SIAM J. NUMER. ANAL.
Vol. 13, No. 4, September 1976

A CONSTRUCTIVE PROOF OF THE BROUWER FIXED-POINT THEOREM AND COMPUTATIONAL RESULTS*

R. B. KELLOGG†, T. Y. LI‡ AND J. YORKE¶

Abstract. A constructive proof of the Brouwer fixed-point theorem is given, which leads to an algorithm for finding the fixed point. Some properties of the algorithm and some numerical results are also presented.

1. Introduction. Let D be an open bounded convex set in R^n, and let $F:\bar{D} \to \bar{D}$ be continuous. The Brouwer fixed-point theorem guarantees the existence of a fixed point, a point x such that $x = F(x)$. In this paper, we give a constructive proof of the Brouwer fixed-point theorem, establish the most important properties of the method, and describe the results of our computational experience. A preliminary announcement of the results with a description of the algorithm, is contained in [6].

Several proofs of the Brouwer fixed-point theorem are based on its equivalence with the no retraction theorem; that is, the nonexistence of a continuous map $H:\bar{D} \to \partial D$ such that $H(x) = x$ for all $x \in \partial D$. (See, e.g., [8].) One of the origins of our method lies in a paper of Pontryagin [12] which contains a study of the existence of smooth essential maps $h: S^n \to S^{n-1}$, where S^n is the n-sphere. The results and techniques of Pontryagin are much more complicated than we require, but the setting of his study is quite similar. (See [1] for an application of framed bordism to differential equations.) Pontryagin used Sard's theorem and its corollary fact that for a.e. $p \in S^{n-1}$, $h^{-1}(p)$ is the finite union of nonintersecting closed curves. Applying these ideas in the special setting of the existence of a retraction $H:\bar{D} \to \partial D$, a contradiction is obtained almost immediately, as was shown by Hirsch [5]. (See also Milnor [10] for a detailed exposition of Sard's theorem and Hirsch's proof.) Our contribution consists in noting that, even with fixed points, there is a retraction H from a part of D to ∂D, and for a.e. $x^0 \in \partial D$, the set $H^{-1}(x^0)$ contains a curve leading from x^0 to the set of fixed points. Furthermore, a study of the curve reveals a close connection with Newton's method, so our results may be interpreted as giving a kind of global convergence theorem for Newton's method.

The first general computational algorithm for finding fixed points was proposed by Scarf [15] and has since been improved and used by many authors [3], [7]. This method employs a simplicial decomposition of D and uses a systematic search technique to find a simplex of the decomposition which, hopefully, contains or is near a fixed point. The more recent versions of the algorithm employ

* Received by the editors February 13, 1975, and in revised form September 12, 1975.

† Institute for Fluid Dynamics and Applied Mathematics, University of Maryland, College Park, Maryland 20742. The work of this author was supported in part by the National Science Foundation under Grant GP-20555.

‡ Mathematics Department, University of Utah, Salt Lake City, Utah 84112. The work of this author was supported in part by the National Science Foundation under Grant GP-31386X2 and MPS.

¶ Institute for Fluid Dynamics and Applied Mathematics, University of Maryland, College Park, Maryland 20742. The work of this author was supported in part by the National Science Foundation under Grant GP-31386X2 and MPS.

an infinite complex on a semi-infinite cylinder. The method has the advantage of being formulated for functions F which are continuous but not necessarily differentiable. On the other hand, our method has advantages of conceptual simplicity, of not requiring a simplicial decomposition, and of bearing a close relationship with Newton's method. The method has been used to solve problems with dimension up to $n = 60$. Some of our computational experience is described in § 5.

2. The basic method. Let D be a convex bounded open subset of R^n, and let $F : \bar{D} \to D$ be continuous. We assume that F is twice continuously differentiable in D, and that the first and second derivatives of F have continuous extensions to $\bar{D}$. We denote by F' the matrix of first derivatives of F. Let $C = \{x \in \bar{D} : F(x) = x\}$ be the set of fixed points of F. We define a map $H : \bar{D} \setminus C \to \partial D$ as follows. For $x \in \bar{D} \setminus C$, we denote by $L(x)$ the ray from $F(x)$ passing through x, and we let $H(x)$ denote the intersection of this ray with ∂D. (See the Fig. 1.) Since $F(x) \neq x$, the ray is well-defined. It is easy to see that H is continuous on $\bar{D} \setminus C$.

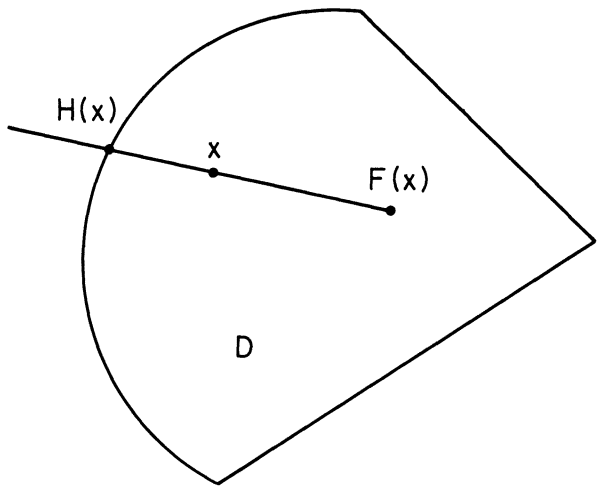

FIG. 1

We first describe the set $H^{-1}(x^0)$, for $x^0 \in \partial D$. For this we require some regularity of ∂D. We do this by introducing an open set U with the property that $U \cap \partial D$ is a C^2 surface.

THEOREM 2.1. *For a.e. $x^0 \in U \cap \partial D$ (a.e. in the sense of $(n - 1)$-dimensional measure), the set $H^{-1}(x^0) \cap D$ consists of a number of connected components, each one of which is a C^1 diffeomorphic image of a circle or open interval.*

Proof. By extending F, we may assume that it is defined and twice continuously differentiable on an open set $D_1 \supset \bar{D}$. Furthermore, we may pick D_1 so that $F(x) \in D$ for $x \in D_1 \cap U$. In the open set $D \cup (D_1 \cap U) \setminus C$, we define $H(x)$ to be the intersection of the ray $L(x)$ with ∂D. For $H(x) \in U \cap \partial D$, $H(x)$ is the intersection of a ray and an $(n - 1)$-dimensional surface. Hence $H(x)$ is twice continuously differentiable for $H(x) \in U \cap \partial D$. Since ∂D is an $(n - 1)$-dimensional surface, $H'(x)$ is a singular matrix. By Sard's theorem [4] there is a set $Z \subset \partial D \cap U$, having $(n - 1)$-dimensional measure $= 0$, and such that for $y \in (\partial D \cap U) \setminus Z$, and $H(x) =$

y, $H'(x)$ has rank $n-1$. Fix a point $x^0 \in \partial D \cap U$, $x^0 \notin Z$. Then for $x \in H^{-1}(x^0)$, $H'(x)$ has rank $n-1$, and so has at least one nonzero $(n-1)$st order minor. Hence there is an open set $V \supset H^{-1}(x^0)$ such that for $x \in V$, $H'(x)$ has rank $n-1$. We shall show that there is an open set $V \supset H^{-1}(x^0)$ and a continuous vector function $u(x) \neq 0$, $x \in V$, such that $H'(x)u(x) = 0$, $x \in V$. For this, let $N(y)$ denote the outward pointing unit normal to $y \in \partial D$. Evidently, we may find an index i and an open set W of x^0 such that for $y \in W \cap \partial D$, $N_i(y) \neq 0$. Since $H^{-1}(W \cap \partial D)$ is an open set, we may suppose that the open set $V \supset H^{-1}(x^0)$ has been chosen so that for some i, $N_i(H(x)) \neq 0$, $x \in V$. Now for any vector v, $H'(x)v$ is tangent to ∂D at $H(x)$. Hence

$$H'(x)^T N(H(x)) = 0. \tag{2.1}$$

We let $r^j(x)$ be the jth row of $H'(x)$. Then from (2.1) we see that for $x \in V$, $r^i(x)$ is a linear combination of the remaining $n-1$ rows, so the rows $r^1(x), \cdots, r^{i-1}(x)$, $r^{i+1}(x), \cdots, r^n(x)$ are linearly independent. Setting $r^j(x) = [r^j_1(x), \cdots, r^j_n(x)]$, we define $(n-1)$-dimensional column vectors $a^k(x)$, $1 \leqq k \leqq n$ by

$$a^k(x) = [r^1_k(x), \cdots, r^{i-1}_k(x), r^{i+1}_k(x), \cdots, r^n_k(x)]^T.$$

Then we define an n-dimensional column vector $u(x)$ by

$$u(x) = \sum_{j=1}^{n} (-1)^j \det (a^1(x), \cdots, a^{j-1}(x), a^{j+1}(x), \cdots, a^n(x))e^j,$$

where e^j is the jth unit column vector. Evidently, $u(x)$ is a continuous function of $x \in V$. Since $H'(x)$ has rank $n-1$, at least one of the determinants is nonzero, so $u(x) \neq 0$. Finally one may verify that $u(x)$ is orthogonal to each of the vectors $r^j(x)$, so $H'(x)u(x) = 0$ and $u(x)$ is the desired vector field.

Pick a $\bar{x} \in H^{-1}(x^0)$. Since $H'(\bar{x})$ has rank $n-1$, we see from the implicit function theorem [1, p. 265] that there is an open neighborhood O of $\bar{x}$ and a C^1 curve $x(t)$, defined for $-t_1 < t < t_2$, $t_1 > 0, t_2 > 0$, such that $H^{-1}(x^0) \cap O$ consists of the points $x(t)$, $-t_1 < t < t_2$, and $dx/dt \neq 0$, $-t_1 < t < t_2$. Since $H(x(t)) = x^0$,

$$H'(x(t))\frac{dx}{dt} = 0, \qquad -t_1 < t < t_2,$$

so dx/dt is proportional to $u(x(t))$. By redefining the parameter t, we may suppose that $x(t)$ satisfies the system

$$\begin{aligned} \frac{dx}{dt}(t) &= u(x(t)), \\ x(0) &= \bar{x} \end{aligned} \tag{2.2}$$

in a neighborhood of the point $t = 0$. Thus (2.2) has one and only one solution in a neighborhood of the point $t = 0$. The solution $x(t)$ of (2.2) may be continued provided $x(t) \in D \setminus C$. Let the curve exist for $0 \leqq t < T$. Then again using the implicit function theorem, it is seen that x provides a local homeomorphism of $(0, T)$ into $D \setminus C$. There are 2 cases. Suppose first that $x(t_1) \neq x(t_2)$ for all $t_1 \neq t_2$, $t_1, t_2 \in (0, T)$. Then x provides a homeomorphism of $(0, T)$ onto its image in $D \setminus C$.

Next, suppose $x(t_1) = x(t_2)$ for some $t_1 \neq t_2$. Let

$$L = \inf\{t_2 - t_1 : x(t_1) = x(t_2); 0 \leqq t_1 < t_2 < T\}.$$

We show that $L > 0$. For otherwise, we may find sequences $t_1^j \in (0, T)$, $t_2^j \in (0, T)$ such that

$$t_1^j \to t_1,$$
$$t_2^j \to t_1, \qquad t_2^j > t_1^j,$$
$$x(t_2^j) = x(t_1^j).$$

Since $H(x(t_1)) = x^0$, there is a neighborhood of $x(t_1)$ such that in this neighborhood, $H^{-1}(x^0)$ is the homeomorphic image of an open interval. Hence the curve $x(t)$ is 1–1 for t near t_1, which contradicts $L = 0$. Hence $L > 0$. Let $t_1^j \to t_1$, $t_2^j \to t_2$, $t_2^j - t_1^j \to L$, $x(t_1^j) = x(t_2^j)$. Then taking the limit, we have $x(t_1) = x(t_1 + L)$. By uniqueness of solutions of (2.2), we have $x(t) = x(t + L)$. It is easily shown that if $x(t) = x(s)$, then $t - s = kL$ for some integer k. Hence in this case, the curve is the homeomorphic image of a circle.

To prove that the curve is a closed subset of $H^{-1}(x^0) \cap D$, let $z \in D \setminus C$, $H(z) = x^0$, and let $x(t_j)$ be points on the curve with $x(t_j) \to z$. By the implicit function theorem, there is neighborhood W of z such that $H^{-1}(x^0) \cap W$ is the homeomorphic image of an interval. Hence $z = x(t)$ is on the curve, and the curve is a closed subset of $H^{-1}(x^0) \cap D$. Now $H^{-1}(x^0) \cap D$ has been written as the disjoint union of a number of curves, each one of which is closed in $H^{-1}(x^0) \cap D$. Therefore each curve is a connected component of $H^{-1}(x^0) \cap D$, and the theorem is proved.

Regarding the connected components of $H^{-1}(x^0) \cap D$ we have

LEMMA 2.1. *Each connected component of $H^{-1}(x^0) \cap D$ is a closed subset of $D \setminus C$ in the relative topology of $D \setminus C$.*

Proof. Let $x(t)$ be a curve whose image is a connected component of $H^{-1}(x^0) \cap D$. Set $z \in D \setminus C$, and let $x(t_j) \to z$. Then $H(z) = \lim H(x(t_j)) = x^0$, so $z \in H^{-1}(x^0) \cap D$. Since the curve is a closed subset of $H^{-1}(x^0) \cap D$, we see that $z = x(t)$ lies on the curve.

We shall let R denote the set of $x^0 \in \partial D$ given by Theorem 2.1. The points $x^0 \in R$ are known as *regular* values of H. For $x^0 \in R$, we let $x(s, x^0)$ be the solution of the initial value problem

(2.3a) $$H'(x(s))\dot{x}(s) = 0, \qquad \dot{x}(s) = dx/ds,$$

(2.3b) $$x(0) = x^0,$$

(2.3c) $$x(s) \in D \quad \text{for } s > 0,$$

(2.3d) $$(\dot{x}(s), \dot{x}(s)) = 1.$$

The condition (2.3c) fixes the sign of the initial tangent vector $\dot{x}(0)$, and requires it to point into the set D. The condition (2.3d) implies that the parameter s is the arc length along the curve. Other parametrizations of the curve will be considered later.

Since $x(s, x^0) \to x^0 \in \partial D$ as $s \to 0$, the curve $x(s, x^0)$ is a homeomorphic map of an open interval into $D \setminus C$. Let the interval be given by $0 < s < L(x^0)$, where

$L(x^0) \leqq \infty$. The next theorem shows the existence of a fixed point and gives the basis for our numerical method.

THEOREM 2.2. *Let $x^0 \in R$, and let s_j be a sequence with $s_j \to L(x^0)$. Then $x(s_j, x^0)$ contains a subsequence which converges to a fixed point of F, so C is not empty. Furthermore, for $\varepsilon > 0$, there is a $s(\varepsilon, x^0) < L(x^0)$ such that for $s(\varepsilon, x^0) < s < L(x^0)$,*

$$\operatorname{dist}(x(s, x^0), C) < \varepsilon.$$

Proof. Since $x(s_j, x^0) \in \bar{D}$, we may, by choosing a subsequence, suppose that $x(s_j, x^0) \to z \in \bar{D}$. If $z \in \partial D$, then $H(z) = x^0$ so $z = x^0$. Since, in a neighborhood of x^0, $H^{-1}(x^0)$ consists of the homeomorphic image of an open interval, we must have $x(s_j, x^0) = x(s, x^0)$ for some j sufficiently large and some s close to 0. This contradicts Theorem 2.1, so $z \notin \partial D$. If $z \in D \setminus C$, then $H(z)$ is defined and $H'(z)$ has rank $n - 1$. Hence $x(s, x^0)$ may be continued to $s > L(x^0)$, which is a contradiction. Hence $z \in C$, so C is not empty. Let $W_\varepsilon = \{x \in \bar{D} : \operatorname{dist}(x, C) < \varepsilon\}$. Since $C \cap \partial D$ is empty, for ε sufficiently small, $W_\varepsilon \subset D$, and $\bar{D} \setminus W_\varepsilon$ is a compact subset of R^n. Let $A = \{x(s, x^0), 0 \leqq s < L(x^0)\}$. Then $B = A \cap (\bar{D} \setminus W_\varepsilon)$ is a compact subset of R^n. For let $x(t_j)$ be an infinite sequence of points of B. Since $B \subset \bar{D} \setminus W_\varepsilon$, we may, by picking a subsequence, suppose that $x(t_j) \to z \in \bar{D} \setminus W_\varepsilon$. If $z \in \partial D$, then $z = H(z) = x^0$, so $z \in B$. If $z \notin \partial D$, then $z \in D \setminus C$, so by Lemma 2.1, $z \in A$ and $z \in B$. Hence B is compact as asserted. The function $\psi : x(s) \to s$ is continuous on B. The continuity of ψ at a point $x \neq x^0$ of B follows from Theorem 2.1. To show the continuity of ψ at x^0, we recall that F and H have smooth extensions to an open neighborhood of x^0. Denoting these extensions again by F and H, we may find an open neighborhood V of x^0 such that F and H are defined in V and $F(x) \in D$, $x \in V$. From the implicit function theorem, the set $H^{-1}(x^0) \cap V$ then consists of a curve $x(s)$, $-s_1 < s < s_2$, $s_1 > 0$, $s_2 > 0$, such that $x(s) \in D$ if and only if $s > 0$. Hence a point $x \in V \cap B \cap D$ if and only if $x = x(s, x^0)$ for $0 < s < s_2$. Then if $x^j \in B$, $x^j \to x^0$, we see that $x^j = x(s^j, x^0)$, with $s^j = \psi(x^j) \to 0$, so ψ is continuous at x^0. Hence ψ assumes a maximum, $s(\varepsilon)$. Then for $s > s(\varepsilon)$, $x(s, x^0) \in W_\varepsilon$, and the theorem is proved.

Remarks. (i) Our assumption that F maps $\bar{D}$ into the open set D can be weakened to $F : \bar{D} \to \bar{D}$ with the additional assumption that if x and $F(x)$ are in ∂D, the line segment between x and $F(x)$ lies in D. Notice also that $\bar{D}$ can be a simplex.

(ii) The form of Sard's theorem given in [4] is for smooth functions. However, it may be verified that, since in our case H maps a manifold of dimension n into a manifold of dimension $n - 1$, the proof of [4] is valid for twice continuously differentiable functions.

3. Relation with Newton's method. We give a condition which guarantees that any x^0 in a smooth portion of ∂D is a regular value of H, and hence may be taken as a starting value for the continuation curve. As a consequence we relate the method described here with Newton's method.

Since $H(x)$ lies on the ray drawn from $F(x)$ and passing through x, we may write

$$H(x) = (1 - \mu(x))x + \mu(x)F(x), \qquad \mu(x) < 0. \tag{3.1}$$

The function $\mu : \bar{D} \setminus C \to R^1$. If $H(x)$ lies on a continuously differentiable portion of ∂D, then $\mu(x)$ is of class C^1. We let $\mu'(x)$ denote the gradient of μ. Then the matrix $H'(x)$ acts on a vector v by the formula

$$H'(x)v = [(1 - \mu(x))I + \mu(x)F'(x)]v + (\mu'(x), v)(F(x) - x). \tag{3.2}$$

The following lemma is in the spirit of the Woodbury formula for calculating the inverse of a rank one perturbation of a matrix.

LEMMA 3.1. *Let U be an open set such that $U \cap \partial D$ is a continuously differentiable surface. Let $z \in D \setminus C$ be such that $H(z) \in U \cap \partial D$, and suppose that the matrix $F'(z)$ has no eigenvalues which lie on* $(1, \infty)$. *Then $H'(z)$ has rank $n - 1$, and the null-space of $H'(z)$ consists of vectors of the form* $[(1 - \mu(z))I + \mu(z)F'(z)]^{-1}(F(z) - z)$.

Proof. Set $A = \mu(z)F'(z) + (1 - \mu(z))I$. Then since $\mu(x) < 0$, the eigenvalue condition on $F'(z)$ implies that A is nonsingular. Also, $\mu(x)$ is continuously differentiable near $x = z$. The matrix $H'(z)$ is singular so let u be a null vector of $H'(z)$. Then from (3.2),

$$Au = -(\mu'(z), u)(F(z) - z),$$

so u is proportional to $A^{-1}(F(z) - z)$. Hence $H'(z)$ has rank $n - 1$ and the null-space is of the asserted form.

From Lemma 3.1 and Theorem 2.2 we obtain

THEOREM 3.1. *Let $F : \bar{D} \to D$ and suppose that for each $x \in D$, the matrix $F'(x)$ has no eigenvalues in the interval* $(1, \infty)$. *Then each x^0 on a C^1 part of ∂D is a regular value of H, and the continuation curve starting at x^0 is well-defined and goes to the set C.*

Remark. It is reasonable to suppose that the eigenvalue condition of Theorem 3.1 implies some restrictions on the fixed point set C. In particular, we conjecture that C has only one connected component which can be joined to ∂D by a path in $D \setminus C$. In this regard, if for each $x \in D$, $F'(x)$ does not have eigenvalue 1, then using Hopf's theorem on the index of vector fields [10], we can show that F has a unique fixed point.

Suppose that $F'(x)$ never has eigenvalues on $(1, \infty)$, for $x \in D$. Then from Lemma 3.1,

$$-[F'(x) + (\mu(x)^{-1} - 1)I]^{-1}(F(x) - x)$$

is a null vector of $H'(x)$, so after a reparametrization the continuation curve may be described by

$$\begin{aligned} &x(0) = x^0 \in \partial D, \\ &\frac{dx}{dt} = -[F'(x) - I + \mu(x)^{-1}I]^{-1}(F(x) - x). \end{aligned} \tag{3.3}$$

Since $F(x) - x \to 0$ along the curve, it is easy to see that $\mu(x) \to -\infty$ as x moves along the curve $x(t)$. Hence for $x(t)$ close to the fixed point, the curve (3.3) is given approximately by

$$\frac{dx}{dt} = -[F'(x) - I]^{-1}(F(x) - x). \tag{3.4}$$

Solving (3.4) gives rise to the continuous Newton's method for solving the equation $F(x) - x = 0$ [9]. Thus the continuation curve given by (3.3) may be regarded as a modified form of the continuous Newton curve for solving the equation $F(x) - x = 0$.

If one solves (3.4) approximately by using Euler's method with a step of $\Delta t = 1$, Newton's method is obtained. Similarly, if a step size $\Delta t = 1$ and Euler's method is used with (3.3), one obtains the iterative method

$$\begin{aligned} &x^{k+1} = x^k - [F'(x^k) - I + \mu(x^k)^{-1}I]^{-1}(F(x^k) - I), \\ (3.5) \qquad &\mu(x^k) < 0, \\ &(1 - \mu(x^k))x^k + \mu(x^k)F(x^k) \in \partial D. \end{aligned}$$

The method (3.5) may be regarded as a modified Newton's method with a specific rule for choosing the modifications.

4. Relation with M-functions. In this section, we show that the method of continuation may be applied to solving equations involving M-functions, and that as a result, we obtain a modified form of the continuous Newton's method with guaranteed convergence and interesting monotonicity properties. For another method of solving equations involving M-functions, that is motivated by Scarf's method, see [14].

We recall the definition of an M-function [11], [13], and an M-matrix [16]. A function $G : R^n \to R^n$ is an M-function if G satisfies the following three conditions:

(a) for each j, and $t \in R^1$, the quantity

$$G_j(x_1, \cdots, x_j + t, \cdots, x_n)$$

is a strictly increasing function of t;

(b) for each $j \neq k$, and $t \in R^1$, the quantity

$$G_j(x_1, \cdots, x_k + t, \cdots, x_n)$$

is a nonincreasing function of t;

(c) for $x, y \in R^n$, if $G_j(x) \leqq G_j(y)$, $1 \leqq j \leqq n$, then $x_j \leqq y_j$, $1 \leqq j \leqq n$.

A matrix $A = (a_{jk})$ is called an M-matrix if: $a_{jk} \leqq 0$ for $j \neq k$; A is nonsingular and each entry of A^{-1} is nonnegative. It is, for example [13], known that if G is of class C^1 and $G'(x)$ is nonsingular, $x \in R^n$, then G is an M-function if and only if $G'(x)$ is an M-matrix, $x \in R^n$.

As is customary in the theory of M-functions, we shall use the inequality $x < y$ to mean $x_j < y_j$, $1 \leqq j \leqq n$. Suppose G is a C^1 M-function and suppose we are given points x^1, x^2 with

$$G(x^1) < 0, \qquad G(x^2) > 0.$$

Then there is a unique x^* with $G(x^*) = 0$, and furthermore, $x^1 < x^* < x^2$ [13]. We want to calculate x^*. The points x^1 and x^2 are called, respectively, a subsolution and a supersolution of the problem.

We first modify the function G. We define $\bar{G} : R^n \to R^n$ by

$$\bar{G}(x) = G(x + \tfrac{1}{2}(x^1 + x^2)).$$

Setting $\bar{x} = \frac{1}{2}(x^2 - x^1) > 0$, we have

$$\bar{G}(\bar{x}) = G(x^2) > 0,$$
$$\bar{G}(-\bar{x}) = G(x^1) < 0.$$

Since $\bar{G}$ is also an M-function, there is a unique $\tilde{x}$ with $\bar{G}\tilde{x} = 0$, and furthermore, $-\bar{x} < \tilde{x} < \bar{x}$. We let $\bar{Q}$ denote the set

$$\bar{Q} = \{x : -\bar{x} \leqq x \leqq \bar{x}\}.$$

We set $F(x) = x - \alpha\bar{G}(x)$, and we pick $\alpha > 0$ so small that

$$0 < F(\bar{x}) < \bar{x}, \tag{4.1}$$

$$0 > F(-\bar{x}) > -\bar{x}, \tag{4.2}$$

$$\alpha^{-1} > \text{each diagonal entry of } \bar{G}'(x), \qquad x \in \bar{Q}. \tag{4.3}$$

Since $\bar{G}'(x)$ is an M-matrix, the condition (4.3) gives a positive upper bound for α. We then have

LEMMA 4.1. *If α satisfies* (4.1), (4.2), (4.3), $F(\bar{Q}) \subset \bar{Q}$.

Proof. We first show that F is isotone; that is $x \leqq y \Rightarrow F(x) \leqq F(y)$. Let $y = x + h$, $h \geqq 0$, and let $x(t) = x + th$. Using (4.3) and the fact that $\bar{G}'(x(t)) = (a_{ij})$ is an M-matrix, we have

$$\begin{aligned} 0 \leqq \alpha h_i &\leqq a_{ii}^{-1} h_i \\ &\leqq a_{ii}^{-1} h_i - \alpha \sum_{j \neq i} a_{ij} a_{ii}^{-1} h_j, \end{aligned}$$

so,

$$\alpha\bar{G}'(x(t))h \leqq h.$$

Now

$$\frac{d}{dt} F(x(t)) = h - \alpha\bar{G}'(x(t))h \geqq 0,$$

so upon integrating this inequality,

$$F(y) = F(x(1)) \geqq F(x(0)) = F(x).$$

Using the isotonicity of F and (4.1), (4.2), if $-\bar{x} \leqq x \leqq \bar{x}$, then we have

$$-\bar{x} < F(-\bar{x}) \leqq F(x) \leqq F(\bar{x}) < \bar{x}$$

so $F(\bar{Q}) \subset \bar{Q}$, proving the lemma.

Using Lemma 4.1, we may formulate our continuation method as a method for solving the equation $Gx^* = 0$, and using the spectral properties of M-matrices, we may apply Theorem 3.1. Motivated by these considerations, we shall formulate a continuation method for solving $Gx^* = 0$, and give a direct convergence proof for the method. The continuation method will also be seen to have some useful monotonic convergence properties.

We define, for $G_1(x) \neq 0$,

$$K(x) = x - \mu(x)G(x),$$
$$\mu(x) = \frac{x_1 - x_1^1}{G_1(x)}.$$

Note that the first component of $K(x)$ satisfies $K_1(x) = x_1^1$. Thus, the first row of $K'(x)$ is 0 and $K'(x)$ is a singular matrix. However we have

LEMMA 4.2. *If $\mu(x) < 0$, $K'(x)$ has rank $n - 1$, and a null vector of $K(x)$ is given by the formula*

$$u(x) = [I - \mu(x)G'(x)]^{-1}G(x). \tag{4.4}$$

Proof. The vector $K'(x)v$ is given by

$$K'(x)v = [I - \mu(x)G'(x)]v - (\mu'(x), v)G(x).$$

As in the proof of Lemma 3.1, if $I - \mu(x)G'(x)$ is nonsingular, then $K'(x)$ has rank $n - 1$, and (4.4) holds. If $I - \mu(x)G'(x)$ is singular, then $G'(x)$ has an eigenvalue $1/\mu(x)$, which contradicts the fact that all the eigenvalues of an M-matrix have positive real part [16, p. 87].

We consider the curve $x(t)$ defined by

$$\begin{aligned} K(x(t)) &= x^1, \qquad t \geqq 0, \\ x(0) &= x^1, \end{aligned} \tag{4.5}$$

where x^1 is a subsolution of the problem. Then we have

THEOREM 4.1. *If $G'(x(t))$ is irreducible, then the curve $x(t)$ is well-defined for $0 \leqq t < \infty$ and satisfies*

$$\dot{x}(t) > 0, \tag{4.6}$$

$$G(x(t)) < 0, \tag{4.7}$$

$$\mu(t) = -t, \tag{4.8}$$

$$\lim_{t \to \infty} x(t) = x^*. \tag{4.9}$$

Proof. Using Lemma 4.2 and the theory of ordinary differential equations, we see that $x(t)$ is well-defined on some interval $0 \leqq t < T \leqq \infty$, and may be taken to be the solution of the initial value problem

$$\dot{x} = -[I - \mu(x)G'(x)]^{-1}G(x), \tag{4.10}$$

$$x(0) = x^1. \tag{4.11}$$

Since $\mu(x(0)) = 0$ and x^1 is a subsolution, $\dot{x}(0) = -G(x^1) > 0$. Moreover, by (4.10)

$$\dot{x} - \mu(x)\dot{G}(x) = -G(x). \tag{4.12}$$

In particular $\dot{x}_1 - \mu(x)\dot{G}_1(x) = -G_1(x)$. Therefore if $G_1(x) \neq 0$,

$$\dot{\mu} = \frac{G_1(x)x_1 - (x_1 - x_1^1)\dot{G}_1(x)}{G_1^2(x)} = \frac{\dot{x}_1 - \mu(x)\dot{G}_1(x)}{G_1(x)} = -1.$$

Clearly, (4.6) and (4.7) hold for t sufficiently small, while (4.8) holds at least as long as (4.7) does. Let $t_1 < T$ be the smallest number such that one of (4.6) or (4.7) does not hold. If $G_1(x(t_1)) = 0$, then $-\infty = \lim_{t \to t_1} \mu(x(t)) = \lim_{t \to t_1} - t$, which contradicts $t_1 < \infty$. Thus $G_1(x(t_1)) < 0$, and since $I - \mu(x(t_1))G'(x(t_1))$ is an irreducible M-matrix, $\dot{x}(t_1) > 0$. Now if $G_j(x(t_j)) = 0, j \neq 1$, we have from (4.12) that $\dot{G}_j(x(t_1)) = -\dot{x}_j(t_1)/t < 0$ which contradicts $G_j(x(t)) < 0$ for $t < t_1$. This proves that (4.6), (4.7) and (4.8) hold for $0 \leqq t < T$. To complete the proof we note that by (4.6) and the fact that $x(t) \leqq x^*$, $z \equiv \lim_{t \to T} x(t) \leqq x^*$ necessarily exists. However, T must be infinite since otherwise we could continue $x(t)$ beyond z. If $G_j(z) \neq 0$ for some j, then $K_j = -\infty$ contradicting (4.5). Thus $G(z) = 0$; that is, $z = x^*$. This completes the proof of the theorem.

5. Numerical experiments. A series of numerical experiments were made on calculating fixed points of a function $F : \bar{D} \to \bar{D}$ where $D = \{x : 0 < x_j < 1\}$ is the unit cube in R^n. If $H(x)$ lies on the face of D given by $x_1 = 0$, then

$$H(x) = x - \frac{x_1}{x_1 - F_1(x)}(x - F(x)).$$

If the initial guess x^0 satisfies $x_1^0 = 0$, we may use this formula to write the differential equation for the solution curve $H^{-1}(x^0)$. Using Euler's method to solve approximately the differential equation, an algorithm has been developed for the approximate calculation of the fixed point. A description of the algorithm is given below; for more details, see [6].

In the numerical experiments, each component function $F_j(x)$ was chosen to be of the form

$$F_j(x) = a + bx_{i_1}x_{i_2}x_{i_3};$$

the integers i_1, i_2, i_3 and the positive numbers a, b were chosen, for each j, from a random number generator. Using the algorithm, 5 problems were run on a Univac 1106 for $n = 20, 30, \cdots, 60$. The results of these 25 problems are summarized in Table 1. It should be noted in this table that the term "function call" refers to the number of times the scalar-valued functions $F_j(x)$ were evlauated. The problems were continued until $|F(x^k) - x^k| < 10^{-6}$.

TABLE 1
Results of numerical experiments 5 problems for each value of n

n	Steps per problem	Scalar function calls per problem	time per problem in seconds
20	20–28	1,160–1,500	3.3
30	24–35	1,980–2,700	8
40	23–41	2,560–4,040	17
50	31–50	4,100–6,000	35
60	28–40	4,500–6,120	55

Description of algorithm.

1. Pick $x^0 \in R^n$ such that $x_1^0 = 0$, and pick numbers $h > 0$, $\varepsilon_1 > 0$, $\varepsilon_2 > 0$.
2. Set $i = 0$.
3. Compute $H'(x^i)$.
4. Compute v such that $H'(x^i)v = 0$ and $v_1 > 0$.
5. Replace v by $v/\|v\|$.
6. Set $x^{i+1} = x^i + hv$.
7. Compute $F(x^{i+1})$.
8. If $F_1(x^{i+1}) \leqq 0$, replace h by $h/2$ and go to step 6. Otherwise, go to next step.
9. Compute $H(x^{i+1})$.
10. If $|H(x^{i+1}) - H(x^i)| \geqq \varepsilon_1$, replace h by $h/2$ and go to step 6. Otherwise, go to next step.
11. If $|F(x^{i+1}) - F(x^i)| < \varepsilon_2$, exit from the program with x^{i+1} as the approximate solution. Otherwise, replace i by $i + 1$ and go to step 3.

Remark. The method given in step 10 of reducing the step size h is one of several that we have considered.

REFERENCES

[1] J. ALEXANDER AND J. A. YORKE, *Global bifurcation of periodic orbits*, preprint.

[2] J. DIEUDONNÉ, *Foundations of Modern Analysis*, Academic Press, New York, 1960.

[3] B. C. EAVES AND R. SAIGAL, *Homotopies for computation of fixed points on unbounded regions*, Math. Programming, 3 (1972), pp. 225–237.

[4] M. GOLUBITSKY AND V. GUILLEMIN, *Stable Mappings and their Singularities*, Springer-Verlag, New York, 1973.

[5] M. HIRSCH, *A proof of the nonretractability of a cell onto its boundary*, Proc. Amer. Math. Soc., 14 (1963), pp. 364–365.

[6] R. B. KELLOGG, T. Y. LI AND J. YORKE, *A method of continuation for calculating a Brouwer fixed point*, Computing Fixed Points with Applications, S. Karamadian, ed., Academic Press, New York, to appear.

[7] H. W. KUHN, *Simplicial approximation of fixed points*, Proc. Nat. Acad. Sci. U.S.A., 61 (1968), pp. 1238–1242.

[8] K. KURATOWSKI, *Topology, Volume I*, Academic Press, New York, 1966.

[9] G. H. MEYER, *On solving nonlinear equations with a one-parameter operator imbedding*, this Journal, 5 (1968), pp. 739–752.

[10] J. W. MILNOR, *Topology from the Differentiable Viewpoint*, University Press of Virginia, Charlottesville, 1965.

[11] J. MORÉ AND W. RHEINBOLDT, *On P- and S-functions and related classes of n-dimensional nonlinear mappings*, Linear Algebra and Appl., 6 (1973), pp. 45–68.

[12] L. S. PONTRYAGIN, *Smooth manifolds and their applications in homotopy theory*, Amer. Math. Soc. Transl., 11 (1959), pp. 1–114.

[13] W. RHEINBOLDT, *On M-functions and their application to nonlinear Gauss-Seidel iterations, and network flows*, J. Math. Anal. Appl., 22 (1971), pp. 274–307.

[14] W. RHEINBOLDT AND C. K. MESZTENYI, *A combinatorial search process for M-functions*, Tech. Rep. TR-296, Univ. of Maryland Computer Science Center, College Park, 1974.

[15] H. SCARF, *The approximation of fixed points of a continuous mapping*, SIAM J. Appl. Math., 15 (1967), pp. 1328–1343.

[16] R. S. VARGA, *Matrix Iterative Analysis*, Prentice-Hall, Englewood Cliffs, N.J., 1962.