;, \tag{7}$$

where $v(s)$ is the volume of the n dimensional sphere.

Proof: From the previous lemma, E is image of the unit sphere S under the transformation $T(y) = \bar{x} + J^T y$. Therefore,

$$\begin{aligned} v(E) &= |\det(J^T)|\; v(s) \\ &= |\det D^{\frac{1}{2}}|\; |\det P|\; v(s) \\ &= \prod_{j=1}^{n} \lambda_j^{\frac{1}{2}}\, v(s) \;, \end{aligned}$$

since $D^{\frac{1}{2}}$ is diagonal and P is orthogonal.

This lemma will prove useful in two places, the first in establishing the formulas for a minimal volume ellipsoid encapsulating a given region, and also in computing the decrease in successive volumes as the algorithm progresses.

Turning attention to the sphere, we define a "truncated sphere" $S(e^1,\delta)$ as

$$S(e^1,\delta) \; = \; \{z: \; ||z|| \leq 1, \quad z_1 \geq \delta\} \; ,$$

where $e^1 = (1,0,\dots,0)^T$ and $0 \leq \delta \leq 1$. Figure 3 illustrates this definition. We wish to encapsulate $S(e^1,\delta)$ in an ellipsoid $\tilde{E}$ of minimal volume.

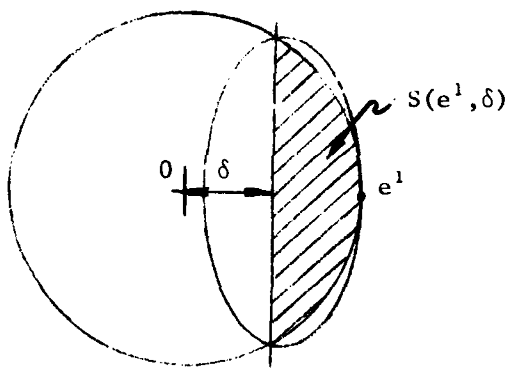


Figure 3.--A truncated sphere.

We will restrict our attention to candidate ellipsoids of the form $\tilde{E}(\tilde{B},\tilde{z}) = \{z: \tau_1(z_1-\xi)^2 + \tau_2 z_2^2 + \dots + \tau_n z_n^2 \leq 1\}$, which contain $S \cap \{z: z_1 = \delta\}$ and e^1 on $\partial\tilde{E}$.

Lemma 3: The ellipsoid $\tilde{E}(\tilde{B},\tilde{z}) = \{z:(\tau_1-\xi)^2 + \tau_2 z_2^2 + \dots + \tau_n z_n^2 < 1\}$ of minimal volume which contains the surface of the sphere $\{z: ||z|| = 1, \; z_1 = \delta\}$ and the point e^1 has

$$\tilde{z} = \left(\frac{n\delta+1}{n+1}, \; 0, \; \dots, \; 0\right)^T$$

and

$$\tilde{B} = \operatorname{diag}\left\{\left(\frac{n(1-\delta)}{n+1}\right)^2, \left(\frac{n^2(1-\delta^2)}{n^2-1}\right), \ldots, \left(\frac{n^2(1-\delta^2)}{n^2-1}\right)\right\} .$$

Proof: We will first characterize the values of $\tau_1,\ldots,\tau_n$ and ξ, which yield a minimal volume ellipsoid containing only the points e^1 and the $n-1$ points $(\delta, \sqrt{1-\delta^2}, 0, \ldots, 0), \ldots, (\delta, 0, \ldots, 0, \sqrt{1-\delta^2})$. We will then show that the resulting $\tilde{E}$ contains all of $S(e^1,\delta)$.

Since $e^1 \in \partial\tilde{E}$, we have

$$\tau_1(1-\xi)^2 = 1 ,$$

while $(\delta, 0, \ldots, 0, \sqrt{1-\delta^2}, 0, \ldots, 0) \in \partial\tilde{E}$ implies

$$\tau_1(\delta-\xi)^2 + \tau_k(1-\delta^2) = 1 , \quad \text{for } k=2,\ldots,n ;$$

i.e.,

$$\tau_1 = \frac{1}{(1-\xi)^2}$$

and

$$\tau_k = \frac{(1-\xi)^2 - (\delta-\xi)^2}{(1-\delta^2)(1-\xi)^2} , \quad k=2,\ldots,n .$$

Thus, using (7), we have

$$v(\tilde{E}) = |(1-\xi)|\left(\sqrt{\left|\frac{(1-\delta^2)(1-\xi)^2}{(1-\delta)^2 - (\delta-\xi)^2}\right|}\right)^{n-1} v(s) .$$

It is easier to work with $v^2(\tilde{E})$, and since this is minimized at the same value of ξ as is $v(\tilde{E})$, we lose no generality. Thus,

$$v^2(\tilde{E}) = (1-\xi)^2\left[\frac{(1-\delta^2)(1-\xi)^2}{(1-\xi)^2 - (\delta-\xi)^2}\right]^{n-1} v^2(s)$$

$$= \frac{(1-\xi)^{2n}(1-\delta^2)^{n-1}}{[(1-\xi)^2 - (\delta-\xi)^2]^{n-1}}\, v^2(s) .$$

This quantity is minimized when its log is minimized, so, taking the derivative of the log of $v^2(\tilde{E})$, and setting it to zero, we obtain

$$\xi = \frac{n\delta+1}{n+1} \tag{8}$$

as the (unique) point which satisfies the resulting equation. A check of the second derivative of the log of $v^2(\tilde{E})$ establishes that the above ξ minimizes.

We thus have

$$\tau_1 = \left(\frac{n+1}{n(1-\delta)}\right)^2$$

and

$$\tau_k = \frac{\left(1 - \frac{1}{n^2}\right)}{(1-\delta^2)}, \quad k=2,\ldots,n \tag{9}$$

Equations (8) and (9) define the minimum volume ellipsoid containing $(1,0,\ldots,0)$ and the $(n-1)$ points of the form $(\delta, 0, \ldots, 0, \sqrt{1-\delta^2}, 0, \ldots, 0)$.

Finally, we must show that *all* of $S(e^1,\delta)$ is contained in $\tilde{E}$. Set

$$\tau = \frac{n^2 - 1}{n^2(1-\delta^2)} \quad (= \tau_2 = \ldots = \tau_n),$$

so that

$$\tilde{B}^{-1} = \text{diag}\{\tau_1, \tau, \ldots, \tau\}.$$

Noting that

$$\tau_1 = \tau + \frac{2(n+1)(\delta n+1)}{n^2(1-\delta)(1-\delta^2)},$$

we have

$$
\begin{aligned}
(z-\bar{z})^T\tilde{B}^{-1}(z-\bar{z}) &= z^T\tilde{B}^{-1}z - 2\bar{z}\tilde{B}^{-1}z + \bar{z}^T\tilde{B}^{-1}z \\
&= \tau||z||^2 + \frac{2(n+1)(\delta n+1)}{n^2(1-\delta)(1-\delta^2)} z_1^2 - 2\tau_1\bar{z}_1 z_1 + \tau_1(\bar{z}_1)^2 \\
&= \left(\frac{n^2-1}{n^2}\right)\left(\frac{1}{1-\delta^2}\right)||z||^2 + \frac{2(n+1)(\delta n+1)}{n^2(1-\delta)(1-\delta^2)} z_1^2 \\
&\quad - 2\left(\frac{(n+1)(\delta n+1)}{n^2(1-\delta)^2}\right)z_1 + \left(\frac{\delta n+1}{n(1-\delta)}\right)^2 .
\end{aligned}
$$

Now, adding and subtracting coefficients of $||z||^2$, z_1^2 , and z_1 , we obtain

$$
\begin{aligned}
(z-\bar{z})^T\tilde{B}^{-1}(z-\bar{z}) &= \left(\frac{n^2-1}{n^2}\right)\left(\frac{1}{1-\delta^2}\right)(||z||^2-1) + \frac{2(n+1)(\delta n+1)}{n^2(1-\delta)(1-\delta^2)} z_1(z_1-1) \\
&\quad + 2\left(\frac{(n+1)(\delta n+1)}{n^2(1-\delta)^2}\right)\left(\frac{1}{1+\delta} - 1\right)(z_1-1) \\
&\quad + 2\left(\frac{(n+1)(\delta n+1)}{n^2(1-\delta)^2}\right)\left(\frac{1}{1+\delta} - 1\right) + \left(\frac{n^2-1}{n^2}\right)\left(\frac{1}{1-\delta^2}\right) \\
&\quad + \left(\frac{\delta n+1}{n(1-\delta)}\right)^2 .
\end{aligned}
$$

But the first three terms are less than or equal to zero, while the last three terms sum to

$$
\frac{2(n+1)(\delta n+1)(-\delta) + (n^2-1)(1-\delta) + (\delta n+1)^2(1+\delta)}{n^2(1-\delta)^2(1+\delta)} = 1 ,
$$

so that

$$
(z-\bar{z})^T\tilde{B}^{-1}(z-\bar{z}) \le 1 ,
$$

i.e.,

$$
S(e^1,\delta) \subset \tilde{E} ,
$$

and the proof is complete.

Having now worked out the formulas for enclosing a truncated sphere, the next step is to translate our results to apply to a truncated ellipsoid.

Suppose we have an ellipsoid $E = E(B,\bar{x})$, and suppose $a\cdot\bar{x} > \beta$. We shall assume E contains points in E satisfying $a\cdot x \leq \beta$. As e^1 was the point of tangency of the hyperplane $\{z: z_1 = 1\}$ to the sphere in z-space, we wish to locate the point $\bar{t}$ which is the point of tangency of the hyperplane $\{x: a\cdot x = a\cdot\bar{t}\}$ and the ellipsoid E (see Figure 4).

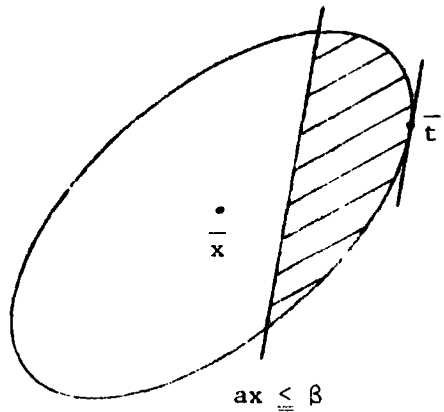


Figure 4.--Definition of $\bar{t}$.

Lemma 4: Given an ellipsoid $E(B,\bar{x})$ and a half space $\{x: a\cdot x \leq \beta\}$, the point $\bar{t}$ which lies in this half space and is tangent to ∂E and the hyperplane $\{x: a\cdot x = a\cdot\bar{t}\}$ is

$$\bar{t} = \bar{x} - \left(\frac{1}{\|Ja\|}\right) J^T Ja \,. \tag{10}$$

Proof: $\bar{t}$ solves the problem

$$\text{minimize} \quad a\cdot x$$
$$\text{subject to} \quad (x-\bar{x})^T B^{-1}(x-\bar{x}) \geq 1 \,.$$

This is a convex program with feasible region having a nonempty interior, so the first order optimality conditions are necessary and sufficient. Thus there is a $u > 0$ such that

$$a + u\left(2B^{-1}(\bar{t}-\bar{x})\right) = 0 \ . \tag{11}$$

Thus

$$(\bar{t}-\bar{x})^T a + 2u(\bar{t}-\bar{x})^T B^{-1}(\bar{t}-\bar{x}) = 0$$

and, since $\bar{t} \in \partial E$,

$$(\bar{t}-\bar{x})^T a + 2u = 0 \ .$$

Also, from (11),

$$Ba + 2u(\bar{t}-\bar{x}) = 0 \ ,$$

so that

$$a^T Ba + 2u(-2u) = 0 \ ,$$

i.e.,

$$u = \tfrac{1}{2} \sqrt{a^T Ba} \ ,$$

or, since $B = J^T J$,

$$u = \tfrac{1}{2} ||Ja|| \ .$$

Thus, from (11) again,

$$\bar{t} = \bar{x} - \left(\frac{1}{||Ja||}\right) J^T Ja \ .$$

We now have the ingredients to derive the equations (1), (2), and (3) defining the Method of Contracting Ellipsoids. We need only take care that the transformations introduced to transform ellipsoids into spheres and back again are properly applied.

Lemma 5: Let $E(B,\bar{x})$, a , β , and $\bar{t}$ be as given in Lemma 4. Then the transformations

$$y = (J^T)^{-1}(x-\bar{x}) \tag{12}$$

and

$$z = Qy \ , \tag{13}$$

where $Q^T Q = I$, $(Q)_1 = -\left(\frac{1}{||Ja||}\right)(Ja)^T$, take E onto the unit sphere in z-space, with $\bar{t} \to e^1$, and $\{x: a \cdot x < \beta\} \to \{z: z_1 > \delta\}$, where $\delta = \frac{a \cdot x - \beta}{||Ja||}$. (See Figure 5.)

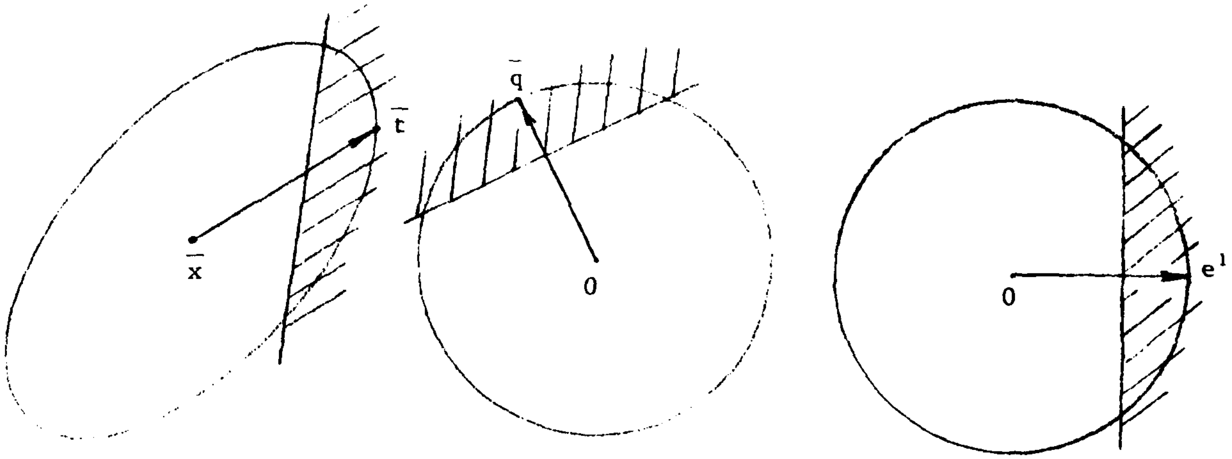


Figure 5.--Transforming $\bar{t}$ into e^1 .

<u>Proof</u>: The first transformation was derived in Lemma 1. The image of $\bar{t}$ under this transformation is

$$\bar{q} = (J^T)^{-1}(\bar{t}-\bar{x})$$

$$= (J^T)^{-1} \left(- \frac{1}{||Ja||}\right) J^T Ja$$

$$= -\left(\frac{1}{||Ja||}\right) Ja \ .$$

We now wish to construct Q so that

$$Q\bar{q} = e^1 \ .$$

Since Q will be orthogonal, and since $||\bar{q}|| = 1$, we may set its first row

$$(Q)_1 = \bar{q}^T = - \frac{1}{||Ja||} (Ja)^T \ ,$$

and the other rows $(Q)_j$ $(j \neq 1)$ are arbitrary, except that they must be mutually orthogonal, have norm 1, and be orthogonal to $(Q)_1$.

With these two transformations now defined, we wish to investigate the image of the hyperplane $a \cdot x = \beta$ in z-space. Now

$$a \cdot x = \beta$$

if and only if

$$a^T(J^T y + \bar{x}) = \beta$$

from (12) above. Then, from (13), the above is true if and only if

$$a^T J^T Q^T z = \beta - a \cdot \bar{x} . \qquad (14)$$

Since the first column of Q^T is $\bar{q} = \left(-\frac{1}{||Ja||}\right) Ja$, the coefficient of z_1 above is

$$a^T J^T \left(-\frac{1}{||Ja||}\right) Ja = -||Ja|| .$$

Also, since columns 2 through n of Q^T are orthogonal to (Ja) , the coefficients of z_2 through z_n are zero. It follows that the hyperplane $a \cdot x = \beta$ is transformed into the hyperplane

$$-||Ja||z_1 = \beta - a \cdot \bar{x}$$

i.e.,

$$z_1 = \frac{a \cdot \bar{x} - \beta}{||Ja||} .$$

Finally, since the image of $\bar{t}$ is e^1 , and since $a \cdot \bar{t} \leq \beta$, it must be that $1 \geq \frac{a \cdot \bar{x} - \beta}{||Ja||}$, i.e.,

$$\{x: a \cdot x \leq \beta\} \text{ is transformed into } \{z: z_1 \geq \delta\} ,$$

where $\delta = \frac{a \cdot \bar{x} - \beta}{||Ja||}$.

We are now in a position to derive the formulas for the Method of Contracting Ellipsoids. To that end, we suppose that a feasible point of the set $F = \{x: Ax \leq b\}$ is contained in an ellipsoid $E = E(B,\bar{x})$. The center $\bar{x}$ is assumed to be infeasible; i.e., we assume that at least one of the constraints $a \cdot x \le \beta$ of the system $Ax \le b$ is violated by $\bar{x}$. We then surround the set $E \cap \{x: a \cdot x < \beta\}$ in an ellipsoid $E^+ = E^+(B^+, x^+)$ of minimal volume. As E^+ must contain a point feasible to F, and, as we shall show, E^+ has a smaller volume than E; it is more likely that x^+ is feasible to F.

All that remains is to derive the proper formulas for B^+ and x^+.

<u>Theorem 1</u>: Assume $F \subset E(B,\bar{x})$. Then $F \subset E^+(B^+,x^+)$, where x^+ and B^+ are given by

$$\text{(i)} \quad \delta = \frac{a \cdot \bar{x} - \beta}{||Ja||},$$

$$\text{(ii)} \quad x^+ = \bar{x} - \left(\frac{n\delta+1}{n+1}\right) \frac{J^T Ja}{||Ja||},$$

$$\text{(iii)} \quad J^+ = n\left(\frac{1-\delta^2}{n^2-1}\right)^{\frac{1}{2}} \left[I - \left(1 - \left(\frac{(n-1)(1-\delta)}{(n+1)(1+\delta)}\right)^{\frac{1}{2}}\right) \frac{Ja(Ja)^T}{||Ja||^2}\right] J,$$

where $a \cdot x \le \beta$ is violated by $\bar{x}$ and $B^+ = (J^+)^T J^+$.

<u>Proof</u>: Using the transformations

$$y = (J^T)^{-1}(x-\bar{x})$$

$$z = Qy,$$

where $Q^T Q = I$ and $(Q)_1 = \left(-\frac{1}{||Ja||}\right)(Ja)^T$, we map the set $E \cap \{x: a \cdot x \le \beta\}$ into the set $\{z: z_1 \ge \delta\}$ (Lemma 5). Clearly this set contains the image of a point feasible to F.

In z-space, the ellipsoid $\tilde{E}(\tilde{B},\tilde{z})$, where

$$\tilde{B} = \operatorname{diag}\left\{\left(\frac{n(1-\delta)}{n+1}\right)^2, \left(\frac{n^2(1-\delta^2)}{n^2-1}\right), \ldots, \left(\frac{n^2(1-\delta^2)}{n^2-1}\right)\right\}$$

and

$$\tilde{z} = \left(\frac{n\delta+1}{n+1}, 0, \ldots, 0\right)^T$$

contains $\{z: z_1 > \delta\}$, and is of minimal volume (Lemma 3). All that remains is to transform $\breve{E}$ back into x-space, using the transformation of Lemma 1, and the Q of Lemma 5.

Working on x^+ first, we have

$$\begin{aligned} y^+ &= Q^T\tilde{z} \\ &= \left(\frac{n\delta+1}{n+1}\right)\left(-\frac{1}{||Ja||}\right) Ja \end{aligned}$$

from Lemma 5. Then

$$\begin{aligned} x^+ &= \bar{x} + J^T y^+ \\ &= \bar{x} - \left(\frac{n\delta+1}{n+1}\right)\frac{J^T Ja}{||Ja||} , \end{aligned}$$

which is formula (ii).

Now working on B^+ , we have

$$(z-\bar{z})^T(\tilde{B})^{-1}(z-\bar{z}) \leq 1$$

if and only if

$$(y-\bar{y})^T Q^T(\tilde{B})^{-1}Q(y-\bar{y}) \leq 1$$

if and only if

$$(x-\bar{x})^T J^{-1}Q^T(\tilde{B})^{-1}Q(J^{-1})^T(x-\bar{x}) \leq 1 ;$$

so

$$\begin{aligned} B^+ &= \left[J^{-1}Q^T(\tilde{B})^{-1}Q(J^{-1})^T\right]^{-1} \\ &= J^T Q^T \operatorname{diag}\left\{\left(\frac{n+1}{n(1-\delta)}\right)^2, \frac{n^2(1-\delta^2)}{n^2-1}, \ldots, \frac{n^2(1-\delta^2)}{n^2-1}\right\} QJ . \end{aligned}$$

Now

$$\operatorname{diag}\left\{\left(\frac{n+1}{n(1-\delta)}\right)^2, \frac{n^2(1-\delta^2)}{n^2-1}, \ldots, \frac{n^2(1-\delta^2)}{n^2-1}\right\} =$$

$$\frac{n^2(1-\delta^2)}{n^2-1}\left[I - \operatorname{diag}\left\{1 - \left(\frac{(1-b)(n-1)}{(1+b)(n+1)}\right), 0, \ldots, 0\right\}\right],$$

so that

$$B^+ = \left(\frac{n^2(1-\delta^2)}{(n^2-1)}\right) J^T Q^T \left[I - \operatorname{diag}\left\{1 - \left(\frac{(1-b)(n-1)}{(1+b)(n+1)}\right), 0, \ldots, 0\right\}\right] QJ .$$

But since

$$J^T Q^T \operatorname{diag}\{t,0,\ldots,0\} QJ = J^T (Q_1)^T (t) Q_1 J ,$$

where Q_1 is the first row of Q, we obtain

$$J^T Q^T \operatorname{diag}\{t,0,\ldots,0\} QJ = J^T \left(-\frac{1}{||Ja||}\right) Ja(t) \left(-\frac{1}{||Ja||}\right) a^T J^T J$$

$$= \left(\frac{t}{||Ja||^2}\right)(Ba)(Ba)^T ;$$

i.e.,

$$B^+ = \left(\frac{n^2(1-\delta^2)}{n^2-1}\right)\left[B - \left(1 - \left(\frac{(1-b)(n-1)}{(1+b)(n+1)}\right)\right) \frac{(Ba)(Ba)^T}{||Ja||^2}\right].$$

It is easily checked that the J^+ of formula (iii) gives

$$(J^+)^T J^+ = B^+ ,$$

as above, so we are done.

4. Properties of the Method

Having derived the method, we now turn to some of its properties. One of the most useful observations is that of the decrease in the volume of the encapsulating ellipsoids.

Theorem 2: At each iteration of the method, the volume of the encapsulating ellipsoid decreases by the factor

$$(1-\delta)\left(1-\delta^2\right)^{\frac{n-1}{2}}\left(\frac{n}{n+1}\right)\left(\frac{n^2}{n^2-1}\right)^{\frac{n-1}{2}} .$$

Proof: Returning to Lemma 5, we note that

$$\tilde{E} = Z(E^+) \quad \text{and} \quad S = Z(E) ,$$

where Z is the composite transformation $z = Q(J^T)^{-1}(x-\bar{x})$, and S is the unit sphere. Since affine transformations do not change the proportion of the volume, we have

$$\frac{v(E^+)}{v(E)} = \frac{v(\tilde{E})}{v(S)} .$$

But

$$v(\tilde{E}) = \frac{n(1-\delta)}{(n-1)}\left[\frac{n^2(1-\delta^2)}{n^2-1}\right]^{\frac{n-1}{2}} v(S) ,$$

so

$$\frac{v(E^+)}{v(E)} = (1-\delta)\left(1-\delta^2\right)^{\frac{n-1}{2}}\left(\frac{n}{n+1}\right)\left(\frac{n^2}{n^2-1}\right)^{\frac{n-1}{2}} . \tag{15}$$

Corollary: $v(E^+) < v(E) \cdot \left(e^{-\frac{1}{2(n+1)}}\right)(1-\delta)\left(1-\delta^2\right)^{\frac{n-1}{2}}$.

Proof: Since e^t is a convex function, $1+t \leq e^t$. In particular,

$$\frac{n}{n+1} = 1 - \frac{1}{n+1} < e^{-\frac{1}{n+1}}$$

and

$$\frac{n^2}{n^2-1} = 1 + \frac{1}{n^2-1} < e^{\frac{1}{n^2-1}} .$$

It should be noted here that the factor

$$\left(\frac{n}{n+1}\right)\left(\frac{n^2}{n^2-1}\right)^{\frac{n-1}{2}}$$

is close to 1 for large n , producing relatively little decrease in successive volumes. The factor $(1-\delta)\left(1-\delta^2\right)^{n-1/2}$, however, is dominated by

$$e^{-\delta-\delta^2\left(\frac{n-1}{2}\right)} ,$$

which, for $\delta > 0$, is *small* for large n , resulting in significant volume decrease. Recall δ is a measure of the infeasibility of the current estimate. If δ is determined by the formula

$$\delta = \max_i \left\{\frac{A_i\bar{x} - b_i}{||JA_i||}\right\} ,$$

the largest decrease in volume per iteration is obtained. Note $\delta = 0$ implies that the current point is feasible, whilst $\delta = 1$ implies that the next point will be feasible. The next table illustrates the joint effects of δ and n on expression (15).

5. Polynomial Boundedness of the Method

The most startling property of this method, and Khatchian's main contribution, is that the amount of work required to execute the technique is bounded by a polynomial function of the problem data. By contrast, the popular Simplex Method is *not* polynomially bounded, as Klee and Minty [5] pointed out in 1972.

TABLE I

RATIOS OF SUCCESSIVE VOLUMES FOR VARIOUS SETTINGS OF δ AND n

δ \ n	2	4	8	16	32	64	128	256	512	1024	2048
0	0.770	0.881	0.939	0.969	0.984	0.992	0.996	0.998	0.999	1.000	1.000
.1	0.690	0.781	0.816	0.809	0.758	0.651	0.474	0.249	0.069	0.005	0.000
.2	0.603	0.663	0.651	0.571	0.418	0.219	0.060	0.004	0.000	0.000	
.3	0.514	0.536	0.473	0.334	0.160	0.036	0.002	0.000			
.4	0.423	0.407	0.306	0.157	0.040	0.002	0.000				
.5	0.333	0.286	0.172	0.056	0.006	0.000					
.6	0.246	0.180	0.079	0.014	0.000						
.7	0.165	0.096	0.027	0.002							
.8	0.092	0.038	0.005	0.000							
.9	0.034	0.007	0.000								
1.0	0.000	0.000									

5.1 The Klee-Minty example

With the usual pivot selection rule, the problem

$$\text{maximize} \quad \sum_{j=1}^{n} 10^{(n-j)} x_j$$

$$\text{subject to} \quad x_i + 2 \sum_{j<i} 10^{(i-j)} x_j \leq 10^{2i-2} \;, \qquad i=1,\ldots,n$$

$$x_j \geq 0$$

takes 2^{n-1} pivot steps to execute.

Figure 6 (not to scale) illustrates the sequence of vertices visited by the Simplex Method.

This example is, of course, only valid under the assumption that the usual pivot selection rule is employed. Jeroslow [3], however, has shown that similar examples can be constructed to illustrate that an exponential amount of work may be required under *any* entering variable selection rule.

It should be emphasized that the aforementioned examples only show that the *worst* case could occur. Over 40 years of experience with the Simplex Method have shown that such worst cases seldom, if ever, occur in practice, and, indeed, the method has proven to be remarkably efficient for a tremendously wide variety of problems.

The proof of the polynomial boundedness property of the Method of Contracting Ellipsoids was first established by Khachian [4] and later refined by Gács and Lovasz [2]. Since their papers contain a very detailed proof of the polynomial boundedness property, we shall not repeat it here, but rather sketch the reasoning behind it.

Suppose we wish to solve (or determine infeasibility thereof) the system of strict inequalities

$$Ax < b \;, \tag{16}$$

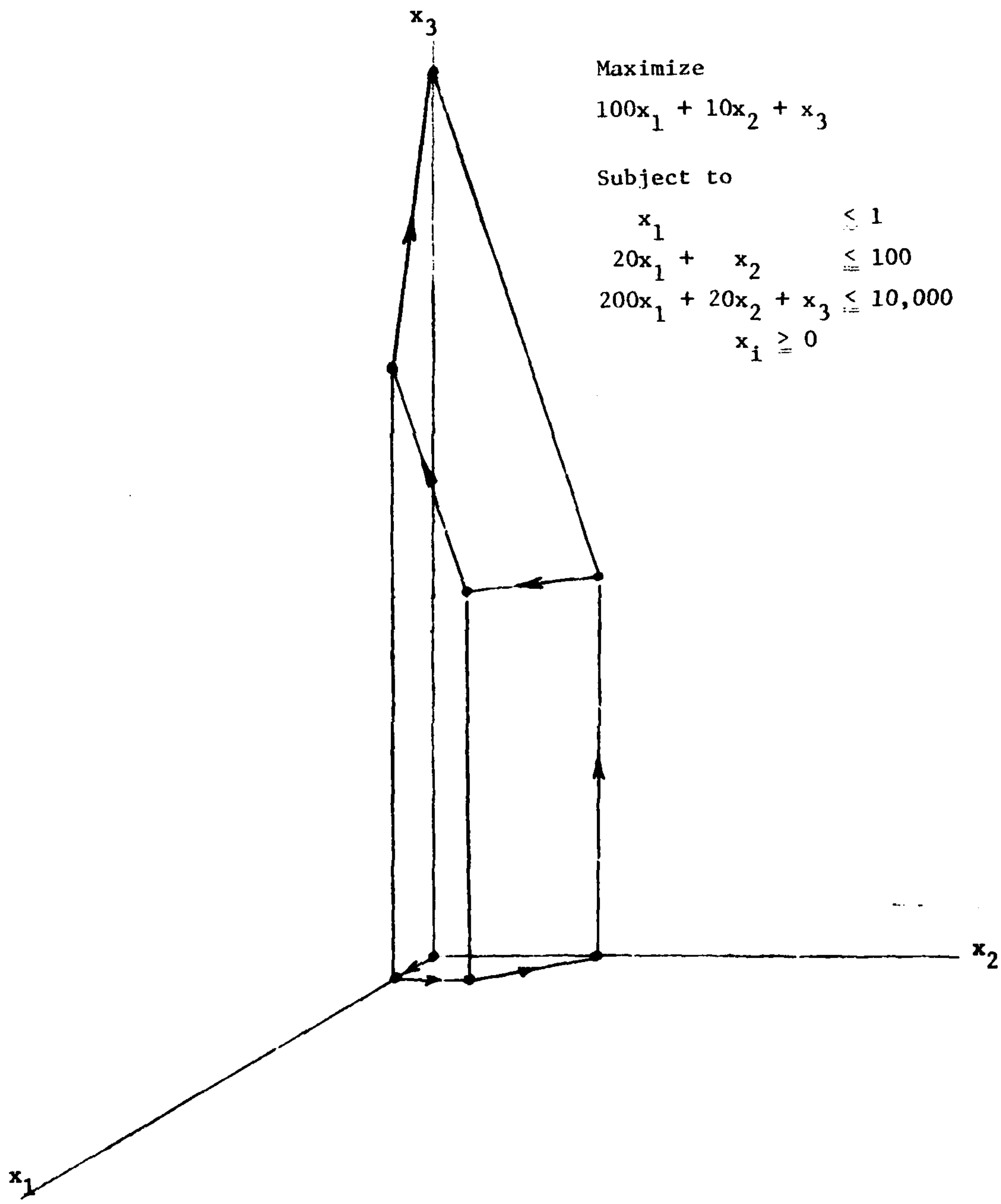


Figure 6.--The Simplex Method may visit every vertex.

where *the entries of* A *and* b *are integers.* Define the constant

$$L = \left[\sum_{i=1}^{m} \sum_{j=1}^{n} \log_2(|a_{ij}| + 1) + \sum_{i=1}^{m} \log_2(|b_i| + 1) + \log_2 mn \right] + 1 .$$

The following lemma is used in the proofs of the theorems which follow.

Lemma 6: Let B be a square, nonsingular submatrix of the matrix $[A|b]$. Then

$$|\det B| < 2^L/n .$$

The proof of this lemma is a result of the fact that $|\det B|$ is bounded above by the product of the norms of the rows of B . This fact follows by reducing BB^T to its diagonal form, and applying the arithmetic-geometric inequality to $B_i B_i^T$, where B_i is the *ith* row of B .) The product of the norms of the rows of B is easily shown to be bounded by $2^L/n$.

This lemma is used in the proof of the following pair of theorems. The first theorem establishes an upper bound on norms of any solutions of (16), while the second establishes a lower bound on the volume of the set of all solutions.

Theorem 3: Any vertex v of the system

$$Ax \leq b$$

satisfies

$$|v_j| < 2^L/n .$$

The proof of this theorem simply looks at v_j as expressed by Cramer's rule, as the quotient of two determinants. The previous lemma yields an upper bound on the numerator of this quotient, while the denominator is at least one in absolute value, since the elements of the determinant are integers.

Theorem 4: If the system

$$Ax < b \tag{16}$$

has a solution, then the volume of its interior inside the ball $\{x: ||x|| \leq 2^L\}$ is at least 2^{-nL}.

The proof of this theorem begins by establishing the existence of a nondegenerate simplex generated by vertices $v^0,\ldots,v^n$ of the system

$$Ax \leq b .$$

Thus, the volume of the solution set described by (16) is bounded below by the volume of the simplex, which is the absolute value of

$$\frac{1}{n!} \det \begin{pmatrix} 1 & \ldots & 1 \\ v^0 & \ldots & v^n \end{pmatrix} .$$

But this is shown to be bounded below by 2^{-nL}.

The following theorem establishes the polynomial boundedness of the method for the system (16).

Theorem 5: If (16) has a solution, then the Method of Contracting Ellipsoids will locate a solution within $[2n(n+1)(2L+1)]$ steps, if one starts with an ellipsoid $E^0 = \{x: ||x|| \leq 2^L\}$.

By Theorem 4, if (16) has a solution, then the set of all solutions has volume at least 2^{-nL}. By Theorem 5, an initial ball of radius 2^L will contain the solution of (16), if any. With δ set to 0 (the most conservative of settings) the corollary to Theorem 2 yields a per step decrease of an amount $e^{-[2(n+1)/1]}$. Therefore, after k steps, the volume of the enclosing ellipsoid is

$$v(E^k) = e^{-\frac{k}{2(n+1)}} v(E^0) .$$

We wish to determine a value of k such that the volume of the enclosing ellipsoid is *less* than the volume of the solution space.

Noting that the volume of the initial sphere is less than $(2\times 2^L)^n$, the volume of the smallest enclosing box whose sides are of length 2×2^L , it suffices to determine an integer k such that

$$e^{-\frac{k}{2(n+1)}} \cdot 2^{n(L+1)} < 2^{-nL} ;$$

i.e.,

$$-\frac{k}{2(n+1)} + n(L+1)\,\ell n2 < -nL\,\ell n2 .$$

A value of k which satisfies this is

$$\bar{k} = [2n(n+1)(2L+1)] .$$

Thus, if no feasible point is found after $\bar{k}$ iterations, there cannot be a feasible point.

Note that the expression for $\bar{k}$ is a polynomial in n and L . Hence the algorithm is polynomially bounded. Note also that the expression for $\bar{k}$ is not particularly tight, as a much tighter, but more complicated bound would result if the exact expression for the contraction of volumes were used, as well as the exact expression for the volume of the initial sphere. To obtain this bound, we would seek the smallest k such that

$$\left[\left(\frac{n}{n+1}\right)\left(\frac{n^2}{n^2-1}\right)^{\frac{n-1}{2}}\right]^k \left(\frac{\pi^{\frac{n}{2}}\, 2^{nL}}{\Gamma\left(\frac{n}{2}+1\right)}\right) < \frac{1}{n!}\, 2^{-nL} \cdot n^n ,$$

where the right-hand side is a better bound on the volume of the solution space. While this is a much tighter bound than derived before, the polynomial boundedness property is not apparent.

5.2 Is polynomial boundedness desirable?

An algorithm which is known to be polynomially bounded may or may not be superior to one which is not. Polynomial boundedness is a desirable characteristic of an algorithm, but may not accurately depict performance over the long run.

Dantzig [1] compares the estimated running time of the ellipsoidal method to the actual running time of a sophisticated implementation of the Simplex Method on a particular case study. He concludes that the Simplex Method wins by a factor of 30 minutes to 50,000,000 years! This is, however, computed using the loose upper bound and consistently setting $\delta = 0$.

McCall [8] studied the ellipsoidal method in more detail. Among other negative results, he concluded that "Khachian's algorithm is of no practical value for real-world LP problems."

A more useful statistic for measuring performance might be the *expected* number of steps required to find a feasible solution to a system of linear inequalities. Years of observing the Simplex Method at work suggest that this number is usually less than $4m$, where m is the number of constraints [1].

In spite of criticism, the practical import of this method has yet to be evaluated. Considerable work is on-going. As of February 6, 1980, Wolfe [12] had compiled a list of 23 research references relating to the new algorithm, and this list is surely out-dated by now (March 25, 1980).

6. How to Solve Linear Programs

Gács and Lovász [2] suggest the following scheme to set a linear program in the required setting. The problem

$$\left.\begin{array}{rl} \text{minimize} & cx \\ \text{subject to} & Ax > b \\ & x > 0 \end{array}\right\} P$$

and its dual

$$\begin{array}{rl} \text{maximize} & ub \\ \text{subject to} & uA < c \\ & u > 0 \end{array}$$

have solutions if and only if the system

$$\begin{aligned} Ax &\geq b \\ uA &\leq c \\ x &\geq 0 \\ u &\geq 0 \\ ub &\geq cx \end{aligned}$$

has a solution. The latter is in the required form.

A potential difficulty in solving the latter system is that it generally will have a unique solution. Small perturbations will yield approximate solution sets, but with small volumes.

Another variant, employing fewer variables and fewer constraints than the above scheme is first to locate a point x^0 feasible to P using the ellipsoidal method. Then, continue the method by adding the constraint $cx \leq cx^0$ to the constraints $Ax \geq b$, $x \geq 0$. Contracting ellipsoids will continue to encapsulate that portion of the feasible region which contains any feasible point offering better values to cx than the current one.

6.1 An example

The following simple example may be used to illustrate the method. Locate $x = (x_1, x_2)$ such that

$$\begin{aligned} -x_1 \qquad &\leq -1 \\ - x_2 &\leq -1 \\ x_1 + x_2 &\leq \;\; 3 \;. \end{aligned}$$

Table II records the data generated by starting with

$$\begin{aligned} x^0 &= (0,0) \\ B^0 &= \text{diag}\{100,100\} \;, \end{aligned}$$

with δ set equal to zero. The modified method is tabulated in Table

TABLE II

CONVERGENCE IN EXAMPLE, $\delta = 0$

Iteration #	x_1	x_2	Row a	b_{11}	$b_{12} = b_{21}$	b_{22}
0	0.000	0.000	(-1,0)	100	0	100
1	3.333	0.000	(0,-1)	44.44	0	133.3
2	3.333	3.849	(1,1)	59.26	0	59.26
3	1.518	2.035	(1,1)	52.67	-26.34	52.67
4	0.309	0.825	(-1,0)	58.53	-46.82	58.53
5	2.859	-1.215	(0,-1)	26.01	-20.81	44.74
6	1.822	1.014	feasible			

III. Figure 7 illustrates the sequence of points generated by both schemes, while Figures 8a through 9g illustrate the sequence of ellipsoids generated by both the basic and the "deep cut" schemes.

It is interesting to note that the latter scheme requires one more iteration than does the unrefined method. The more significant decrease in volumes is apparent in the second scheme, however.

TABLE III

CONVERGENCE IN EXAMPLE, δ VARIABLE

Iteration #	x_1	x_2	Row a	δ	b_{11}	$b_{12} = b_{21}$	b_{22}
0	0.000	0.000	(-1,0)	0.100	100	0	100
1	4.000	0.00	(0,-1)	0.087	36.00	0	132.0
2	4.000	4.496	(1,1)	0.559	47.64	0	48.90
3	0.576	0.981	(-1,0)	0.086	24.13	-20.02	24.20
4	2.496	-0.612	(0,-1)	0.400	8.953	-7.428	16.26
5	1.391	1.807	(1,1)	0.075	6.771	-1.189	2.603
6	0.582	1.602	(-1,0)	0.191	4.767	-2.644	3.182
7	1.586	1.044	feasible				

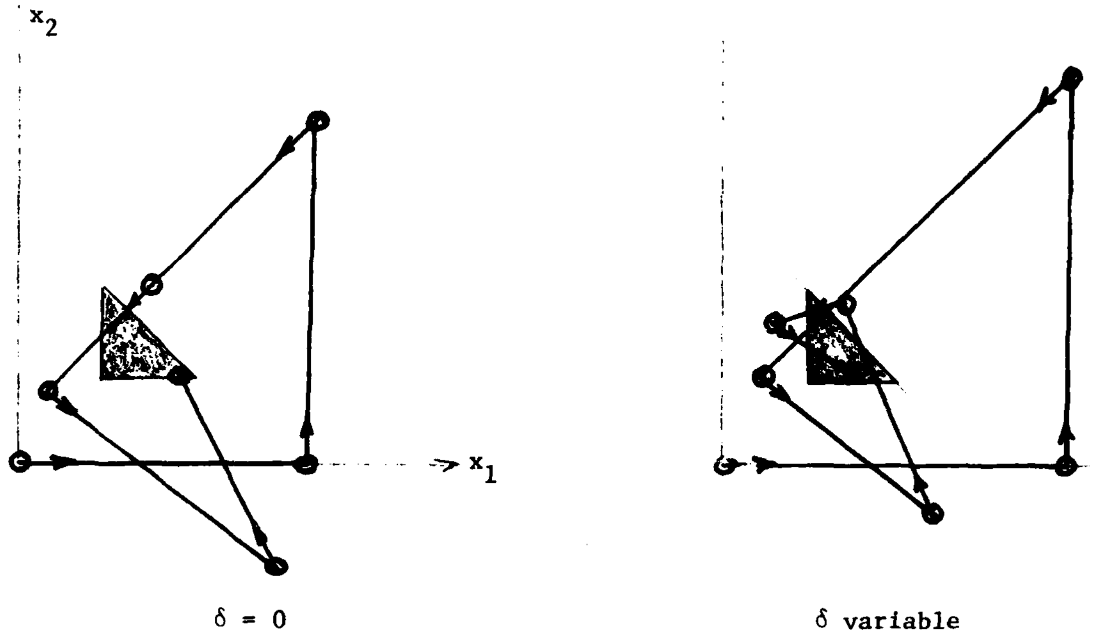
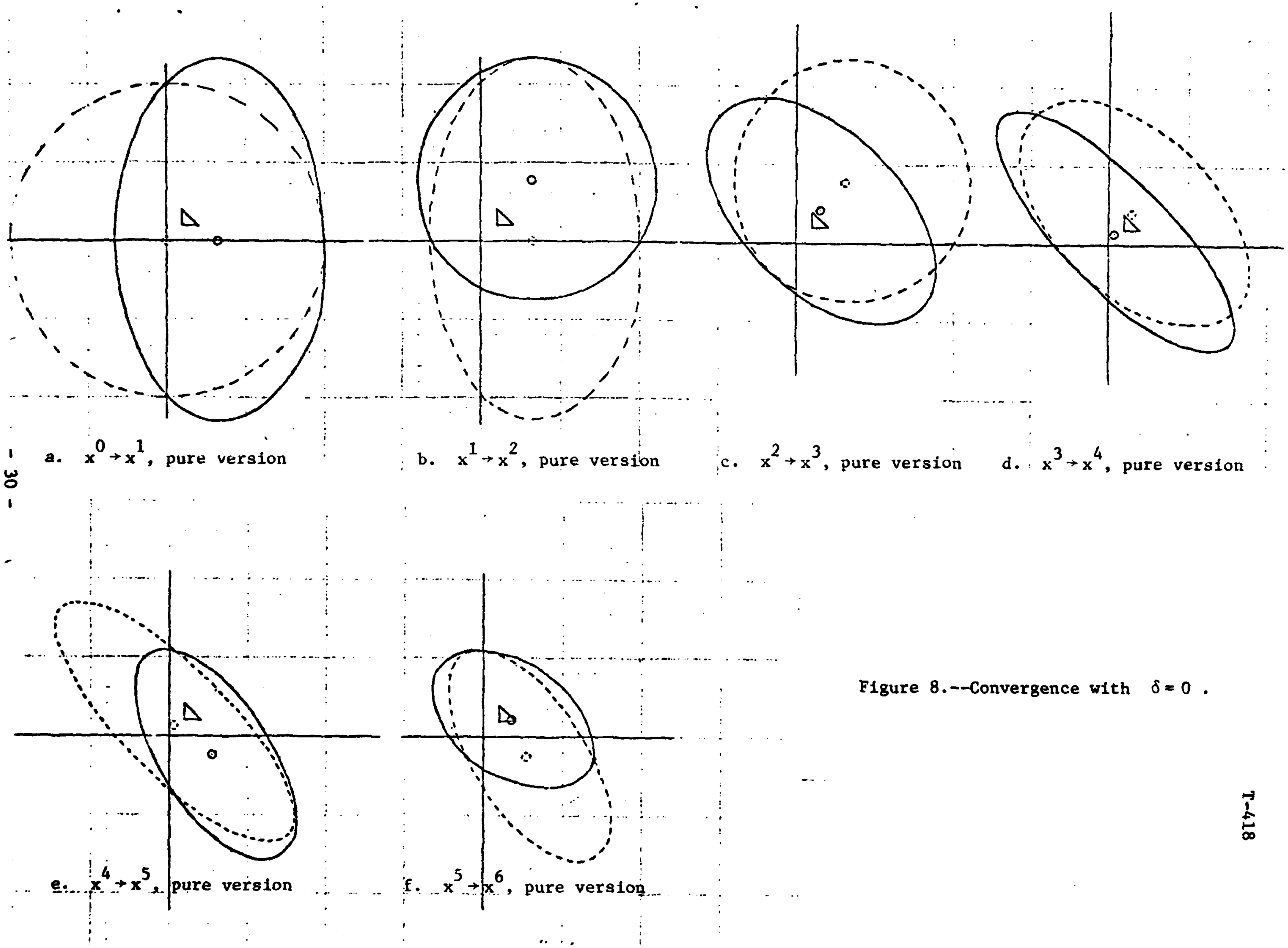


Figure 7.--Convergence of pure and refined methods.



a. $x^0 \to x^1$, pure version

b. $x^1 \to x^2$, pure version

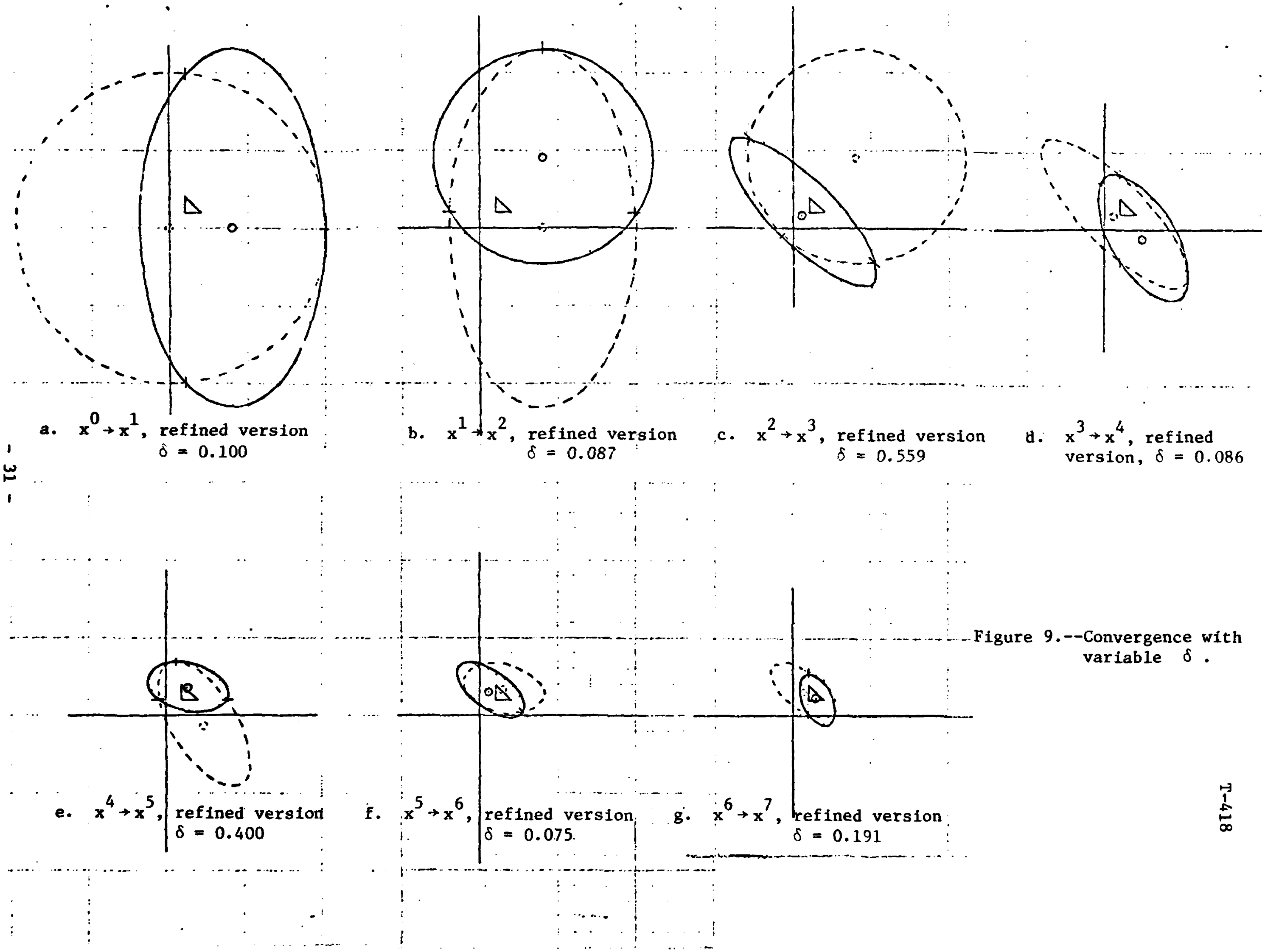
c. $x^2 \to x^3$, pure version

d. $x^3 \to x^4$, pure version

e. $x^4 \to x^5$, pure version

f. $x^5 \to x^6$, pure version

Figure 8.--Convergence with $\delta = 0$.



a. $x^0 \rightarrow x^1$, refined version $\delta = 0.100$

b. $x^1 \rightarrow x^2$, refined version $\delta = 0.087$

c. $x^2 \rightarrow x^3$, refined version $\delta = 0.559$

d. $x^3 \rightarrow x^4$, refined version, $\delta = 0.086$

e. $x^4 \rightarrow x^5$, refined version $\delta = 0.400$

f. $x^5 \rightarrow x^6$, refined version $\delta = 0.075$

g. $x^6 \rightarrow x^7$, refined version $\delta = 0.191$

Figure 9.--Convergence with variable δ .

REFERENCES

[1] DANTZIG, G. B. (1979). Comments on Khachian's algorithm for linear programming. Report SOL 79-22, Department of Operations Research, Stanford University (November).

[2] GÁCS, P. and L. LOVÁSZ (1979). Khachian's algorithm for linear programming. Report CS750, Computer Science Department, Stanford University (July).

[3] JEROSLOW, R. G. (1973). The simplex algorithm with the pivot rule of maximizing criterion improvement. *Discrete Mathematics*, **4** 367-378.

[4] KHACHIAN, L. G. (1979). A polynomial algorithm in linear programming. *Soviet Mathematica Doklady*, **20** (1) 1093-1096 (January-February). Translation for American Mathematical Society by A. Rosa.

[5] KLEE, V. and G. J. MINTY (1972). How good is the simplex algorithm?, in O. Shisha (ed.), *Inequalities III*, 159-175, Academic Press, New York.

[6] KOLATA, GINA BARI (1979). Mathematicians amazed by Russian's discovery. *Science*, **206** 545-546 (November).

[7] LEVIN, A. YU. (1965). An algorithm for the minimization of convex functions. *Dokl. Akad. Nauk SSSR*, **160** (6).

[8] McCALL, E. H. (1980). A study of the Khachian algorithm for real-world linear programming. Sperry Univac POB 43942, St. Paul, Minnesota 55164 (January).

[9] SHOR, N. Z. (1970). Utilization of the operation of space dilation in the minimization of convex functions. *Kibernetika*,

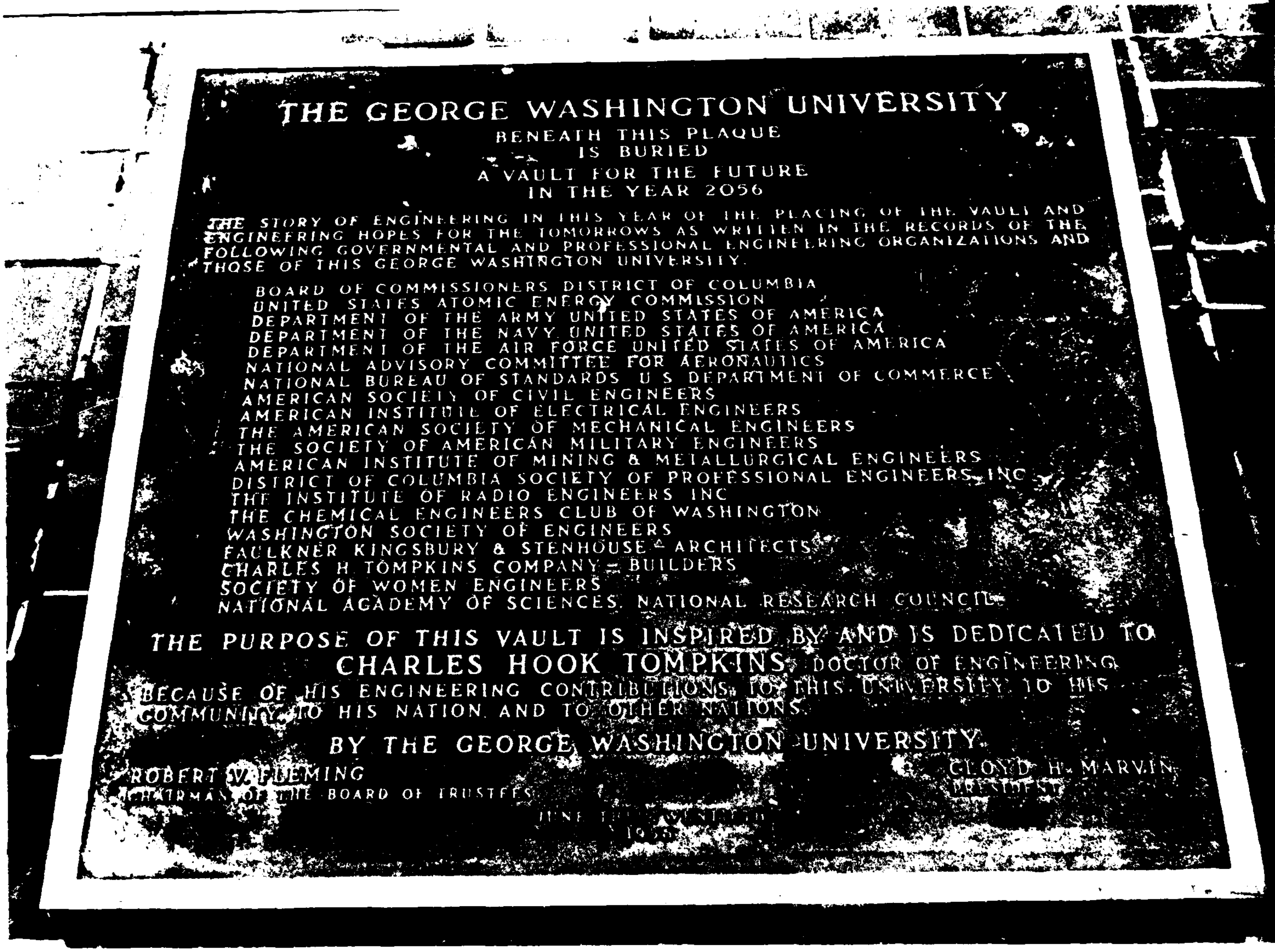
6 (1). Translated in *Cybernetics*, **6** 7-15.

[10] SHOR, N. Z. (1977). Cut-off method with space extension in convex programming problems. Translated in *Cybernetics*, **13** 94-96.

[11] STEEN, LYNN ARTHUR (1979). Linear programming: solid new algorithm. *Science News*, **116** 234-236 (October).

[12] WOLFE, PHILLIP (1980). A bibliography for "ellipsoidal" algorithms. Unpublished notes (February).

[13] YUDIN, D. B. and A. S. NEMIROVSKII (1976). Informational complexity and effective methods for the solution of convex extremal problems. *Ékonom. i Mat. Metody*, **12** (2).



To cope with the expanding technology, our society must be assured of a continuing supply of rigorously trained and educated engineers. The School of Engineering and Applied Science is completely committed to this objective.