(2)

SYSTEMS OPTIMIZATION LABORATORY
DEPARTMENT OF OPERATIONS RESEARCH
STANFORD UNIVERSITY
STANFORD, CALIFORNIA 94305-4022

Converting a Converging Algorithm into a Polynomially Bounded Algorithm

by

George B. Dantzig

TECHNICAL REPORT SOL 91-5

March 1991

Research and reproduction of this report were partially supported by the Office of Naval Research Grant N00014-89-J-1659 and the National Science Foundation Grants DMS-8913089 and ECS-8906260.

Converting a Converging Algorithm

into a

Polynomially Bounded Algorithm

by

George B. Dantzig

Abstract: We consider the general Phase I linear programming problem with a convexity constraint which can be written after some algebraic manipulation in the form:

$$\text{Find } x_j \geq 0, \ \sum_1^n P_j x_j = 0, \ \sum_1^n x_j = 1$$

where P_j are m-vectors satisfying $\|P_j\| = 1$. If feasible, von Neumann's Center of Gravity Algorithm generates a sequence $t = 1, 2, \ldots$ of approximate solutions $\sum P_j x_j^t = b^t$, $\sum x_j^t = 1$, $x_j^t \geq 0$ which converges in the limit as $t \to \infty$ to a feasible solution to the Phase I problem. We assume that all perturbed problems $\sum_1^n P_j x_j = \hat{b}$, $\sum x_j = 1$, $x_j \geq 0$ are feasible for all $\|\hat{b}\| < r$ where $r > 0$ is given. We apply this algorithm to $m+1$ perturbed problems with right hand sides $\hat{b} = \hat{b}^i$, $i = 1, 2, \ldots, m+1$ to obtain an exact solution to the unperturbed problem with $\hat{b} = 0$ in $T < 4r^{-2}(m+1)^3$ iterations. Each iteration consists of $m(n+3)\delta$ multiplications and additions where δ is the non-zero coefficient density.

Von Neumann* in 1948 proposed the first interior algorithm for solving a general Phase I linear program with a convexity constraint. We will reproduce his proof that in $t < 1/\rho^2$ iterations an approximate solution $\sum P_j x_j^t = b^t$ will be generated with $\|b^t\| < \rho$. When applied to a perturbed problem $b = \hat{b} \neq 0$, we will show that in $t < 4/\rho^2$ iterations an approximate solution will be generated with $\|b^t - \hat{b}\| < \rho$.

---

* verbal communication

Geometrically, in the m-space of the columns, since $\|P_j\| = 1$, all points P_j lie on the surface of the m-dimensional hypersphere S_0 of unit radius with center at the origin. We are given r the radius of a concentric hypersphere $S_1 \subseteq S_0$ centered at the origin that lies in the convex hull of the points P_j. Thus r is a measure of how deeply the origin is embedded in the set of b such that $b = \sum P_j x_j$, $x_j \geq 0$, $\sum x_j = 1$ is feasible.

To generate the $m+1$ different *finite* sequences (x^t, b^t) whose b^t approach $m+1$ different points $\hat{b}^i$, the $\hat{b}^i$ are prechosen. These can be the vertices of any simplex lying in the set of feasible b that contains the origin as an interior point. We choose $\hat{b}^i$ to be the vertices of an $(m+1)$ *equilateral simplex* whose center is the origin and whose vertices are located at distances $r \cdot m/(m+1)$ from the origin; for example the coordinates of $\hat{b}^i$ may be chosen as follows:

$$
\begin{array}{lllllll}
\hat{b}^{m+1} & = & [\ 0 & 0 & \cdots & 0 & m a_m]^T \\
\hat{b}^{m} & = & [\ 0 & 0 & & (m-1)\,a_{m-1} & -a_m]^T \\
\hat{b}^{m-1} & = & [\ 0 & 0 & & -a_{m-1} & -a_m]^T \\
\vdots & \vdots & \vdots\ \vdots & \vdots & & \vdots & \vdots \\
\hat{b}^{3} & = & [\ 0 & +2a_2 & \cdots & -a_{m-1} & -a_m]^T \\
\hat{b}^{2} & = & [\ a_1 & -a_2 & \cdots & -a_{m-1} & -a_m]^T \\
\hat{b}^{1} & = & [\ -a_1 & -a_2 & \cdots & -a_{m-1} & -a_m]^T
\end{array}
\tag{1}
$$

where $a_i = r\sqrt{\frac{m}{m+1}} \cdot \sqrt{\frac{1}{i(i+1)}}$.

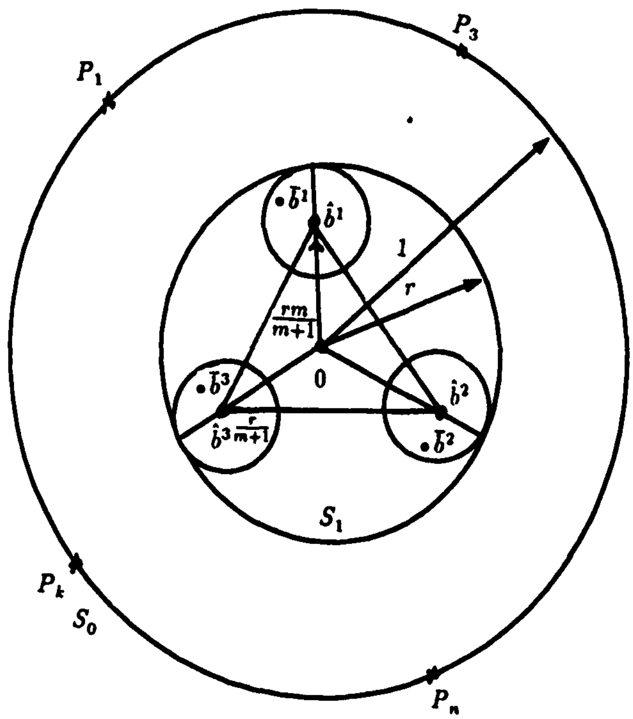

Figure 1. The Iterations Converge to $\hat{b}^i$ Instead of the Origin 0.

When the i^{th} sequence (x^t, b^t) (which is converging towards $\hat{b}^i$) reaches a point $b^t = \bar{b}^i$ such that $\|\bar{b}^i - \hat{b}^i\| < r/(m+1)$, the sequence for that i is terminated. Note that all interior points of Ball$_i$ of radius $\rho = r/(m+1)$ centered at $\hat{b}^i$ lie inside the hypersphere $S_1 \subseteq S_0$. We will show $b^t = \bar{b}^i \subset \text{Ball}_i$ is attainable by the iterative process. Associated with $\bar{b}^i$ is the approximate solution $\bar{x}^i = x^t$ that generated it. Thus an upper bound to generate all $m+1$ approximate solutions $\left(\bar{x}^i, \bar{b}^i\right)$ whose $\bar{b}^i$ lie strictly in $m+1$ ρ-balls centered at $\hat{b}^i$ can be done in

(2) $$\text{iteration count} < 4(m+1)/\rho^2 = 4(m+1)^3/r^2 \text{ , } \rho = r/(m+1),$$

iterations. The final step is to generate the feasible solution $\bar{x}$ to the Phase I problem by finding weights $\bar{\lambda}_i > 0$, $\bar{x} = \sum \lambda_i \bar{x}^i \geq 0$, $\sum \bar{x}_j = 1$, 1, $\sum P_j \bar{x}_j = 0$. These weights $\bar{\lambda} = (\bar{\lambda}_1, \bar{\lambda}_2, \ldots, \bar{\lambda}_{m+1})$ are found by solving the $(m+1) \times (m+1)$ system

(3) $$\begin{aligned} \sum \bar{b}^i \bar{\lambda}_i &= 0 \\ \sum \bar{\lambda}_i &= 1. \end{aligned}$$

We will prove that this system has a unique solution $\bar{\lambda} = (\bar{\lambda}_1, \ldots, \bar{\lambda}_{m+1}) > 0$.

We now describe the detailed steps of von Neumann's algorithm for finding an approximate solution to a perturbed problem $\sum P_j x_j = \hat{b}$, $\sum x_j = 1$, $x \geq 0$ and give a proof of the rate of convergence of the i-th sequence to some $\hat{b} = \bar{b}^i \subset B_i$. We initiate the sequence of iterations by $x = x^1 = (1, 0, \ldots, 0)$, $b^1 = P_1$. Inductively let x^{t-1}, b^{t-1} be the $t-1$ approximation. We use it to generate x^t, b^t.

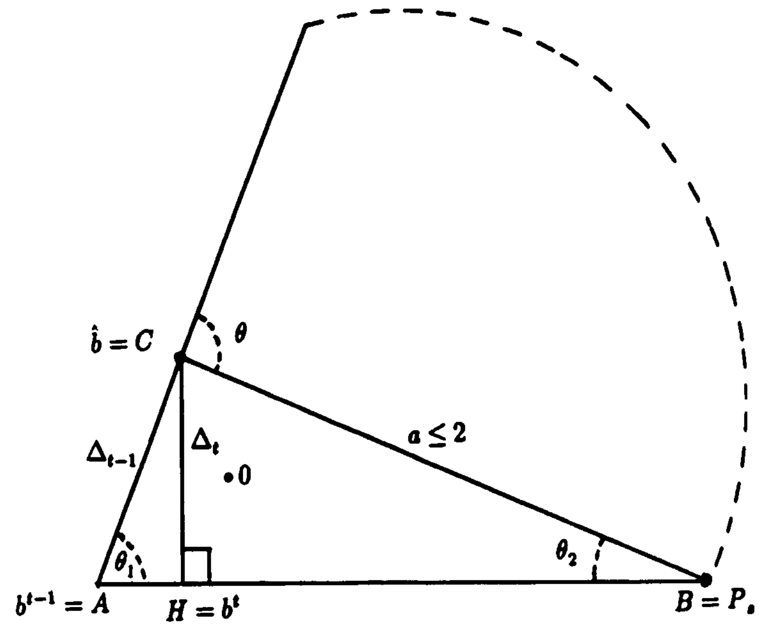

Figure 2. The Von Neumann Iterative Step

Referring to Figure 2, P_s is selected as that P_j such that $P_j - \hat{b}$ makes the sharpest angle θ with direction $\hat{b} - b^{t-1}$, namely

(4) $$s = \underset{j}{\text{ARGMAX}}\ \|[\hat{b} - b^{t-1}]^{\mathrm{T}}[P_j - \hat{b}]/\|P_j - \hat{b}\|\|.$$

which can be carried out in $m(n+3)$ operations assuming $\|P_j - \hat{b}\|$ is preprocessed. The triangle b^{t-1}, P_s, $\hat{b}$ will be labeled ABC. The next approximation point $H = b^t$ is the foot of perpendicular dropped from C onto the side AB of the triangle ABC. From the figure, it is clear that H is a weighted convex combination of A and B with weights proportional to $\cos\theta_2$ and $\cos\theta_1$, i.e.,

(5) $$b^t = (\cos\theta_2 \cdot b^{t-1} + \cos\theta_1 \cdot P_s)/(\cos\theta_2 + \cos\theta_1),$$

$$x^t = (\cos\theta_2 \cdot x^{t-1} + \cos\theta_1 \cdot U_s)/(\cos\theta_2 + \cos\theta_2),$$

where U_s is the unit n vector with 1 in component s. $\cos\theta_1$ and $\cos\theta_2$ are computed by

(6) $$\cos\theta_2 = \frac{(\hat{b} - P_s)^{\mathrm{T}}(b^{t-1} - P_s)}{\|\hat{b} - P_s\|\ \|b^{t-1} - P_s\|}, \quad \cos\theta_1 = \frac{(P_s - b^{t-1})^{\mathrm{T}}(\hat{b} - b^{t-1})}{\|P_s - b^{t-1}\|\ \|\hat{b} - b^{t-1}\|}.$$

In order to determine the rate of convergence, note $\theta \le \pi/2$ because if, on the contrary, $\theta > \pi/2$ then all points P_j would lie on one side of the hyperplane through $\hat{b}$ orthogonal to $b^{t-1} - \hat{b}$ implying that $\hat{b} = \hat{b}^i$ for the i-th sequence lies outside the convex hull of the P_j's contrary to our assumption that all points located at a distance r or less from the origin are in the set of feasible b (i.e., $\hat{b}^i$ by construction lies in the interior of the set of feasible $\hat{b} \subset S_1$ at a distance $r/(m+1)$ from the boundary of S_1. To simplify the notation, let

$$\Delta_{t-1} = \|b^{t-1} - \hat{b}\| \text{ and } \Delta_t = \|b^t - \hat{b}\|,$$

then

(7) $$\Delta_t = \Delta_{t-1} \sin\theta_1 \text{and } \Delta_t = \|P_s - \hat{b}\| \sin\theta_2.$$

Therefore, noting $\theta_1 + \theta_2 = \theta \le \pi/2$,

$$\left(\frac{\Delta_t}{\Delta_{t-1}}\right)^2 + \left(\frac{\Delta_t}{\|P_s - \hat{b}\|}\right)^2 = \sin^2\theta_1 + \sin^2\theta_2 \le 1.$$

Recalling that diameter of the hypersphere is 2, it follows that $\|P_s - \hat{b}\| < 2$ and therefore for $\tau = 2, 3, \ldots, t$:

(8) $$\left(\frac{\Delta_\tau}{\Delta_{\tau-1}}\right)^2 + \left(\frac{\Delta_\tau}{2}\right)^2 < 1.$$

Comment: These inequalities can be made tighter when $\hat{b} = 0$ because $\|P_s - \hat{b}\| = \|P_s\| = 1$. If so, (8) can be replaced by $(\Delta_\tau/\Delta_{\tau-1})^2 + \Delta_\tau^2 \le 1$ and the development that follows can be modified accordingly with the conclusion that if the von Neumann iterative process is applied to the case $\hat{b} = 0$ instead of to $\hat{b}^i \neq 0$ an approximation b^t such that $\|b^t\| < \rho$ can be attained in less that $1/\rho^2$ iterations (instead of less than $4/\rho^2$ iterations).

Dividing (8) through by $(\Delta_\tau)^2$ for $\tau = 2, \ldots, t$:

$$\begin{array}{rcl} (1/\Delta_{t-1})^2 + (1/4) & < & (1/\Delta_t)^2 \\ (1/\Delta_{t-2})^2 + (1/4) & < & (1/\Delta_{t-1})^2 \\ \vdots & \vdots & \vdots \\ (1/\Delta)^2 \;\; + (1/4) & < & (1/\Delta_2)^2. \end{array}$$

Summing the above, canceling terms common to both sides of the sum and, recalling $\Delta_1 < 2$, we have

$$(1/\Delta_t)^2 > (1/4) + (t-1)/4 = t/4. \tag{10}$$

We conclude that $t < 4/\Delta_t^2$ iterations, i.e. less than $4/\rho^2$ iterations would be needed for the i^{th} sequence to terminate by reaching $b^t = \bar{b}^i$, an interior point of the ρ-ball centered at $\hat{b}^i$. Since $\rho = r/(m+1)$ and there are $(m+1)$ ρ-balls, the upper bound on

$$\text{iteration count} < 4(m+1)^3/r^2. \tag{11}$$

What remains to show is that the $(m+1) \times (m+1)$ system (3) can be solved, that the solution $\bar{\lambda}$ is unique, and that $\bar{\lambda} = (\bar{\lambda}_1, \bar{\lambda}_2, \ldots, \bar{\lambda}_{m+1}) > 0$.

Existence of Separating Hyperplanes: Let $y = (y_1, y_2, \ldots, y_m)$ represent a general point in R^m. The equation of any hyperplane through the origin has the form $a^T y = 0$. This hyperplane is said to *separate* y^1 from y^2 if $a^T y^1$ and $a^T y^2$ are of opposite signs.

Fact 1. Each hyperplane $(\hat{b}^i)^T y = 0$ for $i = 1, 2, \ldots, m$ separates any point in the ρ-ball centered at $\hat{b}^i$ from any point lying in any of the other ρ-balls centered at $\hat{b}^j$.

Proof: Because of the $m+1$ fold symmetry of the equilateral simplex it is sufficient to demonstrate that the hyperplane $(\hat{b}^{m+1})^T y = 0$ separates $\bar{b}^{m+1}$ from $\bar{b}^m$ where $\|\bar{b}^{m+1} - \hat{b}^{m+1}\| < r/(m+1)$ and $\|\bar{b}^m - \hat{b}^m\| < r/(m+1)$. The coordinates of $\hat{b}^{m+1}$ and $\hat{b}^m$ defined by (1) are $\hat{b}^{m+1} = (0, 0, \ldots, rm/(m+1))^T$ and $\hat{b}^m = (0, 0, \ldots, r\sqrt{m-1}/\sqrt{m+1}, -r/(m+1))^T$. The hyperplane $(\hat{b}^{m+1})^T y = 0$ reduces to $(0, \ldots, 1)y = U_m^T y = 0$. Letting

$\bar{b}^{m+1} = \hat{b}^{m+1} + u$ where $\|u\| < r/(m+1)$, we have $U_m^T\ \bar{b}^{m+1} = \bar{b}_m^{m+1} = \hat{b}_m^{m+1} + u_m > rm/(m+1) - r/(m+1) > 0$ since $\|u_{m+1}\| < r/(m+1)$. Letting $\bar{b}^m = \hat{b}^m + v$ where $\|v\| < r/(m+1)$, we have $U_m^T\ \bar{b}^m = \bar{b}_m^m + v_m < -r/(m+1) + r/(m+1) = 0$. Thus $U_m \bar{b}^{m+1}$ and $U_m\ \bar{b}^m$ have opposite signs and so the hyperplane $U_m y = 0$ separates $\bar{b}^{m+1}$ from $\bar{b}^m$. ∎

The Separating Hyperplanes Theorem below states conditions which imply that the points $\bar{b}^1, \bar{b}^2, \ldots, \bar{b}^{m+1}$ are the vertices of a simplex containing the origin in its interior. That these conditions are satisfied follows from Fact 1.

Separating Hyperplanes Theorem: Given (1) that $(\hat{b}^1, \hat{b}^2, \ldots, \hat{b}^{m+1})$ are any $(m+1)$ vertices of an m-dimensional simplex $\hat{T}$ containing the origin; given (2) that $a^i y = 0$ for $i = 1, 2, \ldots, m+1$ are the equations of $m+1$ hyperplanes separating $\hat{b}^i$ from $\hat{b}^j$ for all $j \neq i$; and given (3) any $m+1$ points $\bar{b}^1, \bar{b}^2, \ldots, \bar{b}^{m+1}$ such that each hyperplane $a^i y = 0$ separates $\bar{b}^i$ (on the same side as $\hat{b}^i$) from $\hat{b}^j$ for all $j \neq i$; then $\bar{b}^1, \bar{b}^2, \ldots \bar{b}^m$ are the vertices $\bar{T}$ of an m-dimensional simplex that contains the origin as an interior point.

Proof: Since the simplex associated with $\hat{T}$ contains the origin, we know there exist $\hat{\lambda}_i \geq 0, \bar{\lambda}_i \geq 0$ such that

$$\sum \hat{b}^j \hat{\lambda}_j + \sum \bar{b}^i \bar{\lambda}_i = 0 \tag{13.1}$$

$$\sum \hat{\lambda}_i + \sum \bar{\lambda}_i = 1. \tag{13.2}$$

Before continuing with the proof, we show two more facts:

Fact 2. If $(\hat{\lambda}, \bar{\lambda})$ is a feasible solution to (13.1), (13.2), then $\hat{\lambda}_i + \bar{\lambda}_i > 0$ for all i.

Suppose, on the contrary, $\hat{\lambda}_k = 0$, $\bar{\lambda}_k = 0$ for some k. Multiply (13.1) on the left by a^k; recall, by assumption, $a^k \hat{b}^j < 0$ and $a^k \bar{b}^j < 0$ for all $j \neq k$. We have

$$\sum_{1 \neq k} (a^k \hat{b}^i) \hat{\lambda}_i \quad + \sum_{j \neq k} (a^k \bar{b}^i) \bar{\lambda}_i \quad = 0 \tag{14.1}$$

$$\sum_{j \neq k} \hat{\lambda}_i \quad + \sum_{j \neq k} \bar{\lambda}_i \quad = 1, \tag{14.2}$$

implying, that (14.1) is the sum of non-negative terms (not all zero by (14.2), a contradiction. ∎

Fact 3. If T is any simplex containing the origin whose vertices i are separated from the remaining vertices $j \neq i$ by a hyperplane $a^i y = 0$ for each i, then T contains the origin strictly in its interior. ∎

Fact 3 follows from Fact 2 by setting $\bar{b}^i = \hat{b}^i$ for all i.

Continuing with the proof of the separating hyperplanes theorem, define $\mathfrak{B}$ and U_{m+1} by

$$(15) \qquad \mathfrak{B} = \begin{bmatrix} \hat{b}^1 & \hat{b}^2 & \dots & \hat{b}^{m+1} \\ 1 & 1 & & 1 \end{bmatrix}, \qquad U_{m+1} = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}.$$

Since $\hat{T}$ are the vertices of an m-dimensional simplex by assumption, it means that $\mathfrak{B}$ is non-singular and that $\mathfrak{B}\hat{\lambda} = U_{m+1}$ can be solved for $\hat{\lambda}$ and, when solved, $\hat{\lambda} \geq 0$. From Fact 3 it follows that $\hat{\lambda} > 0$. We view $\mathfrak{B}$ as a feasible non-degenerate basis and consider $\begin{bmatrix} \bar{b}^1 \\ 1 \end{bmatrix}$ as an incoming non-basic column. We assert it will replace $\begin{bmatrix} \hat{b}^1 \\ 1 \end{bmatrix}$ in the basis because, on the contrary, if it replaced some column $k \neq 1$ in the basis, it would imply after the replacement that both $\bar{\lambda}_k$ and $\hat{\lambda}_k$are 0 in a feasible solution, contrary to Fact 2. By replacing in turn basis columns $\begin{bmatrix} \hat{b}^2 \\ 1 \end{bmatrix}$ by $\begin{bmatrix} \hat{b}^2 \\ 1 \end{bmatrix}$, $\begin{bmatrix} \hat{b}^3 \\ 1 \end{bmatrix}$ by $\begin{bmatrix} \bar{b}^3 \\ 1 \end{bmatrix}$, etc., we arrive at the conclusion that $\bar{T}$ are the vertices of a simplex containing the origin. It then follows from Fact 3 that this simplex contains the origin as a strictly interior point. ∎

This completes the proof that the $(m+1)$ sequences converge to $m+1$ points $\bar{b}^i$ in less than $4(m+1)^3/r^2$ iterations. By applying the weights $\bar{\lambda}_i > 0$ to the corresponding $\bar{x}^i$, we generate the exact solution x to the Phase I linear program.

One final remark: Just because an algorithm is polynomial does not necessarily make it practical. The von Neumann algorithm has a poor convergence rate. Like the simplex method each of its iterations requires about $mn\delta$ multiplications and additions where δ is the density of non-zero coefficients. When applied to $(m+1)$ perturbed problems as we do in this paper, we obtain an upper bound of $4(m+1)^3/r^2$ iterations where $0 < r < 1$. The moral of this tale is that, like gunners, we may do better by first bracketing the target and then applying a final correction.

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. AGENCY USE ONLY *(Leave blank)*

2. REPORT DATE
March 1991

3. REPORT TYPE AND DATES COVERED
Technical Report

4. TITLE AND SUBTITLE
Converting a Converging Algorithm into a Polynomially Bounded Algorithm

5. FUNDING NUMBERS
ONR - N00014-89-J-1659

6. AUTHOR(S)
George B. Dantzig

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
Department of Operations Research - SOL
Stanford University
Stanford, CA 94305-4022

8. PERFORMING ORGANIZATION REPORT NUMBER
1111MA

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)
Office of Naval Research - Department of the Navy
800 N. Quincy Street
Arlington, VA 22217

10. SPONSORING/MONITORING AGENCY REPORT NUMBER
SOL 91-5

12a. DISTRIBUTION/AVAILABILITY STATEMENT
UNLIMITED

12b. DISTRIBUTION CODE
UL

13. ABSTRACT *(Maximum 200 words)*

We consider the general Phase I linear programming problem with a convexity constraint which can be written after some algebraic manipulation in the form:

$$\text{Find } x_j \geq 0, \quad \sum_1^n P_j x_j = 0, \quad \sum_1^n x_j = 1,$$

where P_j are m-vectors satisfying $\|P_j\| = 1$. If feasible, von Newmann's Center of Gravity Algorithm generates a sequence $t = 1, 2, \ldots$ of approximate solutions $\sum P_j x_j^t = b^t$, $\sum x_j^t = 1$, $x_j^t \geq 0$ which converges in the limit as $t \to \infty$ to a feasible solution to the Phase I problem. We assume that all perturbed problems $\sum_1^n P_j x_j = \hat{b}$, $\sum x_j = 1$, $x_j \geq 0$ are feasible for all $\| \hat{b} \| < r$ where $r > 0$ is given. We apply this algorithm to $m+1$ perturbed problems with right hand sides $\hat{b} = \hat{b}^i$, $i = 1, 2, \ldots, m+1$ to obtain an exact solution to the unperturbed problem with $\hat{b} = 0$ in $T < 4r^{-2}(m+1)^3$ iterations. Each iteration consists of $m(n+3)\delta$ multiplications and additions where δ is the non-zero coefficient density.

14. SUBJECT TERMS
Linear Programming; Polynomial Algorithm; Phase I

15. NUMBER OF PAGES
7 pp.

16. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT
UNCLASSIFIED

18. SECURITY CLASSIFICATION OF THIS PAGE

19. SECURITY CLASSIFICATION OF ABSTRACT

20. LIMITATION OF ABSTRACT
SAR