DISCUSSION PAPER NUMBER 311

# ON PIECEWISE LINEAR APPROXIMATIONS TO SMOOTH MAPPINGS*

by

R. Saigal

November, 1977



*This research was partially supported by Grant No. MCS77-03472 from the National Science Foundation.

# ON PIECEWISE LINEAR APPROXIMATIONS TO SMOOTH MAPPINGS*

R. Saigal
Northwestern University
Evanston, Illinois

## ABSTRACT

Given a continuously differentiable mapping $f$ from $R^n$ into $R^n$, in this work we study piecewise linear approximations to it on certain subdivisions of $R^n$. It is shown that several properties of the subdivision are critical when the Jacobeans of the pieces of linearity of the approximation are required to be close to the Jacobeans of $f$. In addition, it is shown that even under arbitrary scaling of the triangulations used in fixed point algorithms, good approximations of the derivatives result.

*This research was partially supported by Grant No. MCS77-03472 from the National Science Foundation.

ON PIECEWISE LINEAR APPROXIMATIONS TO SMOOTH MAPPINGS

R. Saigal

## 1. INTRODUCTION

Given a continuously differentiable mapping $f \equiv (f_1, f_2, \ldots, f_n)$ from $R^n$, the n-dimensional Euclidean space into itself, we consider here the problem of generating piecewise linear approximations to f on some given subdivisions of $R^n$. In particular, we will study the properties of the subdivisions which ensure that both f and its Jacobean Df are "close" to the piecewise linear approximation and its derivative. This study is related to the recent study of Kojima [3], who used some fundamental theorems of Whitehead [15] to establish such approximations. Our approach is to get explicit bounds under the assumption that Df is Lipschitz continuous.

Throughout this paper we use the $\ell_2$ norm $||x|| = (x^t x)^{\frac{1}{2}}$ and for a n x n matrix, $||A|| = \max_{||x||=1} ||Ax||$. The results established in this paper have implications for the recent algorithms that compute fixed points, namely those of Eaves and Saigal [2] and Merrill [5]. Some of these results have been used in Saigal [9], Saigal and Todd [11].

In section 2 we establish some basic results used in the subsequent sections; in section 3 we establish some basic properties of piecewise linear approximations; in section 4 we study approximations on scaled triangulations; and in section 5, an application is given.

## 2. NOTATION, DEFINITIONS AND PRELIMINARY RESULTS

Throughout this work we will assume that f is differentiable, with its derivative Df Lipschitz continuous, that is, there is a $\alpha>0$ such that for each x and y in $R^n$

$$||Df(x) - Df(y)|| \leqq \alpha||x-y||$$

We shall frequently use the following result in this paper, and its proof can be found in [6, 3.2.12].

Theorem 2.1: Let f be differentiable, and its derivative Df be Lipschitz continuous on an open convex set W. Then, there exists a mapping e: $W \times W \to R^n$ such that for all x and y in W,

$$f(y) = f(x) + Df(x)(y-x) + e(y,x)$$

and that $||e(y,x)|| \leqq \frac{1}{2}\alpha||x-y||^2$, where $\alpha>0$ is the Lipschitz constant.

Given n + 1 affinely independent points, called vertices, $v^1, v^2, \ldots v^{n+1}$ in $R^n$, we call their convex hull a n-dimensional simplex $\sigma = (v^1, v^2, \ldots, v^{n+1})$. For an n-simplex , define a n x n matrix

$$V = (v^2 - v^1, v^3 - v^1, \ldots, v^{n+1} - v^1).$$

Since $v^i$ are affinely independent, V is nonsingular, and hence invertible. Also, we define

$$\varepsilon = \max_{x,y\in\sigma} ||x - y||$$

as the diameter of $\sigma$ and for $\bar{v} = \frac{1}{n+1}\Sigma v^i$

$$\rho = \min_{x\in\partial\sigma} ||x - \bar{v}||$$

the radius of $\sigma$, where $\partial\sigma$ is the boundary of the simplex $\sigma$. Also, we define $\theta = \rho/\varepsilon$ as the thickness of the simplex $\sigma$.

We can then prove:

Theorem 2.2:

$$\frac{1}{||V^{-1}||} \cdot \frac{1}{(n+1)\sqrt{n}} \leqq \rho \leqq \frac{1}{||V^{-1}||} \frac{\sqrt{n}}{n+1}$$

for any n-simplex $\sigma$. The lower inequality is tight, in the sense that there are simplexes for which equality is attained.

Before we prove this theorem, we have a Lemma:

Lemma 2.3: Let $\gamma = \max \{||eV^{-1}||, \max_i ||u_i{}^t V^{-1}||\}$ where $u_i$ is the $i^{th}$ unit vector, and $e = (1, 1, \ldots, 1)$. Then

$$\rho = \frac{1}{(n+1)\gamma}$$

Proof:

Since $\sigma$ is a n-simplex, its boundary $\delta\sigma$ can be written as the union of (n-1)-simplexes $\sigma^i$, where $\sigma^i = (v^1, \ldots, v^{i-1}, v^{i+1}, \ldots, v^{n+1})$. Hence

$$\min_{x\varepsilon\delta\sigma} ||x-\bar{v}|| = \min_i \left\{ \min_{x\varepsilon\sigma^i} ||x-\bar{v}|| \right\}$$

Define $V^i = (v^1-v^i, v^2-v^i, \ldots, v^{i-1}-v^i, v^{i+1}-v^i, \ldots, v^{n+1}-v^i)$ for each $i = 1, \ldots, n+1$. Then

$$V^i = VM^iP^i \qquad (2.1)$$

where

$$M^i = \begin{bmatrix} 1 & & & & & & & \\ & 1 & & & & & & \\ & & \ddots & & & & & \\ -1 & -1 & \cdots & -1 & \cdots & & & -1 \\ & & & & \ddots & & & \\ & & & & & & & 1 \end{bmatrix}$$

and $P^i$ a permutation matrix.

For $x \varepsilon \sigma^i$ we have $\lambda_j \geqq 0$, $j \neq i$, $\Sigma\lambda_j = 1$, and $\sum \lambda_j v^j = x$. Then

$$x-\bar{v} = (\lambda_1 - \frac{1}{n+1})v^1 + \ldots + (\lambda_{i-1} - \frac{1}{n+1})v^{i-1} + (\lambda_{i+1} - \frac{1}{n+1})v^{i-1} + \ldots + (\lambda_{n+1} - \frac{1}{n+1})v^{n+1} - \frac{1}{n+1} v^i = V^i_y$$

where $y \in \rho = \{x: \sum_{i=1} x_i = \frac{1}{n+1}, -\frac{1}{n+1} \leqq x_i \leqq \frac{n}{n+1}\}$. And, in view of (2.1), $x-\bar{v} = VM^iP^iy$.

Now, $\rho_1 = \min_{x\varepsilon\sigma} ||x-\bar{v}|| = \min_{y\varepsilon P} ||V_y|| = \frac{1}{(n+1)||eV^{-1}||}$.

and, $\rho_i = \min_{x\varepsilon\sigma^i} ||x-\bar{v}|| = \min_{y\varepsilon P} ||VM^iP^iy|| = \min_{y\varepsilon M^iP^iP} ||Vx||$

$= \frac{1}{(n+1)||u_i{}^tV^{-1}||}$, and thus, we have our result.

Proof of Theorem 2.2

To see the upper bound, let x be such that $||V^{-1}|| = ||V^{-1}x||$. Then $||V^{-1}|| = ||(u_1{}^tV^{-1}x, u_2{}^tV^{-1}x, \ldots, u_n{}^tV^{-1}x)||$ $\leqq \sqrt{n} \max_i ||u_i{}^tV^{-1}|| \leqq \sqrt{n}\,\gamma = \frac{\sqrt{n}}{(n+1)\rho}$.

To see the lower bound, note that $||u_i{}^tV^{-1}|| \leqq ||V^{-1}||$ and $||eV^{-1}|| \leqq \sqrt{n}\,||V^{-1}||$. Thus, $\gamma \leqq \sqrt{n}\,||V^{-1}||$ and so we have our result.

To see the remaining half of the theorem, it can be readily confirmed that for the simplex $\sigma = (v^i, \ldots, v^{n+1})$ with

$v^1 = 0$

$v^{i+1} = u_i \qquad i = 1, \ldots, n$

for the lower bound, equality is attained.

## 3. PIECEWISE LINEAR APPROXIMATIONS

We shall now study piecewise linear approximations to a function f generated by a subdivision of $R^n$.

Triangulations: Given a collection K of subsets of $R^n$, we will say that it triangulates $R^n$ with vertices $K^0$ if

(1) members of K are n-simplexes, with vertices in $K^0$.

(2) each x in $R^n$ belongs to at least one n-simplex of K.

(3) if any two n-simplexes meet, they do so on a common face.

Given an arbitrary triangulation K of $R^n$ we say its grid is $\varepsilon > 0$ if the diameter of any simplex of K is bounded by $\varepsilon$, and its thickness is $\theta$ if the thickness of any simplex in K is larger than $\theta$.

Linear Approximations: Given a triangulation K of $R^n$, and a simplex $\sigma = (v^1, \ldots, v^{n+1})$ in K, we say the mapping Ax-a is a linear approximation to f on $\sigma$ if

$$Av^i - a = f(v^i) \text{ for each } i = 1, \ldots, n+1.$$

We shall denote this linear approximation by $f_\sigma$.

For linear approximations, we can prove:

Theorem 3.1: Let K be a triangulation of $R^n$ of grid $\varepsilon > 0$. And let $\sigma$ in K be a n-simplex. Then, for each $x \varepsilon \sigma$, $||f(x) - f_\sigma(x)|| \leqq \frac{1}{2}\alpha\varepsilon^2$.

Proof:

Since $x \varepsilon \sigma = (v^1, \ldots, v^{n+1})$, there exist $0 \leqq \lambda_i \leqq 1$, $i = 1, \ldots, n+1$, $\Sigma\lambda_i = 1$ such that $x = \Sigma\lambda_i v^i$. Using Theorem 2.1, we have

$$f(v^i) = f(x) + Df(x)(v^i - x) + e(v^i, x)$$

Also,

$$f_\sigma(x) = \sum_{i=1}^{n+1} \lambda_i f(v^i) = f(x) + \sum_{i=1}^{n+1} \lambda_i e(v^i, x)$$

And so

$$||f_\sigma(x) - f(x)|| \leqq \frac{1}{2} \sum_{i=1}^{n+1} \lambda_i \alpha ||v^i - x||^2$$
$$\leqq \frac{1}{2}\alpha\varepsilon^2$$

and we have our result.

Thus, a piecewise linear approximation $f_K$ to $f$ on $K$ is then generated by setting $f_K|\sigma \equiv f_\sigma$ (i.e., $f_K$ restricted to $\sigma$ is the linear mapping $f_\sigma$).

Let $f_\sigma(x) = A_\sigma x - a$. Then, we can prove:

Theorem 3.2: Let $K$ be a triangulation of $R^n$ with grid $\varepsilon > 0$ and thickness $\theta > 0$. Then, for any $\sigma$ in $K$ and $x\varepsilon\sigma$

$$||Df(x) - A_\sigma|| \leqq \frac{n\alpha}{n+1} \frac{\varepsilon}{\theta}.$$

Proof:

Let $\sigma = (v^1, \ldots, v^{n+1})$ and $x \varepsilon \sigma$. Using Theorem 2.1, we have

$f(v^i) = f(x) + Df(x)(v^i - x) + e(v^i,x)$. Hence,

$f(v^i) - f(v^1) = Df(x)(v^i - v^1) + e(v^i,x) - e(v^1,x)$,

Or,

$$A_\sigma V = Df(x)V + E$$

where $||E_i|| \leqq \alpha\varepsilon^2$ and $E_i$ is the $i^{th}$ column of $E$. Hence

$$||A_\sigma - Df(x)|| \leqq ||E||.||V^{-1}||$$

And, from Theorem 2.2, if $\rho$ is the radius of $\sigma$,

$$\leqq \frac{n\alpha}{n+1} \frac{\varepsilon^2}{\rho}$$

$$= \frac{n\alpha}{n+1} \frac{\varepsilon}{\theta}$$

and we have our result.

From Theorem 3.2 it is evident that if the thickness of a simplex $\sigma$ in K is very small (i.e., it is relatively long and skinny), for $x \in \sigma$, $A_\sigma$ may be a "poor" approximation of $Df(x)$. To demonstrate this, consider the function

$$f(x_1,x_2) = \begin{bmatrix} x_1 + x_1^2 + x_2^2 \\ \\ x_2 + x_1^2 + x_2^2 \end{bmatrix}$$

and, for some $\Delta > 0$, the simplex $\sigma = (v^1, v^2, v^3)$ whose vertices are

$v^1 = (0,0)$, $v^2 = (\Delta,0)$, $v^3 = (\Delta^2,\Delta^4)$

A simple calculation results in:

$$A_\sigma = \begin{bmatrix} 1 + \Delta & -\frac{1}{\Delta} + 1 + \Delta^4 \\ \\ \Delta & -\frac{1}{\Delta} + 2 + \Delta^4 \end{bmatrix}$$

and for $0 \in \sigma$, $Df(0) = I$ and $||I - A_\sigma|| = \frac{1}{\Delta} - 1 - \Delta^4$ and for $\Delta$ sufficiently small, $A_\sigma$ is a very poor approximation of I. For $\Delta = 0.20$, the simplex $\sigma$ is shown in Figure 3.1, and the thickness of $\sigma$ is less than $\frac{\Delta^3}{3}$.

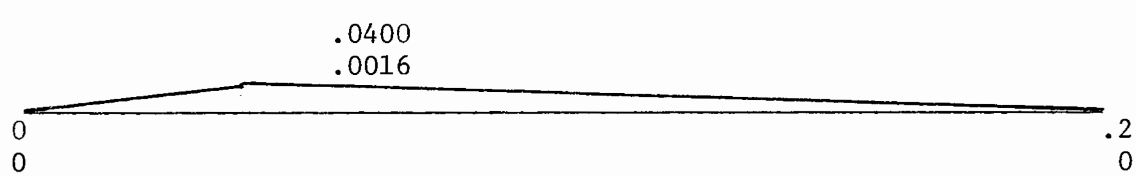



FIGURE 3.1

The above example indicates that there may be a fundamental problem in scaling the triangulations in the fixed point algorithms. In the next section we will show that even for arbitrarily scaled triangulations that are popularly used in fixed point computing, we get reasonable approximations to the derivatives. We note that arbitrary scaling can make the thickness of triangulations very small.

## 4. APPROXIMATIONS ON SCALED TRIANGULATIONS

In this section we consider scaled versions of the popular triangulations I and J3, [8], [9], [13], [14], used in fixed point computing. We now give a brief description of these triangulations.

The (unscaled) triangulation I of $R^n$ is generated as follows: given a positive number d, the vertices $I^0$ of this triangulation are all vectors in $R^n$ whose coordinates are integer multiples of d. Then, each simplex $\sigma$ in I has a unique representation $(v,\pi)$ where v is in $I^0$ and $\pi$ is a permutation of $\{1, . . ., n\}$. The vertices $v^1, . . .,$ $v^{n+1}$ are then generated as follows:

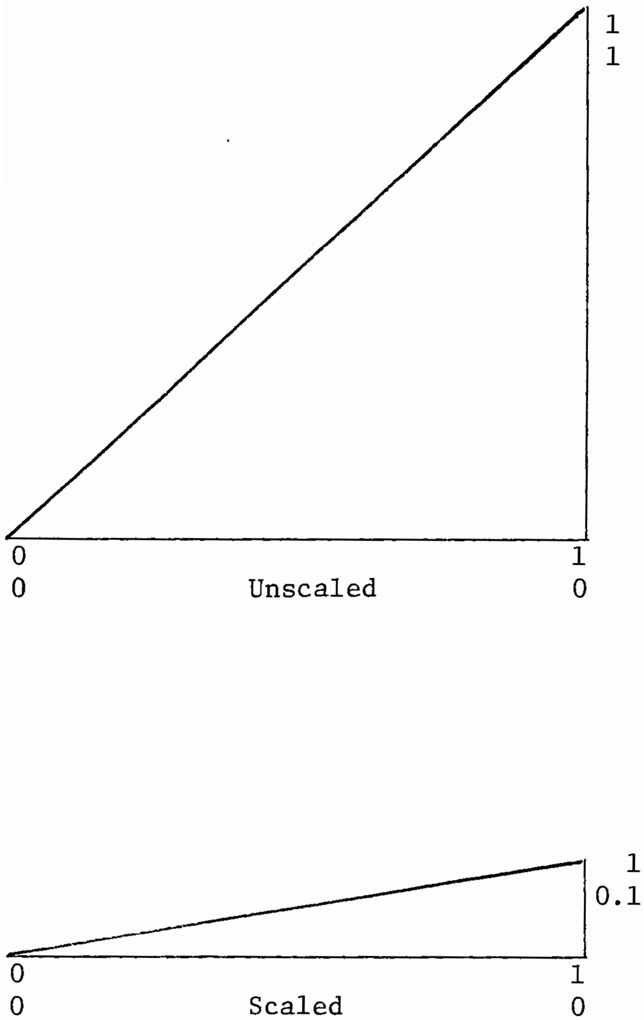
$$v^1 = v$$

$$v^{i+1} = v^i + du_{\pi(i)} \qquad i = 1, . . ., n$$

where $u_i$ is the $i^{th}$ unit vector in $R^n$.

A scaled version SI of triangulation I is generated by picking a diagonal matrix S whose diagonal entries $S_{ii}$ are the scale parameters. The vertices of SI are $SI^0 = \{Su: \; u \varepsilon I^0\}$ and the simplexes of SI are $S\sigma = \{Sx: x \varepsilon \sigma\}$ for some simplex $\sigma$ in I. Figure 4.1 shows a scaled and unscaled simplex in SI and I, respectively, where $S_{11} = 1$, $S_{22} = .1$ and $d = 1$. We note that the scaled simplex is long and skinny.

Another triangulation most frequently used in fixed point computing is J3. This is the usual one on which the Eaves-Saigal algorithm is implemented. This triangulation can be described as follows. Given an arbitrary positive number d, J3 triangulates $R^n \times (0,d]$, in such a manner that the vertices lie in the sections $R^n \times \{d.2^{-k}\}$ for k = 0, 1, 2, . . .. Also, if $v = (v_1, . . ., v_{n+1})$ in $R^n \times (0,d]$ is a vertex, $v_i/v_{n+1}$ are integers. For a vertex v, if $v_i/v_{n+1}$ is an odd integer, it is called a central vertex. The piecewise linear approximations to f are created on the induced subdivision of $R^n \times \{d.2^{-k}\}$. We shall study




$d = 1$, $S_{11} = 1$, $S_{22} = 0.1$, $v = (0,0)$, $\pi = (1,2)$

FIGURE 4.1

this induced subdivision called J1 for $k = 0$.

Now, if an n-simplex $\sigma = (v^1, \ldots, v^{n+1})$ lies in $R^n \times \{d\}$, then it can be uniquely expressed as a triplet $(v,\pi,s)$ where v is a central vertex of J3, $\pi$ a permutation of $\{1, \ldots, n\}$ and s an n-vector with $s_i \varepsilon \{-1,1\}$ for each $i = 1, \ldots, n$.

The vertices of $\sigma$ are then generated by

$v^1 = v$

$v^{i+1} = v^i + d.s_i.u_{\pi(i)}$

where, as before, $u_i$ is the $i^{th}$ unit vector.

As in I, a scaled version, SJ1 of J1 can be generated by choosing a diagonal matrix S with $S_{ii}$ the scale parameters. We now show how scaling affects the thickness:

<u>Lemma 4.1</u>: Let S be the scaling matrix, and let $\bar{s} = \max S_{ii}$ and $\underline{s} = \min S_{ii}$. Then the thickness of SI or SJ1 is bounded by $\frac{\underline{s}}{\bar{s}}$ .

<u>Proof</u>:

It is readily confirmed that the diameter of any simplex is at least $\bar{s}$, and that the radius is at most $\underline{s}$ and we have our result.

The Lemma 4.1 indicates that by picking the scaling parameters such that $\underline{s}/\bar{s}$ is very small, in view of Theorem 3.2, we may get very poor approximations to the derivative. We now show that this is not the case. The following lemma can be readily established.

<u>Lemma 4.2</u>: Let $\sigma = (v^1, v^2, \ldots, v^{n+1})$ be a simplex in SI or SJ1, and $V = (v^2-v^1, v^3-v^2, \ldots, v^{n+1}-v^n)$. Then $V = dQPS$ where P is a permutation matrix, S the scaling matrix, Q a diagonal matrix with $Q_{ii} \varepsilon \{-1, +1\}$ and $d$ the positive real number determining the grid of the triangulation.

We now prove our main theorem.

Theorem 4.3: Let $\sigma$ be a simplex in SI or SJ1, and let $d > 0$ and S (the scaling matrix) be arbitrary. Then, for $x \in \sigma$, $||A_\sigma - Df(x)|| \leqq \frac{3}{2}\alpha\sqrt{n}\,\hat{s}d$, where $\hat{s} = (S_{11}^{\ 2} + S_{22}^{\ 2} + \ldots + S_{nn}^{\ 2})^{\frac{1}{2}}$. Hence, for sufficiently small d, we have good approximations.

Proof: It can be readily confirmed that the grid of SI or SJ1 is sd. Now, for each i, applying Theorem 2.1,

$$f(v^{i+1}) = f(v^i) + Df(v^i)(v^{i+1} - v^i) + e(v^{i+1}, v^i)$$

or,

$$f(v^{i+1}) - f(v^i) = Df(x)(v^{i+1} - v^i) + (Df(v^i) - Df(x))(v^{i+1} - v^i) + e(v^{i+1}, v^i)$$

and letting $E_i = (Df(v^i) - Df(x))(v^{i+1} - v^i) + e(v^{i+1}, v^i)$, the $i^{th}$ column of E, we get

$$A_\sigma V = Df(x)V + E$$

Hence,

$$||A_\sigma - Df(x)|| \leqq ||EV^{-1}|| \leqq \frac{1}{d}\,||ES^{-1}||$$

But, if $\hat{E} = ES^{-1}$, $\hat{E}_i = \frac{1}{S_{ii}} E_i$. Hence,

$$||\hat{E}_i|| \leqq \frac{1}{S_{ii}}\,[d^2\,\alpha\hat{s}S_{ii} + \tfrac{1}{2}\alpha d^2 S_{ii}^{\ 2}]$$

$$\leqq \frac{3}{2}\,\alpha d^2\hat{s}.$$

Hence, $||\hat{E}|| \leqq \frac{3}{2}\sqrt{n}\;\alpha\, d^2\hat{s}$, and we have our result.

## 5. AN APPLICATION

In this section we will consider the problem of finding a zero of F, or, equivalently, finding an x such that $f(x) = 0$. In particular, we wish to study the application of the fixed point algorithms [2], [5], which generate a sequence of approximate solutions by creating piecewise linear approximations of f on triangulations of K. In such an application, zeros of piecewise linear approximations are found. Several important uses of these results have already been made in Saigal [9], Saigal and Todd [11]. We now establish some basic lemmas, and then prove the main result.

Proposition 5.1: Let A be a n x n matrix with $||A|| < 1$. Then $\det (I + A) > 0$.

Proof: Consider the matrix $A(t) = I + tA$, $t \in [0,1]$. Det $(A(0)) > 0$. Now, assume $\det (A(t_0)) = 0$ for some $t_0 \in [0,1]$. Then, for some $x_0 \neq 0$ $A(t_0)x_0 = 0$, or $x_0 + t_0Ax_0 = 0$ and so $-\frac{1}{t_0}$ is an eigenvalue of A. Hence $||A|| \geqq \frac{1}{t_0} \geqq 1$, which is a contradiction.

Proposition 5.2: Let $C = A - B$ and $||A^{-1}C|| < 1$. Then $\det (AB) > 0$. Also $(1-t)A + tB$ is non-singular for each t in [0,1].

Proof: Now, $\det (A^{-1}B) = \det (A^{-1}(A-C))$

$$= \det (I - A^{-1}C),$$

and we have the first result from Proposition 5.1. Also, since $(1-t)A + tB = A [I - tA^{-1}C]$ we have the second result using arguments similar to those of Proposition 5.1.

We now prove the main result in this section.

Theorem 5.3: Let x be such that Df(x) is non-singular. Let K be a triangulation of $R^n$ with grid $\varepsilon$ and thickness $\theta$. If $\frac{\alpha\beta\varepsilon}{\theta} < \frac{n+1}{n}$ where $||Df(x)^{-1}|| \leqq \beta$, then det $(A_\sigma Df(x)) > 0$ where $x \in \sigma$.

Proof: Follows readily from Proposition 5.2 and Theorem 3.2.

## 6. Acknowledgements

The author is grateful to Professor M.J. Todd for pointing out the upper bound of Theorem 2.2, which corrected an error in an earlier proof of Theorem 4.3.

## REFERENCES


[1] Charnes, A., C.B. Garcia and C.E. Lemke, "Constructive Proofs of Theorems relating to: F(x) = y, with Applications," Mathematical Programming, 12 (1977), 328-345.

[2] Eaves, B.C. and R. Saigal, "Homotopies for Computation of Fixed Points in Unbounded Regions," Mathematical Programming, 3 (1972), 225-37.

[3] Kojima, M., "Studies on PL Approximations of $C^1$-Mappings in Fixed Point and Complementarity Theory." To appear in Mathematics of Operations Research.

[4] Kuhn, H.W., "Simplicial Approximation of Fixed Points," Proceedings of the National Academy of Sciences, 61 (1968), 1238-42.

[5] Merrill, O.H., "Applications and Extensions of an Algorithm that Computes Fixed Points of Certain Upper Semi-continuous Points to Set Mappings," Tech. Report No. 71-7, Department of Industrial Engineering, University of Michigan, Ann Arbor, Michigan, 1971.

[6] Ortega, J.M. and Rheinboldt, W., Iterative Solutions of Nonlinear Equations in Several Variables, Academic Press, New York, 1970.

[7] Saigal, R., "Investigations into the Efficiency of the Fixed Point Algorithms," Fixed Points--Algorithms and Applications, S. Karamardian, Ed., Academic Press, New York, 1977.

[8] Saigal, R., "Fixed Point Computing Methods," Encyclopedia of Computer Science and Technology, Vol. 8, Marcel Dekker, Inc., New York, 1977.

[9] Saigal, R., "On the Convergence Rate of Algorithms for Solving Equations that are Based on Methods of Complementary Pivoting," Mathematics of Operations Research, 2 (1977), 108-24.

[10] Saigal, R., "On Paths Generated by Fixed Point Algorithms," Mathematics of Operations Research, 1 (1976), 359-80.

[11] Saigal, R. and M.J. Todd, "Efficient Acceleration Techniques for Fixed Point Algorithms," to appear in SIAM Journal on Numerical Analysis.

[12] Scarf, H., "The Approximation of Fixed Points of Continuous Mappings," SIAM Journal of Applied Mathematics, 15 (1967), 1328-48.

[13] Todd, M.J., "Union Jack Triangulations," Fixed Points--Algorithms and Applications, S. Karamardian, Ed., Academic Press, New York, 1977.

[14] Todd, M.J., The Computation of Fixed Points and Applications, Springer-Verlag, Berlin, 1976, 129 pp.

[15] Whitehead, J.H.C., "On $C^1$-Complexes," Annals of Mathematics, 41 (1940), 809-24.