AD-A043 089 CHICAGO UNIV ILL F/G 12/1

RELATIONS BETWEEN COMPLEMENTARY PIVOTING ALGORITHMS AND LOCAL A--ETC(U)

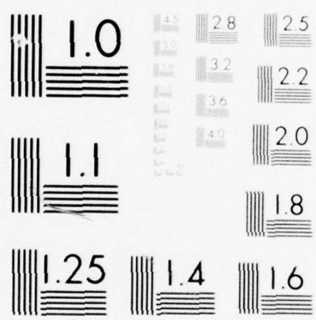
1977 C B GARCIA, F J GOULD N00014-75-C-0495

UNCLASSIFIED NL

1 OF 1

AD
A043 089

END
DATE
FILMED
9-77
DDC



MICROCOPY RESOLUTION TEST CHART
NATIONAL BUREAU OF STANDARDS-1963-A

NR 047-095
code 434

1

B. S.

6

RELATIONS BETWEEN COMPLEMENTARY PIVOTING ALGORITHMS AND LOCAL AND GLOBAL NEWTON METHODS,

by

11 1977

10

C. B. Garcia [redacted] F. J. Gould*
University of Chicago

12 36p.

ABSTRACT

This paper relates a simplicial pivot algorithm ~~of the authors~~ to Merrill's algorithm, Newton's method, and a "global Newton method" presented by Smale, Some computational results are given.

1. Introduction

This paper is expository in nature. The primary goal is to present recently observed relations between a simplicial pivot algorithm of the authors [10], Merrill's algorithm [19], Newton's method, and a "global Newton method" presented by Smale [25]. The underlying concept in Smale's algorithm is the solution of an ordinary differential equation of the form

$$F'(x(\theta))\dot{x}(\theta) = -\lambda(\theta)\, F(x(\theta)), \quad x(0) = x^0 \tag{1}$$

where $F: R^n \to R^n$ is continuously differentiable, $\lambda(\theta)$ is a real valued function of θ such that $\operatorname{sgn} \lambda = \operatorname{sgn} \det F'(x(\theta))$, x^0 is specified in R^n, $\dot{x}(\theta)$ denotes the derivative $\frac{dx}{d\theta}$, and $F'(x)$ is the matrix of partial derivatives

15 *The work of the second author was supported in part by ONR Grant No. N00014-75-C-0495 [redacted] NSF Grant No. ENG-76-81058

082450

LB

$$F'(x) = \left(\frac{\partial F_i}{\partial x_j}\right) . \tag{2}$$

Under suitable hypotheses on F, the solution of the differential equation will be a trajectory $x(\theta)$ which tends to a vector $(x_1^*, x_2^*, \ldots, x_n^*)$ satisfying

$$F(x^*) = 0 . \tag{3}$$

As examples, Smale shows that the method relates to the Arrow-Block-Hurwicz dynamics of price adjustment and Scarf's algorithm for finding economic equilibria.

The computation of the solution to (1) is normally achieved by discretization of the differential equation and then solving the discretized form by, for example, the method of continuation (see Kellog, Li, and Yorke [14] or Davidenko [2], and Meyer [20] for discussions related to this approach). This implementation could lead to problems in convergence due to, say, cycling or singularities of $F'(x)$.

In [10], we proposed a simplicial pivot algorithm that in a precise limiting sense follows points x which satisfy the equation

$$F(x) = tF(x^0) \quad \text{for some} \quad t \in [0, 1] \tag{4}$$

starting from a prescribed x^0 (with $t = 1$ and $F'(x^0)$ nonsingular). This method is an improvement of a method proposed in [8]. We shall show in a later section this simplicial pivot algorithm is closely related to Newton's method and, in fact, may be viewed as a globalized version of it. We shall show that the limiting sequence of points generated by this method will contain a solution to the ordinary differential equation (1) and, hence, the method, in part, may be viewed as a constructive procedure for finding the solution to Smale's differential equation.

The present paper serves three objectives. The first is to exposit relations between the simplicial pivot method in [10], Merrill's algorithm, Newton's method, and the method of continuation. The second objective is to exposit the relationship between the simplicial pivot method of [10] and the global Newton method of Smale. In conjunction with this exposition an example is presented which illustrates the power of starting at "infinity." And the last objective is to show some computational results.

We are indebted to Herbert Scarf for referring us to the paper [25].

2. On the Instability of Merrill's Algorithm Under a Transformation

We here consider an algorithm for solving $F(x) = 0$ which is based upon the concept of complementary pivoting on a labeled triangulation. The seminal work on complementary pivoting and simplicial approximation is due to Lemke [17], Lemke and Howson [18], and Scarf [23], [24]. For extensions see the papers [3], [4], [5], [6], [7], [9], [12], [13], [14], [16], [19], and the numerous additional references provided therein. We shall specifically consider an algorithm due to Merrill [19]. It was shown in [1] that this algorithm generates in a limiting sense, the solution to $t\,F(x) + (1 - t)(x - x^0) = 0$, $t \in [0, 1]$, $x \in R^n$. Equivalently, letting $\lambda = -\frac{(1 - t)}{t}$ this equation is written as

$$F(x) = \lambda(x - x^0), \quad \lambda \leq 0, \quad x \in R^n . \tag{5}$$

The algorithm is initiated at a point $x^0 \in R^n$ (when $t = 0$, or $\lambda = -\infty$). Here $F\colon R^n \to R^n$ is continuous, and λ is a real variable.

Merrill's algorithm may be used for finding zeroes of F, i.e., an $x \in R^n$ satisfying

$$F(x) = 0 . \tag{6}$$

Indeed, it is remarkable that the Leray-Schauder theorem below, which is ordinarily proved by arguments based on degree [22], can be shown constructively by letting $F(x) = x - G(x)$ and $x^0 = 0$ in Merrill's algorithm (see [1]):

> Leray-Schauder Theorem: Let C be an open bounded set in R^n containing the origin and $G\colon \overline{C} \to R^n$ a continuous mapping. If $G(x) \neq \lambda x$ whenever $\lambda > 1$ and $x \in \partial C$, then G has a fixed point.

The Brouwer Fixed Point Theorem, which states that any continuous function from a compact convex set to itself has a fixed point, may be easily seen to follow from the theorem above.

Although the Merrill algorithm is useful in proving classical and central theorems in fixed point theory by this new constructive approach, there are serious difficulties in implementing the approach for finding zeroes of F. (For example, the algorithm need not converge even for linear systems.) An analysis of system (5) shows that the method is unstable due to its dependence on the indexing of the components of F and the signs of the F_i. For example, consider the one-dimensional problem

$$F(x) = x^2 - w^2 = 0 \tag{7}$$

where $w > 0$ is a given constant. The zeros of F are $x = \pm w$. Merrill's algorithm, when started at x^0 to the right of $-w$, will move towards w, but when started to the left of $-w$, will move towards negative infinity. Replace $F(x) = 0$ with $-F(x) = 0$ and the situation is reversed: With $x^0 < w$, Merrill's algorithm will move towards the solution $x = -w$, but with $x^0 > w$, Merrill's algorithm diverges to infinity. Hence, convergence is affected when both sides of the equation are multiplied by -1, and the particular root obtained is also affected.

A way out of this difficulty is to select x^0 such that $\det F'(x^0) \neq 0$ and then to apply Merrill's algorithm to the function

$$G(x) = F'(x^0)^{-1} F(x) \tag{8}$$

as proposed by Fisher, Gould and Tolle [7]. This Modified Merrill algorithm finds solutions, in a limiting sense, to

$$\left.\begin{array}{l} t\,F(x) + (1-t)F'(x^0)(x - x^0) = 0, \quad t \in [0, 1], \quad x \in R^n, \\ \text{or, alternatively,} \\ F(x) = \lambda\, F'(x^0)(x - x^0), \quad \lambda \leq 0, \quad x \in R^n \end{array}\right\}. \tag{9}$$

This algorithm therefore is unaffected by a relabeling or change of signs of the components of F. Using (9) above, it is easy to see that for linear systems, $F(x) = Ax + b$, A nonsingular, the solution to (9) will be the line segment joining x^0 to the solution $-A^{-1}b$, for any x^0. For the problem (7), the Modified Merrill algorithm, when applied to $F(x) = 0$ or $-F(x) = 0$, yields the same solution, namely the solution for

$$(x^2 - w^2) = 2\lambda x^0(x - x^0), \quad \lambda \leq 0$$

so that for $x^0 < 0$, the method moves towards $x = -w$ and for $x^0 > 0$, the method moves towards $x = w$. Observe that the method is still dependent on the sign of x_i. In the simple case where $x^0 = 0$ and $F'(x^0) = I$, for any $\lambda < 0$, a point x such that $x_i = 0$ satisfies (9) if and only if $F_i(x) = 0$. That is, the algorithm can move from an orthant where $x > 0$ to an orthant where $x_i < 0$ only in a neighborhood of points x for which $F_i(x) = 0$ crosses the plane $x_i = 0$. Moreover, for any x satisfying (9), in the simple case with $x^0 = 0$ and $F'(x^0) = I$, $F_i(x)$ must always have a sign opposite x_i, whenever $F_i(x)$ and x_i are nonzero.

It has been shown that if $F'(x^*)$ is nonsingular, then the Modified Merrill method is locally convergent. The underlying reason for the local convergence behavior of the modified scheme is related to the fact that at the initial point x^0 the algorithm moves in the Newton direction.

To verify this initial direction, let us suppose F is continuously differentiable in an open set containing x^0, with $F'(x^0)$ nonsingular. Then note that (9) can be written as

$$H(x, t) \stackrel{\Delta}{=} t\,F(x) + (1 - t)F'(x^0)(x - x^0) = 0$$

where the function H is continuously differentiable on an open set in R^{n+1} containing the point $x = x^0$, $t = 0$. Now we note that the derivative of H with respect to x is

$$t\, F'(x) + (1 - t)F'(x^0)$$

which, when evaluated at x^0, equals $F'(x^0)$ (for any t). Then, by the assumption that $F'(x^0)$ is nonsingular, the Implicit Function Theorem guarantees the existence of an open set $A \subset R^1$ containing $t = 0$ and an open set $B \subset R^n$ containing x^0 such that for each $t \in A$ there is a unique differentiable function $x(t)$ with values in B such that $H(x(t), t) = 0$. Thus

$$t\, F(x(t)) + (1 - t)F'(x^0)(x(t) - x^0) \equiv 0, \quad \forall\, t \in A\,. \tag{10}$$

Now differentiating with respect to t, and using (9),

$$[t\, F'(x(t)) + (1 - t)F'(x^0)]\dot{x}(t) = -F(x(t)) + F'(x^0)(x(t) - x^0) = \frac{1}{t-1} F(x(t)) \tag{11}$$

Now set $x = x^0$ and $t = 0$ to obtain (noting $x(0) = x^0$)

$$F'(x^0)\dot{x}(0) = -F(x(0)), \quad \text{or} \quad \dot{x}(0) = -(F'(x^0))^{-1} F(x(0))\,.$$

Since the algorithm is initiated with t increasing from 0, $\dot{x}(0)$ is the initial tangent direction, and it is seen to be identical with the Newton direction. However, for values of $t \neq 0$ the tangent directions as given by (11) are not Newtonian. Without the nonsingularity assumption on $F'(x^0)$ the directions associated with the original Merrill algorithm applied to F (as opposed to G defined by (8)) are easily shown to satisfy

$$[t\, F'(x(t)) + (1 - t)I]\dot{x}(t) = \frac{1}{t - 1} F(x(t))\,,$$

and these directions are not Newtonian, even when $t = 0$. Thus, the modification of Merrill's algorithm at least initially forces the path to take a Newton direction.

In the next section an algorithm is demonstrated with the property that it moves in either the Newton direction or the negative of the Newton direction at each point. The attractiveness of this property will become clear in the sequel.

3. On a Constructive Global Newton Procedure

Let us now turn to the simplicial pivoting method proposed in [10] which employs a scalar labeling on a triangulation of R^n to follow points x which in a limiting sense satisfy

$$F(x) = t\, F(x^0) \quad \text{for some} \quad t \in [0, 1] \tag{12}$$

starting from a prescribed $x^0 \in R^n$ with $t = 1$, where $F\colon R^n \to R^n$ is continuous and $F'(x^0)$ exists and is nonsingular. The motivation for solving (12) can be explained briefly as follows. Consider the classical problem of how, given an initial $x^0 \in R^n$, one might find a solution x^* to $F(x) = 0$. Let us assume for simplicity (just for the moment) that the preimage of the line segment $[0, F(x^0)]$ is a path $x(t)$, $0 \leq t \leq 1$, such as when F is a homeomorphism on $F^{-1}[0, F(x^0)]$. If in addition $F'(x(t))$ exists and is nonsingular for $t \in [0, 1]$, then this preimage is a differentiable path $x(t)$, $0 \leq t \leq 1$. (See Meyer [20] for a more detailed discussion). Thus

$$F(x(t)) = t\, F(x^0)\ ,\ 0 \leq t \leq 1\ ,$$

where $x(t)$ is differentiable. We thus obtain

$$F'(x(t))\dot{x}(t) = F(x^0)\ ,\quad x(1) = x^0$$

or

$$\dot{x}(t) = F'(x(t))^{-1} F(x^0)$$
$$= F'(x(t))^{-1} \frac{F(x(t))}{t} \quad \text{for } t > 0 .$$

This indicates that at any point x on the path, at which t is decreasing, the algorithm moves in the Newton direction. In particular, since the algorithm is initiated at $t = 1$, with t decreasing, the initial direction of the algorithm is $-(F'(x^0))^{-1} F(x^0)$. This scheme is illustrated in Figure 1, where

$$x^1 = x^0 - \dot{x}(1) = x^0 - F'(x^0)^{-1} F(x^0) . \tag{13}$$

In our context, Newton's method may therefore be viewed as an attempt to follow the preimage of $[0, F(x^0)]$ by taking the next point x^1 to be in the direction of the tangent vector to the preimage curve at $t = 1$.

As is well known, a sequence of Newton iterations can be guaranteed to converge only when x^0 is close to x^* and $F'(x^*)$ is nonsingular. Continuation methods may be used to widen the domain of convergence of Newton's method. This approach has been known for the last century and has been rediscovered many times since. Essentially, the interval $[0, 1]$ is partitioned into subintervals

$$\left[\frac{k}{T}, \frac{k+1}{T}\right], \qquad k = 0, 1, \ldots T - 1 \tag{14}$$

so that for large positive integers T, the paths $x(t)$ on each subinterval are "small enough" so that convergence is obtained when Newton's method is applied on each subinterval.

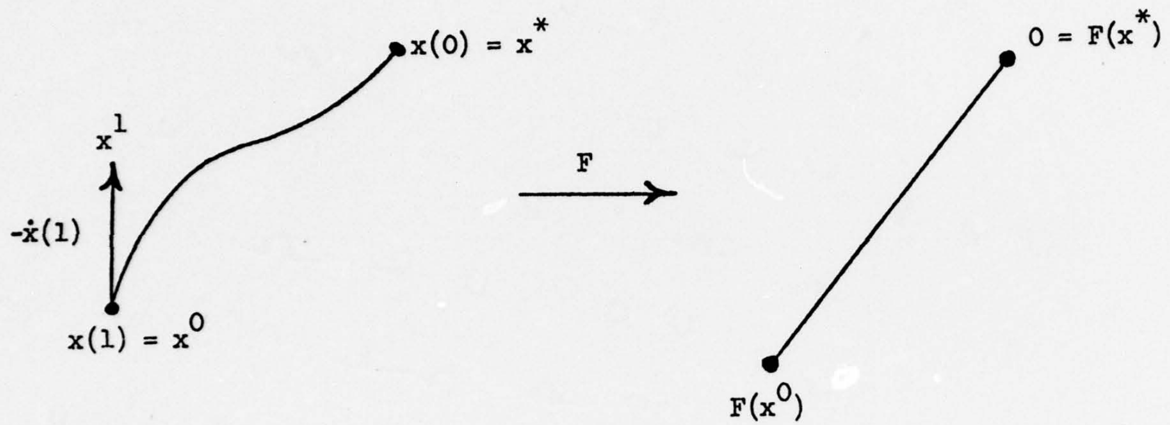


Figure 1

One known variation [22] of the method works as follows. Consider the values $t = 1, \frac{T-1}{T}, \frac{T-2}{T}, \ldots, 0$ and the associated path points $x(1) = x^0, x(\frac{T-1}{T}) = x^1, x(\frac{T-2}{T}) = x^2, \ldots, x(0) = x^T = x^*$. If each x^k is on the path (12), then

$$F(x^1) = \frac{T-1}{T} F(x^0),\ F(x^2) = \frac{T-2}{T} F(x^0),\ \ldots, \text{ and } F(x^k) = \frac{T-k}{T} F(x^0) \ .$$

Step 1. Start at x^0. Take the first Newton step for solving

$$F(x) - F(x^1) = 0 \ .$$

This produces a point

$$y^1 = x^0 - F'(x^0)^{-1} \ (F(x^0) - F(x^1)) = x^0 - F'(x^0)^{-1} \left(F(x^0) + \left(\frac{1-T}{T}\right) F(x^0)\right) \ .$$

Step 2. Start at y^1. Take the first Newton step for solving

$$F(x) - F(x^2) = 0 \ .$$

This produces a point

$$y^2 = y^1 - F'(y^1)^{-1} \ (F(y^1) - F(x^2)) = y^1 - F'(y^1)^{-1} \left(F(y^1) + \left(\frac{2-T}{T}\right) F(x^0)\right) .$$

$\vdots$

Step T. Start at y^{T-1}. Take the first Newton step for solving

$$F(x) - F(x^T) = 0, \text{ or, since } F(x^T) = 0, \ F(x) = 0 \ .$$

This produces a point

$$y^T = y^{T-1} - F'(y^{T-1})^{-1} F(y^{T-1}) \ .$$

Thus we have the following sequence of iterates, where $y^0 = x^0$,

$$y^{k+1} = y^k - F'(y^k)^{-1} \left(F(y^k) + \left(\frac{1+k-T}{T}\right) F(x^0)\right), \ k = 0, 1, \ldots, T-1 \ . \quad (15)$$

As Figure 2 indicates, at any "initial point" y^k the Newton step is the initial

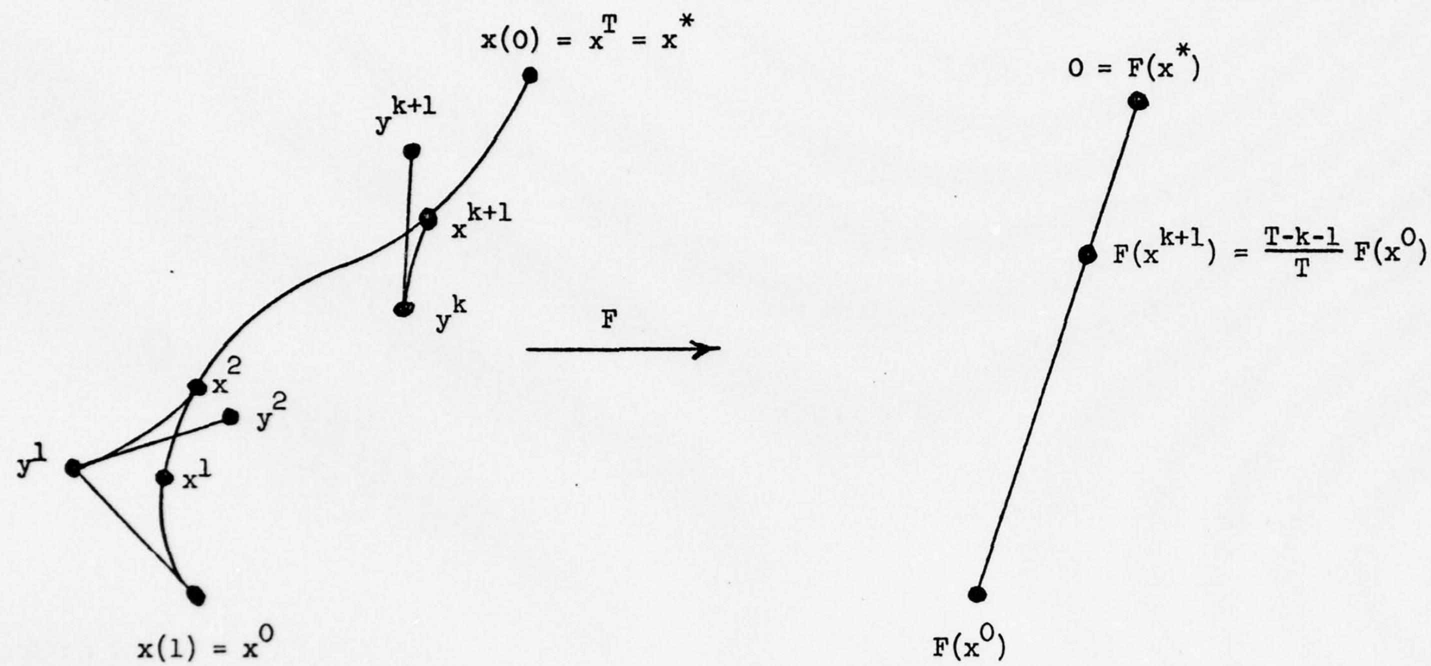


Figure 2

direction of the path

$$F(x(t)) = (1 - t)F(x^{k+1}) + t\, F(y^k), \quad 0 \leq t \leq 1 .$$

The objection to this approach is that in order to assure convergence to x^*, starting from x^0, nonsingularity of the Jacobian of F is required at every point of the path.

The purpose of [10] was to show that the preimage of $[0, F(x^0)]$ can be followed under less stringent conditions on F. It can be shown that:

<u>Theorem 1</u>. If $F: R^n \to R^n$ is continuously differentiable at x^0, $F'(x^0)$ is nonsingular, $F^{-1}[0, F(x^0)]$ is compact, and $F(x) = F(x^0)$ iff $x = x^0$, then there is an x such that $F(x) = 0$.

This theorem is constructively shown in [8]. The difficult part in the proof of the theorem above is to show that the scheme has a unique starting simplex at $t = 1$. But this is assured in [12, Proposition 3.5] which shows that for triangulations with mesh size small enough, the continuous simplicial mapping used to approximate F will also have a unique zero in a neighborhood of x^0, under certain "regularity" assumptions on F and the triangulation.

The scheme was improved in [10] by presenting a scalar labeling version of the method. The computational burden was improved to that of order n per pivot as opposed to an order $(n + 1)^2$ per pivot (see Table 1). A similar comparison can be made between the Merrill algorithm [19] and the scalar version introduced in [9].

Observe that the path generated by (12) and described in [10] is a substantial improvement to Newton's method in terms of global convergence. Newton's method could cycle, or it could "blow up" when the Jacobian at some point is singular. Scheme [10] on the other hand will succeed in generating

TABLE 1

AMOUNT OF ARITHMETIC OPTIONS PER PIVOT

	Algorithms [8] and [19]	Algorithms [9] and [10]
Work/Pivot	n fn evaluation (for F) $(n + 1)^2$ multiplications to get the labels $B^{-1}\begin{bmatrix} F \\ 1 \end{bmatrix} + (n + 1)^2$ $+ (n + 1)$ multiplications to update the linear system	n fn evaluation (for F) + n divisions to get $\frac{F_i(x)}{F_i(x^0)}$ or $\frac{F_i(x)}{x_i^0}$
Total Work/Pivot	n fn evaluations + $(n + 1)(2n + 3)$ multiplications	n fn evaluations + n divisions

an approximate solution even when these cases occur, so long as the conditions of Theorem 1 above are satisfied. In fact, the condition $F(x) = F(x^0)$ iff $x = x^0$ is also needlessly strong, and this assumption will be dispensed with in Theorems 7 and 8 of the next section.

Scheme [10] has several other properties of interest. The method is Jacobian invariant and norm-reducing. The former property is desirable since we would have convergence unaffected by, say, a relabeling of components of F. By the latter property we mean that

$$||F(x^k)|| < ||F(x^{k-1})|| \tag{16}$$

where x^k is the approximate solution at the $k^{\underline{th}}$ iteration. This norm-reducing property seems to be a property not shared by other simplicial pivot procedures.

There is another possibility that is furnished by scheme [10]. If x^k is a point on the $(n - 1)$-simplex generated after k pivots, then the vector $x^{\bar{k}} - x^k$, where $\bar{k}$ is a "few" iterates after k, is approximately the Newton direction (or the negative Newton direction). In fact

$$\begin{aligned} F(x^k) &\cong t_k F(x^0) \\ F(x^{\bar{k}}) &\cong t_{\bar{k}} F(x^0) \end{aligned} \tag{17}$$

for some $0 \le t_k, t_{\bar{k}} \le 1$ so that if $t_{\bar{k}} \ne t_k$ it can be shown that

$$\bar{x} = \frac{t_{\bar{k}} x^k - t_k x^{\bar{k}}}{t_{\bar{k}} - t_k} \tag{18}$$

will be an approximate Newton step. Thus, <u>without much additional</u> work, we can consider taking the Newton step if $||F(\bar{x})||$ is significantly smaller than either of $||F(x^k)||$ or $||F(x^{\bar{k}})||$, and restarting the scheme at $\bar{x}$, where $\bar{x}$ is defined by (18).

4. Relation to Smale's Global Newton Method

In this section we relate the set of solutions to

$$F(x) = tF(x^0), \quad x \in R^n, \quad t \in R \tag{19}$$

to the set of solutions to

$$F'(x(\theta))\dot{x}(\theta) = -\lambda(\theta)\, F(x(\theta)),\ x(0) = x^0 \quad \text{where } \operatorname{sgn} \lambda = \operatorname{sgn} \det F'(x(\theta)). \tag{20}$$

Equation (20) is the differential equation introduced by Smale in [25]. Here λ is some real function of θ, such as, for example, $\lambda = \det F'(x(\theta))$.

Let us make a quick observation on the solution to (20). Assume the solution is a one-dimensional curve $x(\theta)$, $0 \le \theta \le 1$ where $x(0) = x^0$ is some given starting point, and $x(1) = x^*$ is a zero of F. Then starting from x^0 the curve $x(\theta)$ is such that at any regular point x in the path (i.e., $\det F'(x) \neq 0$) the vector $\dot{x}(\theta)$, the tangent to the curve, is either the Newton direction (if λ is positive) or the negative Newton direction (if λ is negative). The change of direction occurs at points $x(\theta)$ where $\lambda(\theta)$ becomes zero and hence where the Jacobian becomes singular. Indeed, one should note that the solution trajectory for equation (20) is a globalization of Newton's method in the sense that the curve is "continued" when singularities of the Jacobian are encountered.

Now consider equation (19). This equation generalizes the simplicial pivot algorithms of [8, 10] in the sense that the parameter t is no longer restricted to the unit interval. We will show that for C^2 functions F with $F'(x^0)$ nonsingular a piece of the solution to (19) is the solution to (20). We dropped the condition $0 \le t \le 1$, which was assumed in [10], because this enables us to better deal with the case where the preimage of $F(x^0)$ is not unique, and, moreover, under appropriate conditions it allows us to continue the paths so as possibly to obtain all zeroes of F.

Let us consider a C^2 (twice continuously differentiable) function $H: R^{n+1} \to R^n$. Given $y \in R^n$, let

$$H^{-1}(y) = \{x \in R^{n+1} | H(x) = y\} \tag{21}$$

and

$$C = \{x \in R^{n+1} | \text{rank } H'(x) < n\} \tag{22}$$

where H' is the Jacobian matrix $\left(\frac{\partial H_i}{\partial x_j}\right)$ of H with respect to $x \in R^{n+1}$. The set C is said to be the set of <u>crticial points</u> of H, and $H(C)$ the set of <u>critical values</u>. $R^n \sim H(C)$ is the set of <u>regular values</u>. Sard's Theorem [26] states that:

> <u>Theorem 2</u>. Let $H: R^{n+1} \to R^n$ be a C^2 map. Then $H(C)$ has measure zero.

Thus, as a corollary, the set of regular values is dense in R^n. Let us henceforth throughout this paper assume that 0 is a regular value of H. The following lemma will be used [21]:

> <u>Lemma 3</u>. Let $H: R^{n+1} \to R^n$ be a C^2 map and let 0 be a regular value of H. Then $H^{-1}(0)$ is a C^1 one-dimensional manifold.

Now, recall that any connected C^1 one-dimensional manifold is diffeomorphic to a circle or an interval (open, closed, or half-open). Thus, each (connected) component of $H^{-1}(0)$ can be described by a curve $x(\theta)$ which is diffeomorphic to a circle or an interval. Furthermore, for any $x(\bar{\theta}) \in H^{-1}(0)$, we have

$$\text{rank } H'(x(\bar{\theta})) = n \tag{23}$$

and $\dot{x}(\bar{\theta})$ is a unique nonzero vector. Consequently, we can differentiate $H(x(\theta)) \equiv 0$ with respect to θ to obtain

$$H'(x(\theta)) \ \dot{x}(\theta) = 0 \ . \tag{24}$$

For a particular $\bar{\theta}$, $\dot{x}(\theta)$ is a vector tangent to the curve at $\theta = \bar{\theta}$ and spans the kernel of $H'(x(\theta))$.

For any $i = 1, 2, \ldots, n + 1$, let $\dot{x}_i(\theta)$ and $H_i(x(\theta))$ denote the i^{th} component of $\dot{x}(\theta)$ and the i^{th} column of $H'(x(\theta))$, respectively, and let $\dot{x}^i(\theta)$, $H^i(x(\theta))$ be the remaining components of $\dot{x}(\theta)$ and columns of $H'(x(\theta))$, respectively. The following theorem for C^2 maps is related to and motivated by a theorem of Eaves and Scarf for piecewise linear maps [4]. The proof of the theorem appears in [11].

<u>Theorem 4</u>. Let $H: R^{n+1} \to R^n$ be a C^2 map and 0 a regular value of H. Then for any component $x(\theta)$ of $H^{-1}(0)$ we have for all $i = 1, 2, \ldots, n + 1$:

$$\operatorname{sgn} \dot{x}_i(\theta) = \operatorname{sgn} \det H^i(x(\theta)) \quad \text{all } \theta \tag{25}$$

or

$$\operatorname{sgn} \dot{x}_i(\theta) = -\operatorname{sgn} \det H^i(x(\theta)) \quad \text{all } \theta \tag{26}$$

(where $\operatorname{sgn} 0 \triangleq 0$)..

Note that the theorem holds if H is restricted to say, $R^n \times [0, 1]$. In most applications to $R^n \times [0, 1]$ a further restriction would be required on the boundaries $R^n \times 0$ {0} and $R^n \times \{1\}$--namely, nonsingularity of the $n \times n$ submatrix $H_x(x, t)$ at points (x, t) in the boundary for which $H(x, t) = 0$. The condition assures that all loops which occur are contained in $R^n \times (0, 1)$. An interesting corollary to Theorem 4 is the following monotonicity theorem.

<u>Corollary 5</u>. Let $H: R^{n+1} \to R^n$ be a C^2 map and 0 a regular value of H. Suppose for some i, $H^i(x)$ is nonsingular for all x in a particular <u>component</u> $x(\theta)$ of $H^{-1}(0)$. Then, on that component of $H^{-1}(0)$, $x_i(\theta)$ is either monotone increasing or monotone decreasing as a function of θ.

Observe that under the assumptions of Corollary 5 the distinguished component of $H^{-1}(0)$ cannot be diffeomorphic to a circle.

Let us now turn to equation (19), where F is assumed to be a C^2 map with $F'(x^0)$ nonsingular. We associate with (19) the homotopy $H: R^{n+1} \to R^n$ defined by

$$H(x, t) = F(x) - tF(x^0) , \quad x \in R^n , \quad t \in R \tag{27}$$

which is of course also C^2. By Lemma 3, $H^{-1}(0)$ is a C^1 one-dimensional manifold, so that the component of $H^{-1}(0)$ containing the initial point $(x^0, 1)$ may be described by a curve

$$(x(\theta), t(\theta)) , \quad (x(0), t(0)) = (x^0, 1) .$$

Let us visualize the complementary pivoting scheme [10] to be tracking this curve in terms of increasing θ. Observe that equation (24) in this special case reduces to:

$$F'(x(\theta))\dot{x}(\theta) = \dot{t}(\theta)F(x^0) = \frac{\dot{t}(\theta)}{t(\theta)} F(x(\theta)) \tag{28}$$

if $t(\theta) \neq 0$. Thus if $\dot{t}(\theta) \neq 0$, it follows from Theorem 4 that $\det F'(x(\theta)) \neq 0$, so that

$$\dot{x}(\theta) = \frac{\dot{t}(\theta)}{t(\theta)} F'(x(\theta))^{-1} F(x(\theta)) . \tag{29}$$

An examination of (29) yields the following conclusions:

i) If $t(\theta)\dot{t}(\theta) > 0$, then scheme [10] is moving in the <u>negative Newton direction</u>.

ii) If $t(\theta)\dot{t}(\theta) < 0$, then scheme [10] is moving in the <u>Newton direction.</u>

Figure 3 below gives a graphical picture of this result.

As an illustration, let $(x(\theta), t(\theta))$ be the curve shown in Figure 3 where θ is increasing in the direction of the arrow, and $(x(0), t(0)) = (x^0, 1)$. Since $t(0)\dot{t}(0) < 0$, we are initially moving along the Newton direction. We continue movement along the Newton direction until we get to a point a, where $\dot{t}(\theta) = 0$. From point a to point b, since $\dot{t}(\theta)$ is now positive, and $t(\theta)$ is still positive, we are moving along the negative Newton direction. From b to c, with $\dot{t}(\theta)$ negative, we move along the Newton direction again. At point c, where we have a zero of F, we again reverse direction since $t(\theta)$ changes sign. We again reverse direction at points d and e.

Thus, the direction reverses if either $t(\theta)$ or $\dot{t}(\theta)$ changes in sign, but not both. The value $t(\theta)$ changes sign at a zero of $F(x)$. It follows from Theorem 4 that $\dot{t}(\theta)$ changes sign only when $\det F'(x(\theta)) = 0$.

We now use Theorem 4 to relate completely equations (19) and (20). In this special case, the theorem reduces to:

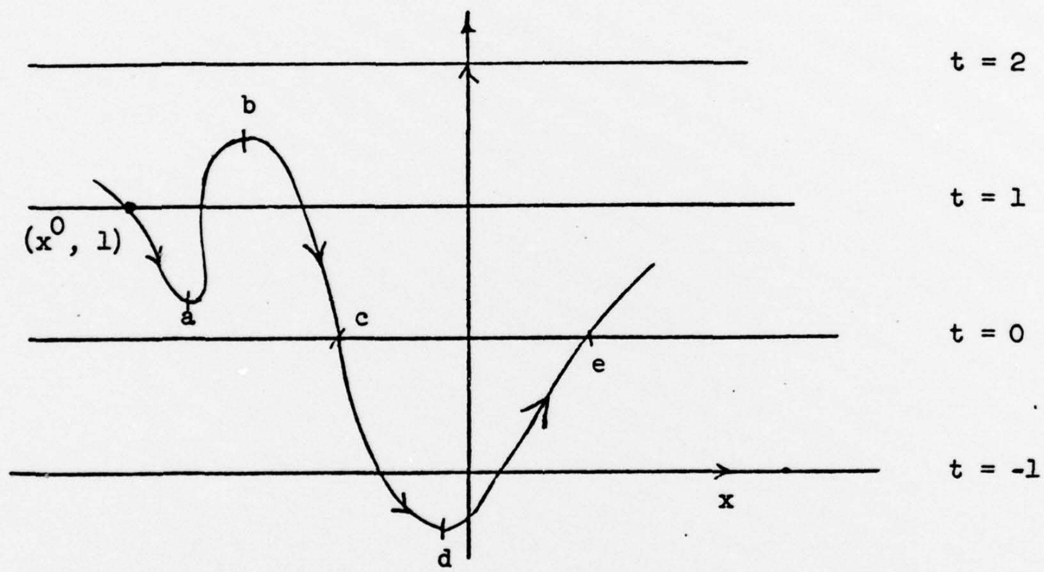


Figure 3

Theorem 6. $\text{sgn}\ \dot{t}(\theta) = \text{sgn}\ \det F'(x(\theta))$ all θ

or $\text{sgn}\ \dot{t}(\theta) = -\text{sgn}\ \det F'(x(\theta))$ all θ.

Observe that the differential equation (28) corresponding to (19) is equation (20) if $\text{sgn}\ \frac{\dot{t}(\theta)}{t(\theta)} = -\text{sgn}\ \det F'(x(\theta))$. To see that this equality in fact holds, recall that $t(0) = 1$, and suppose we initially move in a direction with t decreasing from 1 if $\det F'(x^0) > 0$. Otherwise, if $\det F'(x^0) < 0$, we let t initially increase from 1. (See Figure 3.) Then $\text{sgn}\ \frac{\dot{t}(0)}{t(0)} = -\text{sgn}\ \det F'(x^0)$. If we do indeed choose this initial direction to move, Theorem 9 tells us that $\text{sgn}\ \dot{t}(\theta) = -\text{sgn}\ \det F'(x(\theta))$ all θ. Thus $\text{sgn}\ \frac{\dot{t}(\theta)}{t(\theta)} = -\text{sgn}\ \det F'(x(\theta))$ for all θ where $t(\theta) > 0$, and for such values of θ our scheme, equation (19), is in fact a solution to Smale's equation (20). This actually proves that a piece of the solution to (19) is the solution to (20), for Smale's algorithm terminates the moment the first zero is encountered $(t(\theta) = 0)$. In Figure 3, the path from $(x^0, 1)$ to point c is the Smale path.

The reader should now relate the behavior of the path to the Jacobian $F'(x(\theta))$. If $\det F'(x^0) > 0$, then we move in the Newton direction for all θ where $t(\theta) \det F'(x(\theta)) > 0$ and in the negative Newton direction for $t(\theta) \det F'(x(\theta)) < 0$. The $\det F'(x(\theta))$ changes sign as $\dot{t}(\theta)$ changes. Thus in Figure 3, $\det F'(x(\theta)) > 0$ for points from $(x^0, 1)$ to a, and from b to d, and $\det F'(x(\theta)) < 0$ for points a to b, and from d to e. Finally, $\det F'(x(\theta)) = 0$ at points a, b and d.

In his paper [25] Smale presented a new and interesting convergence result. He considers an open bounded set C of R^n together with its boundary ∂C. Assume that C and ∂C are connected sets, and that ∂C is smooth. Suppose $F: \bar{C} \to R^n$ is a given C^2 function, and assume that F satisfies a boundary condition, e.g.:

<u>Boundary Condition.</u> For each $x \in \partial C$, $\det F'(x) > 0$ and $F'(x)^{-1}F(x)$ intersects ∂C transversally at x (i.e., $F'(x)^{-1}F(x)$ is not tangent to ∂C at x).

This implies that on ∂C the Newton direction is "everywhere pointing into C" or "everywhere pointing out of C." Smale then gives the following theorem, a different proof of which is here included for the sake of completeness.

<u>Theorem 7</u>. Let $F: \bar{C} \to R^n$ be a C^2 function where F and the open bounded $C \subseteq R^n$ satisfy the boundary condition, and where C and ∂C are connected with ∂C smooth. Then for $x^0 \in \partial C$ such that 0 is a regular value of H defined by (27), the (connected) component of $H^{-1}(0)$ containing $(x^0, 1)$ will contain a zero of F.

<u>Proof</u>: Since $F: \bar{C} \to R^n$ we have $H: \bar{C} \times R \to R^n$ where $H(x, t) = F(x) - tF(x^0)$.

Let

$$H^{-1}(0) \triangleq \{(x, t) \in \bar{C} \times R: H(x, t) = 0\} .$$

Since 0 is a regular value of H (i.e., $[H_x, H_t]$ has rank n $\forall (x, t) \in H^{-1}(0)$), and since $H^{-1}(0)$ is closed, the component of $H^{-1}(0)$ containing $(x^0, 1)$ may be described by a curve $(x(\theta), t(\theta))$, where $(x(0), t(0)) = (x^0, 1)$ and $\theta \in [0, 1]$. This component cannot be a loop. That is $(x(1), t(1)) \neq (x^0, 1)$. This conclusion follows from the fact that the path $(x(\theta), t(\theta))$ is C^1, ∂C is smooth, and hence if the path were a loop the transversality boundary condition would be violated. But if the path is not a loop then the endpoint $(x(1), t(1))$ must lie in $\partial(\bar{C} \times R)$. Otherwise, since 0 is a regular value of H, rank $[H_x(x(1), t(1)), H_t(x(1), t(1))] = n$, and by the Implicit Function Theorem $x(1), t(1)$ must be interior to $H^{-1}(0)$. Thus, the point $x(1)$ must lie in ∂C. We now show that this implies $t(1) < 0$ and, since $t(0) = 1$, there is a $\theta^* \in (0, 1)$ for which $t(\theta^*) = 0$ and hence $F(x(\theta^*)) = t(\theta^*) F(x^0) = 0$.

To see that $t(1) < 0$, recall the boundary assumption $\det F'(x) > 0 \; \forall \; x \in \partial C$. By Theorem 6, and the convention that $\dot{t}$ is initially negative (i.e., $\dot{t}(0) < 0$), we have $\dot{t}(\theta) < 0 \; \forall \; \theta$ such that $x(\theta) \in \partial C$. In particular, $\dot{t}(1) < 0$. The convention that $\dot{t}(0) < 0$ is tantamount to assuming that the Newton direction must point inward at each $x \in \partial C$, including, in particular, the point $x(1)$. This means that at $\theta = 1$ the tangent $\dot{x}(\theta)$ is in the negative Newton direction. This implies, by the remarks following (29), that:

$$t(1)\dot{t}(1) > 0 .$$

Since $\dot{t}(1) < 0$ it must also be true that $t(1) < 0$. #

Let us drop the condition that $\det F'(x) > 0$ for all $x \in \partial C$, and then add the regularity condition that:

$$F'(x) \text{ is nonsingular for each } x \in F^{-1}(0) . \tag{30}$$

Note that (30) guarantees the finiteness of zeroes of F in C. In fact we can show the parity of the number of zeroes of F in the component of $H^{-1}(0)$ containing $(x^0, 1)$.

Theorem 8. Let $F: \bar{C} \to R^n$ be smooth where $C \subset R^n$ is open and bounded, C and ∂C are connected. Then for $x^0 \in \partial C$ such that $F(x^0) \neq 0$ and 0 is a regular value of H defined by (27), the connected component $(x(\theta), t(\theta))$, $0 \leq \theta \leq 1$ of $H^{-1}(0)$ with $(x(0), t(0)) = (x^0, 1)$ will contain an odd (even) number of zeroes of F if and only if $t(1) < 0$ (> 0).

Proof: If the component of $H^{-1}(0)$ containing $(x^0, 1)$ is the singleton $(x^0, 1)$, the theorem holds trivially. Otherwise, as in the case of the last theorem, the component is a curve $(x(\theta), t(\theta))$ with $(x(0), t(0)) = (x^0, 1)$, $x(\theta) \in C$, $0 < \theta < 1$, and $x(1) \in \partial C$, where here we allow the possibility of a loop, i.e., $(x(0), t(0)) = (x^0, 1) = (x(1), t(1))$. Now let $0 < \bar{\theta} < 1$ be such that $F(x(\bar{\theta})) = 0$ (so that $t(\bar{\theta}) = 0$). Since $F'(x(\bar{\theta}))$ is nonsingular by assumption (30), it follows from Theorem 6 that $\dot{t}(\bar{\theta}) \neq 0$. Hence there is an $\varepsilon > 0$ such that $t(\bar{\theta} - \alpha) \cdot t(\bar{\theta} + \alpha) < 0$ for $0 < \alpha < \varepsilon$. Thus $t(\theta)$ changes sign on some interval if and only if a zero of F is encountered on the interval. Hence, if $t(1) < 0$, since $t(0) > 0$, the sign of $t(\theta)$ must have changed signs an odd number of times implying an odd number of zeroes of F on the curve. Otherwise, if $t(1) > 0$, the sign of $t(\theta)$ changed signs an even (possibly zero) number of times implying an even number of zeroes of F on the curve.

5. Examples

Example 1: Comparison of Global Newton and Modified Merrill

We recall that scheme [10] requires $o(n)$ arithmetic operations per pivot while the Modified Merrill vector labeling algorithm requires $o(n + 1)^2$ operations per pivot. Furthermore, the two schemes follow quite different paths but initially start in the same direction. An example is shown in Figure 4. Here the Modified Merrill diverges to infinity while scheme [10] converges to a solution. Observe the dependence of the former algorithm on the sign of x_i. In fact a hyperplane $x_i = 0$ cannot be crossed by the limiting path of the Modified Merrill algorithm except at a point on the hyperplane where $F_i(x) = 0$ also. The scheme [10] requires $t \in [0, 1]$ so that the solution is bounded by the intersection of sets $0 \leq F_i(x) \leq F_i(x^0)$ and is not dependent on the signs of x_i. Observe too that the solution to the more general equation

$$F(x) = tF(x^0), \quad t \text{ real}$$

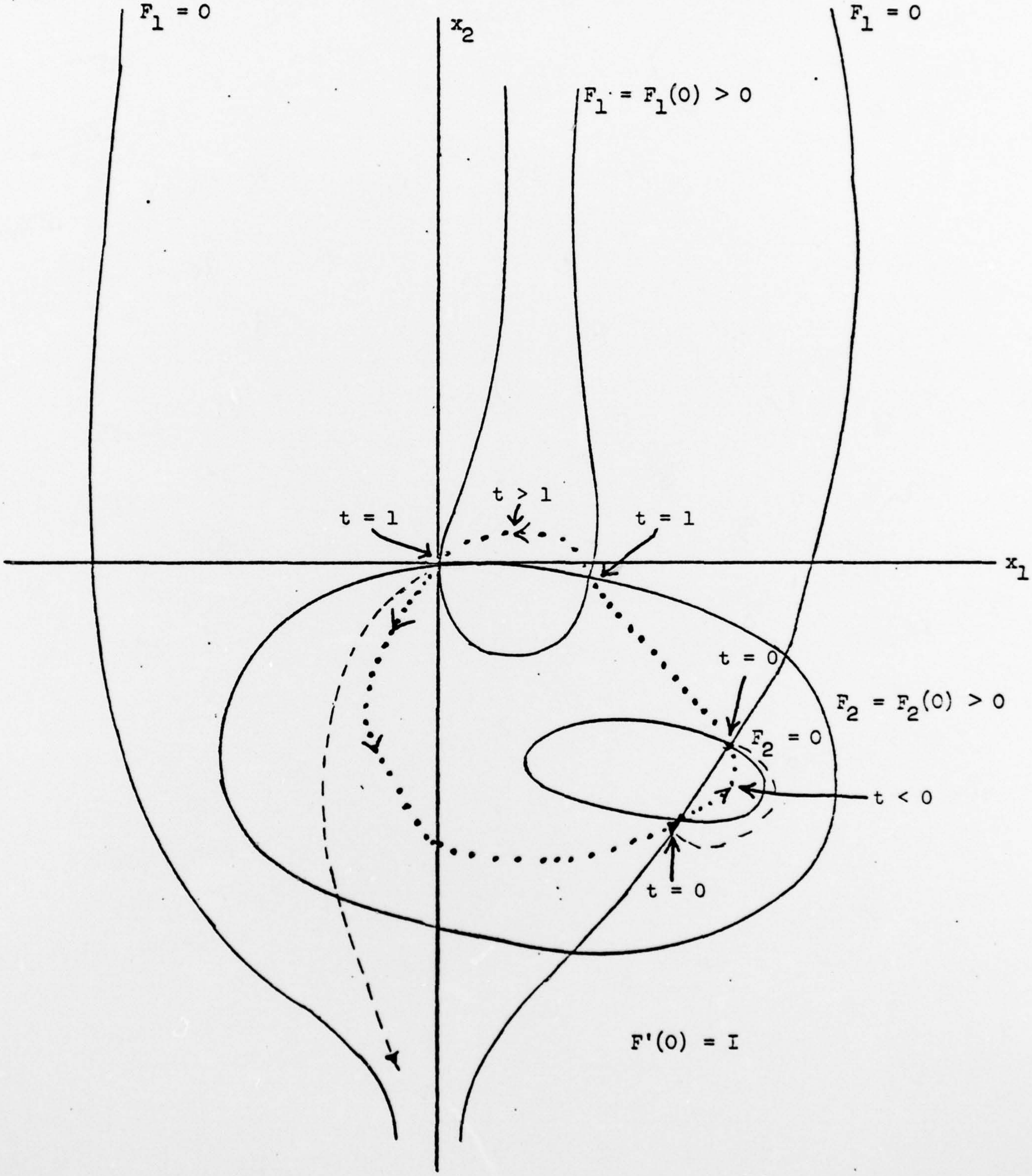
is a loop that connects all preimages of $F(0)$ and all zeros, so that, if the algorithm is started at the origin, it will find all zeros of F.

Example 2: Starting at Infinity

Figure 5 raises the interesting possibility that for certain types of functions convergence of G^2 may be assured by starting sufficiently far away from a zero. Concerning this figure, we make the following observations:

(i) G^2 fails when initiated from w (a loop is generated).

(ii) G^2 succeeds when initiated far away, e.g., from x^0, and moreover all zeros are found.



The Modified Merrill follows the dashed line. Scheme [10] follows the dotted line.

Figure 4

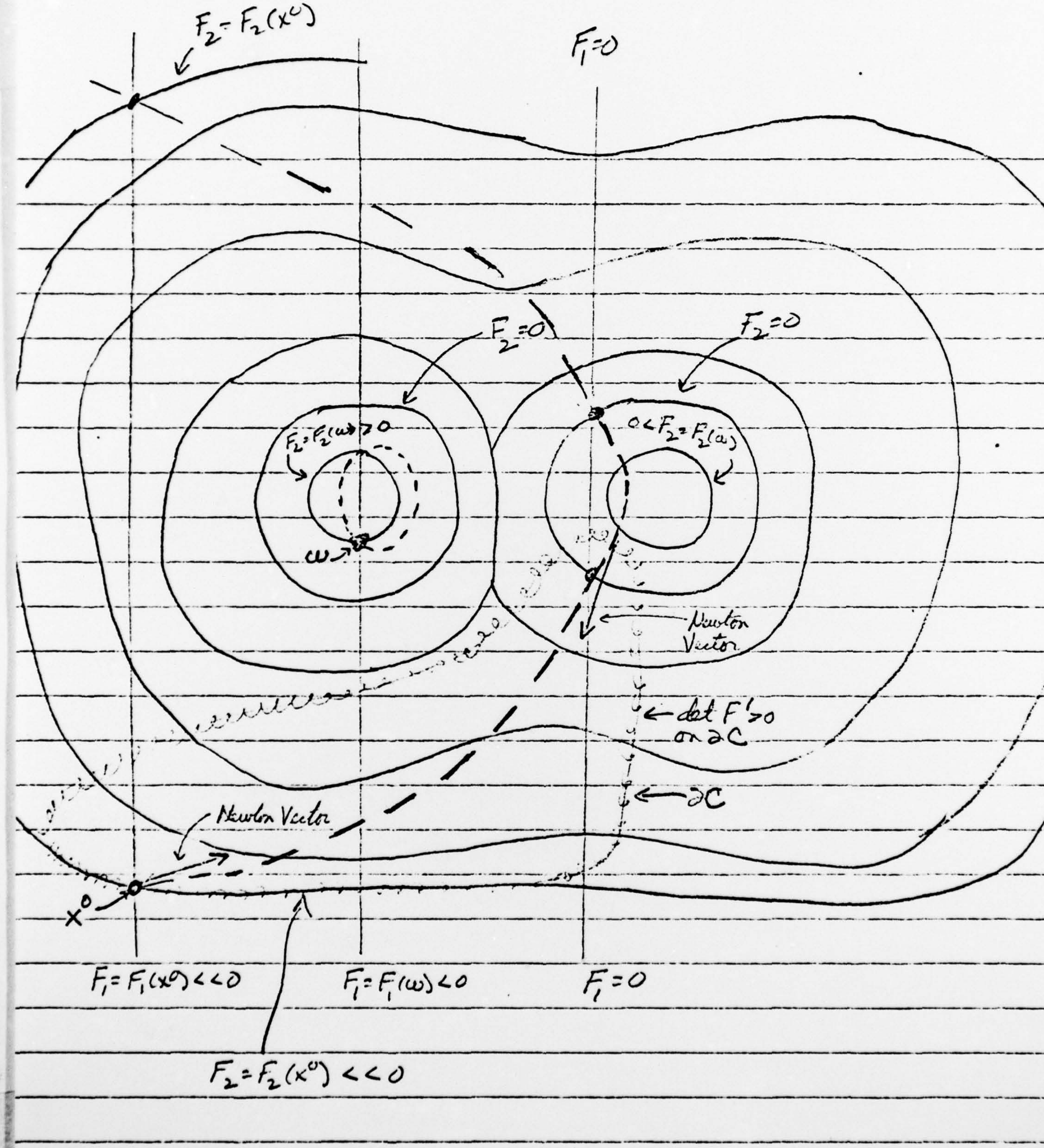

$F_2 = F_2(x^0)$
$F_1 = 0$
$F_2 = 0$
$F_2 = 0$
$F_2 = F_2(\omega) > 0$
$0 < F_2 = F_2(\omega)$
ω
Newton Vector
$\det F' > 0$ on ∂C
∂C
Newton Vector
x^0
$F_1 = F_1(x^0) << 0$
$F_1 = F_1(\omega) < 0$
$F_1 = 0$
$F_2 = F_2(x^0) << 0$

Figure 5

(iii) The set C with the curly border satisfies the hypotheses of Smale (Theorem 7). In particular $\det F'(x) > 0$ for all $x \in \partial C$ and the Newton direction points into C at each $x \in \partial C$.

6. Empirical Comparisons

We have compared the Modified Merrill vector labeling method with G^2, the scalar labeling method from [10], on three nonlinear problems. All tests were done on an HP 2000 Time-Sharing Basic.

The first example is to solve:

$$F_1(x_1, x_2) = x_2^3 - x_1 - 1{,}000 = 0$$

$$F_2(x_1, x_2) = x_1^3 - x_2 - 1{,}000 = 0 \quad .$$

An approximate solution to the problem is $x_1 = x_2 = 10.0333$ with $||F(x)||_\infty = 1.29 \times 10^{-4}$. Here Merrill's algorithm (without modification) did not converge using a mesh size of one unit and starting point $x = (20, 20)$ (after 351 iterations, the method generated the point $x = (-66, 77)$ and the simplicial path is diverging to infinity). Both the Modified Merrill and scheme [10] traverse the line segment joining the starting point and the solution.

TABLE 2 (Number of Pivots)

Starting point $x^0 = (20, 20)$

Mesh size	.01	.1	1	5
Modified Merrill	4,151*	699	69	13
Scheme [10]	3,985	397	37	5

*The Modified Merrill was stopped after the indicated number of simplicial pivots without reaching the terminal simplex.

The result of the test is shown in Table 2. In every case, scheme [10] took fewer pivots than the Modified Merrill to reach the terminal simplex. This is explained by the fact that Modified Merrill operates in R^{n+1} while G^2 operates in R^n. Moreover, the M^2 path (projected onto R^n) will generally differ from the G^2 path, as Example 1 shows. This also will cause the two algorithms to generate different numbers of total pivots. Due to the significant difference in the computational burden of the Modified Merrill ($o(n + 1)^2$ operations per pivot) over scheme [10], ($o(n)$ operations per pivot), the total computational work in the latter scheme, given approximately the same number of pivots, is significantly less than the former scheme.

It is interesting to note that for mesh size equal to ϵ, scheme [10] takes about $4 \cdot \frac{20 - 10.03}{\epsilon}$ pivots to reach the terminal simplex, which is roughly the number of pivots required to reach the terminal simplex in the quickest way (see Figure 6).

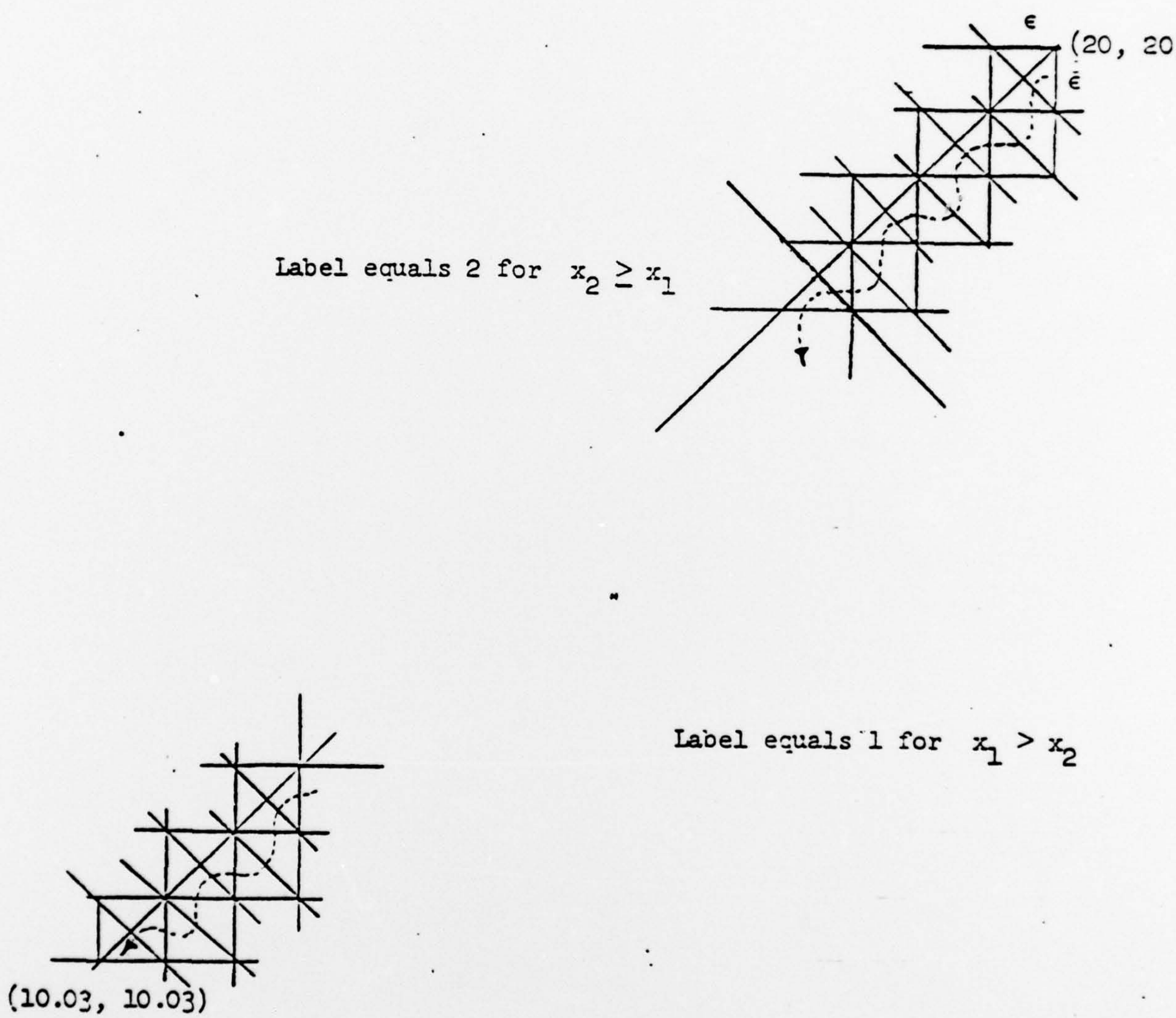

ε
(20, 20)
ε
Label equals 2 for $x_2 \geq x_1$
Label equals 1 for $x_1 > x_2$
(10.03, 10.03)

Figure 6

We ran both schemes on the same problem but using as initial point $x = (20, 30)$. The result of the test is exhibited in Table 3

TABLE 3 (Number of Pivots)

Starting point $x^0 = (20, 30)$

Mesh size	.1	1	5	10
Modified Merrill	1,099	109	21	10
Scheme [10]	597	51	9	3

The second problem tested is:

$$F_1(x_1, x_2) = 2x_1 + \sin x_1 - x_2 - 101 = 0$$

$$F_2(x_1, x_2) = 2x_2 + \sin x_2 - x_1 - 102 = 0 \quad .$$

It was shown in [10] that in a 5-dimensional problem of this form both Newton's method and the continuation method diverged. An approximate solution to this problem is $x = (100.899, 101.157)$ with $||F(x)||_\infty = 2.4 \times 10^{-5}$. Again, as shown in Table 4, scheme [10] is superior to the Modified Merrill in this example.

TABLE 4 (Number of Pivots)

Starting point $x^0 = (20, 20)$

Mesh size	1	5	10
Modified Merrill	569	115	59
Scheme [10]	323	51	33

The third and last problem tested is:

$$F_i(x) = x_i^3 - \sum_{j \neq i} x_j - 1{,}000 = 0 \qquad i = 1, 2, \ldots, n .$$

We ran both schemes on the above for $n = 5$, 10 and $x_i^0 = 20$ all i. The results are shown in Tables 5 and 6.

TABLE 5 (Number of Pivots)

$n = 5$ Starting point $x_i^0 = 20$ $i = 1, 2, \ldots, 5$

Mesh size	1	5	10
Modified Merrill	496	96	46
Scheme [10]	369	73	24

TABLE 6 (Number of Pivots)

$n = 10$ Starting point $x_i^0 = 20$ all i

Mesh size	5	10
Modified Merrill	541	266
Scheme [10]	453	204

REFERENCES

[1] Charnes, A.; Garcia, C. B.; and Lemke, C. E. "Constructive Proofs of Theorems Relating to $F(x) = y$, with Applications," to appear in Mathematical Programming.

[2] Davidenko, D. "On the Approximate Solution of a System of Nonlinear Equations," Ukrain. Mat. ž., 5 (1953), 196-206.

[3] Eaves, B. C. "Homotopies for Computation of Fixed Points," Mathematical Programming 3, 1 (1972), 1-22.

[4] _______, and Scarf, H. "The Solution of Systems of Piecewise Linear Equations," Mathematics of Operations Research 1, 1 (1976), 1-27.

[5] Fisher, M. L., and Gould, F. J. "A Simplicial Algorithm for the Nonlinear Complementarity Problem," Mathematical Programming 6 (1974), 281-300.

[6] _______; _______; and Tolle, J. W. "A New Simplicial Approximation Algorithm with Restarts: Relations between Convergence and Labelling," Fixed Points Algorithms and Applications, edited by S. Karamardian in collaboration with C. B. Garcia (Academic Press, 1977), 41-58.

[7] _______; _______; and _______. "A Simplicial Approximation Algorithm for Solving Systems of Nonlinear Equations," Instituto Nazionale Di Alta Matematica: Symposia Mathematica 19 (1976), 73-90.

[8] Garcia, C. B. "A Global Existence Theorem for the Equation $Fx = y$," Center for Mathematical Studies in Business and Economics Report No. 7527, University of Chicago, June 1975.

[9] _______, and Gould, F. J. "An Algorithm Based on the Equivalence of Vector and Scalar Labels." Submitted to Mathematical Programming.

[10] _______, and _______. "An Improved Scalar Generated Homotopy Path for Solving $f(x) = 0$," submitted to Mathematical Programming.

[11] _______, and _______. "A Theorem on Homotopy Paths," Center for Mathematical Studies in Business and Economics Report No. 7734, University of Chicago, June 1977.

[12] Garcia, C. B., and Zangwill, W. I. "Determining All Solutions to Certain Systems of Nonlinear Equations," submitted to Mathematics of Operations Research.

[13] Gould, F. J., and Tolle, J. W. "A Unified Approach to Complementarity in Optimization," Discrete Mathematics 7 (1974), 225-271.

[14] _______, and _______. "An Existence Theorem for Solutions to $f(x) = 0$," Mathematical Programming 11 (1976), 252-262.

[15] Kellogg, R. B.; Li, T. Y.; and Yorke, J. "A Method of Continuation for Calculating a Brouwer Fixed Point," Fixed Points Algorithms and Applications, edited by S. Karamardian in collaboration with C. B. Garcia (Academic Press, 1977), 133-148.

[16] Kuhn, H. W., and MacKinnon, J. G. "Sandwich Method for Finding Fixed Points," J. Opt. Theory and Appl. 17 (1975), 189-204.

[17] Lemke, C. E. "Bimatrix Equilibrium Points and Mathematical Programming," Management Science 11 (1965), 681-89.

[18] _______, and Howson, J. T., Jr. "Equilibrium Points of Bimatrix Games," Society of Industrial and Applied Math 12 (1964), 413-23.

[19] Merrill, O. H. "Applications and Extensions of an Algorithm That Computes Fixed Points of Certain Upper Semi-Continuous Point-to-Set Mappings" (University of Michigan: Ph.D. dissertation, 1972).

[20] Meyer, G. H. "On Solving Nonlinear Equations With a One-Parameter Operator Imbedding," Siam J. Numer. Anal., Vol. 5, No. 4, December 1968, pp. 739 to 752.

[21] Milnor, J. Topology From the Differentiable Viewpoint (University Press of Virginia, 1965).

[22] Ortega, J. M., and Rheinboldt, W. C. Iterative Solutions of Nonlinear Equations in Several Variables (Academic Press, 1970).

[23] Scarf, H. E. "The Approximation of Fixed Points of Continuous Mappings," Society of Industrial and Applied Mathematics 15 (1967), 1328-43.

[24] _________, The Computation of Economic Equilibria (In collaboration with T. Hansen), Yale University Press, New Haven 1973.

[25] Smale, S. "Convergent Process of Price Adjustment and Global Newton Methods," Cowles Foundation Report, Yale University (1975).

[26] Sternberg, S Lectures on Differential Geometry, Prentice-Hall, 1964.