Intern. J. Computer Math., Vol. 28, pp. 171–181
Reprints available directly from the publisher
Photocopying permitted by license only

THE IMPLICIT FUNCTION THEOREM FOR SOLVING SYSTEMS OF NONLINEAR EQUATIONS IN $\mathbb{R}^2$

THEODOULA N. GRAPSA

Department of Mathematics, University of Patras, Gr-261.10 Patras, Greece

MICHAEL N. VRAHATIS

Department of Mathematics, White Hall, Cornell University, Ithaca, NY 14853, USA

(Received 9 February 1988)

A new method for solving systems of two simultaneous nonlinear algebraic and/or transcendental equations in $\mathbb{R}^2$, which is based on reduction to simpler one-dimensional non-linear equations is presented. This method to approximate a component of the solution does not require any information about the other component in each iteration. It generates a sequence of points in $\mathbb{R}$ which converges quadratically to one component of the solution and afterwards it evaluates the other component using one simple computation. Moreover, it does not require a good initial guess of the solution for both components and it does not directly need function evaluations. A proof of convergence is given.

KEY WORDS: Implicit function theorem, nonlinear equations, Newton's method, reduction to one-dimensional equations, nonlinear SOR method, m-step SOR-Newton method, quadratic convergence, numerical solution, zeros.

C.R. CATEGORY: G.1.5.

1. INTRODUCTION

Suppose that $F=(f_1,\ldots,f_n)^T: D\subset\mathbb{R}^n\to\mathbb{R}^n$ is a Fréchet differentiable mapping on an open neighborhood $D^*\subset D$ of a solution $X^*\in D$ of the system of nonlinear equations

$$F(X)=\Theta=(0,0,\ldots,0)^T. \tag{1.1}$$

Starting with an initial guess X^0, Newton's method for the attainment of an approximation of X^* is given by

$$X^{p+1}=X^p-F'(X^p)^{-1}F(X^p),\quad p=0,1,\ldots \tag{1.2}$$

If, in addition, the Fréchet derivative at X^* is nonsingular and $F'(X)$ is Lipschitz continuous, then the Newton iterates will converge to X^* provided the initial

guess X^0 is sufficiently close to X^*. In this case the convergence of Newton's method at X^* is quadratic [9], that is

$$\|X^{p+1}-X^*\| \leqq C\|X^p-X^*\|^2, \quad \text{for some } C, \quad p=0,1,\ldots \tag{1.3}$$

The quadratic convergence of Newton's method is attractive. However, the method depends on a good initial approximation [4], and it requires in general n^2+n function evaluations per iteration, besides the solution of an $n \times n$ linear system.

There is a class of methods for solving systems of nonlinear equations which arise from iterative methods used for systems of linear equations, [7–10, 12]. These methods to solve (1.1) use reduction to simpler one-dimensional nonlinear equations of the components $f_1, f_2, \ldots, f_n$. The best-known method of this type is the nonlinear successive overrelaxation, (SOR), method which solves the equation

$$f_i(X_1^{p+1}, \ldots, X_{i-1}^{p+1}, X_i, X_{i+1}^p, \ldots, X_n^p)=0 \tag{1.4}$$

for X_i and then sets

$$X_i^{p+1}=X_i^p+\omega(X_i-X_i^p), \quad i=1,\ldots,n, \quad p=0,1,\ldots \tag{1.5}$$

provided that $\omega \in (0,1]$. Now, if the solutions of (1.4) are computed exactly then we obtain the exact nonlinear SOR process. Moreover if m steps of the one-dimensional Newton's method are applied to (1.4) to yield an approximate solution, then we obtain the m-step SOR-Newton process.

Suppose, now, that $F: D \subset \mathbb{R}^n \to \mathbb{R}^n$ is continuously differentiable on an open neighborhood $D^* \subset D$ of a point $X^* \in D$ for which $F(X^*)=\Theta$. Assume further that $F'(X^*)$ is an M-matrix [9]. Then the iterates of the nonlinear SOR, exact nonlinear SOR and m-step SOR-Newton will converge to X^* provided that $\omega \in (0,1]$. In this case the convergence of the above processes at X^* is linear [9].

In this paper we implement the well-known Implicit Function Theorem [3, 9] to obtain a method for solving systems of two-dimensional nonlinear equations. This method although uses reduction to simpler one-dimensional nonlinear equations, as the previous methods use, yet it generates a sequence of points in $\mathbb{R}$ which converges quadratically to one component of the solution, while the other component is evaluated by one simple computation. Moreover this method does not directly need any function evaluation, unless it is required for the computation of the various partial derivatives.

In Section 2 the derivation and the description of the new method are presented. In Section 3 a proof of convergence is given. Lastly, in Section 4, numerical applications are given.

2. DERIVATION OF THE METHOD

NOTATION 2.1 In what follows, $\partial_i f$ denotes the partial derivative of f with respect to the ith variable, $\bar{S}$ denotes the closure of the set S, $\operatorname{diag}(a_1,\ldots,a_n)$

defines the diagonal matrix with elements $a_1,\dots,a_n$ and $f(X_1,\cdot)$ defines the mapping derived by holding X_1 fixed.

In the development of our analysis we shall implement the following theorem.

THEOREM 2.1 (Implicit Function Theorem) *Suppose that $f: D\subset\mathbb{R}^2\to\mathbb{R}$ is defined and continuous, together with its partial derivatives $\partial_1 f$ and $\partial_2 f$, on an open neighborhood $D_0\subset D$ of a point $X^0=(X_1^0,X_2^0)^T$ for which $f(X^0)=0$ and $\partial_2 f(X^0)\neq 0$. Then there exist open neighborhoods $J_1\subset\mathbb{R}$ and $J_2\subset\mathbb{R}$ of the points X_1^0 and X_2^0 respectively such that for any $X_1\in\bar{J}_1$, the equation $f(X)=0$ has a unique solution $X_2=\phi(X_1)\in\bar{J}_2$. Moreover the mapping $\phi: J_1\to\mathbb{R}$ is continuous, it has a continuous derivative in J_1 and*

$$\phi'(X_1)=-\partial_1 f(X_1,\phi(X_1))/\partial_2 f(X_1,\phi(X_1)).$$

Proof See [3, pp. 184–188; 9, pp. 128–129].

Remark 2.1 The X_1 and X_2 of the above theorem can change roles. Thus, if at some point $\partial_1 f\neq 0$ then in a neighborhood of it the equation $f(X)=0$ can be expressed by the same equation in the form $X_1=g(X_2)$ [3].

Suppose, now, that $F=(f_1,f_2)^T: D\subset\mathbb{R}^2\to\mathbb{R}^2$ is twice-continuously differentiable on an open neighborhood $D^*\subset D$ of a solution $X^*=(X_1^*,X_2^*)^T\in D$ of the system of nonlinear equations

$$F(X)=\Theta=(0,0)^T. \tag{2.1}$$

Our interest lies in obtaining an approximate solution of (2.1). To do this we consider the sets S_i, $i=1,2$ to be those connected components of $f_i^{-1}(0)$ containing X^* on which $\partial_2 f_i\neq 0$ for $i=1,2$ respectively. Then according to the above theorem there exist open neighborhoods $J_1^*\subset\mathbb{R}$ and $J_{2,i}^*\subset\mathbb{R}$, $i=1,2$ of the points X_1^* and X_2^* respectively such that for any $X_1\in\bar{J}_1^*$ the equations $f_i(X)=0$, $i=1,2$ define X_2 as a single-valued function of X_1, thus

$$X_2=\phi_i(X_1)\in\bar{J}_{2,i}^*,\quad i=1,2. \tag{2.2}$$

Furthermore the functions $\phi_i: J_1^*\to\mathbb{R}$ are continuous and have continuous derivatives in J_1^* which are given by

$$\phi_i'(X_1)=-\partial_1 f_i(X_1,\phi_i(X_1))/\partial_2 f_i(X_1,\phi_i(X_1)),\quad i=1,2. \tag{2.3}$$

Suppose now that $X^0=(X_1^0,X_2^0)^T$ is an initial approximation of the solution where $X_1^0\in\bar{J}_1^*$, then using Taylor's formula we can expand the $\phi_i(X_1)$, $i=1,2$ about X_1^0. So, we get

$$\phi_i(X_1)\simeq\phi_i(X_1^0)+(X_1-X_1^0)\phi_i'(X_1^0),\quad i=1,2, \tag{2.4}$$

or, using (2.2) and (2.3) we can form the following linear system

$$X_1\partial_1 f_i(X_1^0, X_2^{0,i})/\partial_2 f_i(X_1^0, X_2^{0,i}) + X_2 = X_2^{0,i} + X_1^0\partial_1 f_i(X_1^0, X_2^{0,i})/\partial_2 f_i(X_1^0, X_2^{0,i}), \quad i=1,2 \tag{2.5}$$

where $X_2^{0,i}=\phi_i(X_1^0)$, $i=1,2$ are the corresponding solutions of the equations $f_i(X_1^0, \cdot)=0$. Equivalently we can write the above linear system with the following form

$$A_0(X\text{-diag}(1,0)X^0)=V_0, \tag{2.6}$$

where the matrix A_0 and the column vector V_0 are defined as follows

$$A_0=[a_{ij}]=[\partial_j f_i(X_1^0, X_2^{0,i})], \quad i,j=1,2,$$

$$V_0=[v_i]=[X_2^{0,i}\partial_2 f_i(X_1^0, X_2^{0,i})], \quad i=1,2$$

Suppose now that $\det A_0 \neq 0$ then the solution X of (2.6) gives a new approximation of the solution X^* of (2.1). Thus in general we can obtain the following iterative scheme

$$X^{p+1}=\text{diag}(1,0)X^p + A_p^{-1}V_p, \quad p=0,1,\ldots \tag{2.7}$$

where

$$A_p=[a_{ij}]=[\partial_j f_i(X_1^p, X_2^{p,i})], \quad i,j=1,2,$$

$$V_p=[v_i]=[X_2^{p,i}\partial_2 f_i(X_1^p, X_2^{p,i})], \quad i=1,2,$$

provided that the matrix A_p is nonsingular.

Note that the above process does not require the expressions ϕ_i but only the values $X_2^{p,i}$ which are given by the solution of the equations $f_i(X_1^p, \cdot)=0$. So, by holding the X_1 fixed, we solve the equations

$$f_i(X_1^p, r_i^p)=0, \quad i=1,2 \tag{2.8}$$

for r_i^p in the interval $(a, a+\beta)$ with an accuracy δ. Of course, we can use any one of the well-known one-dimensional methods [9, 11, 12, 14] to solve the above equations. Here we use the traditional one-dimensional bisection (see [2, 13] for a discussion of usefulness), since frequently the steps β are long and also a few significant digits for the computations of the roots of the equations (2.8) are required. So, to solve an equation of the form

$$\Psi(t)=0, \tag{2.9}$$

where $\Psi:[\gamma_1,\gamma_2]\subset\mathbb{R}\to\mathbb{R}$ is continuous, we recommend the following iterative scheme, which is a simplified version of the bisection method [15–17]

$$t_{n+1}=t_n+\operatorname{sgn}\Psi(t_0)\cdot\operatorname{sgn}\Psi(t_n)\cdot h/2^{n+1},\quad n=0,1,\ldots \tag{2.10}$$

with $t_0=\gamma_1$ and $h=\gamma_2-\gamma_1$ and where for any real number a,

$$\operatorname{sgn} a=\begin{cases}-1 & \text{if } \alpha<0\\ 0 & \text{if } \alpha=0\\ 1 & \text{if } \alpha>0.\end{cases} \tag{2.11}$$

Of course, (2.10) converges to a root t^* in (γ_1,γ_2) if for some t_n, $n=1,2,\ldots$ holds that

$$\operatorname{sgn}\Psi(t_0)\cdot\operatorname{sgn}\Psi(t_n)=-1. \tag{2.12}$$

Also, the minimum number of iterations μ, that are required in obtaining an approximate root t' such that $|t'-t^*|\leqq\varepsilon$, for some $\varepsilon\in(0,1)$ is given by

$$\mu=\lceil\log_2(h\cdot\varepsilon^{-1})\rceil, \tag{2.13}$$

where the notation $\lceil v\rceil$ refers to the least integer that is not less than the real number v.

Lastly, from the (2.7) we can easily see that the computation of the first component of X^{p+1} does not require any information about X_2^p. Then, obviously we can obtain from (2.7) the following process

$$\begin{aligned}X_1^{p+1}&=X_1^p+W_pV_p,\quad p=0,1,\ldots,m,\\ X_2^{m+1}&=(X_2^{m,1}+X_2^{m,2})/2,\end{aligned} \tag{2.14}$$

where

$$W_p=[\partial_2 f_2(X_1^p,X_2^{p,2})/\det A_p,\ -\partial_2 f_1(X_1^p,X_2^{p,1})/\det A_p].$$

Moreover, since $X_2^*=\phi_1(X_1^*)=\phi_2(X_1^*)$ we can also take either $X_2^{m,1}$ or $X_2^{m,2}$ to approximate the X_2^*.

Instead of the previous process, we can obtain an alternative process if we take into consideration that the components X_1 and X_2 can change roles. Thus, if S_i, $i=1,2$ are those connected components of $f_i^{-1}(0)$, containing X^* on which $\partial_1 f_i\neq 0$ for $i=1,2$ respectively, then there exist open neighborhoods $J_{1,i}^*\subset\mathbb{R}$, $i=1,2$ and $J_2^*\subset\mathbb{R}$ of the points X_1^* and X_2^* respectively, such that for any $X_2^*\in\bar{J}_2^*$ the equations $f_i(X)=0$, $i=1,2$ have correspondingly the unique solutions $X_1=g_i(X_2)\in\bar{J}_{1,i}^*$, $i=1,2$. Moreover the functions $g_i: J_2^*\to\mathbb{R}$ are continuous and have continuous derivatives in J_2^* which are given by

$$g_i'(X_2)=-\partial_2 f_i(g_i(X_2),X_2)/\partial_1 f_i(g_i(X_2),X_2),\quad i=1,2. \tag{2.15}$$

Consequently, we can derive the following alternative process

$$X^{p+1}=\operatorname{diag}(0,1)X^p+B_p^{-1}U_p, \tag{2.16}$$

where

$$B_p=[b_{ij}]=[\partial_j f_i(X_1^{p,i},X_2^p)],\quad i,j=1,2$$

$$U_p=[u_i]=[X_1^{p,i}\,\partial_1 f_i(X_1^{p,i},X_2^p)],\quad i=1,2,$$

provided that the matrix B_p is nonsingular. The $X_1^{p,i}$ in (2.16) denotes the solution of the equation $f_i(\cdot,X_2^p)=0$.

Obviously, the computation of X_2^{p+1} in (2.16) does not require any information about X_1^p. So we conclude with the following process

$$\begin{gathered}X_2^{p+1}=X_2^p+Z_pU_p,\quad p=0,1,\ldots,k,\\ X_1^k=(X_1^{k,1}+X_1^{k,2})/2,\end{gathered} \tag{2.17}$$

where

$$Z_p=[-\partial_1 f_2(X_1^{p,2},X_2^p)/\det B_p,\ \partial_1 f_1(X_1^{p,1},X_2^p)/\det B_p].$$

Moreover, since $X_1^*=g_1(X_2^*)=g_2(X_2^*)$ we can also take either $X_1^{k,1}$ or $X_1^{k,2}$ to approximate the X_1^*.

Lastly, we give a geometric interpretation of the new method and a corresponding illustration of the main differences between Newton's method and new method. To do this, we write Newton's method as

$$(X_1^{p+1}-X_1^p)\,\partial_1 f_i(X^p)+(X_2^{p+1}-X_2^p)\,\partial_2 f_i(X^p)+f_i(X^p)=0,\quad i=1,2. \tag{2.18}$$

Now, the equations

$$X_3=(X_1-X_1^p)\,\partial_1 f_i(X^p)+(X_2-X_2^p)\,\partial_2 f_i(X^p)+f_i(X^p),\quad i=1,2 \tag{2.19}$$

represent planes in the (X_1,X_2,X_3)-space which are tangent to the surfaces $X_3=f_i(X)$, $i=1,2$ at the points $(X_1,X_2,f_i(X^p))$, $i=1,2$ respectively. Then the point X^{p+1} which is determined from (1.2) is the point of intersection of these two planes with the (X_1,X_2)-plane [9].

Instead of the above planes we consider the following ones.

$$X_3=(X_1-X_1^p)\,\partial_1 f_i(X_1^p,X_2^{p,i})+(X_2-X_2^{p,i})\,\partial_2 f_i(X_1^p,X_2^{p,i}),\quad i=1,2, \tag{2.20}$$

which are tangent to the surfaces $X_3=f_i(X)$, $i=1,2$ at the points $(X_1,X_2^{p,i},0)$, $i=1,2$ respectively. Then, the point X^{p+1} which is obtained using (2.7) is the point of intersection of these two planes with the (X_1,X_2)-plane. Or equivalently, X^{p+1} is the intersection of the following lines in (X_1,X_2)-space

$$(X_1 - X_1^p)\,\partial_1 f_1(X_1^p, X_2^{p,i}) + (X_2 - X_2^{p,i})\,\partial_2 f_i(X_1^p, X_2^{p,i}) = 0, \quad i = 1, 2. \tag{2.21}$$

The geometric interpretation of the process (2.16) is analogous to the previous one.

3. A PROOF OF CONVERGENCE

In this section we shall give a proof of convergence of the new method. We shall examine the convergence of the iterates of the process (2.14) which, obviously, is similar to the convergence of the iterates of (2.17).

The results of the following theorem will be needed in the sequel.

THEOREM 3.1 *Suppose that $f: E \subset \mathbb{R} \to \mathbb{R}$ is twice-continuously differentiable on an open neighborhood $E^* \subset E$ of a point $X^* \in E$ for which $f(X^*) = 0$ and $f'(X^*) \neq 0$. Then the iterates X^p, $p = 0, 1, \ldots$ of Newton's method*

$$X^{p+1} = X^p - f(X^p)/f'(X^p) \tag{3.1}$$

will converge to X^* provided the initial guess is sufficiently close to X^*. Moreover the order of convergence will be two.

Proof See [11, 14].

Now, we proceed with the following convergence result.

THEOREM 3.2 *Suppose that $F = (f_1, f_2)^T: D \subset \mathbb{R}^2 \to \mathbb{R}^2$ is twice-continuously differentiable on an open neighborhood $D^* \subset D$ of a point $X^* = (X_1^*, X_2^*)^T \in D$ for which $F(X^*) = \Theta = (0, 0)^T$. Let S_i, $i = 1, 2$ be those connected components of $f_i^{-1}(0)$ containing X^* on which $\partial_2 f_i \neq 0$ for $i = 1, 2$ respectively and that $\partial_1 f_i(X^*) \neq 0$, $i = 1, 2$. Then the iterates of (2.14) will converge to X^* provided the initial guess X_1^0 is sufficiently close to X_1^*. Moreover the iterates X_1^p, $p = 0, 1, \ldots$ have order of convergence two.*

Proof Obviously from (2.14) we can obtain

$$X_1^{p+1} = X_1^p + \partial_2 f_1(X_1^p, X_2^{p,1})\,\partial_2 f_2(X_1^p, X_2^{p,2})(X_2^{p,1} - X_2^{p,2})/\det A_p, \quad p = 0, 1, \ldots, \tag{3.2}$$

or, after some algebraic manipulations

$$X_1^{p+1} = X_1^p - (X_2^{p,1} - X_2^{p,2})/\{[-\partial_1 f_1(X_1^p, X_2^{p,1})/\partial_2 f_1(X_1^p, X_2^{p,1})]$$

$$-[-\partial_1 f_2(X_1^p, X_2^{p,2})/\partial_2 f_2(X_1^p, X_2^{p,2})]\}, \quad p = 0, 1, \ldots \tag{3.3}$$

Using (2.2) and (2.3) from the above relationship we get

$$X_1^{p+1} = X_1^p - [\phi_1(X_1^p) - \phi_2(X_1^p)]/[\phi_1'(X_1^p) - \phi_2'(X_1^p)], \quad p = 0, 1, \ldots \tag{3.4}$$

Consider now the mapping

$$\Phi: \bar{J}_1^* \subset \mathbb{R} \to \mathbb{R}, \quad \text{by} \quad \Phi(X_1) = \phi_1(X_1) - \phi_2(X_1). \tag{3.5}$$

Then obviously, for the above mapping the conditions of Theorem 3.1 are fulfilled. Consequently the iterates X_1^p, $p=0,1,\ldots$ of (2.14) converge to X_1^* and the order of convergence is two.

Suppose now, that for some p we have $X_1^p = X_1^*$ then from (2.14) we get

$$X_2^{p+1} = [\phi_1(X_1^*) + \phi_2(X_1^*)]/2 \tag{3.6}$$

or

$$X_2^{p+1} = X_2^*. \tag{3.7}$$

Thus the theorem is proven. □

4. NUMERICAL APPLICATIONS

To test the method of this paper, we applied the method (2.14) to the following systems

$$F(X) = \begin{bmatrix} f_1(X_1, X_2) \\ f_2(X_1, X_2) \end{bmatrix} = \begin{bmatrix} X_1^2 - 4X_2 \\ X_2^2 - 2X_1 + 4X_2 \end{bmatrix} = \Theta \tag{4.1}$$

$$F(X) = \begin{bmatrix} f_1(X_1, X_2) \\ f_2(X_1, X_2) \end{bmatrix} = \begin{bmatrix} X_1^3 - 3X_1X_2^2 \\ 3X_1^2X_2 - X_2^3 \end{bmatrix} = \Theta \tag{4.2}$$

$$F(X) = \begin{bmatrix} f_1(X_1, X_2) \\ f_2(X_1, X_2) \end{bmatrix} = \begin{bmatrix} (10^{5-X_1X_2}+1)(10^{-X_1} - 10^{X_1-14})/X_2 - 0.002 \\ 10^{-X_1/2}/2(1+10^{-X_1/2}) - 10^{-X_1-1} + \log X_2 \end{bmatrix} = \Theta \tag{4.3}$$

The system (4.1) is a well-known test case [5, 6]. It has two roots $(0,0)^T$ and $(1.69541521\ldots, 0.71860819\ldots)^T$ and its Jacobian at them is not singular. The Jacobian of the system (4.2) is singular at the root $(0,0)^T$. Also system (4.2) has a particular difficulty since the values of the function F even for points which are not very close to the root $(0,0)^T$, tend to be close to zero. The system (4.3) appears in some chemical applications [1, p. 57] because it calculates the pH of a weak acid from its concentration. The system (4.3) has a particular difficulty since the values of the function F for points which are not very close to the root $(3.891743658\ldots, 0.9872169279\ldots)^T$ cannot be computed with high accuracy. The Jacobian of the system (4.3) at its solution is not singular.

In Tables 1, 2 and 3 we present the number of iterations N which are required in obtaining an approximate solution of the systems (4.1), (4.2) and (4.3) correspondingly, for requiring accuracy 10^{-8} and 10^{-17}, respectively by using the iterative schemes (2.14) and (2.10), for several starting points X_1^0. Further, we

Table 1 Results for system (4.1)

Scheme (2.14)							*Newton's method*					
X_1^0	$\varepsilon = 10^{-8}$			$\varepsilon = 10^{-17}$			X_1^0	X_2^0	$\varepsilon = 10^{-8}$		$\varepsilon = 10^{-17}$	
	N	FE	AS	N	FE	AS			N	FE	N	FE
−1	5	20	50	7	28	70	−1	−1	5	30	6	36
1	6	24	60	8	32	80	1	−1	6	36	7	42
2	5	20	50	6	24	60	2	−1	*D*	*D*	*D*	*D*
3	5	20	50	7	28	70	3	−1	9	54	10	60
4	6	24	60	8	32	80	4	−1	8	48	9	54
5	7	28	70	10	40	100	5	−1	8	48	9	54
6	6	24	60	9	36	90	6	−1	9	54	10	60
7	7	28	70	10	40	100	7	−1	8	48	9	54
8	9	26	90	11	44	110	8	−1	9	54	10	60

Table 2 Results for system (4.2)

Scheme (2.14)							*Newton's method*					
X_1^0	$\varepsilon = 10^{-8}$			$\varepsilon = 10^{-17}$			X_1^0	X_2^0	$\varepsilon = 10^{-8}$		$\varepsilon = 10^{-17}$	
	N	FE	AS	N	FE	AS			N	FE	N	FE
−5	7	28	140	8	32	160	−5	−1	47	282	98	588
−4	5	20	100	6	24	120	−4	−1	47	282	98	588
−3	4	16	80	5	20	100	−3	−1	46	276	97	582
−2	2	8	40	3	12	60	−2	−1	45	270	96	576
−1	2	8	40	3	12	60	−1	−1	43	258	94	564
1	2	8	40	3	12	60	1	−1	43	258	94	564
2	2	8	40	3	12	60	2	−1	45	270	96	576
3	4	16	80	5	20	100	3	−1	46	276	97	582
4	5	20	100	6	24	120	4	−1	47	282	98	588
5	7	28	140	8	32	160	5	−1	47	282	98	588

present also in these tables the corresponding results obtained by the Newton's method for the same points X_1^0 while retaining the second component X_2^0 fixed. So, in the following tables "ε" indicates the requiring accuracy, "N" indicates the number of iterations, "FE" indicates the number of function evaluations, "AS" indicates the total number of algebraic signs that are required for applying the iterative scheme (2.10), and "D" indicates divergence.

From these results we observe that, although the second component X_2^0 of the starting points is near to the second component X_2^* of the solution for the cases (4.1), (4.2) and (4.3), yet iterative formula (2.14) using only the X_1^0 is seen to be superior to the Newton's method for the cases (4.2) and (4.3). The results for the method based on (2.17) for all the cases (4.1)–(4.3) are analogous to the previous one.

Table 3 Results for system (4.3)

Scheme (2.14)							*Newton's method*					
X_1^0	$\varepsilon=10^{-8}$			$\varepsilon=10^{-17}$			X_1^0	X_2^0	$\varepsilon=10^{-8}$		$\varepsilon=10^{-17}$	
	N	*FE*	*AS*	*N*	*FE*	*AS*			*N*	*FE*	*N*	*FE*
1	10	40	200	15	60	300	1	0.1	41	246	77	462
2	8	32	160	13	52	260	2	0.1	39	234	75	450
3	6	24	120	11	44	220	3	0.1	37	222	73	438
4	5	20	100	10	40	200	4	0.1	34	204	71	426
5	7	28	140	12	48	240	5	0.1	35	210	72	432
6	8	32	160	13	52	260	6	0.1	35	210	72	432
7	8	32	160	13	52	260	7	0.1	37	222	73	438

5. CONCLUDING REMARKS

The method we have analysed in this paper compares favourably with the Newton's method when the Jacobian of F at the root of the system (1.1) is singular or when it is difficult to evaluate the components of F accurately.

Also although our method uses reduction to simpler one-dimensional equations, yet it converges quadratically to one component of the solution, say the X_1 one, while the other component of the solution is evaluated by the simple computation $X_2=(\phi_1(X_1)+\phi_2(X_2))/2$. So it does not require a good estimate of both solution components as the Newton's method does.

Moreover, the method does not need function evaluation directly, and also using the iterative scheme (2.10) it requires only their algebraic signs to be correct in finding the $\phi_i(X_1)$.

References

[1] J. N. Butler and D. G. Bobrow, *The Calculus of Chemistry*, W. A. Benjamin Inc., New York, 1965.

[2] A. Eiger, K. Sikorski and F. Stenger, A bisection method for systems of nonlinear equations *ACM Trans. Math. Software* **10** (1984), 367–377.

[3] G. M. Fikhtengol'ts, *The Fundamentals of Mathematical Analysis*, Vol. [illegible], Pergamon Press, Frankfurt, 1965.

[4] S. Gorn, Maximal convergence intervals and a Gibb's-type phenomenon [illegible] Newton's approximation procedure, *Ann. Math.* **59** (1954), 463–476.

[5] C. Harvey and F. Stenger, A two-dimensional analogue to the method of bisections for solving nonlinear equations, *Quart. Appl. Math.* **33** (1976) 351–368.

[6] W. M. Kincaid, A two-point method for the numerical solution of systems of simultaneous equations, *Quart. Appl. Math.* **18** (1961), 313–324.

[7] J. Moré, Nonlinear generalizations of matrix diagonal dominance with application to Gauss–Seidel iterations, *SIAM J. Numer. Anal.* **9** (1972), 357–378.

[8] J. Ortega and M. Rockoff, Nonlinear difference equations and Gauss–Seidel type iterative methods, *SIAM J. Numer. Anal.* **3** (1966), 497–513.

[9] J. Ortega and W. C. Rheinboldt, *Iterative Solution of Nonlinear Equations in Several Variables*, Academic Press, New York, 1970.

[10] J. Ortega and W. C. Rheinboldt, Local and global convergence of generalized linear iterations,

Numerical Solution of Nonlinear Problems, J. Ortega and W. C. Rheinboldt (eds.), SIAM (1970), 122–143.

[11] A. Ostrowski, *Solution of Equations in Euclidean and Banach Spaces*, Third ed., Academic Press, London, 1973.

[12] W. C. Rheinboldt, *Methods for Solving Systems of Nonlinear Equations*, SIAM, Philadelphia, 1974.

[13] K. Sikorski, Bisection is optimal, *Numer. Math.* **40** (1982), 111–117.

[14] J. F. Traub, *Iterative Methods for the Solution of Equations*, Prentice-Hall Inc., Englewood Cliffs, NJ, 1964.

[15] M. N. Vrahatis, *The Topological Degree for the Generalized Method of Bisection*, Technical Report T.R. No. 6, Department of Mathematics, University of Patras, Patras, October 1981.

[16] M. N. Vrahatis and K. I. Iordanidis, A rapid generalized method of bisection for solving systems of nonlinear equations, *Numer. Math.* **49** (1986), 123–138.

[17] M. N. Vrahatis, Solving systems of nonlinear equations using the nonzero value of the topological degree, *ACM Trans. Math. Software* **14** (1988), in press.