1968

Iterative techniques for systems of nonlinear equations

Charles W. Schelin
The University of Montana

/ - 9

ITERATIVE TECHNIQUES FOR SYSTEMS

OF NONLINEAR EQUATIONS

By

Charles W. Schelin

B.S., University of Minnesota, 1965

Presented in partial fulfillment of the requirements

for the degree of

Master of Arts

UNIVERSITY OF MONTANA

1968

Approved by:

Chairman, Board of Examiners

Dean, Graduate School

NOVEMBER 14, 1968
Date

ACKNOWLEDGEMENTS

I wish to express my gratitude to Professor Charles A. Bryan for his encouragement and guidance during the preparation and writing of this thesis. I also wish to thank Professors Robert P. Banaugh and William M. Myers for their helpful suggestions and critical reading of the manuscript. Finally, I wish to thank my wife, Virginia, for her patience in typing the manuscript.

TABLE OF CONTENTS

INTRODUCTION

A basic problem of elementary algebra is solving a system of two linear equations in two unknowns. The general problem involving a system of n-linear equations in n unknowns can be solved in a variety of ways. Direct methods of solution such as Cramer's rule, Gaussian elimination, or straight forward matrix inversion and multiplication may be employed. Also iterative methods such as the Jacobi, Gauss-Seidel, or overrelaxation techniques may be used to solve the system.

Non-linear systems of equations are usually studied as equations involving nonlinear operators on abstract spaces. For example, a system of three nonlinear homogenous equations in three real unknowns can be thought of as the equation $F(x) = 0$, where $F: R^3 \to R^3$ is composed of three nonlinear functions $f_i: R^3 \to R$, $i = 1, 2, 3$. A solution x^* to the equation $F(x) = 0$ is thus a solution to the system $f_i(x) = 0$, $i = 1, 2, 3$.

In chapter one, the concepts of functional analysis necessary for analyzing operator equations on abstract spaces are developed. Newton's method and related iterative techniques for nonlinear systems are examined for convergence using the unified approach of Rheinboldt and Ortega in chapter two. Finally, in chapter three conditions are found to insure the convergence of the extended Gauss-Seidel iterative technique in R^2.

It is assumed that the reader has a modest knowledge of linear algebra, including the basic vector space properties found in Halmos 7 , chapters one and two.

CHAPTER 1

Concepts of Functional Analysis

The subject matter collected under the title functional analysis includes major portions of analysis, topology, and linear algebra. The development of all the concepts to be used in this paper would provide ample material for a textbook. Thus, only a few basic definitions are stated and the less common results, crucial in the examination of iterative techniques for nonlinear operator equations are developed in detail.

We shall reserve R to denote the real number system, and begin with a listing of basic properties.

Definition 1.1: A vector space is a set V associated with a field $(F, +, \cdot)$ such that the operations $\phi: V \times V \to V$ and $\psi: F \times V \longrightarrow V$, where $\phi(x,y)$ is denoted $x + y$ and $\psi(\alpha, x)$ is denoted by αx, satisfy the following:

1) $(V,+)$ is a commutative group.
2) $\alpha \in F$, $x,y \in V$ implies that $\alpha(x+y) = \alpha x + \alpha y$.
3) $\alpha, \beta \in F$, $x \in V$ implies that $(\alpha + \beta)x = \alpha x + \beta x$.

4) $\alpha, \beta \varepsilon F$, $x \varepsilon V$ implies that $(\alpha\beta)x = \alpha(\beta x)$.

5) $x \varepsilon V$ implies that $1x = x$.

If $F = R$, then V is said to be a real vector space.

Definition 1.2: A real vector space V is said to be n-dimensional if there exists a set $\{ x_1, \ldots, x_n \} \subset V$ such that $c_i \varepsilon R$, $i = 1, \ldots, n$, $\sum_{i=1}^{n} c_i x_i = 0$, implies that $\sum_{i=1}^{n} c_i^2 = 0$, and for any subset $\{ y_1, \ldots, y_{n+1} \}$ of V containing n + 1 elements, there exists $d_i \varepsilon R$, $i = 1, \ldots, n + 1$, such that $\sum_{i=1}^{n+1} d_i x_i = 0$, and $\sum_{i=1}^{n+1} d_i^2 > 0$. A vector space V is said to be finite dimensional if there exists a positive integer n such that V is n-dimensional.

Definition 1.3: Suppose M and N are vector spaces associated with the same field F. Suppose $T:M \rightarrow N$. Then T is said to be a linear operator if $T(\alpha x + \beta y) = \alpha T(x) + \beta T(y)$ for all $x,y \varepsilon M$, $\alpha, \beta \varepsilon F$. In the special case where $N = F = R$, T is said to be a linear functional.

Definition 1.4: A real vector space V is said to be normed if there exists $\phi: V \rightarrow R$, where $\phi(x)$ is denoted by $||x||$, satisfying the following:

1) $x \varepsilon V$ implies that $||x|| \geq 0$ and $||x|| = 0$ if and only if $x = 0$.

2) $x \varepsilon V$, $\alpha \varepsilon R$ implies that $||\alpha x|| = |\alpha| \, ||x||$.

3) $x \varepsilon V$, $y \varepsilon V$ implies that $||x + y|| \leq ||x|| + ||y||$.

Definition 1.5: If M and N are real normed vector spaces, and if $T:M \rightarrow N$ is a linear operator, then T is said to be bounded if there exists $K \varepsilon R$, $K > 0$ such that

(1.51) $||T(x)|| \leq K||x||$ for all $x \varepsilon M$. The least such K satisfying (1.51) is said to be the norm of T and is denoted by $||T||$.

Note that if $||x|| = 1$, then (1.51) becomes $||T(x)|| \leq M$, and hence $||T|| = \sup_{\substack{x \varepsilon M \\ ||x||=1}} ||T(x)||$.

Definition 1.6: A real normed vector space V is said to be complete if whenever $\{x_i\}$, a sequence in V, is such that $\lim_{m,n \to \infty} ||x_m - x_n|| = 0$, then there exists an $x \varepsilon V$ such that $\lim_{n \to \infty} ||x_n - x|| = 0$. i.e., A real normed vector space V is complete if every Cauchy sequence in V is convergent in V.

Definition 1.7: A complete real normed vector space is said to be a Banach space or simply a B-space.

Theorem 1.1: Suppose X and Y are B-spaces. Then we shall let $[X \to Y]$ denote the set of all bounded linear operators mapping X into Y. If we define $||F||$ as above, $(F + G)(x) = F(x) + G(x)$, and $(\alpha F)(x) = \alpha(F(x))$ for all F, $G \varepsilon [X \to Y]$, $\alpha \varepsilon R$, and $x \varepsilon X$, then $[X \to Y]$ is a B-space.

Proof: Clearly $[X \to Y]$ is a real normed vector space. Suppose $\{U_n\}$ is a Cauchy sequence of elements of the space $[X \to Y]$. Then given $\varepsilon > 0$, there exists an $N \varepsilon R^+$ such that $n, m, > N$ implies $||U_n - U_m|| < \varepsilon$. Thus for any fixed $x \varepsilon X$, $||U_n(x) - U_m(x)|| < \varepsilon ||x||$, and so the sequence $\{U_n(x)\}$ of elements of Y is Cauchy. Y is complete, hence there exists

$U(x) = \lim_{n\to\infty} U_n(x)$ for $x \varepsilon X$. U is linear because $U(x + y) = \lim_{n\to\infty} U_n(x + y) = \lim_{n\to\infty} \left(U_n(x) + U_n(y)\right) = U(x) + U(y)$ and $U(\alpha x) = \lim_{n\to\infty} U_n(\alpha x) = \alpha\left(\lim_{n\to\infty} U_n(x)\right) = \alpha U(x)$. Also, $\|U(x) - U_m(x)\| = \lim_{n\to\infty} \|U_n(x) - U_m(x)\| \leq \varepsilon \|x\|$ for $m > N$, so the operator V, where $V(x) = U(x) - U_N(x)$ for $x \varepsilon X$, is an element of the space $[X \to Y]$. Thus $U = V + U_N \varepsilon [X \to Y]$, so $\|U - U_n\| \leq \varepsilon$ for $n \geq N$ and $\{U_n\}$ converges to $U \varepsilon [X \to Y]$. Therefore $[X \to Y]$ is complete, and hence a B-space.

Remarks: 1) In the particular case where $Y = R$, $[X \to R]$ corresponds to the adjoint space X^* of X. (See Halmos [7]).

2) If X is a B-space, then we may define a second operation on $[X \to X]$ as follows: if $U, V \varepsilon [X \to X]$, then $UV = W$ where $W \varepsilon [X \to X]$ such that $W(x) = U(V(x))$ for all $x \varepsilon X$. In this case $\|W(x)\| = \|U(V(x))\| \leq \|U\|\|V(x)\| \leq \|U\|\,\|V\|\,\|x\|$ so $\|W\| \leq \|U\|\|V\|$. With the two operations defined on $[X \to X]$, $[X \to X]$ becomes a ring over R or $[X \to X]$ is a real algebra. In the general case, given $U \varepsilon [Y \to X]$ and $V \varepsilon [X \to Y]$ we shall again define $UV = W$ where $W \varepsilon [X \to X]$ such that $W(x) = U(V(x))$ for all $x \varepsilon X$.

<u>Theorem 1.2:</u> (Banach) Let X be a B-space and let $U \varepsilon [X \to X]$. If $\|U\| \leq q < 1$ then the operation $(I - U)^{-1}$ exists and is an element of $[X \to X]$. Further, $\|(I - U)^{-1}\| \leq \frac{1}{1 - q}$.

Proof: Let $I \epsilon [X \to X]$ be the identity mapping and define $U^0 = I$, $U^n = U^{n-1}U$ for $n = 1, 2, \ldots$. In view of the above remark, we have $\|U^n\| \leq \|U\|^n$ for $n = 0, 1, \ldots$. Then $V = \sum_{j=0}^{\infty} U^j$ is convergent since $\|V\| \leq \sum_{j=0}^{\infty} \|U^j\| < \sum_{j=0}^{\infty} \|U\|^j < \sum_{j=0}^{\infty} q^j = \dfrac{1}{1 - q}$. Further, $V(I - U) = \sum_{j=0}^{\infty} U^j (I - U) = \sum_{j=0}^{\infty} U^j - \sum_{j=0}^{\infty} U^{j+1} = U^0 = I$, and similarly $(I - U)V = I$. Hence $V = (I - U)^{-1}$, so $(I - U)^{-1}$ exists and $\|(I - U)^{-1}\| \leq \dfrac{1}{1 - q}$.

In view of the fact that every n-dimensional real vector space is isomorphic to R^n (see Halmos p. 15, [7]), we shall pay particular attention to this space. If $x \epsilon R^n$, then $x = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$ and we shall define $\|x\| = \max_{i=1,\ldots,n} |x_i|$. With this norm, usually called the Tchebycheff norm, R^n is a B-space. Recalling that $T \epsilon [R^n \to R^m]$ implies that T may be represented by an $n \times m$ matrix, (see Halmos 7), we have $\|T\| = \left\| \begin{pmatrix} a_{11} & \cdots & a_{1m} \\ \vdots & & \vdots \\ a_{n1} & \cdots & a_{nm} \end{pmatrix} \right\|$

$= \max_{i=1,\ldots,n} \sum_{j=1}^{m} |a_{ij}|$.

<u>Definition 1.8</u>: Suppose X and Y are B-spaces, A is an open subset of X, $x_0 \epsilon A$, and $T: A \to Y$. Suppose that there exists $U \epsilon [X \to Y]$ such that for every $x \epsilon X$,

(1.81) $\lim_{t \to 0} \dfrac{T(x_0 + tx) - T(x_0)}{t} = U(x)$. Then the linear operator U is said to be the Gateaux derivative of T at the point x_0, denoted $U = T'(x_0)$. The element $U(x)$ is called the Gateaux differential. If the convergence relationship is satisfied uniformly for all $x \in X$ such that $||x|| = 1$, then U and U(x) are called the Frechet derivative and Frechet differential respectively.

<u>Theorem 1.3</u>: The operator T has a Frechet derivative at the point x_0 if and only if there exists $U \in [X \to Y]$ such that for every $\varepsilon > 0$, there exists a $\delta > 0$, such that $\Delta x \in X$, $||\Delta x|| < \delta$ implies $||T(x_0 + \Delta x) - T(x_0) - U(\Delta x)|| \leq \varepsilon||\Delta x||$.

Proof: Let $\Delta x = tx$ where $||x|| = 1$. Then $||\Delta x|| = t$ and statement (1.81) is equivalent to

(1.82) $\lim\limits_{||\Delta x|| \to 0} \dfrac{T(x_0 + \Delta x) - T(x_0) - U(\Delta x)}{||\Delta x||} = 0$. Now statement (1.82) converges uniformly if and only if given $\varepsilon > 0$ there exists a $\delta > 0$ such that $\Delta x \in X$, $||\Delta x|| < \delta$ implies $\left\|\dfrac{T(x_0 + \Delta x) - T(x_0) - U(\Delta x)}{||\Delta x||}\right\| \leq \varepsilon$, or $||T(x_0 + \Delta x) - T(x_0) - U(\Delta x)|| \leq \varepsilon||\Delta x||$, yielding the desired result.

<u>Corollary 1.31</u>: If the operator F has a Frechet derivative at the point x_0, then F is continuous at x_0.

Proof: $F'x_0 \in [X \to Y]$ implies that there exists $M \in R^+$ such that $||F'x_0(x)|| \leq M||x||$ for all $x \in X$. Given $\varepsilon > 0$, there exists $\delta > 0$ such that $\delta < \dfrac{\varepsilon}{\varepsilon + M}$, and $\Delta x \in X$, $||\Delta x|| < \delta$ implies $||F(x_0 + \Delta x) - F(x_0)|| - ||F'x_0(\Delta x)|| \leq ||F(x_0 + \Delta x) - F(x_0) - F'x_0(\Delta x)|| \leq \varepsilon||\Delta x||$. Then $||F(x_0 + \Delta x) - F(x_0)|| \leq \varepsilon||\Delta x|| + M||\Delta x|| = (\varepsilon + M)||\Delta x|| \leq \varepsilon$.

A useful tool of real analysis is the mean value theorem for differential calculus. The next theorem gives the generalized form of this result.

Theorem 1.4: Suppose X and Y are B-spaces, $F: X \longrightarrow Y$, and F is Gateaux differentiable on the convex subset A of X. Then if $x, y \in A$,

(1.41) $$\|Fy - Fx\| \leq \|y - x\| \sup_{0<\theta<1} \|F(\theta x + (1-\theta)y)\|.$$

Proof: Let g be any functional in Y^* and let $\phi(t) = g\big(F(x + t(y - x))\big)$. Then $\phi(t)$ has a derivative in the interval $[0, 1]$ because: $\lim_{\Delta t \to 0} \frac{\phi(t + \Delta t) - \phi(t)}{\Delta t} =$

$$\lim_{\Delta t \to 0} g\left(\frac{F\big(x + t(y - x) + \Delta t(y - x)\big) - F(x + t(y - x))}{\Delta t}\right) =$$

$$g\left(\lim_{\Delta t \to 0} \frac{F(x + t(y - x) + \Delta t(y - x)) - F(x + t(y - x))}{\Delta t}\right) =$$

$g\big(F'(x + t(y - x))(y - x)\big)$. i.e., $\phi'(t) = g\big(F'(x + t(y - x)\big)$ for $t \in [0, 1]$.

Now, applying the mean value theorem of differential calculus to $\phi(t)$, we get $\phi(1) - \phi(0) = \phi'(\theta)$ for some θ, $0 < \theta < 1$. Thus substituting we get $g\big(F(y) - F(x)\big) = g\big(F'(x + \theta(y - x))(y - x)\big)$, so $\|g\big(F(y) - F(x)\big)\| \leq \|g\| \sup_{0<\theta<1} \|F(x + \theta(y - x))\| \|y - x\|$. By invoking a corollary of the Hahn-Banach theorem (See Vainberg [12], p. 11) we may choose g to be that functional in Y^* such that $\|g\| = 1$ and $g\big(F(y) - F(x)\big) = \|F(y) - F(x)\|$. Then upon substitution we obtain $\|F(y) - F(x)\| \leq \|y - x\| \sup_{0<\theta<1} \|F(\theta x + (1 - \theta)y)\|$.

<u>Corollary 1.41</u>: Suppose X and Y are B-spaces, $F: D \subset X \to Y$, F is Gateaux differentiable on some convex set $D_0 \subset D$, and $\|F'(x)\| \leq \gamma$ for $x \in D_0$. Then $\|F(x) - F(y)\| \leq \gamma \|x - y\|$ for all $x, y \in D_0$.

Proof: By the mean value theorem we know that $\|F(x) - F(y)\| \leq \sup_{0<t<1} \|F'(ty + (1 - t)x)\| \|x - y\|$. D_0 is a convex set, $x, y \in D_0$, thus $ty + (1 - t)x \in D_0$ for all t such that $0 \leq t \leq 1$. Therefore $\sup_{0<t<1} \|F'(ty + (1 - t)x)\| \leq \gamma$, so $\|F(x) - F(y)\| \leq \gamma \|x - y\|$.

To obtain conditions for convergence in chapter three we shall be dealing with $F'(x)$ where $F: R^2 \to R^2$. Therefore it is necessary to further investigate $F'(x_0)$ where $F: R^n \to R^m$.

Let $x = \begin{pmatrix} \xi_1 \\ \vdots \\ \xi_n \end{pmatrix} \in R^n$ and $x_0 = \begin{pmatrix} \xi_1^{(0)} \\ \vdots \\ \xi_n^{(0)} \end{pmatrix} \in R^n$. If $F: R^n \to R^m$, and if $\begin{pmatrix} \eta_1 \\ \vdots \\ \eta_n \end{pmatrix} = F(x)$, then there exist ϕ_i, $i = 1, \ldots, m$, such that $\phi_i : R^n \to R$ and $\eta_i = \phi_i(\xi_1, \ldots, \xi_n)$ for $i = 1, \ldots, m$. If $F'(x_0)$ exists then $F'(x_0) \in [R^n \to R^m]$, and if $F'(x_0)(x) = \begin{pmatrix} \zeta_1 \\ \vdots \\ \zeta_m \end{pmatrix}$ then $F'(x_0)$ may be represented by an $n \times m$ matrix $\begin{pmatrix} a_{11} & \cdots & a_{1m} \\ \vdots & & \vdots \\ a_{n1} & \cdots & a_{nm} \end{pmatrix}$ such that

$\zeta_i = \sum_{j=1}^{n} a_{ij} \xi_i$ for $i = 1, \ldots, m$. Now substituting into $\lim_{t \to 0} \dfrac{F(x_0 + tx) - F(x_0)}{t} = F'(x_0)(x)$ we obtain

$$\lim_{t \to 0} \frac{\phi_i(\xi_1^{(0)} + t\xi_1, \ldots, \xi_n^{(0)} + t\xi_n) - \phi_i(\xi_1^{(0)}, \ldots, \xi_n^{(0)})}{t} = \sum_{K=1}^{n} a_{iK} \xi_K$$

for $i = 1, \ldots, m$, and for all $x = \begin{pmatrix} \xi_1 \\ \vdots \\ \xi_n \end{pmatrix} \epsilon R^n$. If in particular we let x_i be the vector with i-th component equal to one and all other components equal to zero, for $i = 1, \ldots, n$, we find that the functions ϕ_i, $i = 1, \ldots, m$ have partial derivatives with respect to ξ_i, $i = 1, \ldots, n$. Hence we have

$$\frac{\partial \phi_i(\xi_1^{(0)}, \ldots, \xi_n^{(0)})}{\partial \xi_k} = \lim_{t\to 0} \frac{\phi_i((\xi_1^{(0)}, \ldots, \xi_n^{(0)}) + (0, \ldots, 0, t\xi_k, 0, \ldots, 0))}{t}$$

$$- \lim_{t\to 0} \frac{\phi_i(\xi_1^{(0)}, \ldots, \xi_n^{(0)})}{t} = a_{ik} \quad \text{for } i = 1, \ldots, m,\ k = 1, \ldots, n .$$

i.e., $F'(x_0)$ is represented by the matrix of partials of $\phi_1, \ldots, \phi_m$. This matrix is called the Jacobian matrix of F at x_0.

Recalling the definition of the Gateaux derivative, it is clear that if X and Y are B-spaces and $F: X \to Y$ such that F exists on the open subset A of X, then F may be thought of as a mapping from A into $[X \to Y]$. With this in mind we make the following generalization.

Definition 1.9: Suppose $F: A \subset X \to Y$ and F has a Gateaux derivative on A, where X and Y are B-spaces and A is an open subset of X. Suppose $x_0 \epsilon A$ and there exists $U \epsilon [X \to [X \to Y]]$ such that

(1.91) $\lim_{t\to 0} \frac{F'(x_0 + tx) - F'(x_0)}{t} = U(x)$ for all $x \epsilon X$. Then

U is said to be the second Gateaux derivative of F at x_0, denoted $F''(x_0)$. Similarly if F has a Frechet derivative on

A, and the convergence in (1.91) is uniform for all $x \varepsilon X$ of unit norm, then $U = F''(x_0)$ is said to be the second Frechet derivative of F at x_0.

In order to examine elements of the space $[X \to [X \to Y]]$ we introduce the concept of a bilinear operator, and show that any element of $[X \to [X \to Y]]$ may be regarded as a bounded bilinear operator.

<u>Definition 1.10:</u> Suppose X and Y are B-spaces and suppose $T: X \times X \to Y$. Then T is said to be a bounded bilinear operator if:

1) Given $x_1, x_2, \bar{x}_1, \bar{x}_2 \varepsilon X$ and $\alpha, \beta \varepsilon R$, $T(\alpha x_1 + \beta x_2, z) = \alpha T(x_1, z) + \beta T(x_2, z)$ for all $z \varepsilon X$ and $T(w, \alpha \bar{x}_1 + \beta \bar{x}_2) = \alpha T(w, \bar{x}_1) + \beta T(w, \bar{x}_2)$ for all $w \varepsilon X$.

2) There exists an $M \varepsilon R$, $M > 0$ such that $||T(x_1, x_2)|| \leq M||x_1||\,||x_2||$ for all $x_1, x_2 \varepsilon X$.

As with linear operators, we define $||T||$ to be the least M satisfying 2). If $||x_1|| = ||x_2|| = 1$, then $||T(x_1, x_2)|| \leq M$. So $||T|| = \sup\limits_{\substack{x_1, x_2 \varepsilon X \\ ||x_1||=||x_2||=1}} ||T(x_1, x_2)||$. Similarly, if we define $(T_1 + T_2)(x_1, x_2) = T_1(x_1, x_2) + T_2(x_1, x_2)$ then $[X^2 \to Y] = \{T \mid T \text{ is a bounded bilinear operator mapping } X^2 \text{ into } Y\}$ is a B-space.

In the finite dimensional case we find that a bilinear operator may be represented by a finite collection of matrices. This can be seen by letting X and Y be B-spaces of dimension n and m respectively, and letting $T \varepsilon [X^2 \to Y]$. Suppose $x_i \varepsilon X$ has zero components, except for the i-th entry which is one, for $i = 1,\ldots,n$, and suppose $T(x_i, x_j) = (a_{ij}^{(1)},\ldots,a_{ij}^{(m)})$ for $i,j = 1,\ldots,n$. Then for $x = \begin{pmatrix} \xi_1 \\ \vdots \\ \xi_n \end{pmatrix} \varepsilon X$, $\overline{x} = \begin{pmatrix} \overline{\xi}_1 \\ \vdots \\ \overline{\xi}_n \end{pmatrix} \varepsilon X$ we have $T(x,\overline{x}) =$

$T(\sum_{i=1}^n \xi_i x_i, \sum_{j=1}^n \overline{\xi}_j x_j) = \sum_{i,j=1}^n \xi_i \overline{\xi}_j T(x_i, x_j)$. So if $T(x,\overline{x}) = z$, where $z = \begin{pmatrix} \gamma_1 \\ \vdots \\ \gamma_m \end{pmatrix}$, then $\gamma_k = \sum_{i,j=1}^n a_{ij}^{(k)} \xi_i \overline{\xi}_j$ for $k = 1,\ldots,m$.

Clearly, whatever collection of m, n×n matrices $(a_{ij}^{(k)})$, $i,j = 1,\ldots,n$, $k = 1,\ldots,m$ we consider, the operation T defined above will be bilinear. Thus every bilinear operation may be represented by a set of m matrices.

Using the Tchebycheff norm on X and Y we see that $|\gamma_k| = |\sum_{i=1}^n \sum_{j=1}^n a_{ij}^{(k)} \xi_i \overline{\xi}_j$. So if $||x|| = ||\overline{x}|| = 1$, then $|\gamma_k| \leq |\sum_{i=1}^n \sum_{j=1}^n a_{ij}^{(k)}|$ and hence $||T|| = \max_{k=1,\ldots,m} \sum_{i=1}^n \sum_{j=1}^n |a_{ij}^{(k)}|$.

<u>Theorem 1.5</u>: There exists a norm preserving isomorphism between $[X \to [X \to Y]]$ and $[X^2 \to Y]$.

Proof: Define $\psi: [X \to [X \to Y]] \longrightarrow [X^2 \to Y]$ by $\psi(U) = T$ where $U \varepsilon [X \to [X \to Y]]$, and $x' \varepsilon X$ implies that $U(x') = Vx' \varepsilon [X \to Y]$ and $T(x,x') = Vx'(x) \varepsilon [X^2 \to Y]$. Clearly ψ is additive and homogeneous. Suppose $T \varepsilon [X^2 \to Y]$. Then there exists $T^* \varepsilon [X \to Y]$ such that $T^*x'(x) = T(x,x')$ for x' fixed and for all $x \varepsilon X$. Let $W(x') = T^*x'$ for all $x' \varepsilon X$. Then $\psi(W) = T$, so ψ is onto. Now $||T(x,x')|| = ||Vx'(x)|| \leq ||Vx'|| \, ||x|| \leq ||U|| \, ||x|| \, ||x'||$, so $||T|| \leq ||U||$.

Also, $\|U(x')\| = \sup_{\substack{x\varepsilon X\\ \|x\|=1}} \|Vx'(x)\| = \sup_{\substack{x\varepsilon X\\ \|x\|=1}} \|T(x,x')\| \leq \|T\|\|x'\|,$

which implies $\|U\| \leq \|T\|$. Hence $\|U\| = \|\psi(U)\|$, and so ψ is norm preserving and one-to-one. Therefore, ψ is an isomorphism.

Therefore, the second Gateaux derivative of an operator F mapping X into Y, being an element of $[X \longrightarrow [X \longrightarrow Y]]$, may be thought of as an element of $[X^2 \longrightarrow Y]$. Regarding $F''(x_0)$ as a bilinear operator we find that for any $x, x' \varepsilon X$,

(1.92) $F''(x_0)(x,x') = \lim_{t\to 0} \dfrac{F'(x_0 + tx')x - F'(x_0)x}{t}$. As before, if the convergence in (1.92) is uniform, then $F''(x_0)$ is the second Frechet derivative of F at x_0. Again we are concerned with the form of $F''(x_0)$ in the case where X and Y are finite dimensional.

Let $F:X\longrightarrow Y$ where X and Y are B-spaces of dimension n and m respectively, and let $F''(x_0)$ exist and be defined as a bilinear operator represented by the m matrices $(a_{ij}^{(k)})$, $i,j = 1,\ldots,n$, $k = 1,\ldots,m$. Further let $x = \begin{pmatrix}\xi_1\\ \vdots\\ \xi_n\end{pmatrix} \varepsilon X$, $x' = \begin{pmatrix}\xi_1'\\ \\ \xi_n'\end{pmatrix} \varepsilon X$, and $F(x) = \begin{pmatrix}\phi_1(\xi_1,\ldots,\xi_n)\\ \vdots\\ \phi_m(\xi_1,\ldots,\xi_n)\end{pmatrix} \varepsilon Y$. Then substituting our earlier results for first derivatives into (1.92) we obtain:

(1.93) $\sum_{i,j=1}^{n} a_{ij}^{(k)} \xi_i \xi_j' = \lim_{t\to 0}\left[\dfrac{\sum_{i=1}^{n} \frac{\partial\phi_k}{\partial\xi_j}(\xi_1^{(0)} + t\xi_1' + \ldots + t\xi_n')\xi_i}{t} - \dfrac{\sum_{i=1}^{n} \frac{\partial\phi_k}{\partial\xi_i}(\xi_1^{(0)},\ldots,\xi_n^{(0)})\xi_i}{t}\right]$, for $k = 1,\ldots,m$. Now if we let x and x' be the elements with all zero entries except the i-th and j-th components respectively, which are equal to 1, we find that ϕ_k, $k = 1, \ldots, m$ have second order partial derivatives, and

(1.94) $a_{ij}^{(k)} = \dfrac{\partial^2 \phi_k(\xi_1^{(0)}, \ldots, \xi_n^{(0)})}{\partial \xi_i \partial \xi_j}$ for $i, j = 1, \ldots, n$, $k = 1, \ldots, m$.

A final concept in the general theory of operators on B-spaces of which we shall make limited use is that of integration.

Definition 1.11: Suppose $F: [a, b] \subset R \to X$, X a B-space, and suppose $a = t_0 < t_1 < t_2 \ldots < t_n = b$, $\xi_k \varepsilon [t_k, t_{k+1}]$, $k = 0, 1, \ldots, n-1$, and $\lambda = \max_{k=0,\ldots,n-1} |t_{k+1} - t_k|$, then $\int_a^b F(t)dt = \lim_{\lambda \to 0} \sum_{k=0}^{n-1} F(\xi_k)(t_{k+1} - t_k)$ if this limit exists.

Note that $\left\|\int_a^b F(t)dt\right\| = \lim_{\lambda \to 0} \left\|\sum_{k=0}^{n-1} F(\xi_k)(t_{k+1} - t_k)\right\| \leq \lim_{\lambda \to 0} \sum_{k=0}^{n-1} \|F(\xi_k)\|(t_{k+1} - t_k) = \int_a^b \|F(t)\|dt$.

Definition 1.12: Suppose $T: D \subset X \to Y$ where X and Y are B-spaces and D is a convex subset of X. If $x_0, x_1 \varepsilon D$, and t_k, ξ_k, $k = 1, \ldots, n-1$, are as in (1.11), we define:

$$\int_{x_0}^{x_1} T(x)dx = \int_0^1 T(x_0 + t(x_1 - x_0))(x_1 - x_0)dt$$

(1.121) $$= \lim_{\lambda \to 0} \sum_{k=0}^{n-1} T(x_0 + \xi_k(x_1 - x_0))(x_1 - x_0)(t_{k+1} - t_k)$$

if this limit exists.

Clearly if T is continuous the limit (1.121) exists and the integral represents an element of Y.

With the above definition of integration we can derive an analog of the fundamental theorem of calculus.

<u>Theorem 1.6</u>: Suppose $F: D \subset X \longrightarrow Y$ where X and Y are B-spaces, and suppose $F'(x)$ exists and is continuous on the convex set $D_0 \subset D$. Then for $x_0, x_1 \varepsilon D_0$, $\int_{x_0}^{x_1} F'(x)dx$ exists and $\int_{x_0}^{x_1} F'(x)dx = F(x_1) - F(x_0)$.

Proof: Let $x_k = x_0 + t_k(x_1 - x_0)$, $\Delta x_k = (t_{k+1} - t_k)\cdot(x_1 - x_0)$, $\overline{x}_k = x_0 + \xi_k(x_1 - x_0)$, and $\lambda = \max_{k=0,1,\ldots,n-1} |t_{k+1} - t_k|$. Then using definition 1.12 we have

$$\left\|F(x_1) - F(x_0) - \int_{x_0}^{x_1} F'(x)dx\right\| = \left\|F(x_1) - F(x_0) - \lim_{\lambda \to 0} \sum_{k=0}^{n-1} F'(\overline{x}_k)\Delta x_k\right\| = \lim_{\lambda \to 0} \left\| F(x_1) - F(x_0) - \sum_{k=0}^{n-1} F'(\overline{x}_k)\Delta x_k\right\|.$$

Noting that $F(x_1) - F(x_0) = \sum_{k=0}^{n-1}\Big(F(x_{k+1}) - F(x_k)\Big)$ we have

$$\lim_{\lambda \to 0} \left\|F(x_1) - F(x_0) - \sum_{k=0}^{n-1} F'(\overline{x}_k)\Delta x_k\right\| =$$

$$\lim_{\lambda \to 0}\left\|\sum_{k=0}^{n-1} F(x_{k+1}) - F(x_k) - F'(\overline{x}_k)\Delta x_k\right\| \leq$$

$$\lim_{\lambda \to 0} \|x_1 - x_0\| \sum_{k=0}^{n-1} (t_{k+1} - t_k) \sup_{0<\theta<1} \|F'(x_k + \theta\Delta x_k)$$

$- F'(\overline{x}_k)\|$, making use of the mean value theorem. Now since F' is continuous and $\lim_{\lambda \to 0} \sup_{0<\theta<1} \|x_k + \theta\Delta x_k - \overline{x}_k\| = 0$ we have

$\lim_{\lambda \to 0} \sup_{0<\theta<1} \|F'(x_k + \theta\Delta x_k - \overline{x}_k\| = 0$, giving the desired result.

Finally, we use some of the concepts developed in this chapter to verify the following lemmas which will aid in investigating the convergence of iterative techniques in chapter 2.

<u>Lemma 1.1</u>: Suppose X and Y are B-spaces, $F: D \subset X \to Y$, F is Gateaux differentiable on some convex set $D_0 \subset D$, and $\|F'(y) - F'(x)\| \leq \gamma$ for $x, y \in D_0$. Then $\|F(y) - F(x) - F'(z)(y - x)\| \leq \gamma\|y - x\|$ for all $x, y, z \in D_0$.

Proof: Using the mean value theorem we have $\|F(x) - F(y)\| \leq \sup_{0<t<1} \|F'(ty + (1 - t)x)\|\|x - y\|$. Also, $F'(z) \in [X \to Y]$ for all $z \in D_0$. Replacing F by $F - F'(z)$ in the above inequality we get $\|F(x) - F(y) - F'(z)(y - x)\| \leq \sup_{0<t<1} \|\left(F - F'(z)\right)'(ty + (1 - t)x)\|\|x - y\|$. Let $w = ty + (1 - t)x$, then $\left(F - F'(z)\right)'(w) = F'(w) - \left(F'(z)\right)'(w)$ directly from definition 1.8. Further, since $F'(z) \in [X \to Y]$,

$$\left(\left(F'(z)\right)'(w)\right)(\xi) = \lim_{t \to 0} \frac{F'(z)(w + t\xi) - F'(z)(w)}{t} = F'(z)(\xi)$$

for all $\xi \in X$. Thus $\left(F'(z)\right)'(w) = F'(z)$, so we have $\|F(x) - F(y) - F'(z)(y - x)\| \leq \sup_{0<t<1} \|F(ty + (1 - t)x) - F'(z)\|\|x - y\| \leq \gamma\|x - y\|$.

<u>Lemma 1.2</u>: Suppose X and Y are B-spaces, $F: D \subset X \to Y$, F has a Gateaux derivative on some convex set $D_0 \subset D$, and $\|F'(y) - F'(x)\| \leq \gamma\|y - x\|$ for $x, y \in D_0$. Then $\|F(y) - F(x) - F'(x)(y - x)\| \leq \frac{1}{2}\gamma\|y - x\|$ for all $x, y \in D_0$.

Proof: Clearly $\|F'(y) - F'(x)\| \leq \gamma\|y - x\|$ implies that F' is continuous on D_0. Thus $F(y) - F(x) =$

$\int_x^y F'(x)dx = \int_0^1 F'(\theta y + (1-\theta)x)(y - x)d\theta$. Therefore,
$\|F(y) - F(x) - F'(x)(y - X)\| = \left\|\int_0^1 [F'(\theta y + (1-\theta)x) - F'(x)] \circ (y - x)d\theta\right\| \leq \|y - x\|\int_0^1 \|F'(\theta y + (1-\theta)x) - F'(x)\|d\theta \leq \gamma\|y - x\|^2\int_0^1 \theta d\theta = \frac{1}{2}\gamma\|y - x\|^2$.

CHAPTER 2

Solving Nonlinear Operator Equations

In this paper we are primarily concerned with operator equations of the form

(1) $F(x) = 0$, where $F: X \longrightarrow Y$, X and Y are B-spaces. As was noted in the introduction, if X and Y are of dimension n and m respectively, then a solution x^* of $F(x) = 0$ may be regarded as a solution to a system of m equations in n unknowns.

Another equation of interest is

(2) $T(x) = x$, where $T: X \longrightarrow X$, X a B-space. If x^* is a solution to $T(x) = x$, then x^* is said to be a fixed point of the operator T. Clearly, if a solution to the equation $T(x) = x$ is known, then a solution to $F(x)$ is known, where $F(x) = x - T(x)$. Keeping this in mind we shall examine operator equations of forms (1) and (2).

Suppose that X and Y are B-spaces and $F: A \subset X \longrightarrow Y$, where A is an open subset of X. Suppose further that F has a continuous Gateaux derivative in A, and there exists an element $x^* \varepsilon A$ such that $F(x^*) = 0$. Then for $x_0 \varepsilon A$, $F(x_0) = F(x_0) - F(x^*)$ can be approximated by $F'(x_0)(x_0 - x^*)$, so it is reasonable to assume that a solution to the equation $F'(x_0)(x_0 - x) = F(x_0)$ is close to x^*. If x_1 is such a solution and $[F'(x_0)]^{-1}$ exists, we have $x_1 = x_0 - [F'(x_0)]^{-1} F(x_0)$. Continuing this process, assuming $[F'(x_k)]^{-1}$ exists, we obtain a sequence $\{x_k\}$ defined by

(3) $x_{k+1} = x_k - [F'(x_k)]^{-1} F(x_k)$. This iteration process for constructing x_k is called Newton's method. In practice $[F'(x_k)]^{-1}$ may be difficult to compute, or may not even exist. For this reason a cruder sequence of approximation, $\{\bar{x}_k\}$, to x^*, defined by

(4) $\bar{x}_{k+1} = \bar{x}_k - [F'(\bar{x}_0)]^{-1} F(\bar{x}_k)$, is often constructed. The iterative method in (4) is called the modified Newton method. A wealth of variations on Newton 's method have been proposed. A few of these Newton like techniques have been studied by Dennis [5], Altman [1], and Ben-Israel [2].

If either the sequence $\{x_k\}$ or $\{\bar{x}_k\}$, constructed above, converges to an element x^*, and if F is continuous, then x^* is a solution to $F(x) = 0$. We shall therefore devote the remainder of chapter two to determining conditions which insure the convergence of sequences of iterates constructed by (3), (4), or related techniques. The general approach to the convergence of iterative techniques used in this chapter is essentially that of Rheinboldt [11].

<u>Definition 2.1</u>: Suppose $\{x_k\}$ is a sequence in the B-space X. Then a real non-negative sequence $\{t_k\}$ is said

to majorize $\{x_k\}$ if $||x_{k+1} - x_k|| \leq t_{k+1} - t_k$ for $k = 0, 1, \ldots$.

Noting that any majorizing sequence is non-decreasing we make the following observation:

Theorem 2.1: If $\{t_k\}$ majorizes a sequence $\{x_k\}$ in a B-space X and if $\{t_k\}$ is convergent, then $\{x_k\}$ is convergent.

Proof: $||x_m - x_n|| \leq \sum_{j=n}^{m-1} ||x_{j+1} - x_j|| \leq \sum_{j=n}^{m-1} (t_{j+1} - t_j) = t_m - t_n$. Thus the convergence of $\{t_k\}$ implies that $\{x_k\}$ is Cauchy. But X is complete, hence $\{x_k\}$ is convergent.

Definition 2.2: Suppose X is a B-space and D is a closed subset of X. Then $T:D \rightarrow D$ is said to be a contraction mapping if there exists $\alpha \varepsilon R$, $0 < \alpha < 1$, such that $||T(x) - T(x')|| \leq \alpha ||x - x'||$ for all $x, x' \varepsilon D$.

Theorem 2.2: Suppose D is a closed subset of the B-space X, and $T:D \rightarrow D$ is a contraction mapping. Then there exists a unique fixed point x^* of T in D.

Proof: Choose $x_0 \varepsilon D$ and construct $\{x_k\}$ by the successive approximations $x_{k+1} = T(x_k)$. Let $t_0 = 0$ and $t_k = \sum_{j=1}^{k} \alpha^{j-1} ||x_1 - x_0||$, for $k \geq 1$, where α is as in definition (2.2). We now show inductively that $\{t_k\}$ majorizes $\{x_k\}$. If $k = 1$, then $||x_1 - x_0|| \leq ||x_1 - x_0|| = t_1 - t_0$. Assuming $||x_k - x_{k-1}|| \leq t_k - t_{k-1}$ we find that

$||x_{k+1} - x_k|| = ||F(x_k) - F(x_{k-1})|| \leq \alpha ||x_k - x_{k-1}|| =$
$\alpha(\sum_{j=1}^{K} \alpha^{j-1} ||x_1 - x_0|| - \sum_{j=1}^{K-1} \alpha^{j-1}||x_1 - x_0||) = \alpha(\alpha^{K-1}||x_1 - x_0||) =$
$\sum_{j=1}^{K+1} \alpha^{j-1} ||x_1 - x_0|| - \sum_{j=1}^{K} \alpha^{j-1}||x_1 - x_0|| = t_{k+1} - t_k$. Now
$\lim_{t\to\infty} t_k = \frac{\alpha}{1-\alpha}||x_1 - x_0||$. Therefore $\{x_k\}$ converges, say to $x^*\varepsilon D$. Noting that T is a contraction map which implies T is continuous we have $0 = \lim_{k\to\infty} ||x^* - T(x_k)|| =$ $||x^* - T(x^*)||$. Therefore $T(x^*) = x^*$.

Suppose $\overline{\overline{x}}\varepsilon A$ and $T(\overline{\overline{x}}) = \overline{\overline{x}}$. Then $||\overline{\overline{x}} - x^*|| =$ $||F(\overline{\overline{x}}) - F(x^*)|| \leq \alpha||\overline{\overline{x}} - x^*||$ which implies that $||\overline{\overline{x}} - x^*|| = 0$ and $\overline{\overline{x}} = x^*$. Hence x^* is unique.

<u>Corollary 2.21</u>: Suppose $T:A\subset X\to X$ where X is a B-space, and T has a Gateaux derivative on $A_0\subset A$, a closed convex set. If

1) $T(A_0) \subset A_0$ and

2) $\sup_{x\varepsilon A_0} ||T'(x)|| = \alpha < 1$,

then there exists a unique fixed point x^* of T into A_0.

Proof: Suppose $x_1, x_2\varepsilon A_0$. Then by the mean value theorem, $||T(x_2) - T(x_1)|| \leq \sup_{0<\theta<1} ||T'(\theta x_2 + (1-\theta)x_1)|| \circ$ $||x_2 - x_1|| \leq \alpha||x_2 - x_1||$. So T is a contraction mapping and by theorem (2.2), T has a unique fixed point $x^*\varepsilon A$.

Recalling the relationship between equations (1) and (2) we obtain the following result.

<u>Corollary 2.22</u>: Suppose $F: X \longrightarrow X$, X a B-space, F continuous, and suppose $T(x) = x - [F'(x)]^{-1} F(x)$ exists on $A \subset X$ and satisfies the conditions of theorem (2.2). Then there exists a unique $x^* \varepsilon A$ such that $F(x^*) = 0$.

Before the concept of majorizing sequences can be utilized, a method is required to obtain a majorizing sequence $\{t_k\}$ for a given sequence $\{x_k\}$. We have found such a majorizing sequence when $\{x_k\}$ is constructed by $x_{k+1} = G(x_k)$, $k = 0, \ldots,$ and G is a contraction mapping, in theorem 2.2. We shall formulate two lemmas whose proof depends upon the construction of majorizing sequences in a more general setting. To simplify notation we make use of a definition of Rheinboldt [11].

<u>Definition 2.3</u>: A function $\phi: Q \subset R^n \longrightarrow R$ is said to be of class $\Gamma^n(Q)$ if it has the following properties:

1) The domain $Q = J_1 \times J_2 \times \ldots \times J_n$ where $J_i = [0, b_i)$, $b_i \leq \infty$, for $i = 1, \ldots, n$.

2) ϕ is non-negative and isotone. i.e., if $(U_1^{(i)}, \ldots, U_n^{(i)}) \varepsilon Q$, for $i = 1, 2$, and $U_j^{(1)} < U_j^{(2)}$, for $j = 1, \ldots, n$, then $0 \leq \phi(U_1^{(1)}, \ldots, U_n^{(1)}) < \phi(U_1^{(2)}, \ldots, U_n^{(2)})$.

<u>Lemma 2.1</u>: Let X be a B-space and G be a continuous operator such that $G: D \subset X \longrightarrow X$. Suppose there exists a function $\phi \varepsilon \Gamma^2(Q)$ and a point $x_0 \varepsilon D$ such that on some set $D_0 \subset D$ $\|G(y) - G(x)\| \leq \phi(\|y - x\|, \|x - x_0\|) = \phi(u, v)$ for all $x, y \varepsilon D_0$.

Further suppose there exists a continuous function $\psi: [0,\infty) \longrightarrow R$ such that $\phi(u - v, v) = \psi(u) - \psi(v)$, and let the sequence $\{t_k\}$ defined by $t_{k+1} = \Psi(t_k)$, $k = 0, 1, \ldots$, $t_0 = 0$, $t_1 = ||x_1 - x_0||$, exist and be such that $t_{k+1} \geq t_k$, for $k = 0, 1, \ldots$, and $\lim_{k\to\infty} t_k = t^*$. Finally, suppose that $\overline{S}(x_0, t^*) \subset D_0$. Then G has a unique fixed point $x^* \varepsilon \overline{S}(x_0, t^*)$.

Proof: Define $\{x_k\}$ by $x_{k+1} = Gx_k$ for $k = 0, 1, \ldots$. We shall show by induction that $x_k \varepsilon \overline{S}(x_0, t^*)$ and $||x_k - x_{k-1}|| \leq t_k - t_{k-1}$ for $k = 1, 2, \ldots$.

For $k = 1$, $||x_1 - x_0|| \leq ||x_1 - x_0|| - 0 = t_1 - t_0 \leq t^*$, so $x_1 \varepsilon \overline{S}(x_0, t^*)$ and $||x_1 - x_0|| \leq t_1 - t_0$. Assume that $x_k \varepsilon \overline{S}(x_0, t^*)$ and $||x_k - x_{k-1}|| \leq t_k - t_{k-1}$. Then $x_{k+1} = Gx_k$ is defined, and $||x_{k+1} - x_k|| = ||Gx_k - Gx_{k-1}|| \leq \phi(||x_k - x_{k-1}||, ||x_{k-1} - x_0||) \leq \phi(t_k - t_{k-1}, t_{k-1}) = \phi(u - v, v)$, where $u = t_k$ and $v = t_{k-1}$, so $\phi(t_k - t_{k-1}, t_{k-1}) = \psi(t_k) - \psi(t_{k-1}) = t_{k+1} - t_k$. Thus $||x_{k+1} - x_k|| \leq t_{k+1} - t_k$. Also, $||x_{k+1} - x_0|| \leq \sum_{j=0}^{k} ||x_{j+1} - x_j|| \leq \sum_{j=0}^{k} t_{j+1} - t_j = t_{k+1} - t_0 = t_{k+1} \leq t^*$. Therefore, $x_{k+1} \varepsilon \overline{S}(x_0, t^*)$. Hence, $\{t_k\}$ majorizes $\{x_k\}$ and $\lim_{k\to\infty} t_k = t^*$. Therefore, there exists $x^* \varepsilon X$ such that $\lim_{k\to\infty} x_k = x^*$. Since $x_k \varepsilon \overline{S}(x_0, t^*)$ for $k = 0, 1, \ldots$, it follows that $x^* \varepsilon \overline{S}(x_0, t^*)$, and the continuity of G implies that $Gx^* = x^*$.

Suppose there exists $y^* \varepsilon \overline{S}(x_0, t^*)$ such that $Gy^* = y^*$. We shall show by induction that $||y^* - x_k|| \leq t^* - t_k$ for $k = 0, 1, \ldots$. Clearly if $k = 0$, then $||y^* - x_0|| \leq t^* = t^* - t_0$.

Assuming that $||y^* - x_k|| < t^* - t_k$, we have $||y^* - x_{k+1}|| = ||G(y^*) - G(x_k)|| \leq \phi(||y^* - x_k||, ||x_k - x_0||) \leq \phi(t^* - t_k, t_k) = \phi(u - v, v) = \psi(u) - \psi(v)$ where $u = t^*$ and $v = t_k$. So $||y^* - x_k|| \leq \psi(t^*) - \psi(t_k) = t^* - t_{k+1}$. Thus $\lim_{k\to\infty} x_k = y^*$, so $y^* = x^*$ and hence x^* is unique.

<u>Lemma 2.2</u>: Let X be a B-space, $x_0 \varepsilon X$, $k \varepsilon R^+$ and $T: S(x_0, k) \longrightarrow X$. Let there exist a function $\phi \varepsilon \Gamma^5(Q)$ and $M \varepsilon R^+$ such that $J_i = [0, b_i)$, $b_i < M$ for $i = 1, \cdots, 5$ and $b_1 + b_3 < M$, where $Q = J_1 \times J_2 \times \cdots \times J_5$, and such that (2.21): $||T(x) - T(y)|| \leq \phi(||x - y||, ||x - x_0||, ||y - x_0||, ||T(y) - x||, ||T(y) - y||)$ whenever both sides exist. Further let there exist a function $\psi: [0, M) \longrightarrow R$, and a $\rho_0 \varepsilon [0, M)$ such that

(2.22) $||x_0 - T(x_0)|| \leq \psi(\rho_0) - \rho_0$ and

(2.23) $\phi(\rho - \sigma, \rho - \rho_0, \sigma - \rho_0, \psi(\rho) - \rho, \psi(\sigma) - \sigma) \leq \psi(\rho) - \psi(\sigma)$ for $\rho_0 \leq \sigma \leq \rho$ whenever both sides are defined. Let the sequence $\{\rho_n\}$ defined by $\rho_{n+1} = \psi(\rho_n)$ exist and be such that $\rho_{n+1} \geq \rho_n$ and $\lim_{n\to\infty} \rho_n = \rho^* < M$.

Then if $\rho^* - \rho_0 < k$, the sequence $\{x_n\}$ defined by $x_{n+1} = T(x_n)$ is well defined and converges to an element $x^* \varepsilon S(x_0, k)$ such that $T(x^*) = x^*$ and $||x^* - x_n|| \leq \rho^* - \rho_n$.

Proof: We shall show by induction that $x_n \varepsilon S(x_0, k)$, and $||x_{n+1} - x_n|| \leq \rho_{n+1} - \rho_n$. For $n = 0$, $x_0 \varepsilon S(x_0, k)$ and $||x_1 - x_0|| = ||T(x_0) - x_0|| \leq \psi(\rho_0) - \rho_0 = \rho_1 - \rho_0$. Assume that $||x^* - x_n|| \leq \rho^* - \rho_n$ holds for $n \leq k$. Then $x_{k+1} = T(x_k)$ is

defined and $||x_{k+1} - x_0|| \leq \sum_{j=0}^{K} ||x_{j+1} - x_j|| \leq$
$\sum_{j=0}^{K} \rho_{j+1} - \rho_j = \rho_{k+1} - \rho_0 < \rho^* - \rho_0 < k$. Hence $x_{k+1} \varepsilon S(x_0, k)$.
Also, $||x_{k+1} - x_k|| = ||T(x_k) - T(x_{k-1})|| \leq$
$\phi(||x_k - x_{k-1}||, ||x_k - x_0||, ||x_{k-1} - x_0||, ||x_k - x_k||, ||x_k - x_{k-1}||)$
$\leq \phi(\rho_k - \rho_{k-1}, \rho_k - \rho_0, \rho_{k-1} - \rho_0, \rho_k - \rho_k, \rho_k - \rho_{k-1}) =$
$\phi(\rho - \sigma, \rho - \rho_0, \sigma - \rho_0, \psi(\sigma) - \rho, \psi(\sigma) - \sigma) \leq$
$\psi(\rho) - \psi(\sigma)$ where $\rho = \psi(\sigma) = \rho_k$ and $\sigma = \rho_{k-1}$. Therefore,
$||x_{k+1} - x_k|| \leq \psi(\rho_k) - \psi(\rho_{k-1}) = \rho_{k+1} - \rho_k$. Thus $\{x_k\}$ is majorized by $\{\rho_k\}$ and $\lim_{k \to \infty} \rho_k = \rho^*$, so there exists $x^* \varepsilon X$ such that $\lim_{k \to \infty} x_k = x^*$. Further, $||x^* - x_0|| \leq \rho^* - \rho_0 < k$, hence $x^* \varepsilon S(x_0, k)$.

We shall defer application of lemma (2.2) until chapter 3, but we invoke lemma (2.1) immediately.

<u>Theorem 2.3</u>: Suppose X is a B-space and $G: D \subset X \longrightarrow X$ is such that G is Frechet differentiable on D_0 and $||G'(x) - G'(y)|| \leq \gamma ||x - y||$ on D_0 for $D_0 \subset D$, a convex set and $\gamma \varepsilon R^+$. Suppose that for some $x_0 \varepsilon D_0$;

1) $||G'(x_0)|| \leq \delta < 1$.

2) $||x_0 - G(x_0)|| \leq \alpha$ and

3) $h = \dfrac{\gamma\alpha}{(1 - \delta)^2} \leq \dfrac{1}{2}$.

Finally, if $t^* = \dfrac{1 - \sqrt{1 - 2h}}{h} \cdot \dfrac{\alpha}{1 - \delta}$, $t^{**} = \dfrac{1 + \sqrt{1 - 2h}}{h} \cdot \dfrac{\alpha}{1 - \delta}$, and $\bar{S}(x_0, t^*) \subset D_0$, then the iterates $x_{k+1} = G(x_k)$,

$k = 0, 1, \ldots,$ remain in $\bar{S}(x_0, t^*)$ and converge to a fixed point x^* of G which is unique in $D_0 \cap S(x_0, t^{**})$.

Proof: For $x, y \varepsilon D_0$, applying lemma (1.2) we get $\|G(y) - G(x)\| \leq \|G(y) - G(x) - G'(x)(y - x)\| + \|(G'(x) - G'(x_0))(y - x)\| + \|G'(x_0)(y - x)\| \leq \frac{1}{2}\alpha\|y - x\|^2 + \gamma\|y - x_0\|\|y - x\| + \delta\|y - x\| = \phi(\|y - x\|, \|x - x_0\|)$ where $\phi(u,v) = \frac{1}{2}\gamma u^2 + \gamma vu + \delta u$. $\phi \varepsilon \Gamma^2(Q)$ because $\gamma \geq 0$, $\delta \geq 0$. If $\psi(t) = \frac{1}{2}\gamma t^2 + \delta t + \alpha$, then $\phi(u - v, v) = \psi(u) - \psi(v)$. If $t_0 = 0$ and $t_{k+1} = \psi(t_k)$, then $t_1 = \psi(0) = \alpha > 0 = t_0$ and $t_1 = \alpha < t^*$. Assume for $n \leq k$ that $t_n \geq t_{n-1}$ and $t_n \leq t^*$. Then since ψ is isotone, $t_{k+1} = \psi(t_k) \geq \psi(t_{k-1}) = t_k$. Also, because $t_k \leq t^*$ and ψ is both isotone and continuous, $t_{k+1} = \psi(t_k) \leq \psi(t^*) = t^*$. Therefore, $\lim\limits_{k \to \infty} t_k = t^*$. Hence by Lemma (2.1), $x_k \varepsilon \bar{S}(x_0, t^*)$ for $k = 0, 1, \ldots,$ and there exists a unique fixed point $x^* \varepsilon \bar{S}(x_0, t^*)$ of G.

If $h = \frac{1}{2}$, then $t^* = t^{**}$ and x^* is the unique fixed point in $D_0 \cap S(x_0, t^{**})$. Suppose $h < \frac{1}{2}$ and there exists $y^* \varepsilon D_0 \cap S(x_0, t^{**})$ such that $y^* = G(y^*)$. If $\|y^* - x_0\| = S_0 = S_0 - t_0$, and if $\|y^* - x_n\| \leq S_n - t_n$ for $n \leq k$, then $\|y^* - x_{k+1}\| \leq \phi(\|y^* - x_k\|, \|x_k - x_0\|) \leq \phi(S_k - t_k, t_k) = \psi(S_k) - \psi(t_k) = S_{k+1} - t_{k+1}$. So we need only show that $\lim\limits_{k \to \infty} S_k = t^* = \lim\limits_{k \to \infty} t_k$. For $t \varepsilon (t^*, t^{**})$, we have $\psi(t) < t$. Thus $S_{k+1} \leq S_k$ and $t^* \leq S_k$. Applying the same argument that was used for $\{t_k\}$ we have $\lim_{k\to\infty} s_k = t^*$. Hence $\lim_{k\to\infty} x_k = y^*$, and so $x^* = y^*$ which verifies the uniqueness of x^* in $D_0 \cap S(x_0, t^{**})$.

Again exploiting the relationship between equations (1) and (2), this theorem may be applied to the generalized Chord iterative technique;

(2.31) $x_{k+1} = x_k - A^{-1} F(x_k)$, $k = 0, 1, \cdots$, where $A \in [X \to Y]$; to produce the following result.

<u>Corollary 2.3</u>: Suppose X and Y are B-spaces and $F: D \subset X \to Y$ such that F has a Frechet derivative on D_0 and $\|F'(x) - F'(y)\| \leq \gamma \|x - y\|$ for $x, y \in D_0$, where $\gamma \in R^+$ and $D_0 \subset D$ is a convex set. Suppose $A \in [X \to Y]$ has a bounded inverse $A^{-1} \in [Y \to X]$ with $\|A^{-1}\| \leq \beta$. Choose $x_0 \in D_0$ such that $\|I - A^{-1} F'(x_0)\| \leq \delta < 1$, $\|A^{-1} F(x_0)\| \leq \alpha$ and $h = \frac{\beta\gamma\alpha}{(1-\delta)^2} \leq \frac{1}{2}$. Let $t^* = \frac{1 - \sqrt{1 - 2h}}{h} \cdot \frac{\alpha}{1 - \delta}$ and $t^{**} = \frac{1 + \sqrt{1 - 2h}}{h} \cdot \frac{\alpha}{1 - \delta}$. If $\bar{S}(x_0, t^*) \subset D_0$, then the iterates (2.31) remain in $\bar{S}(x_0, t^*)$ and converge to a solution x^* of $F(x) = 0$ which is unique in $D_0 \cap \bar{S}(x_0, t^{**})$.

In the particular case where $F'(x_0)$ and $[F'(x_0)]^{-1}$ exist, and $[F'(x_0)]^{-1} \in [Y \to X]$, by letting $A = F'(x_0)$ we obtain from (2.31) the modified Newton iterative technique. Thus we have specified conditions under which the modified Newton method will produce a sequence of iterates which converge to a solution of $F(x) = 0$. This special case of Corollary (2.3) is essentially the convergence theorem for the modified Newton method given by Kantorovich and Akilov [8].

To formulate conditions under which Newton's method will produce a convergent sequence of iterates, we begin by considering approximate Newton processes of the form

(2.32) $x_{k+1} = x_k - A^{-1}(x_k)F(x_k)$, $k = 0, 1, \cdots$, where $A(x)$ is a linear operator for fixed x.

<u>Theorem 2.4</u>: Let X and Y be B-spaces; and let $F: D \subset X \rightarrow Y$ be such that F is Frechet differentiable on $D_0 \subset D$, where D_0 is a convex set. Suppose $\|F'(x) - F'(y)\| \leq \gamma\|x - y\|$ on D_0, where $\gamma \in R^+$, and suppose that $A: D_0 \subset X \rightarrow [X \rightarrow Y]$ has a bounded inverse $A^{-1}(x) \in [Y \rightarrow X]$ for each $x \in D_0$ such that $\|A^{-1}(x)\| \leq \beta$ and $\|F'(x) - A(x)\| \leq \delta$ for $x \in D_0$. Let $x_0 \in D_0$ be such that $\|A^{-1}(x_0)F(x_0)\| \leq \alpha$ and $h = \frac{1}{2}\beta\gamma\alpha + \beta\delta < 1$. If $\bar{S}(x_0, r) \subset D_0$ where $r = \dfrac{\alpha}{1 - h}$, then the sequence $\{x_k\}$ defined by (2.32) remains in $\bar{S}(x_0, r)$ and converges to a solution x^* of $F(x) = 0$.

Proof: Define $G: D_0 \subset X \rightarrow X$, by $G(x) = x - A^{-1}(x)F(x)$. Then whenever x, $G(x) \in D_0$; $\|G(G(x)) - G(x)\| =$
$\|-A^{-1}(G(x))F(G(x))\| \leq \beta\,\|F(G(x)) - F(x) - F'(x)(G(x) - x)\| +$
$\beta\|(A(x) - F'(x))\cdot(G(x) - x)\| \leq \phi(\|G(x) - x\|)$ where
$\phi(u) = \frac{1}{2}\beta\gamma u^2 + \beta\delta u$. Now in order to construct a majorizing sequence $\{t_k\}$ for the iterates $\{x_k\}$ defined by (2.32), we

must solve the difference equation $t_{k+1} - t_k = \frac{1}{2}\beta\gamma(t_k - t_{k-1})^2 + \beta\delta(t_k - t_{k-1})$ for $k = 0, 1, \cdots$, with $t_0 = 0$ and $t_1 = \alpha$.

By induction we shall show that $t_k \leq \alpha\sum_{j=1}^{k} h^j$, for $k = 1, 2, \cdots$. Clearly $t_0 = 0 \leq \alpha$ and $t_1 = \alpha \leq \alpha + h\alpha$. Now suppose $t_n \leq \alpha\sum_{j=0}^{n} h^j$ for $n \leq k$. Then for $k + 1$: $t_{k+1} - t_k = \frac{1}{2}\beta(t_k - t_{k-1})^2 + \beta\delta(t_k - t_{k-1})$. So $t_{k+1} \leq \frac{1}{2}\beta(\alpha h^k)^2 + \beta\delta(\alpha h) + \alpha\sum_{j=0}^{k} h^j \leq h(\alpha h^k) + \sum_{j=0}^{k} h^j = \alpha\sum_{j=0}^{k+1} h^j$. Therefore $\lim_{k\to\infty} t_k = \dfrac{\alpha}{1 - h} = r$ and so $x_k \in \bar{S}(x_0, r)$ for $k = 0, 1, \cdots$. Also $\{t_k\}$ majorizes $\{x_k\}$ which implies that $\lim_{k\to\infty} x_k = x^* \in \bar{S}(x_0, r) \subset D_0$. Thus $A^{-1}(x^*)$ exists and $x^* = G(x^*)$ implies that $F(x^*) = 0$.

Newton's method is the particular case of this theorem where $A(x) = F'(x)$. So for Newton's method $\delta = 0$ and $h = \frac{1}{2}\beta\gamma$, and the majorizing sequence is constructed by solving: $t_{k+1} - t_k = h(t_k - t_{k-1})^2$, $k = 0, 1, \cdots$, $t_0 = 0$ and $t_1 = \alpha$. By induction we find that

(2.41) $t_{k+1} - t_k < \alpha h^{2^k - 1}$, $k = 0, 1, \cdots$. If $k = 0$, then $t_1 - t_0 = \alpha < \alpha$. Suppose (2.41) is true for $k \leq n$, then $t_{n+2} - t_{n+1} = h(t_{n+1} - t_n)^2 \leq h\alpha^2 h^{2^{n+1} - 2} \leq \alpha h^{2^{n+1} - 1}$. Now $t^* - t_k = \sum_{j=k}^{\infty} t_{j+1} - t_j \leq \sum_{j=k}^{\infty} \alpha h^{2^k - 1} = \dfrac{\alpha h^{2^k - 1}}{1 - h^{2^k}}$. Therefore,

$$\|x^* - x_k\| \leq \frac{\alpha h^{2^k - 1}}{1 - h^{2^k}}.$$

This result with the above error bound is known as the Newton - Mysovskih theorem. (See Kantorovich and Akilov [8]).

For a final result concerning the convergence of Newton's method, we shall prove the Newton - Kantorovich theorem using the method of Ortega [10]. This theorem is more general than the Newton - Mysovskih theorem in that $[F'(x)]^{-1}$ is assumed to exist only at a single point.

Theorem 2.5: Let X and Y be B-spaces and let $F: D \subset X \rightarrow Y$. Suppose on an open convex set $D_0 \subset D$, F is Frechet differentiable and $\|F'(x) - F'(y)\| \leq \gamma \|x - y\|$ for $x, y \in D_0$. For some $x_0 \in D_0$, assume that $\Gamma_0 = [F'(x_0)]^{-1}$ is defined on all of Y and that $h = \beta\gamma\eta \leq \frac{1}{2}$, where $\|\Gamma_0\| \leq \beta$ and $\|\Gamma_0 F(x_0)\| \leq \eta$. Let $t^* = \frac{(1 - \sqrt{1 - 2h})}{\beta\gamma}$, $t^{**} = \frac{(1 + \sqrt{1 - 2h})}{\beta\gamma}$, and suppose that $\bar{S}(x_0, t^*) \subset D_0$. Then the iterates $x_{k+1} = x_k - [F'(x_k)]^{-1} F(x_k)$, $k = 0, 1, \cdots$, are defined, lie in $\bar{S}(x_0, t^*)$ and converge to a solution x^* of $F(x) = 0$ which is unique in $D_0 \cap \bar{S}(x_0, t^{**})$.

Proof: Let $q = \frac{1}{\beta\gamma}$, then for $x \in S(x_0, q)$ we shall show $[F'(x)]^{-1}$ exists and $\left\|[F'(x)]^{-1}\right\| \leq \frac{\beta}{I - \beta\gamma\|x - x_0\|}$. If $x \in S(x_0, q)$, then using remark 2) page 6, $\|I - [F'(x_0)]^{-1} F'(x)\|$
$= \left\|[F'(x_0)]^{-1}(F'(x_0) - F'(x))\right\| \leq \|F'(x_0)\| \| F'(x_0) - F'(x)\| \leq$
$\beta\gamma\|x - x_0\| < 1$. So for $x \in S(x_0, q)$, $(I - (I - [F'(x_0)]^{-1} F'(x)))^{-1}$ exists and $\|(I - (I - [F'(x_0)]^{-1} F'(x)))^{-1}\| < \frac{1}{1 - \beta\gamma\|x - x_0\|}$

by theorem (1.2). Simplifying we have

$(I - (I - [F'(x_0)]^{-1} F'(x)))^{-1} = ([F'(x_0)]^{-1} F'(x))^{-1} =$ $[F'(x)]^{-1} F'(x_0)$ exists for $x \in S(x_0, q)$. $[F'(x_0)]^{-1}$ exists so $[F'(x)]^{-1} F'(x_0) [F'(x_0)]^{-1} = [F'(x)]^{-1}$ exists on $S(x_0, q)$, and $\|[F'(x)]^{-1}\| \leq \|[F'(x)]^{-1} F'(x_0)\| \|[F'(x_0)]^{-1}\| \leq \dfrac{\beta}{1 - \beta\gamma\|x - x_0\|}$.

Letting $N(x) = x - [F'(x)]^{-1} Fx$ where x and $N(x)$ are in $S(x_0, q)$, and noting that $F(x) + F'(x)(N(x) - x) = 0$, we obtain $\|N(N(x)) - Nx\| = \|[F'(N(x))]^{-1} F(N(x))\| \leq$ $\dfrac{\beta}{1 - \beta\gamma\|x_0 - N(x)\|} \| F(N(x)) - F(x) - F'(x)(N(x) - x)\| \leq$ $\dfrac{1}{2} \cdot \dfrac{\beta\gamma\|x - Nx\|^2}{1 - \beta\gamma\|x_0 - Nx\|}$ using lemma (1.3).

Next we shall show that $\{x_k\}$ is well defined and is majorized by $\{t_k\}$ where $t_{k+1} = t_k - \dfrac{\frac{1}{2}\beta\gamma(t_k)^2 - t_k + n}{t_{k-1}}$, $k = 0, 1, \ldots$ and $t_0 = 0$. Note that $\{t_k\}$ is the sequence of Newton iterates for the polynomial $\frac{1}{2}\beta\gamma t^2 - t + n$ with roots t^* and t^{**}. Thus $t_{k+1} \geq t_k$ and $\lim_{k\to\infty} t_k = t^*$. For $k = 1$, $x_1 = x_0 - [F'(x_0)]^{-1} F(x_0)$ and $\|x_1 - x_0\| \leq \eta = t_1 = t_1 - t_0$. Suppose that $x_1, \ldots, x_k$ exist and $\|x_n - x_{n-1}\| \leq t_n - t_{n-1}$ for $n \leq k$. Then $\|x_k - x_0\| \leq t_k - t_0 \leq t^*$ so $x_k \in \bar{S}(x_0, t^*)$ $S(x_0, q)$ and hence x_{k+1} is defined. Also, $\|x_{k+1} - x_k\| =$ $\|N(N(x_{k-1}) - N(x_{k-1})\| \leq \dfrac{\frac{1}{2}\beta\gamma\|x_k - x_{k-1}\|^2}{1 - \beta\gamma\|x_k - x_0\|} \leq \dfrac{\frac{1}{2}\beta\gamma(t_k - t_{k-1})^2}{1 - \beta\gamma t_k} =$ $t_{k+1} - t_k$. Again we have shown $\lim_{k\to\infty} t_k = t^*$, $\{t_k\}$ majorizes $\{x_k\}$, and $\{x_k\} \subset \bar{S}(x_0, t^*)$. Using the same reasoning as in the proof of theorem (2.3) we find that there exists a solution $x^* \in \bar{S}(x_0, t^*) \subset S(x_0, q)$ such that $N(x^*) = x^*$. By the continuity of $F(x)$ and the existence of $[F'(x^*)]^{-1}$ we see that $F(x^*) = 0$. The uniqueness follows again from the reasoning in theorem (2.3).

In a practical situation, the application of Newton's method involves a tremendous amount of work, partially because $[F'(x_k)]^{-1}$ must be computed at each iteration. In chapter three we shall consider the generalized Gauss-Seidel technique which is more efficiently applied to real problems.

CHAPTER 3

Generalized Gauss-Seidel Technique

Two of the most common iterative methods used for solving linear systems of equations are the Jacobi and the Gauss-Seidel techniques. Given the linear system Ax = B where

$$A = \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{n1} & \cdots & a_{nn} \end{pmatrix}, \quad x = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}, \text{ and } B = \begin{pmatrix} b_1 \\ \vdots \\ b_n \end{pmatrix},$$

the Jacobi method is described by the formulas

$$(1) \quad x_i^{(K+1)} = -\sum_{\substack{j=1 \\ j \neq i}}^{n} \frac{a_{ij}}{a_{ii}} x_j^{(K)} + \frac{b_i}{a_{ii}}, \quad i = 1, \cdots, n.$$

For the same system the Gauss-Seidel technique is written as:

$$(2) \quad x_i^{(K+1)} = -\sum_{j=1}^{i-1} \frac{a_{ij}}{a_{ii}} x_j^{(K+1)} - \sum_{j=i+1}^{n} \frac{a_{ij}}{a_{ii}} x_j^{(K)} + \frac{b_i}{a_{ii}}, \quad i = 1, \cdots, n.$$

Sufficient conditions to insure convergence of these techniques may be found in Faddeev and Faddeeva [6].

For solving a nonlinear system of equations $f_i(x_1, \cdots, x_n)$, $i = 1, \cdots, n$, H. M. Lieberstein [9] proposed two iterative techniques which are generalizations of (1) and (2). The generalized Jacobi method has the form:

(3) $x_i^{(k+1)} = x_i^{(k)} - \dfrac{f_i(x_1^{(k)}, \circ\circ\circ, x_n^{(k)})}{f_{ii}(x_1^{(k)}, \circ\circ\circ, x_n^{(k)})}$, $i = 1, 2, \circ\circ\circ, n$; $k = 0,$ $1, \circ\circ\circ$, where $f_{ii} = \dfrac{\partial f_i}{\partial x_i}$, and the generalized Gauss-Seidel technique is described by:

(4) $x_i^{(k+1)} = x_i^{(k)} - \dfrac{f_i(x_1^{(k+1)}, \circ\circ\circ, x_{i-1}^{(k+1)}, x_i^{(k)}, \circ\circ\circ, x_n^{(k)})}{f_{ii}(x_1^{(k+1)}, \circ\circ\circ, x_{i-1}^{(k+1)}, x_i^{(k)}, \circ\circ\circ, x_n^{(k)})}$, $i = 1, \circ\circ\circ,$ n; $k = 0, 1, \circ\circ\circ$.

Bryan [3] has derived conditions sufficient to guarantee the convergence of (3) in a special type of B-space. We shall state this result for R^n with Tchebycheff norm, without proof, and shall employ the methods of Bryan to derive conditions which insure the convergence of (4) in R^2.

<u>Theorem 3.1</u>: Let $f_i(x_1, \circ\circ\circ, x_n)$ be a real valued function defined on a set $A \subset R^n$ for $i = 1, \circ\circ\circ, n$, such that there exists $x_0 = (x_1^{(0)}, \circ\circ\circ, x_n^{(0)})$ belonging to A and positive real numbers r, N, P_0, Q_0, and H_0 so that for $i = 1, 2, \circ\circ\circ,$ n, f_i has continuous second-order partial derivatives for $x \in S(x_0, r) \subset A$,

1) $\max\limits_{i=1,\circ\circ\circ,n} \sum_{j=1}^{n} \sum_{k=1}^{n} |f_{ijk}(x)| \leq N$, for $x \in S(x_0, r)$,

2) $\max\limits_{i=1,\circ\circ\circ,n} \left| \dfrac{1}{f_{ii}(x_0)} \right| \leq P_0$,

3) $\max\limits_{i=1,\circ\circ\circ,n} \left| \dfrac{f_i(x_0)}{f_{ii}(x_0)} \right| \leq Q_0$,

4) $\max\limits_{i=1,\circ\circ\circ,n} \sum_{\substack{j=1 \\ j \neq i}}^{n} \left| \dfrac{f\ \ (x\ )}{f\ \ (x_0)} \right| \leq H_0 < 1$,

5) $b_0 = \dfrac{P_0 Q_0 N}{(1 - H_0)^2} < \dfrac{1}{4}$ and

6) $r_0 = \frac{Q_0}{2b_0(1 - H_0)}(1 - \sqrt{1-4b_0}) < r$. Then the sequence defined by (3) converges to a point $x^* = (x_1^*, \ldots, x_n^*)$ in A such that $f_i(x^*) = 0$ and $\max_{i=1,\ldots,n} |x_i^* - x_i^0| \leq r_0$. Moreover, x^* is the only solution of $f_i(x_1, \ldots, x_n) = 0$, $i = 1, \ldots, n$, in the set $S(x,k) \subset S(x,r)$ where

$k = \frac{Q_0}{2b_0(1 - H_0)} (1 + \sqrt{1 - 4b_0})$.

Let $F: R^2 \to R^2$, $T_1: R^2 \to R^2$, and $T_2: R^2 \to R^2$ where

$$F(x_1,x_2) = \begin{pmatrix} f_1(x_1,x_2) \\ f_2(x_1,x_2) \end{pmatrix}, \quad T_1(x_1,x_2) = \begin{pmatrix} x_1 - \frac{f_1(x)}{f_{11}(x)} \\ x_2 \end{pmatrix} \text{ and}$$

$$T_2(x_1,x_2) = \begin{pmatrix} x_1 \\ x_2 - \frac{f_2(x)}{f_{22}(x)} \end{pmatrix}.$$ Then if $T: R^2 \to R^2$ is defined by $T(x) = T_2(T_1(x))$, the iteration $x_{k+1} = T(x_k)$ becomes the generalized Gauss-Seidel technique. With these operators in mind we formulate the following theorem for the extended Gauss-Seidel technique in R^2.

Theorem (3.2): Let $F: A \subset R^2 \to R^2$, where $F(x) = \begin{pmatrix} f_1(x) \\ f_2(x) \end{pmatrix}$, and let T_1, T_2, and T be defined as above. Suppose there exists $x_0 \in A$ and r_0, N, β_0, Q_0, and $H_0 \in R^+$ such that

1) F is twice Frechet differentiable on $S(x_0, r_0) \cup S(T_1(x_0), r_0) \subset A$, and $\|F''(x)\| \leq N$ for $x \in S(x_0, r_0) \cup S(T_1(x_0), r_0)$.

2) $f_{11}(x_0) \neq 0$, $f_{11}(T_1(x_0)) \neq 0$, $f_{22}(x_0) \neq 0$, and $f_{22}(T_1(x_0)) \neq 0$

3) $\max\left(\left|\frac{1}{f_{11}(x_0)}\right|, \left|\frac{1}{f_{11}(T_1(x_0))}\right|, \left|\frac{1}{f_{22}(x_0)}\right|, \left|\frac{1}{f_{22}(T_1(x_0))}\right|\right) \leq \beta_0$

4) $\max\left(\left|\frac{f_1(x_0)}{f_{11}(x_0)}\right|, \left|\frac{f_2(T_1(x_0))}{f_{22}(T_1(x_0))}\right|\right) \leq Q_0$

5) $\max\left(\left|\frac{f_{12}(x_0)}{f_{11}(x_0)}\right|, \left|\frac{f_{21}(T_1(x_0))}{f_{22}(T_1(x_0))}\right|\right) \leq H_0 < 1$

6) $(1 - H_0)^2 > 4Q_0N\beta_0$ and

7) $\frac{(1 - H_0) - \sqrt{(1 - H_0)^2 - 4Q_0N\beta_0}}{2N\beta_0} < r_0$,

then the iteration $x_{n+1} = T(x_n)$ is well defined and the sequence $\{x_n\}$ will converge to $x^* \varepsilon \overline{S}(x_0, r_0)$ such that $F(x^*) = 0$. Moreover, $\|x_n - x^*\| \leq \frac{\alpha^n}{1 - \alpha} Q_0$ where

$$\alpha = \frac{(H_0 + 1) - \sqrt{(1 - H_0)^2 - 4Q_0N\beta_0}}{(H_0 + 1) + \sqrt{(1 - H_0)^2 - 4Q_0N\beta_0}} < 1.$$

Proof: We begin by showing that $\|T_1(x) - T_1(y)\| \leq \|x - y\|$ for $x, y \varepsilon S(x_0, k) \subset S(x_0, r_0)$.

Define $H_1: R^2 \to R^2$ by $H_1(x_1, x_2) = \begin{pmatrix} \frac{f_1(x_1, x_2)}{f_{11}(x_1, x_2)} \\ 0 \end{pmatrix}$. Then for $x, y \varepsilon S(x_0, r_0)$, $H_1'(x) - H_1'(y) =$

$$\begin{pmatrix} \frac{1}{f_{11}(x_0)} & 0 \\ 0 & \frac{1}{f_{22}(x_0)} \end{pmatrix} \begin{pmatrix} f_{11}(x) - f_{11}(y) & f_{12}(x) - f_{12}(y) \\ 0 & 0 \end{pmatrix},$$

so $\|H_1'(x) - H_1'(y)\| \leq \beta_0 \|F'(x) - F'(y)\| \leq \beta_0 N\|x - y\|$. Let $\overline{H}: [0, \infty) \to [0, \infty)$ be defined by $\overline{H}(t) = \frac{\beta_0 N}{2} t^2$, then $\overline{H}(t)$ is monotone increasing and

(3.21) $\|H_1'(x) - H_1'(y)\| \leqslant \bar{H}'(\|x - y\|)$. Choose $\|x - x_0\| < k$ where $k = \min(r_0, \frac{1}{\beta_0 N})$. Then

$$\left\| \begin{pmatrix} \frac{1}{f_{11}(x_0)} & 0 \\ 0 & \frac{1}{f_{22}(x_0)} \end{pmatrix} \begin{pmatrix} f_{11}(x_0) - f_{11}(x) & 0 \\ 0 & f_{22}(x_0) - f_{22}(x) \end{pmatrix} \right\|$$

$\leqslant \beta_0 N\|x - x_0\|$. So by theorem (1.2), the linear operator $R_1 x =$

$$\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} - \begin{pmatrix} \frac{1}{f_{11}(x_0)} & 0 \\ 0 & \frac{1}{f_{22}(x_0)} \end{pmatrix} \begin{pmatrix} f_{11}(x_0) - f_{11}(x) & 0 \\ 0 & f_{22}(x_0) - f_{22}(x) \end{pmatrix}$$

has a bounded inverse on $S(x_0,k)$ such that $\|R_1 x^{-1}\| \leqslant$

$\frac{1}{1 - N\beta_0\|x - x_0\|}$. Now $\begin{pmatrix} f_{11}(x) & 0 \\ 0 & f_{22}(x) \end{pmatrix} = \begin{pmatrix} f_{11}(x_0) & 0 \\ 0 & f_{22}(x_0) \end{pmatrix} R_1 x$

on $S(x_0,k)$, so $\begin{pmatrix} \frac{1}{f_{11}(x)} & 0 \\ 0 & \frac{1}{f_{22}(x)} \end{pmatrix} = R_1 x^{-1} \begin{pmatrix} \frac{1}{f_{11}(x_0)} & 0 \\ 0 & \frac{1}{f_{22}(x_0)} \end{pmatrix}$.

Therefore $T_1(x)$ is defined on $S(x_0,k)$.

Since we are attempting to verify the inequality $\|T_1(x) - T_1(y)\| \leqslant \|x - y\|$, and because the second components of the vectors $T_1(x) - T_1(y)$ and $x - y$ are equal, we need only consider the first component of the vector $T_1(x) - T_1(y)$. Thus we define $T_{11}: S(x_0,k) \subset R^2 \longrightarrow R^2$ by $T_{11}(x) = \begin{pmatrix} x_1 - \frac{f\ (x)}{f_{11}(x)} \\ 0 \end{pmatrix}$.

Then,

$$(3.22)\quad T_{11}(x)-T_{11}(y) = \begin{pmatrix} \frac{1}{f_{11}(x)} & 0 \\ 0 & \frac{1}{f_{22}(x)} \end{pmatrix} \left[- \begin{pmatrix} f_1(x) \\ 0 \end{pmatrix} - \begin{pmatrix} f_{11}(x) & 0 \\ 0 & f_{22}(x) \end{pmatrix} \begin{pmatrix} y_1 - \frac{f_1(y)}{f_{11}(y)} - x_1 \\ 0 \end{pmatrix} + \begin{pmatrix} f_1(y) \\ 0 \end{pmatrix} + \begin{pmatrix} f_{11}(y) & 0 \\ 0 & f_{22}(y) \end{pmatrix} \begin{pmatrix} -\frac{f_1(y)}{f_{11}(y)} \\ 0 \end{pmatrix} \right] = R_1 x^{-1} g_1, \text{ where } g_1 =$$

$$- \begin{pmatrix} \frac{f_1(x)}{f_{11}(x_0)} \\ 0 \end{pmatrix} - \begin{pmatrix} \frac{f_{11}(x)}{f_{11}(x_0)} & 0 \\ 0 & \frac{f_{22}(x)}{f_{22}(x_0)} \end{pmatrix} \begin{pmatrix} y_1 - \frac{f_1(y)}{f_{11}(y)} - x_1 \\ 0 \end{pmatrix} + \begin{pmatrix} \frac{f_1(y)}{f_{11}(x_0)} \\ 0 \end{pmatrix} + \begin{pmatrix} \frac{f_{11}(y)}{f_{11}(x_0)} & 0 \\ 0 & \frac{f_{22}(y)}{f_{22}(x_0)} \end{pmatrix} \begin{pmatrix} -\frac{f_1(y)}{f_{11}(y)} \\ 0 \end{pmatrix}.$$

Letting

$$g_1(s) = -\frac{f_1(x)}{f_{11}(x_0)} + \frac{f_1(y)}{f_{11}(x_0)} - \frac{1}{f_{11}(x_0)} \left[\frac{f_1\left(x + s\left(y_1 - \frac{f_1(y)}{f_{11}(y)} - x_1\right)\right) - f_1(x)}{s} - \frac{f_1\left(y + s\left(- \frac{f_1(y)}{f_{11}(y)} \right)\right) - f_1(y)}{s} \right],$$

we find that $g_1 = \lim_{s \to 0^+} g_1(s)$. Substituting we have $g_1(s) = -H_1(x) + H_1(y) - \frac{1}{s}\left(H_1(x + s(z_1 - x_1)) - H_1(x) - H_1(y + s(z_1 - y_1)) + H_1(y)\right)$

where $z = T_1(y)$.

Next, if we let $\bar{x} = tx + (1 - t)y$, then the fundamental theorem of calculus implies that $g_1(s) =$

$- \int_0^1 \sum_{j=1}^{2} H_{1j}(\bar{x})(x_j - y_j)dt \;\; - \frac{1}{s}\int_0^1 H_{12}(\bar{x} + s(z_1 - \bar{x}_1))(x_2 - y_2)dt$

$+ \frac{1}{s}\int_0^1 \sum_{j=1}^{2} H_{1j}(\bar{x})(x_j - y_j)dt$. Hence $|g_1(s)| \leq$

$\frac{1}{s}\int_0^1 \sum_{j=1}^{2} |H_{1j}(\bar{x} + s(z_1 - \bar{x}_1)) - H_{1j}(\bar{x})| \; |x_j - y_j| \, dt \; +$

$\int_0^1 |H_{12}(\bar{x}) - H_{12}(x_0)| \, |x_2 - y_2| dt + |H_{12}(x_0)| \, |x_2 - y_2|$, so

$||g_1(s)|| \leq \frac{1}{s}\int_0^1 ||H_1'(\bar{x} + s(z_1 - \bar{x}_1)) - H_1'(\bar{x})|| \, ||x - y|| dt \; +$

$\int_0^1 ||H_1'(\bar{x}) - H_1'(x_0)|| \, ||x - y|| dt + H_0||x - y||$. Now using (3.21),

$||g_1(s)|| \leq \int_0^1 \bar{H}'(s|z_1 - \bar{x}_1|)||x - y|| dt + \int_0^1 \bar{H}'(||\bar{x} - x_0||)||x - y|| dt$

$+ H_0||x - y||$. If we let $\rho_1 = ||x - y||$, $\rho_2 = ||x - x_0||$, $\rho_3 =$

$||y - x_0||$, $\rho_4 = |z_1 - x_1|$, and $\rho_5 = |z_1 - y_1|$, then $||\bar{x} - x_0|| \leq$

$t\rho_1 + \rho_3$, $|z_1 - \bar{x}_1| = |(z_1 - x_1)t + (1 - t)(z_1 - y_1)| \leq \rho_4 t +$

$(1 - t)\rho_5$, and because $\bar{H}'(\rho)$ is increasing, $||g_1(s)|| \leq$

$\frac{1}{s}\int_0^1 \bar{H}'(s[t\rho_4 + (1 - t)\rho_5])\rho_1 dt + \;\; \bar{H}'(t\rho_1 + \rho_3)\rho_1 dt + H_0\rho_1$.

Thus we have $||g_1(s)|| \leq \; N\beta_0[(\rho_4 + \rho_5)(\frac{1}{2}\rho_1) + \rho_3\rho_1 + \frac{1}{2}(\rho_1)^2] +$

$H_0\rho_1$. Since the right hand member of the above inequality

is independent of s, $||g_1|| \leq N\beta_0[(\rho_4 + \rho_5)(\frac{1}{2}\rho_1) + \rho_3\rho_1 + \frac{1}{2}(\rho_1)^2]$

$+ H_0\rho_1$. Therefore, $||T_{11}(x) - T_{11}(y)|| \leq \; ||R_1 x^{-1}|| \, ||g_1|| \; \leq$

$\phi(||x - y||, ||x - x_0||, ||y - x_0||, |z_1 - x_1|, |z_1 - y_1|)$ where

$\phi(\rho_1, \rho_2, \rho_3, \rho_4, \rho_4) = \dfrac{1}{1 - N\beta_0\rho_2}\Big\{\frac{1}{2}N\beta_0(\rho_4 + \rho_5)\rho_1 + N\beta_0(\rho_3\rho_1 +$

$\frac{1}{2}(\rho_1)^2) + H_0\rho_1\Big\}$. Note that since $N\beta_0 > 0$, $\phi \epsilon \Gamma^5(Q)$ where

$J_i = \left[0, \frac{1}{N\beta_0}\right]$, $i = 1, \ldots, 5$, and $Q = J_1 \times J_2 \times \ldots \times J_5$. Let

$\rho_1 = \rho - \sigma$, $\rho_2 = \rho$, $\rho_3 = \sigma$, $\rho_4 = \zeta - \rho$, $\rho_5 = \zeta - \sigma$, then upon

substitution, $\phi(\rho - \sigma, \rho, \sigma, \zeta - \rho, \zeta - \sigma) =$

$\frac{1}{1 - N\beta_0\rho}\left\{\tfrac{1}{2}N\beta_0(2\zeta - \rho - \sigma)(\rho - \sigma) + N\beta_0(\sigma(\rho - \zeta) + \tfrac{1}{2}(\rho - \sigma)^2\right.$

$\left. + H_0(\rho - \sigma)\right\}$. Next let $\overline{T}(\rho) = \rho + \frac{1}{1 - N\beta_0\rho}\left\{Q_0 + N\beta_0\rho^2 - \rho + H_0\rho\right\}$,

and let $\zeta = \overline{T}(\zeta)$. Using the procedure of (3.22) we obtain

$\overline{T}(\rho) - T(\sigma) = \frac{1}{1 - N\beta_0\rho}\left\{\tfrac{1}{2}N\beta_0(2\zeta - \rho - \sigma)(\rho - \sigma)\right.$

$\left. + N\beta_0(\sigma(\rho - \sigma) + \tfrac{1}{2}(\rho - \sigma)^2) + H_0(\rho - \sigma)\right\} =$

$\phi(\rho - \sigma, \rho, \sigma, \overline{T}(\sigma) - \rho, \overline{T}(\sigma) - \sigma)$. So we have

$\|T_{11}(x) - T_{11}(y)\| \leq \overline{T}(\rho) - \overline{T}(\sigma)$. Solving $\overline{T}(\rho) - \rho =$

$\frac{1}{1 - N\beta_0\rho}\left\{Q_0 + N\beta_0\rho^2 - \rho + H_0\rho\right\} = 0$ for $0 \leq \rho < \frac{1}{\beta_0 N}$ we find

that $\frac{(1 - H_0) \pm \sqrt{(1 - H_0)^2 - 4N\beta_0 Q_0}}{2N\beta_0}$ are the fixed points of

$\overline{T}(\rho)$. By assumptions (5) and (7) these roots are real and

positive. Let $\rho^* = \frac{(1 - H_0) - \sqrt{(1 - H\ )^2 - 4Q\beta N}}{2N\beta_0} <$

$\frac{1}{N\beta_0}$ and consider $\rho \varepsilon [0, \rho^*]$.

$\overline{T}'(\rho) = \frac{(1 - N\beta_0\rho)[N\beta_0\rho + H_0] + N\beta_0(Q_0 + N\beta_0\rho^2 - (1 - H_0)\rho)}{(1 - N\beta_0\rho)^2} =$

$\frac{H_0 + Q_0 N\beta_0}{(1 - N\beta_0\rho)^2} > 0$ for $\rho < \rho^* < \frac{1}{\beta_0 N}$, hence $\overline{T}(\rho)$ is increasing

on $[0, \rho^*]$. Further, since $\overline{T}(\rho) = \rho + \frac{(N\beta_0\rho^2 - (1 - H_0)\rho + Q_0)}{1 - N\beta_0\rho}$

and $\frac{N\beta_0\rho^2 - (1 - H_0)\rho + Q_0}{1 - N\beta_0\rho} > 0$, we have $\overline{T}(\rho) > \rho$ for

$\rho < \rho^*$. Also, $\overline{T}''(\rho) = \frac{2N\beta_0(H_0 + Q_0 N\beta_0)}{(1 - N\beta_0\rho)^4} > 0$ for $\rho < \rho^*$ so

$\overline{T}'(\rho)$ is increasing and hence $\max\limits_{0 \leq \rho \leq \rho^*} \overline{T}'(\rho) = \overline{T}'(\rho^*) =$

$$\frac{N\beta_0\rho^* + H_0}{1 - N\beta_0\rho^*} = \frac{(H_0 + 1) - \sqrt{(1 - H_0)^2 - 4N\beta_0 Q_0}}{(H_0 + 1) + \sqrt{(1 - H_0)^2 - 4N\beta_0 Q_0}} = \alpha < 1.$$

Coupling this result with the mean value theorem, we obtain $|\mathbb{T}(\rho) - \mathbb{T}(\sigma)| < \alpha|\rho - \sigma|$. Thus $\mathbb{T}(\rho)$ is a contraction mapping. So we have $||T_{11}(x) - T_{11}(y)|| \leq \mathbb{T}(\rho) - \mathbb{T}(\sigma) < \alpha|\rho - \sigma| = \alpha\Big|\,||x - x_0|| - ||y - x_0||\,\Big| \leq \alpha||x - y|| \leq ||x - y||$. In view of the previous remark, we now have $||T_1(x) - T_1(y)|| \leq ||x - y||$ for $x, y \varepsilon S(x_0, k)$. An immediate result of this inequality is $T_1(x) \varepsilon S(T_1(x_0), k)$ whenever $x \varepsilon S(x_0, k)$.

We now proceed to show that $T(x)$ is defined on $S(x_0, k)$ using reasoning analogous to that which was applied to $T_1(x)$. First we define $H_2: R^2 \to R^2$ by $H_2(x) = \begin{pmatrix} 0 \\ \frac{f_2(x)}{f_2(T_1(x_0))} \end{pmatrix}$. For, $x, y \varepsilon S(x_0, r_0)$ we have $H_2'(T_1(x)) - H_2'(T_1(y)) = \begin{pmatrix} \frac{1}{f_{11}(T_1(x_0))} & 0 \\ 0 & \frac{1}{f_{22}(T_1(x_0))} \end{pmatrix}$.

$$\begin{pmatrix} 0 & 0 \\ f_{21}(T_1(x)) - f_{21}(T_1(y)) & f_{22}(T_1(x)) - f_{22}(T_1(y)) \end{pmatrix},$$

so that $||H_2'(T_1(x)) - H_2'(T_1(y))|| \leq \beta_0 N||T_1(x) - T_1(y)|| \leq \beta_0 N||x - y||$. Further we have $||H_2'(T_1(x)) - H_2'(T_1(y))|| \leq \mathbb{H}'(||x - y||)$ for $x, y \varepsilon S(x_0, r_0)$. Now if $x \varepsilon S(x_0, k)$ where $k = \min\left(r_0, \frac{1}{\beta_0 N}\right)$, then

$$\left\| \begin{pmatrix} \frac{1}{f_{11}(T_1(x_0))} & 0 \\ 0 & \frac{1}{f_{22}(T_1(x_0))} \end{pmatrix} \cdot \begin{pmatrix} f_{11}(T_1(x_0)) - f_{11}(T_1(x)) & 0 \\ 0 & f_{22}(T_1(x_0) - f_{22}(T_1(x)) \end{pmatrix} \right\| \leq$$

$\beta_0 N \, \|T_1(x) - T_1(x_0)\| \leq \beta_0 \, N\|x - x_0\| < 1$. So again applying Banach's theorem, we see that the linear operator

$$R_2x = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} - \begin{pmatrix} \frac{1}{f_{11}(T_1(x_0))} & 0 \\ 0 & \frac{1}{f_{22}(T_1(x_0))} \end{pmatrix} \cdot \begin{pmatrix} f_{11}(T_1(x_0)) - f_{11}(T_1(x)) & 0 \\ 0 & f_{22}(T_1(x_0)) - f_{22}(T_1(x)) \end{pmatrix}$$

has a bounded inverse on $S(x_0,k)$ such that $\|R_2x^{-1}\| \leq$ $\frac{1}{1 - N\beta_0\|x - x_0\|}$. Further

$$\begin{pmatrix} f_{11}(T_1(x)) & 0 \\ 0 & f_{22}(T_1(x)) \end{pmatrix} = \begin{pmatrix} f_{11}(T_1(x_0)) & 0 \\ 0 & f_{22}(T_1(x_0)) \end{pmatrix} \cdot$$

R_2x, so $\begin{pmatrix} \frac{1}{f(T_1(x))} & 0 \\ 0 & \frac{1}{f_{22}(T_1(x))} \end{pmatrix} = R_2x^{-1}$ ·

$\begin{pmatrix} \frac{1}{f_{11}(T_1(x_0))} & 0 \\ 0 & \frac{1}{f_{22}(T_1(x_0))} \end{pmatrix}$ on $S(x_0,k)$. Therefore

$$T(x) = T_2(T_1(x)) = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} - \begin{pmatrix} \frac{f_1(x)}{f_{11}(x)} \\ \frac{f_2(T_1(x))}{f_{22}(T_1(x))} \end{pmatrix}$$ is

defined on $S(x_0, k)$. If we define $T_{22}: S(T_1 x_0, k) \subset R^2 \longrightarrow R^2$ by $T_{22}(u) = \begin{pmatrix} 0 \\ u_2 - \dfrac{f_2(u)}{f_{22}(u)} \end{pmatrix}$, then $T(x) - T(y) =$ $T_{11}(x) - T_{11}(y) + T_{22}(T_1(x)) - T_{22}(T_1(y))$. Further, $T(x) - T(y) = \max\left\{\|T_{11}(x) - T_{11}(y)\|, \|T_{22}(T_1(x)) - T_{22}(T_1(y))\|\right\}$. Letting $u = T_1(y)$ and $w = T_1(x)$ we note that the form of $T_{22}(w) - T_{22}(u)$ is analogous to the form of $T_{11}(x) - T_{11}(y)$, thus by the same reasoning we applied to $T_{11}(x) - T_{11}(y)$ we find that $\|T_{22}(w) - T_{22}(u)\| \leq$ $\phi(\|w - u\|, \|w - T_1(x_0)\|, \|u - T_1(x_0)\|, |\xi_2 - w_2|,$ $|\xi_2 - u_1|)$ where $\xi = T_2(u)$, and $\phi(\rho_1, \rho_2, \rho_3, \rho_4, \rho_5)$ is defined as before. Recalling that $\phi(\rho_1, \rho_2, \rho_3, \rho_4, \rho_5) \epsilon \Gamma^5(Q)$, and noting that; $\|w - u\| = \|T_1(x) - T_1(y)\| \leq \|x - y\|$, $\|w - T_1(x_0)\| \leq$ $\|x - x_0\|$, $\|u - T_1(x_0)\| \leq \|y - x_0\|$, $|\xi_2 - w_2| \leq$ $\|T(y) - x\|$, $|z_1 - x_1| \leq \|T(y) - x\|$, $|\xi_2 - u_2| \leq \|T(y) - y\|$, and $|z_2 - y_2| \leq \|T(y) - y\|$ we have: $\|T_{11}(x) - T_{22}(y)\| \leq$ $\phi(\|x - y\|, \|x - x_0\|, \|y - x_0\|, \|T(y) - x\|, \|T(y) - y\|)$ and $\|T_{22}(T_1(x)) - T_{22}(T_1(y))\| \leq \phi(\|x - y\|, \|x - x_0\|, \|y - x\|,$ $\|T(y) - x\|, \|T(y) - y\|)$. Therefore $\|T(x) - T(y)\| \leq$ $\phi(\|x - y\|, \|x - x_0\|, \|y - x\|, \|T(y) - x\|, \|T(y) - y\|)$.

Continuing as before, except now $\rho_4 = \|T(y) - x\|$ and $\rho_5 = \|T(y) - y\|$, we have $\|x_0 - T(x_0)\| = Q_0 \leq$ $\overline{T}(0) = 0$ and $\overline{T}(\rho) > \rho$ for $\rho < \rho^*$. Thus if $\{\rho_k\}$ is defined by $\rho_{k+1} = \bar{T}(\rho_k)$ we have $\rho_{n+1} \geq \rho_n$, and because $\bar{T}(\rho)$ is a contraction map $\{\rho_k\}$ converges to the unique fixed point $\rho^* < \frac{1}{N\beta_0}$. By assumption (7) $\rho^* < r_0$, so $\rho^* < \min\left(\frac{1}{N\beta_0}, r_0\right)$ and the hypotheses of Lemma (2.2) are satisfied for $T(x)$. Therefore there exists x^* such that $T(x^*) = x^*$ and $||x_n - x^*|| \leq \rho^* - \rho_n \leq \left(\frac{\alpha^n}{1 - \alpha}\right) Q_0$.

Now $T(x^*) = x^*$ implies that $T(x^*) = x^* - \begin{pmatrix} \frac{f_1(x^*)}{f_{11}(x^*)} \\ \frac{f_2(x^*)}{f_{22}(x^*)} \end{pmatrix}$, and because $f_i(x)$, $i = 1, 2$ are continuous, $f_i(x^*) = 0$ for $i = 1, 2$, hence $F(x^*) = 0$.

Therefore, after examining the concept of majorizing sequences and their application to Newton-like techniques, we have imposed conditions on the generalized Gauss-Seidel iterative method which are sufficient to insure convergence of the iterates to a solution of the equation $F(x) = 0$ in R^2.

REFERENCES

1. Altman, M., A Generalization of Newton's Method, Bull. Acad. Polon. Sci., Cl. III 3 (1955), 189-193.

2. Ben-Israel, A., A Newton-Raphson Method for the Solution of Systems of Equations, J. of Math. Anal. and Appl. 15 (1966), 243 - 252.

3. Bryan, C ., An Iterative Method for Solving Nonlinear Systems of Equations, Doctoral thesis, University of Arizona, Tuscon, (1963).

4. Collatz, L., Functional Analysis and Numerical Mathematics, Academic Press, New York, 1966.

5. Dennis, J., Variations on Newton's Method, Doctoral thesis, University of Utah, Salt Lake City, (1966).

6. Faddeev, D., and Faddeeva, V., Computational Methods of Linear Algebra, W. H. Freeman and Co., San Francisco, 1963.

7. Halmos, P., Finite-Dimensional Vector Spaces, Van Nostrand, New York, 1958.

8. Kantorovich, L., and Akilov, G., Functional Analysis in Normed Spaces, MacMillan, New York, 1964.

9. Lieberstein, H., Overrelaxation for Nonlinear Elliptic Partial Differential Probs., MRC Tech. Summary Rept. 80 (1959).

10. Ortega, J., The Newton-Kantorovich Theorem, Am. Math Monthly, 75, number 6 (June - July 1968), 658-660.

11. Rheinboldt, W., A Unified Convergence Theory for a Class of Iterative Processes, Siam J. on Num. Anal. 5, number 1 (March 1968).

12. Vainberg, M., Variational Methods in the Study of Nonlinear Operators, Holden-Day, San Francisco, 1964.