2003

A Study on Convergence of Newton's Method in Real and Interval Number

Rahman, Md. Majedur

University of Rajshahi

A STUDY ON CONVERGENCE OF NEWTON'S METHOD IN REAL AND INTERVAL NUMBER

THESIS SUBMITTED FOR THE DEGREE OF

DOCTOR OF PHILOSOPHY
IN
MATHEMATICS

BY

MD. MAJEDUR RAHMAN

DEPARTMENT OF MATHEMATICS
FACULTY OF SCIENCE
UNIVERSITY OF RAJSHAHI
RAJSHAHI, BANGLADESH

2003

A STUDY ON CONVERGENCE OF NEWTON'S METHOD IN REAL AND INTERVAL NUMBER

THESIS SUBMITTED FOR THE DEGREE OF

DOCTOR OF PHILOSOPHY
IN
MATHEMATICS

BY

MD.MAJEDUR RAHMAN

DEPARTMENT OF MATHEMATICS
FACULTY OF SCIENCE
UNIVERSITY OF RAJSHAHI
RAJSHAHI, BANGLADESH
2003

M. Zulfikar Ali
Doctor of Engineering (Tokushima)
Associate Professor
Department of Mathematics
University of Rajshahi
Rajshahi - 6205, Bangladesh

Phone : (Off.) 750041-49/4108
(Res.) (0721)
Fax : 88-0721-750064
E-mail : rajucc@citechco.net

Date 07.08.2003

Declaration

To best of my knowledge, the thesis entitled "**A Study on Convergence of Newton's Method in Real and Interval Number**" submitted by Mr. Md. Majedur Rahman to the University of Rajshahi, Bangladesh, under my supervision for the fulfillment of the requirements of the degree of **Doctor of Philosophy** in Mathematics is an original one and has not been submitted elsewhere for any degree.

Md. Zulfikar Ali
(Dr. Md. Zulfikar Ali)
Supervisor

Acknowledgement

I wish to express my sincere admiration, appreciation and gratitude to my supervisor Dr. Md. Zulfikar Ali, Associate Professor Department of Mathematics, University of Rajshahi, Rajshahi for his advice, encouragement and guidance through the course of this work.

I am greatly indebted to the Director, Bangladesh Institute of Technology, Rajshahi (BIT) for his kind permission for admission in the University of Rajshahi as a research student.

I am thankful to the University of Rajshahi and in particular to the Department of Mathematics for extending all facilities and co-operation during the course of my Ph.D. program.

Abstract

In order to find the approximate numerical solution to a system of nonlinear equations as well as an integral and a differential operator equations, Newton's algorithm is widely used. L. V. Kantorovich [1948] and Moore [1977] studied the existence and uniqueness of solution to the system of nonlinear equations and their error bounds. M. Urabe [1965] also studied the existence and uniqueness of the solution to nonlinear operator equations (mainly differential operator equations). Kantorovich and Urabe's methods are two variants of Newton's method in some sense. We study the existence and uniqueness of solutions to the nonlinear systems and their error bounds. Our results will be stated in a theorem that ensures the best possible generalized error bound that is different from that given by Kantorovich and Moore.

We also develop a technique that may be applied to find an approximate numerical solution to an algebraic as well as to a system of nonlinear equations both in real and interval number systems.

Finally, we have treated the error estimation for the quasiperiodic solution to the Van der pol type differential operator equation based on Urabe's theorem.

Contents

Introduction and Literature survey

Let us consider the function $f:R^n \to R^n$, where f is Frechet differentiable and consider the equation

$$f(x) = 0 \qquad (1)$$

In case where $R^n = R$, the Newton's iterative method for solving approximately, the equation (1) is given by

$$x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)}, \quad n = 0,1,2,3,\cdots\cdots \qquad (2)$$

where x_0 is an initial guess.

Newton first applied his iterative method (2) in 1669 for solving a cubic equation (i.e., where $f(x)$ is a cubic function of x). The procedure was systematically discussed in print by J. Raphson as early as 1690. Therefore, the method is sometimes referred to as the Newton-Raphson method.

The Newton method has a long history, with contributions by Cauchy, Runge, Faber and Blutel among others. The theorem of Fine in 1916 seems to be the first in n-dimensional space which, under conditions given for an initial approximate solution, asserts the existence of a solution of $f(x) = 0$ to which the iterates in the Newton method converge. In the same year, Bennett (1916) proved a convergence and existence theorem in more general spaces.

The "method of linearization" is often referred [Kantorovich L. B & Akilov G. P. (1964)] to in the literature as Newton's method for systems of nonlinear equations. It is a natural generalization of Newton's method for a single equation and in fact, can be extended to equations in infinite-dimensional (function) spaces. The first such generalization was done in the context of nonlinear operator equations in Banach spaces by Kantorovich in 1948. This generalization is often called the Newton-Kantorovich method. Another important generalization was given by J. Moser (1961), for the case of operators acting on a continuous scale of Banach spaces with properties similar to the properties of Sobolev spaces. The above generalizations also provide useful tools in the study of the solution of nonlinear differential and integral equations.

Our study in chapter 2, has been based on the fundamental theorem of Kantorovich (1948) that generalizes the Newton's method in the context of nonlinear operator equations in Banach space. Kantorovich, in his theorem in fact, used the boundedness of the second derivative of F [Rall L. B.(1969)] in place of the assumption that F' is Lipschitzan; Fenyo I. (1954)

first made the latter modification. Kantorovich error estimate was weaker, the present ones (which are sharp) were given by Dennis J. (1969).

With the hypotheses made by Kantorovich, in his fundamental theorem, we will be able to give a best possible error bound for Newton's algorithm. Our result will be given by a theorem from which we can conclude a special result by a corollary. We can fit our result by an example.

Let $f : D \subseteq R^n \to R^n$ be continuously differentiable in the open domain D. Let us assume that both of f and f' also have continuous inclusion monotonic interval extensions F and F' defined on interval vectors contained in D. The interval version of Newton's algorithm to solve (1) is given by

$$X^{(k+1)} = X^{(k)} \cap N(X^{(k)})$$

with the interval Newton function

$$N(X) = m(X) - V f(m(X)) \qquad (3)$$

$$[m(X) = \text{mid point of the closed interval } X]$$

and where V is an interval matrix containing $[f'(x)]^{-1}$ for all $x \in X$. The Newton's function $N(X)$ can also be given as

$$K(X) = y - Y f(y) + \{I - Y F'(X)\}(X - y) \qquad (4)$$

Using the function given by (4), we will be able to generate the convergent nested intervals $X^{(0)} \supset X^{(1)} \supset X^{(2)} \supset X^{(3)} \supset \cdots\cdots \supset X^{(k)} \supset \cdots\cdots$. That converges to an interval X (say), which contains the solution of equation (1). We also be able to give a bound for the interval $X^{(k)}$. Our findings can be applied to the algebraic and transcendental equations as well as to the system of nonlinear equations to compute numerical solutions.

Lastly, we study the nonlinear differential operator equations with quasi-periodic forcing term. A specific numerical error bound [Mitsui T. 1977] has been generalized for the quasi-periodic solution to Van der Pol type equations:

$$\frac{d^2x}{dt^2} - 2\lambda(1-x^2)\frac{dx}{dt} + x = \sum_{k=1}^{m}(a_k \cos\vartheta_k t + b_k \sin\vartheta_k t).$$

Chapter-1

Mathematical Preliminaries

In this chapter we discuss some basic properties (such as continuity, uniform continuity, convergence, uniform convergence, boundedness etc) of sequence, function, operator in real Banach space.

1.1 Metric Space. Let R be the set of real numbers and $\rho(x,y)$ be a function defined on the set $R \times R$ of all order pairs (x,y) of members of R. The concepts of real convergent sequences and continuous functions from R to R and most of their properties depend only on the following three conditions.

(i) $\rho(x,y) = 0$ if and only if $x = y$,

(ii) $\rho(x,y) = \rho(y,x)$ and

(iii) $\rho(x,y) \leq \rho(x,z) + \rho(z,y)$.

The function $\rho(x,y)$ is called a metric on R.

The set R together with the metric ρ is called a metric space denoted by (R,ρ).

The members of R are called "point" and the function $\rho(x,y)$ is called the distance from the point x and y.

Different choices of metric give rise to different metric spaces.

For example, the metric space R consisting of all real numbers with the metric $\rho(x,y)=|x-y|$ is different from the space consisting of all real numbers with the metric

$$\frac{2|x-y|}{\sqrt{\{1+|x|^2\}}\sqrt{\{1+|y|^2\}}}.$$

The space R^n of all $n-$dimensional real vectors of the form $x=(x_1,x_2,x_3,\cdots\cdots,x_n)$ is a metric space with any of the matrices.

$$\rho_1(x,y)=\sum_{i=1}^{n}|x_i-y_i|$$

$$\rho_2(x,y)=\left\{\sum_{i=1}^{n}(x_i-y_i)^2\right\}^{\frac{1}{2}}$$

$$\rho_\infty(x,y)=\max_{1\le i\le n}|x_i-y_i|$$

Definition 1.1.1. Let (E,ρ) be a metric space. A sequence $\{x_n\}$ of elements x_n in E "converges to x" if for all $\varepsilon>0$, there exists an $n_0\in N$ (set of positive integers) such that for $n\ge n_0$, $\rho\,(x_n,x)\le\varepsilon$.

Let (E,ρ_E) and (F,ρ_F) be two metric spaces. A function $f:E\to F$ is called "continuous" if for every sequence $\{x_n\}$ of elements x_n in E converging to x the sequence $\{f(x_n)\}$ converges to $f(x)$ in F.

Proposition 1.1.1. Let f be a function from E to F, and let $x\in E$. The following Properties are equivalent:

Property-1

(i) For every sequence $\{x_n\}$ converging to x, $f(x_n)$ converges to $f(x)$.

(ii) For all $\varepsilon>0$, there exist $\eta(\varepsilon,x)=\eta$ such that if $\rho_E(x,y)\le\eta$, then $\rho_F(f(x),f(y))\le\varepsilon$.

(iii) For every neighborhood $v\in V(f(x))$ of $f(x)$ there exists a neighborhood $U\in V(x)$ of x such that $f(U)\subset V$.

(iv) For every neighborhood $v\in V(f(x))$, $f^{-1}(v)\subset V(x)$ is a neighborhood of x.

Definition 1.1.2. If a mapping f from E to F satisfies the above equivalent conditions, we say that "f is continuous at x." We say that "f is continuous on E," if f is continuous at each point x of E.

The notion of continuous function is "local" in the sense that it is a notion defined at each point x of E. Hence property-1 (ii) the radius $\eta=\eta(\varepsilon,x)$ of the ball of number x depends on ε and on x. This remark leads us to introduce the notion of uniform continuity.

Definition 1.1.3. A function f from a metric space E to a metric space F is said to be "uniformly continuous" if for all $\varepsilon > 0$, there exist $\eta = \eta(\varepsilon)$ depending on ε and independent of x such that $\rho_F(f(x), f(y)) \leq \varepsilon$ when $\rho_E(x, y) \leq \eta$.

Proposition 1.1.2.

(a) Every uniformly continuous function is continuous.

(b) Every uniformly continuous function maps Cauchy sequences onto Cauchy sequences.

Proposition 1.1.3. If E is a metric space and if $A \subset E$ is nonempty, the function $x \to \rho(x, A)$ is uniformly continuous from E to R.

Proof. The proposition is a consequence of the inequality:

$$|\rho(x, A) - \rho(y, A)| \leq \rho(x, y).$$

This suggests the introduction of the following definition.

Definition 1.1.4. We say that a function f from a metric space E to a metric space F is "Lipschitz" if there exists a constant $\lambda > 0$ such that

$$\rho_F(f(x), f(y)) \leq \lambda \rho_E(x, y) \text{ for all } x, y \in E.$$

We say that a function f is a "contraction" if in addition, $\lambda < 1$.

For example, the function $x \to \rho(x, A)$ is Lipschitz with $\lambda = 1$. We remark from this definition that the following proposition holds.

Proposition 1.1.4. Every Lipschitz function is uniformly continuous.

Definition 1.1.5. Let $\{x_n\}$ be a sequence of elements x_n in E. If x_n converges to x, thus for all $\varepsilon > 0$, there exists an $n_0 > N$ such that for all $n, m \geq n_0$, then it can easily be shown that $\rho(x_n, x_m) \leq \varepsilon$.

The advantage of Definition 1.1.5 over Definition 1.1.1 is that the limit x does not appear in it. If the condition in the above definition is sufficient for convergence, it would allow us to know that the sequence is convergent without needing to know its limit.

Take the case where $E = Q$, and consider the sequence 0.1, 0.101, 0.101001, 0.1010010001,........ This sequence satisfies Definition 1.1.5 but does not converge in Q. However, Definition 1.1.5 is always sufficient for the convergence if $E = R$.

This leads us to introduce the following fundamental definition.

Definition 1.1.6. Let (E, ρ) be a metric space. We say that a sequence $\{x_n\}$ is a "Cauchy sequence" if it satisfies condition of Definition 1.1.5. We shall say that E is a "complete (metric) space" if every Cauchy sequence is convergent.

1.2 Ball in metric space. Consider a metric space (E, ρ). The "open ball with center x and radius ε" is the set

$$B(x,\varepsilon) = \{\ y \in E \quad such\ that\ \rho\ (x,y) < \varepsilon\},$$

and the "closed ball with center x and radius ε" is the set

$$\bar{B}(x,\varepsilon) = \{\ y \in E \quad such\ that\ \rho\ (x,y) \leq \varepsilon\}.$$

If A and B are two sets, the "distance from A to B" is the number

$$\rho(A,B) = \inf_{x \in A} \inf_{y \in B} \rho(x,y),$$

and we get
$$\rho(x,B) = \rho(\{x\},B) = \inf_{y \in B} \rho(x,y).$$

The "diameter of A" is the finite or infinite number

$$\delta(A) = \sup_{x \in A} \quad \sup_{y \in A} \rho(x,y),$$

and we say that a "set A is bounded" if A is nonempty and $\delta(A) < \infty$.

Here we given an example:

In the case of the metric space $(R, |x|)$, the balls $\overline{B}(x,\varepsilon)$ and $B(x,\varepsilon)$ are the intervals $(x-\varepsilon, x+\varepsilon)$ and $[x-\varepsilon, x+\varepsilon]$.

We now discuss some basic concepts of vector and matrix norms without which the error estimation for the solutions to operator equations in Banach space is impossible.

1.3 Norms of Vector and Matrix. Let us consider the linear space (or vector space) X over the field F, whose elements (vectors) denoted by $x, y, z, \cdots$.

Inner product. Let x, y, z be any three vectors in X. The inner product of two vectors x and y in X defined by (x, y) is a scalar satisfying the following axioms:

1. $(x, x) \geq 0$; $(x, x) = 0$ if and only if $x = 0$ (positive definitions)

2. $(x, y) = (y, x)$ (symmetric property)

3. $(\alpha x + \beta y, z) = \alpha(x, z) + \beta(y, z)$ (linearity), where $\alpha, \beta \in F$.

The inner product space is a vector space satisfying the above inner product axioms.

Let $X = R^n$, and $x = (x_1, x_2, x_3, \cdots, x_n) \in R^n$, then the inner product (x, x) is given by $(x, x) = (|x_1|^2 + |x_2|^2 + \cdots\cdots + |x_n|^2)$.

By the norm $\|.\|$ of a vector x of X, we mean a function $\|x\| : X \to \{t : 0 \leq t < \infty\}$ such that

1. $\| x + x' \| \leq \| x \| + \| x' \|$

2. $\| \alpha x \| = | \alpha | \| x \|$

3. $\| x \| = 0$ if and only if $x = 0$.

1.4 Example of Norms and Normed spaces:

(a) Let $U = R^3$: then the usual norm defined on R^3 is $\|x\| = (x_1^2 + x_2^2 + x_3^2)^{1/2}$ for $x = (x_1, x_2, x_3)$. The extension to R^n is obvious.

(b) The quantity $\|.\|_p$ defined by $\|x\|_p = \left[\sum_{i=1}^{n} |x_i|^p\right]^{\frac{1}{p}}$, $(1 \le p < \infty)$ is a norm on R^n. If we let $p \to \infty$ then the quantity $\|.\|_\infty$ defined on R^n by $\|x\|_\infty = \max_{1 \le i \le n} |x_i|$ is also a norm on R^n.

(c) Let $U = L_p(a, b)$ with $1 \le p < \infty$ [the space of measurable functions $u(x)$ defined on domain (a, b) whose Lebesgue integral $\int_{(a,b)} |u(x)|^p \, dx$ is finite]: the L_p -norm is defined by

$$\|u\|_{L_p} = \left[\int_a^b |u(x)|^p \, dx\right]^{\frac{1}{p}}, \quad u \in L_p(a, b)$$

(d) The space $L_\infty(a, b)$ of bounded measurable functions is a normed space, with norm $\|.\|_\infty$ defined by $\|u\|_\infty = \sup |u(x)|$, the supremum being taken over all subsets of (a, b) with non-zero measure.

1.5 Equivalence of norms. Two norms $\|\cdot\|$ and $\|\cdot\|'$ on a linear space X are called equivalent if there are two constants λ_1 , $\lambda_2 > 0$ such that $\lambda_1 \|x\| \le \|x\|' \le \lambda_2 \|x\|$ for each $x \in X$. These inequalities imply that equivalent norms have the same convergent sequences.

1.6 Normed linear space. A linear space X with a norm $\|\cdot\|$ defined on it is called a normed linear space.

Let us denote by $V^{m\times n}$ the vector space of all matrices of order $m\times n$ over the field F.

Definition 1.6.1. Let A be an element of $V^{m\times n}$. We call a real-valued function $\|.\|$ defined on all square matrices in $V^{n\times n}$, a matrix norm of A on $V^{n\times n}$ and denoted by $\|A\|$ if and only if the following axioms are satisfied:

(1) $\|A\|\geq 0$ and $\|A\|=0$ if and only if $A=0$

(2) $\|\alpha A\|=|\alpha|\|A\|$ for any scalar α in F

(3) $\|A+B\|\leq\|A\|+\|B\|$

(4) $\|AB\|\leq\|A\|\|B\|$

where B is another element in $V^{n\times n}$.

There are numerous ways by which matrix norms can be formed .

The following are matrix norms on $V^{n\times n}$

(a) $\|A\|_a=\sum_{i,j}|a_{ij}|$ (sum of all elements)

(b) $\|A\|_1=\max_j\sum_i|a_{ij}|$

(c) $\|A\|_\infty=\max_i\sum_j|a_{ij}|$

(d) $\|A\|_e=\left(\sum_{i,j}|a_{ij}|^2\right)^{\frac{1}{2}}$ (Euclidean norm), where $A=\lfloor a_{ij} \rfloor$

1.7 Convergence of a sequence in a normed space. A sequence $\{u_n\}$ in a subset U of a normed space is convergent if there is a $u \in U$ for which, given any $\varepsilon > 0$, a number N can be found such that

$$\|u_n - u\| < \varepsilon \text{ for all } n > N \tag{1.1}$$

Suppose we know that a sequence $\{u_n(x)\}$ of continuous functions converges to a limit for each $x \in \Omega \subset R^n$. This implies the following: if we fix x, then the sequence of real numbers $u_n(x)$ $(n=1,2,\cdots\cdots)$ converges to a real number $u(x)$, say. In other words, for every $\varepsilon > 0$ there exists a number N such that

$$|u_n(x) - u(x)| < \varepsilon, \text{ whenever } n > N \tag{1.2}$$

Of course N will depend on x and on the number ε. If we now move to another value of x the statement (1.2) may not be true for some N, a situation which is obviously not desirable. This leads to the following definition of convergence:

Definition 1.7.1. A sequence $\{u_n\}$ of functions defined on an open subset Ω of R^n converges pointwise to $u(x)$ if for $\varepsilon > 0$ there exists a number N depending on x and ε such that (1.2) holds. If N does not depend on the value of x, then u_n converges uniformly to u on Ω and we write $\lim_{n\to\infty} u_n = u$ (uniformly).

Consider a sequence $\{u_n\}$ of functions which belong to the normed space $C[a,b]$ with the norm

$$\|u\|_\infty = \sup |u(x)|, \quad x \in [a,b].$$

Suppose that this sequence is convergent in the sup-norm; that is, given any $\varepsilon > 0$ it is possible to find a number N such that

$$\|u_n - u\|_\infty = \sup | u_n(x) - u(x)| < \varepsilon \tag{1.3}$$

whenever $n > N$, for $x \in [a,b]$ then (1.3) implies that

$$|u_n(x) - u(x)| \leq \sup |u_n(x) - u(x)| < \varepsilon .$$

From which we say that convergence in the sup-norm implies uniform convergence.

Conversely, suppose that $\{u_n\}$ is uniformly convergent sequence, so that (1.2) holds. Then ε is an upper bound for $|u_n(x) - u(x)|$, for any $x \in [a,b]$.

But this implies that

$$\|u_n - u\|_\infty \leq \sup |u_n(x) - u(x)| < \varepsilon \quad \text{for} \quad n > N, \; x \in [a,b]$$

or alternatively,

$$\lim_{n \to \infty} [\, \sup |u_n(x) - u(x)| \,] = 0 .$$

That is, uniform convergence implies convergence in the sup-norm. This useful result can be stated (without proof) in the following theorem.

Theorem 1.7.1. A sequence of functions $\{u_n\}$, where $u_n \in C(\overline{\Omega})$ and Ω is a bounded subset of R^n, converges uniformly if and only if

$$\lim_{n\to\infty}[\ \sup|u_n(x)-u(x)|]=0 \text{ for } x\in\overline{\Omega}.$$

<u>**L_p-Convergence.**</u> Consider the normed space $L_p(\Omega)$ with the usual L_p-norm $\|u\|=\left(\int_\Omega |u|^p\,dx\right)^{\frac{1}{p}}$ with $1\le p<\infty$. The Definition 1.7.1 says that a sequence $\{u_n\}\subset L_p(\Omega)$ converges in the L_p-norm to an element $u\in L_p(\Omega)$ if for any given $\varepsilon>0$ it is possible to find a number N such that $\|u_n-u\|_{L_p}<\varepsilon$ whenever $n>N$,

or $$\left[\int_\Omega |u_n(x)-u(x)|^p\right]^{\frac{1}{p}}<\varepsilon \text{ whenever } n>N,$$

or $$\lim_{n\to\infty}\int_\Omega |u_n(x)-u(x)|^p\,dx=0.$$

This type of convergence is referred to as L_p-Convergence, and the case $p=2$ it is referred to as convergence in the mean. It is important to note that while uniform convergence implies L_p-Convergence, the converse is not true.

1.8 Linear operator. Suppose we are given two linear spaces X and Y over the same scalar field F. An operator is a mapping P which map X into Y over the same field F such that for each $x\in X$ there is a uniquely defined $P(x)\in Y$.

The mapping P is said to be linear if it satisfy two conditions:

(i) $x_1\to y_1$, $x_2\to y_2$ implies $x_1+x_2\to y_1+y_2$

(ii) $x \to y$ implies $\alpha.x \to \alpha.y$

where $x_1, x_2 \in X$, $y_1, y_2 \in Y$ and $\alpha \in F$.

1.9 Non-linear operator. An operator P from a linear space X into a linear space Y is said to be non-linear if it is not a linear operator from X into Y. A simple non-linear operator is one that gives, for all $x \in X$, $P(x) = y_0$, where y_0 is a fixed, nonzero elements.

1.10 Inverse operator. Let P be an operator defined on a vector subspace of X. An operator A defined on the ranges of P, $R(P)$ is called the inverse of P if

$PAx=x$ for all $x \in R(P)$ and

$APx=x$ for all $x \in D(P)$, domain of P.

Theorem 1.10.1. If an operator has an inverse then it is unique.

Theorem 1.10.2. If A is a linear mapping from X into Y, then A^{-1} exists if and only if $Ax = 0$ implies $x = 0$.

1.11 Banach Lemma. Suppose L is a bounded linear operator in X, L^{-1} exists if and only if there is a bounded linear operator M in X such that M^{-1} exists, and

$$\| M - L \| < \frac{1}{\| M^{-1} \|}.$$

If L^{-1} exists, then

$$L^{-1} = \sum_{n=0}^{\infty} (1 - M^{-1}L)^n M^{-1}$$

and

$$\| L^{-1} \| \le \frac{\| M^{-1} \|}{1-\| 1-M^{-1}L \|} \le \frac{\| M^{-1} \|}{1-\| M^{-1} \| \| M-L \|}.$$

Definition 1.11.1. The real number field R is itself a one-dimensional vector space over itself. Then any mapping of $(X,Y) \to (F,F)$, where X is a normed space, is called a functional. If the mapping is linear, it is called linear functional.

Bounded operators. The concept of a bounded operator is closely connected with that of a continuous operator. Let U and V be two normed spaces and let $T:U \to V$ be a linear operator: We say that T is bounded if it is possible to find a number $K > 0$ such that

$$\| Tu \| \le K \| u \| \quad \text{for all } u \in U.$$

For all bounded linear operators T,

$\| Tu \| \le \| T \| \, \| u \|$, where $\| T \| = \sup\{ \| Tu \| / \| u \|,\ u \ne 0 \}$. We have the following theorem connecting the boundedness and continuity of operators:

Theorem 1.11.1. A linear operator T from a normed space U to a normed space V is continuous if and only if it is bounded.

An important class of normed linear space which is named after Stefen Banach (1892-1945), plays an important rule in the existence of the limit x^* of an infinite sequence $\{x_m\}$ of elements of normed linear space X. Consider the sequence $\{x_m\}$ of rational numbers defined by

$$x_0 = 1, \ x_m = \frac{1}{2}(x_{m-1} + \frac{2}{x_{m-1}}), \ m = 1, 2, \cdots\cdots \tag{1.4}$$

There is no rational number x^* which can be the limit of this sequence.

(with $\| x \| = | x |$).

However, if (1.4) generates a sequence $\{x_m\}$ in R, it has a limit x^* which is the solution of the nonlinear equation $x^2 = 2$.

Consequently, the space of real numbers has a property with respect to limits which the set of rational number does not. This property is defined precisely in the more abstract setting in a normed linear space by the following fundamental definition.

Definition 1.11.2. (Cauchy or fundamental sequenc) A sequence $\{x_n\}$ of elements of a normed linear space is called a Cauchy sequence if for every $\varepsilon > 0$ there exist a number λ such that $\| x_m - x_n \| < \varepsilon$ for all $m, n > \lambda$.

Theorem 1.11.2. The following conditions are equivalent

(a) $\{ x_n \}$ is a Cauchy sequence

(b) $\| x_{p_n} - x_{q_n} \| \to 0$ as $n \to \infty$, for every pair of increasing sequences of positive integers $\{ p_n \}$ and $\{ q_n \}$

(c) $\|x_{p_{n+1}} - x_{p_n}\| \to 0$ as $n \to \infty$, for every increasing sequence of positive integers $\{p_n\}$.

From the above theorem it is evident that every convergence sequence is a Cauchy sequence. But the converse is not true. For, let $D([0,1])$ be the space of polynomials on $[0,1]$ with $\|p\| = \max_{[0,1]} |p(x)|$.

Define

$$p_n(x) = 1 + x + \frac{x^2}{2!} + \cdots\cdots + \frac{x^n}{n!}, \text{ for } n = 1, 2, \cdots\cdots$$

Then $\{p_n\}$ is a Cauchy sequence, but it does not converge in $D([0,1])$ because its limit is not a polynomial.

We state the following important theorem without proof:

Theorem 1.11.3. If $\{x_n\}$is a Cauchy sequence in a normed linear space, then the sequence of norms $\{\|x_n\|\}$ converges.

Definition 1.11.3. A normed linear space X is said to be complete if every Cauchy sequence of X converges to a limit which is an element of X.

Theorem1.11.4. Let $(E, \|\cdot\|)$ be a normed space. The function ρ defined by

$$\rho(x,y) = \|x - y\|$$

is a distance on E satisfying the following conditions:

(i) $\rho(x+z, y+z) = \rho(x,y)$ (invariance by translation).

(ii) $\rho(\lambda x,\lambda y)=|\lambda|\rho(x,y)$.

According to theorem 1.11.3, every normed space is a metric space . The structure of a normed space combining the structures of a vector space and of a metric space, has a large role both in mathematical theory and in applications. Complete normed space , in particular , play a very important role.

Definition 1.11.4. A complete normed linear space is called a Banach space.

Theorem 1.11.5. A Euclidean space R becomes a normed linear space when equipped with the norm

$$\|x\|=\sqrt{(x,x)},\ (x\in R)$$

1.12 Balls in normed linear space. Let X be a normed linear space. The open and closed balls in X with center $x_0 \in X$ and radius r can be defined respectively by the subsets

$$B(x_0,r)=\{x:x\in X \quad and \ \|x-x_0\|<r\}$$

and

$$\bar{B}(x_0,r)=\{x:x\in X \quad and \ \|x-x_0\|\le r\}.$$

Example. In the case of R^n, consider the balls

$$B_p(x,\varepsilon)=\{\ y\in R^n \ such\ that\ \|x-y\|_p<\varepsilon\}.$$

In the case where $n=2, x=0, \varepsilon=1,$ the balls $B_p(0,1)$ with center 0 and radius 1 are defined by

$$B_1(0,1)=\{y=(y_1,y_2) \quad such\ that |y_1|+|y_2|\le 1\}$$

$$B_2(0,1)=\{y=(y_1,y_2) \quad such\ that |y_1|^2+|y_2|^2\le 1\}$$

$$B_\infty(0,1)=\{y=(y_1,y_2) \quad such\ that\ \sup(|y_1|,|y_2|)\le 1\}$$

A subset of a normed linear space X is said to be bounded if it is contained in the same ball of finite radius.

1.13 Gateaux and Frechet derivatives. Suppose X and Y be two Banach spaces over the field F, consider the operator $T:X\to Y$ with domain $D_T=X$. Suppose x is a fixed point of X. The operator $T:X\to Y$ is said to be Gateaux differentiable at x if there exists a continuous linear operator L such that

$$\lim_{t\to 0}\|\frac{T(x+th)-T(x)}{t}-L(h)\|=0$$

for every $h\in X$, where $t\to 0$ in F. The operator L is called the Gateaux derivative of T at x, and its value at h is denoted by

$$A(h)=dT(x,h)$$

The notation $dT(x,h)$ or $T'(x)h$ is also used.

Let x be a fixed point in a Banach space X . A continuous linear operator $A:X\to Y$ is called the Frechet derivative of the operator $T:X\to Y$ at x if

$T(x+h)-T(x)=Ah+\varphi(x,h)$.

provided $\lim_{\|h\|\to 0}\frac{\|\phi(x,h)\|}{\|h\|}=0$

or , equivalently

$$\lim_{\|h\|\to 0}\frac{\| T(x,h)-T(x)-Ah \|}{\| h \|}=0$$

The Frechet derivative at x is usually denoted by $T'(x)$ or $dT(x)$. We will now state a theorem that relates these two types of derivatives :

Theorem 1.13.1. If a mapping has the Frechet derivative at a point x, then it has Gateaux derivative at that point and both derivatives are equal.

Corollary: If the Frechet derivative exists, it is unique.

Example. If $f:R^2 \to R$ defined by

$$f(x,y)=\begin{cases}\dfrac{x^3 y}{x^4+y^2} & if\ x\neq 0\ and\ y\neq 0,\\ 0\ if & x=y=0.\end{cases}$$

It is easy to check that f is Gatecux differentiable at 0, and the Gatecux derivative at that point is 0. On the other hand, since

$$\frac{|f(x,x^2)|}{\|(x,x^2)\|}=\frac{|x^3x^2|}{(x^4+x^4)}\frac{1}{\sqrt{x^2+x^4}}=\frac{1}{2\sqrt{1+x^2}}\to\frac{1}{2} \qquad as\ x\to 0,$$

f is not Frechet differentiable at $(0,0)$.

The above example suggests the following theorem:

Theorem 1.13.2. The existence of Frechet derivative implies the Gatecux derivative but the converse is generally false.

Chapter-2

On Convergence and error bound of Newton's algorithm

2.1 Introduction. In this chapter we first introduce the Newton's method in real Euclidean and Banach spaces. We also discuss the famous theorem of Kantorovich and its error bound. We finally state and prove our theorem that concerns a new error bound. We will justify the practical applicability of our theorem fitting the error bound by an example.

2.2 Newton's Method in Real Euclidean Space. Let A be an open set in R, let $f: A \to R$ be a Frechet differentiable function and consider the equation $f(x) = 0$. If x_0 is a point of A near to a root of this equation, then a first approximation be the linear equation

$$f(x_0) + f'(x_0)(x - x_0) = 0$$

and this has the solution

$$x = x_0 - [f'(x_0)]^{-1} f(x_0) \tag{2.1}$$

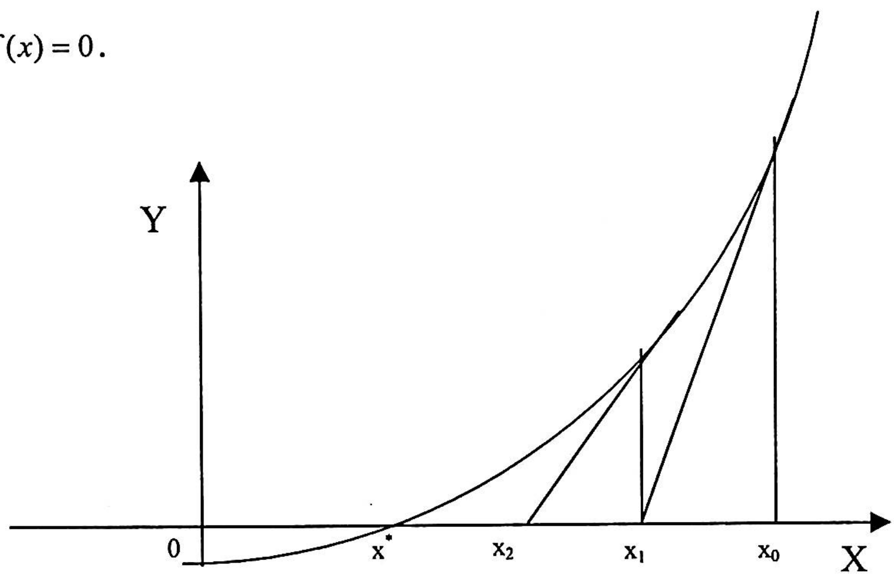
provided that the inverse $[f'(x_0)]^{-1}$ exists. Continuing in this manner, starting from the initial approximation x_0, we obtain points $x_1, x_2, x_3 \cdots\cdots$, given by

$$x_{n+1} = x_n - [f'(x_n)]^{-1} f(x_n) , \qquad n = 0,1,2,3\cdots\cdots \tag{2.2}$$

x_{n+1} being defined so long as $x_1, x_2, x_3 \cdots\cdots x_n \in A$; in effect, the x_n are successive approximations for the equation

$$x = x - [f'(x)]^{-1} f(x).$$

It is intuitive that if we start from a point x_0 for which $f(x_0)$ is sufficiently small, and f' does not vary too much near x_0, then the recurrence relation (2.2) will define a sequence $\{x_n\}$ that converges to a root x^* of the equation $f(x) = 0$.



(Fig-1)

This is particularly transparent in the case of a real-valued function f of a real variable, because the formula (2.2) becomes

$$x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)},$$

so that x_{n+1} is the abscissa of the point where the tangent to the graph of f at x_n meets the x-axis. This case was first considered by Newton and the sequence $\{x_n\}$ given by (2.2) is usually known as Newton sequence for the equation $f(x)=0$.

An alternative possibility is to consider the recurrence relation

$$x_{n+1} = x_n - [f'(x_0)]^{-1} f(x_n), \quad n = 0,1,2,3,\cdots\cdots \tag{2.3}$$

which defines successive approximations for the solution of the equation

$$x = x - [f'(x_0)]^{-1} f(x).$$

The successive approximations using the algorithm (2.3) is known as modifieds Newton method.

2.3 Newton's method in Banach spaces. Let X and Y be (real or complex) Banach spaces, and F be an operator (linear or nonlinear) from X into Y which is twice differentiable in a suitable domain. Starting with an approximate solution x_0 of

$$F(x) = 0 \tag{2.4}$$

we consider the sequence defined by

$$x_{n+1} = x_n - [F'(x_n)]^{-1} F(x_n), \quad n = 0,1,2,3\cdots\cdots \tag{2.5}$$

Kantorovich first proposed to solve the functional equation (2.4) and was able to give theorems concerning the existence, convergence and uniqueness of solution of the equation (2.4). He also gave the error bound for the solution. His two fundamental theorems, the first one guarantees the existence and convergence and the second one guarantees the uniqueness of the solution, are given here. Let us denote by $S(x_0, \rho)$ the open ball

$$\{x \in X : \| x - x_0 \| < \rho\}.$$

Theorem 2.3.1. (Kantorovich; existence and convergence). Suppose the following conditions are satisfied:

(1) $F'(x_0)$ maps X onto Y and has an inverse $[F'(x_0)]^{-1}$ (i.e. one-to-one) for which $\| [F'(x_0)]^{-1} \| \le B$

(2) x_0 is an approximate solution of $F(x) = 0$ such that $\| [F'(x_0)]^{-1} F(x_0) \| \le \eta$, (or equivalently, $\| x_1 - x_0 \| \le \eta$);

(3) F is twice differentiable in the open ball $U_0(x_0, \rho_0)$, and in this ball

$$\| F''(x) \| \le \kappa$$

where κ is a constant and

$$\rho_0 = (1 - \sqrt{1 - 2h})(\frac{\eta}{h})$$

and for constants B, κ, η, satisfying

(4) $h_0 = B\eta\kappa \le \frac{1}{2}$.

Then $F(x) = 0$ has a solution x^* in the closed ball $\overline{U_0}(x_0, \rho_0)$ and the successive approximations defined by (2.5) converge to x^*.

Further

$$\| x_n - x^* \| \le \frac{1}{2^{n-1}} (2h_0)^{2^n - 1} \eta \tag{2.6}$$

Here we can make remark that $h_0 > 0$ holds always , since $h_0 = 0$ if and only if $B = 0$ (impossible for bounded of $F'(x_0)$) or $\eta = 0$ (then x_0 is already a solution). We also see that the restriction on h, viz. $0 < h \le \frac{1}{2}$, gives $1 < (1 - \sqrt{1-2h})(\frac{\eta}{h}) \le 2$. Thus condition (3) of the theorem 2.3.1 holds if $\| F''(x) \| \le \kappa$ in $U_0(x_0, 2\eta)$.

2.3.1 Numerical example of Kantorovich theorem. Here we consider the third polynomial equation $F(x) = x^3 - 3x + 3 = 0$. This equation has only one real root $x^* = \alpha = -2.103803402......$ Let $x_0 = -2.11$ be the approximate solution. Then we have $F(x_0) = -0.06$, $F'(x_0) = 10.35$ and $F''(x_0) = -12.66$.

We can calculate.

$$\| [F'(x_0)]^{-1} \| = |\frac{1}{10.35}| = 0.096618357 \le 0.09662 = B,$$

$$\| [F'(x_0)]^{-1} F(x_0) \| = |\frac{-0.06}{10.35}| = 0.0057971 \le 0.00580 = \eta$$

and

$$\| F''(x_0) \| = |-12.66| \le 12.67.$$

Now we have $h = B\eta k = 0.09662 \times 0.00580 \times 12.67 = 0.00710021732$.

Thus the solution $x^* = -2.103803402.....$ is the only one real root of the above polynomial equation. By the Newton's algorithm,

$$x_1 = x_0 - \frac{F(x_0)}{F'(x_0)} = -2.104202899.$$

So, $\|x_1 - x^*\| = \|-2.104202899 + 2.103803402\| = 0.000399497$, and from the inequality (2.6), $\frac{1}{2^{n-1}}(2h)^{2^n-1}\eta$ becomes 0.01160. Then $\|x_1 - x^*\| \le 0.01160$ is an improvement of the approximation.

The local uniqueness of the solution x^* depends on the bound holding in a larger sphere (or ball):

Theorem 2.3.2. (Kantorovich ; uniqueness) Let the condition (1) to (4) of the theorem 2.3.1 hold with $\| F''(x) \| \le \kappa$ in the open ball $U_0(x_0, \sigma)$. where

$$\sigma = (1 + \sqrt{1-2h})(\frac{\eta}{h}).$$

Then the x^* of theorem 2.3.1 is the unique solution of $F(x) = 0$ in the same ball. Kantorovich in his theorem 2.3.1 used the boundness of the second derivative of the operator. Fenyo I. (1954) first make the assumption of the condition of Lipschitian of F' and gave the modified Kantorovich theorem 2.3.1 as:

Theorem 2.3.3. (Kantorovich; modified) Let $F: X \to Y$, X, Y Banach spaces, be Frechet differentiable function for $x \in U$, an open convex set in X. [Ortega (1968) and Tapia(1971)]

Let $[F'(x_0)]^{-1} \in [Y \to X]$ at some $x_0 \in U$, and

(1) $\| [F'(x_0)]^{-1} \| \le B$,

(2) $\| [F'(x_0)]^{-1} F(x_0) \| \le \eta$,

(3) $\| F'(x) - F'(y) \| \le K \| x - y \|, \quad x, y \in U$,

for constants B, K, η satisfying $h = BK\eta \le \frac{1}{2}$ and

(4) $U_0 \subset U$, where

$$U_0 = \{x : \| x - x_0 \| \le (1 - \sqrt{1-2h})(\frac{\eta}{h})\}$$

then the successive approximations (2.5) of Newton's algorithm are defined for all n, $x_n \in U_0$, $n = 0,1,2,3\cdots\cdots$, and converge to $x^* \in U_0$, which satisfies $F(x^*) = 0$.

Further, $\| x^* - x_n \| \le \frac{\eta}{h} \frac{[1-\sqrt{1-2h}]^{2^n}}{2^n}$, $n = 0,1,2,3,\cdots\cdots$ (2.7)

An example of f is now shown which demonstrates that no stronger claim than the Kantorovich theorem can be made for existence and convergence. In this generality, then, no better theorem can be given.

Let $f(x) = \frac{1}{2}x^2 - x + h$ and $x_0 = 0$. We have $f'(x_0) = -1$ and $f''(x_0) = 1$. Here, we let $B = K = 1$, $\eta = h$ (given), the conditions of the theorem are satisfied.

The roots of f are $1 \pm \sqrt{1-2h}$ if $h \le \frac{1}{2}$.

The smaller root

$$x^* = (1-\sqrt{1-2h}) = \frac{1}{h}(1-\sqrt{1-2h})\eta .$$

The other root is

$$\frac{1}{h}(1+\sqrt{1-2h})\eta$$

and is just excluded from the region. Presumably, Kantorovich (1948) obtained his theorem by comparison with the Newton series for $\frac{1}{2}x^2 - x + h$, for comparison with the Newton series for $x^2 - a$.

Under the same conditions as for theorem 2.3.3, it can be shown that if $\Gamma_0 = [F'(x_0)]^{-1}$ and $h < \frac{1}{2}$ the successive approximation given by the following theorem converge and is known as modified Newton's method.

Theorem 2.3.4. (Modified Newton's method) Under the same conditions as in theorem 2.3.3, if $h < \frac{1}{2}$, the iterations $x_{n+1} = x_n - \Gamma_0 F(x_n)$ where $\Gamma_0 = [F'(x_0)]^{-1}$ are defined for all n, and for any $x_0 \in U$, converge to a root $x^* \in U_0$.

Further, $F(x) = 0$ has a unique root in U_0.

Also

$$\| x_n - x^* \| \le 2(\frac{\eta}{h})[1-\sqrt{1-2h}]^{n+1}, \qquad n = 1,2,3\cdots\cdots$$

2.4 Newton-like or Quasi-Newton method. For a real function f of a real variable x, if we try to find an approximate root of $f(x)=0$ by the sequence of approximations

$$x_{n+1} = x_n - [a_n]^{-1} f(x_n),\ n = 0,1,2,\cdots\cdots,$$

where $\{a_n\}$ is a sequence of real numbers, then the algorithm above is the simplest form of the Newton-like or Quasi-Newton method.

2.5 Illustration of the applicability of Newton's method to nonlinear operator equation:

Let we consider the Hammerstein equation

$$x(t) + \int_0^1 k(s,t) f\big(s, x(s)\big) ds = 0 \tag{2.8}$$

on the space $C[0,1]$. We suppose that $f \in C^2\big([0,1]\times R\big)$ and $k \in C\big([0,1]\times[0,1]\big)$.

If we define $F: C[0,1] \to C[0,1]$ by

$$F(x)(t) = x(t) + \int_0^1 k(s,t) f(s, x(s)) ds,$$

then our problem is to find a root of F. If $x_0 \in C[0,1]$ is an initial approximation, then for any $y \in C[0,1]$,

$$\begin{aligned}\big[F'(x_0)y\big](t) &= y(t) + \int_0^1 k(s,t) f_2'(s, x_0(s)) y(s) ds \\ &= \big(1 + k_0\big) y(t),\end{aligned}$$

where k_0 is the linear integral operator on $C[0,1]$ defined by

$$(k_0 y)(t) = \int_0^1 k(s,t) f_2'(s, x_0(s)) y(s) ds.$$

We suppose that $1+k_0$ is invertible and $\| (1+k_0)^{-1} \| \leq b$. For a given $\delta > 0$, let

$$k_\delta = \sup\{| k(s\ ,t) f_2''(s,u) | :\ s, t \in [\ 0\ , 1],\ |\ u - x_o(s) | \leq \delta\}.$$

Then k_δ serves as a Lipschitz constant for F' on the set

$$S_\delta = \{x \in C[0,1] : \| x - x_0 \| \leq \delta\}.$$

If $\| F(x_0) \| \leq p$, then

$$\| [F'(x_0)]^{-1} F(x_0) \| = \| (1+k_0)^{-1} F(x_0) \| \leq bp\ .$$

Therefore, if $h = b^2 k_\delta p \leq ½$ and $\delta \geq (bk_\delta)^{-1}$, then

$$\{x \in C[0,1] : \| x - x_0 \| \leq t^*\} \subseteq S_\delta,$$

and the following theorem guarantees that the functions x_n given by $x_{n+1} = x_n + y_n$, where the functions y_n are solutions of the linear integral equations

$$y_n(t) + \int_0^1 k(s,t) f_2'(s, x_n(s)) y_n(s) ds = -x_n(t) - \int_0^1 k(s,t) f(s, x_n(s)) ds,$$

converge to a solution of (2.8). Cherles W. Groetsch (1980) has discussed the additional applications of Newton's method.

Theorem 2.5.1. Suppose that C is an open convex subset of a Banach space X and that Y is a Banach space. Let $F:C \to Y$ be differentiable on C and satisfy.

$$\| F'(x) - F'(y) \| \leq k \| x - y \|$$

for $x, y \in C$. Assume that for some $x_0 \in C, G_0 = [F'(x_0)]^{-1}$ exists and that $\| G_0 \| \leq b$ and $\| G_0 F(x_0) \| \leq \eta$, where $h = b\eta k \leq \frac{1}{2}$. Set $t^* = \frac{\eta(1-\sqrt{1-2h})}{h}$ and suppose that

$$S = \{x \in X : \| x - x_0 \| \leq t^*\} \subseteq C$$

Then the Newton's sequence (2.5) is defined, lies in S and converges to a root x^* of F.

Moreover,

$$\| x^* - x_k \| \leq \frac{\eta(1-\sqrt{1-2h})^{2^k}}{(h2^k)} \qquad k = 0,1,2,\cdots\cdots$$

Let X and Y be Banach spaces and D^0 is an open convex subset of X. Also let $F: D^0 \to Y$ be Frechet differentiable on D^0 with $\| F'(x) - F'(x') \| \leq \lambda \| x - x' \|$ for $x, x' \in D^0$.

Let $S(x,r)$ denotes the open ball $\{x' : \| x' - x \| < r\}$ and $\overline{S(x,r)}$ denotes its closure.

Let $x_0 \in D^0$ be such that $[F'(x_0)]^{-1} : Y \to X$ exists, $\| [F'(x_0)]^{-1} \| \leq x$, $\| [F'(x_0)]^{-1} F(x_0) \| \leq \delta$, $h = 2x\lambda\delta \leq 1$ and $S(x, t^*) \subset D^0$, $t^* = \frac{2}{h}(1-\sqrt{1-h})\delta$.

Then

1. The Newton sequence $\{x_n\}$ exists and $x_n \in S(x,t^*) \subset D^0$ for $n \geq 0$

2. $x^* = \lim x_n$ exists, $x^* \in \overline{S(x_0,t^*)} \subset D$ and $F(x^*) = 0$.

3. x^* is the only solution of $F(x) = 0$ in the set $S(x_0,t') \cap D^0$, $t' = \frac{2}{h}(1+\sqrt{1-h})\delta$, if $h < 1$, and in $\overline{S(x_0,t')}$ if $h = 1$.

With the above assumptions and Kantorovich hypotheses, Gragg W.B. and Tapia R.A. (1974) were able to give the following best possible lower and upper bounds for error:

$$\| x^* - x_n \| \leq \frac{4\sqrt{1-h}}{h} \frac{\theta^{2^n}}{1-\theta^{2^n}} \| x_1 - x_0 \|$$

and

$$\frac{2\| x_{n+1} - x_n \|}{1+\sqrt{\frac{1+4\theta^{2^n}}{(1+\theta^{2^n})^2}}} \leq \| x^* - x_n \| \leq \theta^{2^{n-1}} \| x_n - x_{n-1} \|$$

where

$$\theta = \frac{1-\sqrt{1-h}}{1+\sqrt{1-h}} \leq 1.$$

In particular, the bounds for $h = 1$ are

$$\| x^* - x_0 \| \leq 2^{-n+1} \| x_1 - x_0 \|$$

and

$$2(\sqrt{2}-1)\| x_{n+1} - x_n \| \leq \| x^* - x_n \| \leq \| x_n - x_{n-1} \|$$

holds for all $h \le 1$.

In our study, we are trying to give a new best possible error bound constructing a strictly monotonic function that generalizes the numerical bound given in Rall L.B. (1974).

The hypotheses (1), (2) and (4) of theorem 2.3.3 guarantee the existence and convergence of the sequence $\{x_n\}$ obtained by using theorem 2.3.1 for the functional equation $F(x)=0$ to $x^* \in U_0$ such that $F(x^*)=0$.

The Kantorovich hypotheses are

(i) $\| [F'(x^*)]^{-1} \| \le B^*$ (some constant) and the open ball

(ii) $U_0 = \{x : \| x - x^* \| \le \frac{1}{(B^* K)}\} \subset U$ that satisfy (1), (2) and (3) with $h < \frac{1}{2}$ and (4) of theorem 2.3.3.

Rall (1974) had discussed the error bound of the unique solution of Newton's algorithm using Kantorovich hypotheses, where he considered the open ball

$$U_0 = \{x : \| x - x^* \| < (2-\sqrt{2})/(2B^* K)\}$$

Here, we show that the bound $(2-\sqrt{2})/(2B^* K)$ is a special case of our generalized bound given by $(r(\theta)-\sqrt{r(\theta)})/r(\theta)B^* K)$, where $r(\theta) = \frac{1-\sqrt{1-2\theta}}{\theta}$ is a monotone increasing function on $\left(0, \frac{1}{2}\right]$.

For the graphs of the increasing monotone function $r(\theta)$ on the interval $\left(0,\frac{1}{2}\right]$

see fig.1 and fig.2.

We now state and prove our theorem:

Theorem 2.5.2. If x^* be a simple zero of F,

$$\| [F'(x^*)]^{-1} \| \le B^*$$

and

$$U_* = \{x : \| x - x^* \| < \frac{1}{(B^* K)}\} \subset U$$

and if the hypotheses (i) and (ii) of Kantorovich theorem are satisfied at each

$x_0 \in U^*$, where

$$U^* = \left\{x : \| x - x^* \| < (r(\theta) - \sqrt{r(\theta)}) / (r(\theta) B^* K)\right\}.$$

Then h is given by $\frac{(r(\theta) - 1)}{2}$.

Proof. For $x_0 \in U^*$

$$\frac{r(\theta) - \sqrt{r(\theta)}}{r(\theta) B^* K} = \frac{r(\theta) - \sqrt{r(\theta)}}{r(\theta) \frac{1}{Kx^*} K} = (1 - \frac{\sqrt{r(\theta)}}{r(\theta)}) x^*, \text{ where } B^* = \frac{1}{Kx^*}$$

$$x_0 = x^* - (1 - \frac{\sqrt{r(\theta)}}{r(\theta)}) x^* = \frac{x^*}{\sqrt{r(\theta)}},$$

$$\| F'(x_0) - F'(x^*) \| \le K \| x_0 - x^* \| < K \| -(1 - \frac{\sqrt{r(\theta)}}{r(\theta)}) x^* \|$$
$$< \left(r(\theta) - \sqrt{r(\theta)}\right) / \left(r(\theta) B^*\right)$$
$$< \| [F'(x^*)]^{-1} \|^{-1}$$

So that $[F'(x_0)]^{-1}$ exists and by the Banach lemma (sec: 1.11),

$$B = \frac{B^*}{1 - B^* K \| x_0 - x^* \|} = \left(\frac{\sqrt{r(\theta)} K x^*}{r(\theta)} \right)^{-1} \ge \| [F'(x_0)]^{-1} \| \tag{2.9}$$

Using the fundamental theorem of calculus, we have

$$F(x^*) - F(x_0) = \int_0^1 F'(x_0 + t\,(x^* - x_0))(x^* - x_0) dt.$$
$$= F'(x_0)(x^* - x_0) + \int_0^1 [F'(x_0 + t(x^* - x_0)) - F'(x_0)](x^* - x_0) dt$$

As $F(x^*) = 0$,

$$-[F'(x_0)]^{-1} F(x_0) = (x^* - x_0) + [F'(x_0)]^{-1} \int_0^1 [F'(x_0 + t(x^* - x_0)) - F'(x_0)](x^* - x_0) dt$$

Thus

$$\| [F'(x_0)]^{-1} F(x_0) \| \le \left\{ 1 + BK \| x^* - x_0 \| \int_0^1 t\, dt \right\} \| x^* - x_0 \|.$$

From which (2.9) may be used to get

$$\eta = \frac{1 - \frac{1}{2} B^* K \| x^* - x_0 \|}{1 - B^* K \| x^* - x_0 \|} \| x^* - x_0 \| = \frac{\frac{1}{2}(1 - \frac{1}{r(\theta)}) x^*}{\frac{1}{\sqrt{r(\theta)}}} \ge \| [F'(x_0)]^{-1} F(x_0) \|.$$

It follows that from $x_0 \in U^*$

$$h = BK\eta = \left(\frac{K x^*}{\sqrt{r(\theta)}} \right)^{-1} K \frac{\frac{1}{2}(1 - \frac{1}{r(\theta)}) x^*}{\frac{1}{\sqrt{r(\theta)}}} = \frac{1}{2}\left(r(\theta) - 1\right).$$

The following corollary of our theorem is the theorem of Rall(1974):

Corollary. If x^* is a simple zero of F,

$$\| [F'(x_0)]^{-1} \| \leq B^*$$

and

$$U_* = \left\{ x : \| x - x^* \| < \frac{1}{B^* K} \right\} \subset U$$

then the hypotheses (i) and (ii) of Kantorovich theorem are satisfied at each $x_0 \in U^*$, where

$$U^* = \left\{ x : \| x - x^* \| < (\sup_{\theta \in (0,\frac{1}{2}]} r(\theta) - \sqrt{\sup_{\theta \in (0,\frac{1}{2}]} r(\theta)}) / (\sup_{\theta \in (0,\frac{1}{2}]} r(\theta)\, B^* K) = (2 - \sqrt{2}) / (2B^* K) \right\}$$

By the following example, we justify the physical applicability of our theorem.

Example. Consider the quadratic operator

$$F(x) = \frac{1}{2} K (x^2 - x^{*2}), \text{ where } K > 0.$$

Therefore,

$U^* = \{ x : \| x - x^* \| < (r - \sqrt{r}) / (r B^* K) \}$ where r is a positive real number.

We have $\frac{r - \sqrt{r}}{r B^* K} = \frac{r - \sqrt{r}}{r \frac{1}{Kx^*} K} = (1 - \frac{\sqrt{r}}{r}) x^*$ and for

$$x_0 = x^* - (1 - \frac{\sqrt{r}}{r}) x^* = \frac{x^*}{\sqrt{r}}, \text{ where } B^* = \frac{1}{kx^*},$$

$$B = \frac{B^*}{1 - B^* K \| x_0 - x^* \|} = \left(\frac{\sqrt{r} K x^*}{r} \right)^{-1} \geq \| [F(x_0)]^{-1} \|,$$

and

$$\eta = \frac{1 - \frac{1}{2} B^* K \| x_0 - x^* \|}{1 - B^* K \| x_0 - x^* \|} \| x_0 - x^* \| = \frac{\frac{1}{2}(1 - \frac{1}{r}) x^*}{\frac{1}{\sqrt{r}}} \geq \| [F'(x_0)]^{-1} F(x_0) \|.$$

From which

$$h = BK\eta = \frac{\frac{1}{2}(1-\frac{1}{r})}{\frac{1}{r}} = (\frac{r}{2} - \frac{1}{2}).$$

Hence, $h = \frac{1}{2}$ if $r = 2$.

With this value $r = 2$, the best possible error bound found from our theorem is thus justified.

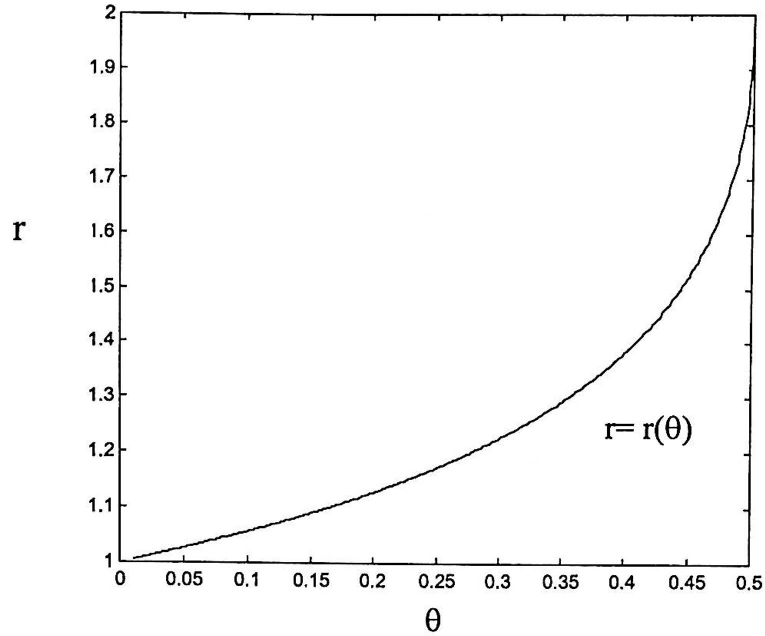


Figure 1. Graph of r(θ)

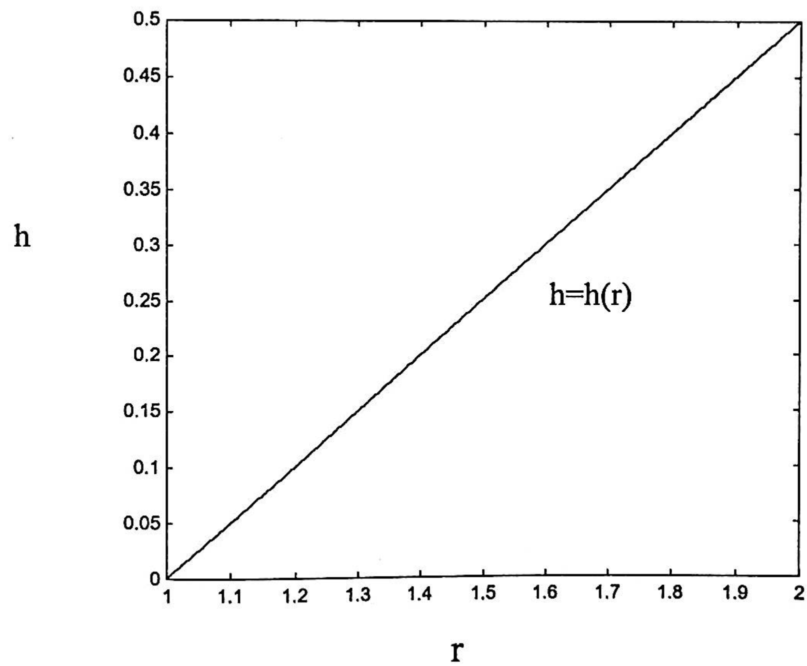


Figure 2. Graph of h(r)

Chapter-3

Interval Number System and Some Fixed Point Theorems.

3.1 Introduction. Many approaches exist for treating the subject of errors in applied mathematics of which we mention the perturbation approach.

Suppose we are given an operator equation.

$$(L + \varepsilon L_1)x(\varepsilon) = f \tag{3.1}$$

Using perturbation technique we seek solution of (3.1) in the form

$$x(\varepsilon) = x_0 + \varepsilon x_1 + \varepsilon^2 x_2 + \cdots\cdots\cdots \tag{3.2}$$

To solve (3.1) recursively for $x_1, x_2, \cdots\cdots$ etc, many problems arise

(i) What guarantees the above expression of $x(\varepsilon)$ as a power series in ε?

(ii) What if ε is big ? Convergence criteria must be investigated.

(iii) What if ε itself is unknown. In tolerance problem for instance, only an upper and lower bound for ε are all what is given.

A subject, which answers the above questions, has appeared as an independent discipline in the mid-sixties, called interval analysis with the pioneering work of Moore (1966) together with the research carried out by his co-workers.

Here we discuss some concepts relating to interval number systems.

3.2 Interval number system. Let the field of real numbers be denoted by R and the numbers of R by lowercases latter $a,b,c,\cdots\cdots,x,y,z$. A subset of R of the form

$$I=[a_1,a_2]=\{t:a_1\le t\le a_2,\quad a_1,a_2\in R\}$$

is called a closed real interval or an interval.

Let the set of all closed real intervals be denoted by $\Im(R)$ and the members of $\Im(R)$ by $I_1,I_2,\cdots\cdots$.

It is seen that the real numbers $x\in R$ may be considered as the special numbers $[x,x]$ from R and this type of intervals are called degenerate intervals.

Two intervals $I_1=[a_1,a_2]$ and $I_2=[b_1,b_2]$ are called equal if $a_1=b_1$ and $a_2=b_2$.

The width of an interval $[a\ ,b]$ is given by $W([a,b])=b-a$ and its magnitude by

$$|[a,b]|=\max(|a|,|b|).$$

Let $*\in\{+,-,\bullet,/\}$ be a binary operation on the set of real numbers R. If $I_1,I_2\in\Im(R)$ then $I_1*I_2=\{t=a*b:a\in I_1,b\in I_2\}$ defines a binary operations on $\Im(R)$.

We now generalize the arithmetic of real numbers with the above binary operations on elements of $\Im(R)$.

Let $I_1 = [a,b]$ and $I_2 = [c, d]$ be two intervals then the operations of addition, subtraction, multiplication and division of I_1 and I_2 are defined by

$$I_1 + I_2 = [a,b] + [c,d] = [a+c, b+d],$$

$$[a,b] - [c,d] = [a-d, b-c],$$

$$I_1 * I_2 = [a,b] * [c,d] = [\min(ac,ad,bc,bd), \max(ac,ad,bc,bd)],$$

and

$$I_1 / I_2 = [a,b] / [c,d] = [a,b] * \left[\frac{1}{d}, \frac{1}{c}\right] \text{ , where } 0 \notin [c,d].$$

Interval addition and interval multiplication are both associative and commutative but the distributive law does not hold for interval arithmetic.

For example,

$$[1, 2]([1,2]-[1,2]) = [1,2]([-1, 1]) = [-2,2]$$

whereas

$$[1,2][1,2] - [1,2][1,2] = [1,4]-[1,4] = [-3,3]$$

For the intervals I, J, K the law $I \cdot (J+K) \subset I \cdot J + I \cdot K$ is referred to as subdistributivity of interval arithmetic.

Interval arithmetic is inclusion monotonic. That is, for the intervals I, J, K and L the following hold:

$$I+J \subset K+L$$
$$I-J \subset K-L$$
$$IJ \subset KL$$
$$I/J \subset K/L \quad (\textit{is } 0 \notin L).$$

Let $f(x_1, x_2, \cdots\cdots, x_n)$ be a real rational function of the variables $x_1, x_2, \cdots\cdots, x_n$ where each x_i occur only once then the corresponding interval extension $F(X_1, X_2, \cdots\cdots, X_n)$ computes the actual range of values of f for x_i belonging to the intervals X_i.

Definition 3.2.1. The distance between two intervals $I_1 = [a, b]$ and $I_2 = [c, d]$ is defined by $\rho(I_1, I_2) = \max\{|a-c|, |b-d|\}$.

It is easy to show that the map ρ introduces a metric in $\Im(R)$, since ρ has the properties

$$\rho(I_1, I_2) \geq 0,$$

and $$\rho(I_1, I_2) = 0 \Rightarrow I_1 = I_2,$$

$$\rho(I_1, I_2) \leq \rho(I_1, I_3) + \rho(I_3, I_2) \text{ (triangle inequality).}$$

We now state two important theorems:

Theorem 3.2.1. The metric space $(\Im(R), \rho)$ with the metric of Definition 3.2.1 is a complete metric space.

(This means that every Cauchy sequence of intervals converges to an interval.)

Theorem 3.2.2. Every sequence of intervals $\{I^{(k)}\}^{\infty}{}_{k=0}$ for which

$$I^{(0)} \supseteq I^{(1)} \supseteq I^{(2)} \supseteq \cdots\cdots$$

is valid, converges to the interval $I = \bigcap_{k=0}^{\infty} I^{(k)}$.

Let (M, ρ) be a metric space with the metric defined by $\rho : M \times M \to R$, then an interval valued function $f : M \to \Im(R)$ on M is continuous at $x_1 \in M$ if and only if for every $\varepsilon > 0$ that is a $\delta > 0$ such that for all $x_2 \in M$, $\rho(x_1, x_2) < \delta$ $\Rightarrow \rho(f(x_1), f(x_2)) < \varepsilon$, and f is uniformly continuous in M if and only if for every $\varepsilon > 0$ there is a $\delta > 0$ such that for all $x_1, x_2 \in M$, $\rho(x_1, x_2) < \delta$ $\Rightarrow \rho(f(x_1), f(x_2)) < \varepsilon$.

3.2.1 Linear Interval Equation. It is often desirable in a variety of applications to obtain a solution to the linear system $Ax = b$ in which A and b are both affected by uncertainties. In that case we are concerned with determining the tolerance in each component x_i, of the solution x knowing the tolerance inherent in each element a_{ij} or b_j.

This type of problem can be solved by characterizing the equation in the interval version:

$$A^I x = b^I$$

where A^I and b^I are an interval matrix and an interval vector respectively having upper and lower bounds, that is $A^I = [\underline{A}, \bar{A}]$, $b^I = [\underline{b}, \bar{b}]$.

By solving for x, we mean to solve the equation

$$Ax = b$$

in which A and b ranges respectively over A^I and

$$b^I \left(\underline{A} \le A \le \bar{A} \quad and \quad \underline{b} \le b \le \bar{b} \right).$$

An $A \in A^I$ and $b \in b^I$ are usually written in the more practical form

$$A^c - \Delta A \le A \le A^c + \Delta A, \qquad b^c - \Delta b \le b \le b^c + \Delta b$$

in which $A^c = \dfrac{\bar{A} + A}{2}$ and $b^c = \dfrac{\bar{b} + b}{2}$, ΔA and Δb are the uncertainties (maximum errors) in A and b, i.e.

$$\Delta A = \frac{\bar{A} - A}{2} \quad \text{and} \quad \Delta b = \frac{\bar{b} - b}{2}.$$

We assume that the matrix A contained in A^I is nonsingular. Then solving the equation

$$A^I x = b^I$$

means that we are able to determine an interval solution x^I, one of the smallest width, enclosing all possible values of the vector $x \in R^n$ satisfying $Ax = b$ when A and b assume all possible combination inside A^I and b^I. In other words, we seek an exact hull to the set $X = \{x : Ax = b, A \in A^I, b \in b^I\}$.

The answer to this problem was first supplied by Oettli and Prager (1964) who gave full characterization of the set X by stating the following:

Theorem 3.2.3. Any solution x to be linear equations $Ax=b$, when $A \in A^I$ and $b \in b^I$ satisfies

$$\left|\sum_{j=1}^{n} a_{ij}^{c} x_j - b_i^c\right| \le \sum_{j=1}^{n} \Delta a_{ij} \left|x_j\right| + \Delta b_i,\ i=1,\cdots\cdots,n,$$

where a_{ij}^c and b_i^c denote respectively mean values of a_{ij} and b_i , Δa_{ij} and Δb_i their range of uncertainties. The vertical bars stand for absolute values.

The proof follows from the perturbed problem

$$\left(A^c + \delta A\right)x = b^c + \delta b$$

in which the errors δA and δb scan the range of uncertainties ΔA and Δb ; that is

$$\delta A \in \left[-\Delta A, \Delta A\right], \quad \delta b \in \left[-\Delta b, \Delta b\right].$$

And by writing the above equation in the form

$$A^c x - b^c = -\delta Ax + \delta b$$

we notice that the right-hand side does not lie outside the interval $[-\Delta A|x|-\Delta b, \Delta A|x|+\Delta b]$ so does the left-hand side, i.e. that

$$-\Delta A|x| - \Delta b \le A^c x - b^c \le \Delta A|x| + \Delta b$$

which is the Oettli-Prager criterion.

Although here we prove necessity only, the Oettli and Prager theorem implies the converse of the above statement too.

For, let the above inequality hold, then $A^c x - b^c = D(\Delta A|x| + \Delta b)$ where D is diagonal and $d_{ii} \in [-1, 1]$, i.e. the existence of

$A \in A^I \left(A = A^c - D\Delta A S, \quad S = sgm(x) \right)$ and $b \in b^I \left(b = b^c + D\Delta b \right)$ hold with $Ax = b$.

Example: For the equations (Hansen (1969))

$$[2\ ,\ 3]\ x_1 + [0\ ,\ 1]\ x_2 = [0\ ,\ 120]$$

$$[1\ ,\ 2]\ x_1 + [2\ ,\ 3]\ x_2 = [60\ ,\ 240]$$

one has

$$A^c = \begin{bmatrix} 5/2 & 1/2 \\ 3/2 & 5/2 \end{bmatrix},\ \Delta A = \begin{bmatrix} 1/2 & 1/2 \\ 1/2 & 1/2 \end{bmatrix},\ b^c = \begin{bmatrix} 60 \\ 150 \end{bmatrix},\ \Delta b = \begin{bmatrix} 60 \\ 90 \end{bmatrix}$$

Thus $$\left|\tfrac{5}{2}x_1 + \tfrac{1}{2}x_2 - 60\right| \le \tfrac{1}{2}|x_1| + \tfrac{1}{2}|x_2| + 60$$

and

$$\left|\tfrac{3}{2}x_1 + \tfrac{5}{2}x_2 - 150\right| \le \tfrac{1}{2}|x_1| + \tfrac{1}{2}|x_2| + 90$$

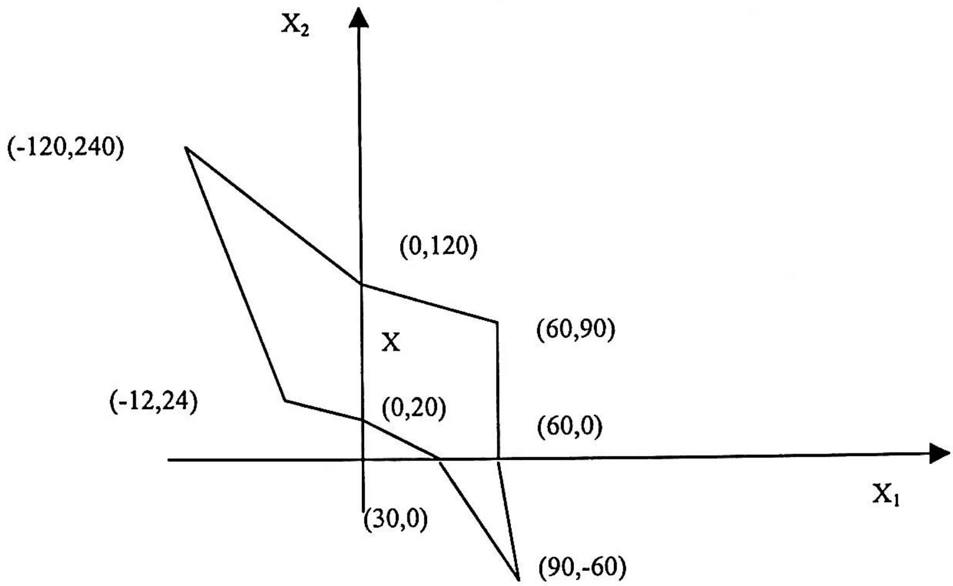
define the set X completely. They read in the first quadrant

$$2x_1 \le 120 \qquad 3x_1 + x_2 \ge 0$$

$$x_1 + 2x_2 \le 240 \qquad 2x_1 + 3x_2 \ge 60$$

by letting $|x_1| = x_1$ and $|x_2| = x_2$. To seek X in the second quadrant, we set $|x_1| = -x_1$, $|x_2| = x_2$ and so on. And although X is convex in any one quadrant, i.e. composed of union of convex sets, it could be generally non-convex.

The figure below demonstrates the various regions of the set X.



Since the interval solution x^I is the narrowest interval containing X, one has from the above figure

$$x^I = ([-120,90][-60,240])^T$$

which is the exact bound for x.

3.2.2 Interval version of Newton's method:

Mean value theorem. We now state the mean-value theorem in interval numbers, which is very important to study the convergence of Newton's method. Let us consider f be a real valued function with a continuous derivative f' on any interval $[a,b]$. If f is a bounded real rational function defined on $[a,b]$ then the mean- value theorem states that for any $x,y \in [a,b]$,

$$f(y) = f(x) + f'(x + \theta(y-x))(y-x)$$

for some $\theta \in [o,1]$. If $\bar{f}'$ is the united extension of f', then it can be concluded that

$$f(y) \in f(x) + \bar{f}'(x + (y-x)[0,1])(y-x)$$

Let $F^{(1)}$ be an interval-valued function $F^{(1)} : \ell_A \to \ell$ defined for $X \subset A$ such that $F^{(1)}(X) \supset \bar{f}'(X)$, then we also have

$$f(y) \in f(x) + F^{(1)}(x + (y-x)[0,1])(y-x)$$

and also, for $X \subset [0,1] \cap A$,

$$\bar{f}(X) \subset f(x) + F^{(1)}(x + (X-x)[0,1])(X-x)$$

From the mean value theorem for the bounded real valued function f on $[a,b]$ which has continuous derivative f' on $[a,b]$, we can write for any $x,y \in [a,b]$:

$$f(x) = f(y) + f'(y + \theta(x-y))(x-y),$$

where $\theta \in [o,1]$. If x is not a zero of f and f' has a constant sign on $[a,b]$, then

$$x = y + \left(\frac{1}{-f'(y+\theta(x-y))} \right) f(y). \tag{3.3}$$

Thus y is a fixed point of the function on the right-hand side of (3.3) if and only if it is a zero of f.

Suppose that F' is a rational interval extension of $f' : \bar{f}'(X) \subset F'(X)$ and

$f'(x) = F'([a,b])$.

For $X \subset [a,b]$, and $m(X)$= the midpoint of X, we define the interval function N (N for 'Newton') by

$$N(X) = m(X) + \left(\frac{1}{-F'(X)}\right) f(m(X)). \tag{3.4}$$

Approximating θ by 0 in (3.3) gives rise to Newton's method for approximating roots by iteration of the function on the right-hand side of (3.3). The equation (3.4) can be used in a similar fashion to provide an "interval version" of Newton's method, namely by choosing X_0 and defining the sequence of intervals $X_1, X_2, \cdots\cdots$ with

$$X_{n+1} = N(X_n) \cap X_n$$

3.3 Fixed Point Theorems. If F is an operator that maps the Banach space X into itself, then any $x \in X$ such that

$$x = F(x) \tag{3.5}$$

is called a fixed point of the operator F.

For example, the operator $F(x) = x^2$ in the space R of real numbers has the fixed points $x = 0$ and $x = 1$.

The linear operator

$$F(x) = x(0) + \int_0^s x(t)dt \tag{3.6}$$

in $C[0,1]$ has any function $x = x(s)$ of the form

$$x(s) = ce^s, \qquad 0 \le s \le 1, \tag{3.7}$$

as a fixed point, where c is a real constant.

The method for finding a fixed point of an equation $f(x)=0$, is an iteration method. This method is based on the principle of finding a sequence $\{x_n\}$, each element of which successively approximates a root x^* of the equation $f(x)=0$ in some interval $[a,b]$. So, there is a deep rooted connection between the study of fixed point theorems and Newton's method of iteration.

The iteration process $x_{n+1} = f(x_n)$ leads to a solution of the equation $x = f(x)$, where f maps the real line into itself if the mapping $f(x)$ is contractive. The Newton's algorithm

$$x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)}$$

for finding the real roots of the algebraic and transcendental equations $f(x)=0$ is also an iterative process.

Let us consider the mapping

$$y = x - \frac{f(x_n)}{f'(x_n)}$$

of the real line into itself, where $\frac{f(x)}{f'(x)}$ is continuous on a closed interval $[a,b]$ and differentiable on the open interval (a,b). If x_1 and x_2 are any two points of $[a,b]$ which map into y_1 and y_2, thus

$$y_1 - y_2 = x_1 - x_2 - \left\{\frac{f(x_1)}{f'(x_1)} - \frac{f(x_2)}{f'(x_2)}\right\}$$
$$= (x_1 - x_2)\frac{f(\zeta)f''(\zeta)}{\{f'(\zeta)\}^2}$$

where ζ lies between x_1 and x_2. Hence if

$$\left|\frac{f(x)f''(x)}{\{f'(x)\}^2}\right| \le \kappa < 1$$

on $[a,b]$, the mapping is a contraction mapping of $[a,b]$ onto a closed interval of the real line. Hence we can conclude from theorem (3.3.1) stated below that if there is a point $x_0 \in (a,b)$ such that $\left|\frac{f(x_0)}{f'(x_0)}\right| < \delta\ |1-\kappa|$, the mapping has a unique fixed point $\alpha \in (x_0-\delta\, , x_0+\delta) \cap [a,b]$ and that Newton's sequence $\{x_n\}$ converges to α.

Here we will present some important theorems relating to the fixed point iteration method:

3.3.1 Contraction mapping principle. A map $F:(X,\rho_1) \to (Y,\rho_2)$ of metric spaces that satisfies $\rho_2(F(x),F(z)) \le L'\rho_1(x,z)$ for some fixed constant L' and $x,z \in X$, is called Lipschitzian; the smallest such L' is called the Lipschitz constant $L(F)$ of F. If $L(F) < 1$, the map F is called contractive with contraction constant $L(F)$; if $L(F) = 1$, the map F is said to be non-expansive.

The equation (3.5) suggests a technique for finding a fixed point. If we know a value x_0 of x such that $F(x_0)$ does not differ greatly from x_0, it is natural to regard

$$x_1 = F(x_0) \tag{3.8}$$

as a possible improvement over x_0, and to generate the sequence $\{x_n\}$ of successive approximations to a fixed point x of F by the relationship

$$x_{n+1} = F(x_n), \qquad n = 0,1,2,\cdots\cdots \tag{3.9}$$

Note that if F is a linear mapping defined on a normed linear space, then F is a contraction if and only if $\| F \| < 1$. It is also worth pointing out that a contraction may have at most one fixed point. Indeed, if x and y are both fixed points of the contraction F,

$$\rho(x,y) = \rho(F(x), F(y)) \leq L\rho(x,y) < \rho(x,y),$$

which is a contraction.

The following result is the basic theorem on contractive mappings, called the Contractive Mapping Principle (also called the Banach - Caccioppoli theorem) that guarantees the existence of a unique fixed point of an operator F:

Theorem 3.3.1. Suppose C is a closed subset of a complete metric space (X,ρ) and $F: C \to C$ is a contraction with contraction constant L. Then F has a unique fixed point $x^* \in C$ and if x_0 is any point of C and $\{x_n\}$ is defined by (3.9), than $x_m \to x^*$ as $m \to \infty$ and the error bound

$$\rho(x_m, x^*) \le L^m (1-L)^{-1} \rho(x_0, x_1)$$

is valid. The Banach principle has a useful local version that involves an open ball B in a complete metric space Y and a contractive map of B into Y which does not displace the center of the ball too far:

Corollary. Let (Y, ρ) be complete and $B = B(y_0, r) = \{y \mid \rho(y, y_0) < r\}$. Let $F: B \to Y$ be contractive with constant $\alpha < 1$. If $\rho(F(y_0), y_0) < (1-\alpha)r$,then F has a fixed point.

We give the application of the contractive mapping principle: Consider the initial value problem

$$\frac{dy}{dx} = f(x, y), \qquad y(x_0) = y_0 \tag{3.10}$$

The following theorem guarantees the existence and uniqueness of the solution to the initial value problem (3.10).

Theorem 3.3.2. Suppose f is continuous on a closed rectangle

$$R = \{(x, y): a \le x \le b, c \le y \le d\}$$

containing the point (x_0, y_0) in its interior and satisfies a Lipschitz condition

$$| f(x, y_1) - f(x, y_2) | \le M \,| y_1 - y_2 | \text{ in } R.$$

Then for $h > 0$ sufficiently small there is a unique solution $g(x)$ to the initial value problem (3.10) defined on $[x_0 - h, x_0 + h]$.

Cluster point. Let E be a metric space with metric ρ. Let $\{x_n\}$ be a sequence in E. We say that "x is a cluster point" of the sequence $\{x_n\}$ if one of the following three equivalent conditions is satisfied

(i) x is the limit point of a subsequence $\{x_{n_k}\}_k$ of the sequence $\{x_n\}_n$.

(ii) for all $\varepsilon > 0$ and n, there exists $m \geq n$ such that $\rho(x, x_m) \leq \varepsilon$,

(ii) for all n, x belongs to the closure of the set say, $A_n = \{x_m\}, \quad m \geq n$.

Compact sets. We say that a subset K of E is "compact" if every infinite sequence $\{x_n\}$ of elements x_n of K has at least one cluster point belonging to K.

We now state, without proof, the Brouwer fixed point theorem, which is the origin of most of the theorems of nonlinear analysis.

Theorem 3.3.3. Let E be a compact convex subset of R^n. Every continuous mapping f from E to itself has a fixed point.

Theorem 3.3.4. (Schauder's fixed point theorem.) Every convex compact subspace of a Banach space is a fixed point space.

Chapter-4

Solution to System of Nonlinear Equations

4.1 Introduction. In the previous chapter, we have used a method for solving a system of linear equations in interval number system. Here, we study an existence and uniqueness theorem of solution to nonlinear system. Our theorem will be applied to find approximate numerical solution to an algebraic as well as to a nonlinear system.

Let $f : D \subseteq R^n \to R^n$ be continuously differentiable in the open domain D. With these assumptions Urabe's theorem guarantees the existence and uniqueness of solution of the nonlinear system

$$f(x) = 0. \qquad (4.1)$$

With further assumptions that both of f and f' have continuous, inclusion monotonic interval extension F and F' defined on interval vectors contained in D.

Let Y be a nonsingular real matrix and let $X = (X_1, X_2, X_3 \cdots\cdots X_n)$ be contained in D where $X_1, X_2, X_3, \cdots\cdots, X_n$ are closed bounded real intervals.

Suppose that we are given n nonlinear equations in n variables whose vector form is given by (4.1).

An interval version of Newton's method to solve (4.1) is given by

$$X^{(k+1)} = X^{(k)} \cap N(X^{(k)})$$

with the interval Newton operator

$$N(X) = m(X) - Vf(m(X)$$

where V is an interval matrix containing $[f'(x)]^{-1}$ for all $x \in X$.

Another form of interval version of Newton's method which does not require the inversion of an interval matrix is given by Krawczyk R. (1969).

Define

$$K(X) = y - Yf(y) + \{I - YF'(X)\}(X - y)$$

where y is a point chosen from X, and Y is an arbitrary nonsingular real matrix.

The mid point of $[a,b]$ is given by $m([a,b]) = \frac{a+b}{2}$.

For the interval vector $X = (X_1, X_2, X_3 \cdots\cdots X_n)$ we define

$$|X| = \max_i |X_i|$$

and

$$w(X) = \max_i w(X_i)\ ,\ m(X) = (m(X_1), m(X_2), m(X_3) \cdots\cdots m(X_n)).$$

For an interval matrix $A = (a_{ij})$ where a_{ij} are interval vector, we define

$$\| A \| = \max_i \sum_{j=1}^{n} | A_{ij} |$$

and $m(A)$ as the real matrix with component $m(A_{ij})$.

We now state and prove two essential lemmas.

Lemma 4.1.1. If $P(x) = x - Yf(x)$ maps X into itself, then $f(x) = 0$ has a solution in X.

Proof. The continuity of P follows from that of f .Since P maps the convex, compact set X into itself, P has a fixed point in X by the Brouwer's fixed point theorem. From the non-singularity of Y, a fixed point of P is a solution of $f(x) = 0$ and the lemma is proved. Q.E.D

Lemma 4.1.2. If A is an interval matrix and X is an interval vector, then $w(A(X - m(X))) \leq \| A \| \, w(X)$.

Proof. Let X and Z be intervals. From $w(Z(X - m(X))) = | Z | \, w(X)$ and $w(X + Z) = w(X) + w(Z)$ it follows that

$$\begin{aligned} w(A(X - m(X))) &= \max_i w\left(\sum_{j=1}^{n} A_{ij} (X - m(X)) \right) \\ &= \max_i \left(\sum_{j=1}^{n} w(A_{ij} (X - m(X))) \right) \\ &= \max_i \sum_{j=1}^{n} | A_{ij} | \, w(X_j) \end{aligned}$$

Since

$$w(X) = \max_j w(X_j) \geq w(X_j),$$

we have

$$w(A(X - m(X))) \leq \left(\max_i \sum_{j=1}^{n} | A_{ij} | \right) w(X) = \| A \| \, w(X). \text{ Q.E.D}$$

Theorem 4.1.1. [Moore R. E. (1977)] Suppose a region (interval vector) $X^{(0)}$, a point $y^{(0)}$ in $X^{(0)}$, and a real matrix $Y^{(0)}$ have been found such that

(i) $K(X^{(0)}) \subseteq X^{(0)}$. Then, there is a solution x in $X^{(0)}$ to the system $f(x) = 0$.

Consider the algorithm

$$X^{(k+1)} = X^{(k)} \cap K(X^{(k)}), \qquad k = 0,1,2,3,\cdots\cdots$$

with

$$K(X^{(k)}) = y^{(k)} - Y^{(k)} f(y^{(k)}) + \{I - Y^{(k)} F'(X^{(k)})\}(X^{(k)} - y^{(k)}),$$

where $y^{(k)}$ and $Y^{(k)}$, $k = 1,2,3,\cdots\cdots\cdots$, are chosen as follows:

$$y^{(k)} = m(X^{(k)}).$$

$$Y^{(k)} = \begin{cases} Y, & \text{an approximation to } [mF'(X^{(k)})]^{-1} \text{ if } \| I - Y F'(X^{(k)}) \| \leq r_{k-1}, \\ & Y^{(k-1)} \quad \text{otherwise} \end{cases}$$

with $r_k = \| I - Y^{(k)} F'(X^{(k)}) \|$, $\qquad k = 0,1,2,3,\cdots\cdots$

If the condition

(ii) $r_0 < 1$ is satisfied, then there is a unique solution x to $f(x) = 0$ in $X^{(0)}$ and the following hold:

(a) $x \in X^{(k)} \subseteq X^{(k-1)}$ for $k = 1, 2, 3, \cdots\cdots$

(b) $w(X^{(k)}) \leq r_0^k w(X^{(0)})$.

Thus, $\{X^{(k)}\}$ is a nested sequence of regions containing and converging at least linearly to the unique solution x in $X^{(0)}$.

Proof. From the definition of P in Lemma 4.1.1, we have

$$\begin{aligned} P(x) &= x - Yf(x) \\ &= y - Yf(y) + x - y - Y(f(x) - f(y)) \end{aligned} \qquad \text{for all } x \text{ in } X.$$

Now, using the mean- value theorem.

$$f(x) = f(y) + \sum_{j=1}^{n} \left. \frac{\partial f}{\partial y_j} \right|_{y=\bar{y}} (x_j - y_j),$$

we have

$$f(x) - f(y) \in F'(X)(x - y) \quad \text{for all } x, y \text{ in } X.$$

Hence we have

$$x - y - Y(f(x) - f(y)) \in x - y - YF'(X)(x - y) \in (I - YF'(X))(X - y).$$

Thus we have

$$P(x) \in y - Yf(y) + \{I - YF'(X)\}(X - y),$$

that is $P(x) \in K(X)$ for all x in X.

Since $K(X^{(0)}) \subseteq X^{(0)}$, P maps $X^{(0)}$ into itself; and by Lemma 4.1.1, $f(x) = 0$ has a solution in $X^{(0)}$.

If $f(x) = 0$ for x in X, then $x = P(x)$ is also in $K(X)$.

Since $X^{(k)} = X^{(k-1)} \cap K(X^{(k-1)})$, the solution x in $X^{(0)}$ is also in $X^{(k)}$ for all $k = 0,1,2,3,\cdots\cdots$ which proves (a).

It remains to show that the inequality (b) holds. From the definition of the algorithm, we have $\quad w\left(X^{(k+1)}\right) \le w(K(X^{(k)}))$.

But we have

$$w\left(K(X^{(k)})\right) = w\left(y^{(k)} - Y^{(k)} f(y^{(k)}) + \{I - Y^{(k)} F'(X^{(k)})\}(X^{(k)} - y^{(k)})\right),$$

where

$$y^{(k)} = m(X^{(k)}) \quad \text{and} \quad Y^{(k)} = [m\left(F'(X^{(k)})\right)]^{-1}.$$

Since

$$w\left(y^{(K)}\right) = 0, \quad w\left(Y^{(k)} f(y^{(k)})\right) = 0,$$

we have

$$w\left(K(X^{(k)})\right) = w\{I - Y^{(k)} F'(X^{(k)})\}(X^{(k)} - y^{(k)}) \le r_k w\left(X^{(k)}\right)$$ from Lemma

4.1.2.

Hence we have

$$w\left(K(X^{(k)})\right) \le r_k \; w(X^{(k)}).$$

Thus we obtain

$$\begin{aligned} w\left(X^{(k)}\right) &\le w(K(X^{(k-1)})) \le r_{k-1} \, w\left(X^{(k-1)}\right) \le r_{k-1} \, w\left(K(X^{(k-2)})\right) \\ &\le r_{k-1} \, r_{k-2} \, w(X^{(k-2)}) \\ &\le r_{k-1} \, r_{k-2} \, r_{k-3} \, w(X^{(k-3)}) \\ &\le r_{k-1} \, r_{k-2} \, r_{k-3} \cdots\cdots r_0 \, w(X^{(0)}) \end{aligned}$$

Since $\{r_k\}$ is non-increasing by constructive, we have $w\left(X^{(k)}\right) \le r_0^k \; w(X^{(0)})$ which proves (b).

Remark. Note that the choice $Y^{(k)} = Y^{(k-1)}$ leads to

$r_k = \| I - Y^{(k-1)} F'(X^{(k)}) \| \le r_{k-1} = \| I - Y^{(k-1)} F'(X^{(k-1)}) \|$ because of the inclusion monotonicity of F'. That is, since

$$X^{(k)} = X^{(k-1)} \cap K(X^{(k-1)}) \subseteq X^{(k-1)},$$

it follows that $F'(X^{(k)}) \subseteq F'(X^{(k-1)})$ and so that sequence $\{r_k\}$ is non-increasing.

Theorem 4.1.2. [Urabe (1965)] Let $F(x)$ be a continuously differentiable function on the domain $D \subset R^n$. Let $x^{(0)} \in D$ and suppose $J(x^{(0)})$ be regular. Also suppose that the following three conditions are satisfied for a positive number δ and a non-negative $\kappa \, (0 \le \kappa < 1)$.

(i) $\Omega_\delta = \left\{ x \in R^n : \| x - x^{(0)} \| \le \delta \right\} \subset D$,

(ii) $\| J(x) - J(x^{(0)}) \| \le \dfrac{\kappa}{M} \qquad \left(x \in \Omega_\delta \right)$,

(iii) $\dfrac{Mr}{1-\kappa} \le \delta$,

where

$$\| F(x^{(0)}) \| \le r, \qquad \| J^{-1}(x^{(0)}) \| \le M .$$

Suppose $x^* \in \Omega_\delta$ be the unique solution of the equation $F(x)=0$ and $J(x^*)$ is regular.

Then the error estimation

$$\| x^{(0)} - x^* \| \le \frac{Mr}{1-\kappa} \quad \text{is satisfied.}$$

4.2 Numerical Example for algebraic equation. Let us consider the third degree polynomial equation $f(x)=x^3-3x+3=0$. This equation has only one real root $x^* = \alpha = -2.103803402.......$ Let $x^{(0)} = -2.11$ be the approximate solution.

We use the theorem 4.1.2 and also assume that $z = x^{(0)}$. Then we get

$$r = \| F(z) \| = 0.06, \ \| [J(z)]^{-1} \| = \| \frac{1}{F'(z)} \| = \frac{1}{10.35} = 0.096618...... \le 0.0967 = M .$$

(i) In the closed domain $\Omega_\delta = \{x \in R; \| x - x^{(0)} \| \le \delta\}$,

$$x + z = (x-z) + 2z \quad \text{leads to } |x+z| \le 2|z| + \delta$$

and

$$\| J(x) - J(z) \| = 3|x+z||x-z| \le (6|z| + 3\delta)\delta$$

so,

(ii) if $(6|z|+3\delta)\delta \le \dfrac{\kappa}{M}$ $(x \in \Omega_\delta)$ is satisfied.

So, $\kappa \ge (6|z|+3\delta)\delta M = (12.66+3\delta)\delta \times 0.0967 = 1.224222\delta + 0.2901\delta^2$.

(iii) But $\dfrac{M\delta}{1-\kappa} \le \delta$ is satisfied if $\delta > \delta(1-\kappa) \ge Mr = 0.0967 \times 0.06 = 0.005802$.

As the closed domain $\Omega_\delta = \{x \in R : \| x - x^{(0)} \| \leq \delta\}$ is the δ-neighborhood of $z = x^{(0)} = -2.11$. If we choose $\delta = 0.10$, then

$$\Omega_\delta = \{x \in R : | x - (-2.11) | \leq 0.10\} = [-2.21, -2.01] \subset D = [-3, -2].$$

Therefore $\kappa \geq 1.224222(0.10) + 0.2901(0.10)^2 = 0.1253232.$

So, if we choose $\kappa = 0.13$, then all the conditions of Urabe's theorem are satisfied. In the closed domain Ω_δ, then $x^* = \alpha = -2.103803402...$ is the unique solution of $f(x) = 0$.

The error estimation gives

$$\| z - x^* \| \leq \frac{Mr}{1-\kappa} = \frac{0.0967 \times 0.06}{1 - 0.13} = 0.0066...... < 0.007 \text{ (say)}$$

$$\text{i.e.} \quad -2.11 - 0.007 < x^* < -2.11 + 0.007.$$

Let $X^{(0)} = [-2.117, -2.103]$.

We shall try to apply Moore's theorem in this interval.

$$y^{(0)} = m(X^{(0)}) = -2.110$$

Horner's Method:

(1) $b^{(-1)}[k] = a_k$, $k = 0,1,2,\cdots\cdots,n$ $\quad (n = 3)$

(2) for $j = 0,1,2,\cdots\cdots$, $b^{(j)}[0] = a_0, b^{(j)}[k] = b^{(j-1)}[k] + b^{(j)}[k-1]y$,

$(k = 1,2,\cdots\cdots,n-j)$

(3) $b^{(0)}[n] = f(y)$, $b^{(1)}[n-1] = f'(y)$.

From the **Table-1** we get

$$f(y^{(0)}) = -0.064 \ , \ f'(y^{(0)}) = 10.356$$

and also

$$Y^{(0)} = [f'(y^{(0)})]^{-1} = 0.09656 .$$

[Note: For the first iteration $\delta = 0.007$, $\kappa = 0.10327857 < 0.10328$.]

From the **Table-2** we get

$$F'(X^{(0)}) = [10.267, 10.446]$$

As

$$K\left(X^{(0)}\right) = y^{(0)} - Y^{(0)} f\left(y^{(0)}\right) + \left\{I - Y^{(0)} F'\left(X^{(0)}\right)\right\}\left(X^{(0)} - y^{(0)}\right)$$
$$= [-2.104 , -2.103] \subseteq X^{(0)} = [-2.117 , -2.103],$$

On the other hand

$$r_0 = \| I - Y^{(0)} F'(X^{(0)}) \| = \| [-0.0087, 0.0087] \| = 0.0087 < 1 .$$

Next:

We have

$$X^{(1)} = K(X^{(0)}) \cap X^{(0)} = K(X^{(0)}) = [-2.104, -2.103]$$

Then

$$y^{(1)} = m(X^{(1)}) = -2.1035$$

and

$$F'(X^{(1)}) = [10.2678, 10.2806]$$

From the **Table-3** we get

$$f(y^{(1)}) = 0.0031, \ f'(y^{(1)}) = 10.2741$$

and also

$$Y^{(1)} = [f'(y^{(1)})]^{-1} = 0.09733\,.$$

[Note: For the second iteration $\delta = 0.0031$, $\kappa = 0.04322580 < 0.04333$.]

From the **Table-4** we get

$$F'(X^{(1)}) = [10.2678, 10.2806]$$

As

$$K\left(X^{(1)}\right) = y^{(1)} - Y^{(1)} f\left(y^{(1)}\right) + \left\{I - Y^{(1)} F'\left(X^{(1)}\right)\right\}\left(X^{(1)} - y^{(1)}\right)$$
$$= [-2.104\,, -2.103] \subseteq X^{(1)} = [-2.104, -2.103]$$

than

$$X^{(2)} = K(X^{(1)}) \cap X^{(1)} = X^{(1)},$$

and

$$X^{(1)} = X^{(2)} = X^{(3)} = \cdots\cdots\cdots \text{ is satisfied.}$$

On the other hand

$$r_1 = \| I - Y^{(1)} F'(X^{(1)}) \| = \| [-0.00062, 0.0007] \| = 0.0007 < r_0 < 1\,.$$

Remark. The satisfied interval has the same upper and lower limit -2.103811. But this is not a valid number because the actual value is $2.103803402\cdots\cdots$ and the interval $X^1, X^2, \cdots$ converges to the interval $X = [2.104, 2.103]$ that contains the unique real root $-2.103803402\cdots$.

Table-1

Calculate $f(y^{(0)})$ & $f'(y^{(0)})$ by Horner's method :

	$k=0$	$k=1$	$k=2$	$k=3$
$b^{(-1)}[k]$ $b^{(0)}[k-1]y$	1	0 -2.110	-3 4.452	3 -3.064
$b^{(0)}[k]$ $b^{(1)}[k-1]y$	1	-2.110 -2.110	1.452 8.904	-0.064
$b^{(1)}[k]$	1	-4.220	10.356	

Table-2

Calculate $F'(X^{(0)})$ by Horner's method :

	$k=0$	$k=1$	$k=2$	$k=3$
$b^{(-1)}[k]$ $b^{(0)}[k-1]y$	[1,1]	0 [-2.117,-2.103]	[-3.-3] [4.422,4.482]	[3,3] [-3.138,-2.990]
$b^{(0)}[k]$ $b^{(1)}[k-1]y$	[1,1]	[-2.117,-2.103] [-2.117,-2.103]	[1.422,1.482] [8.845,8.964]	[-0.138,0.010]
$b^{(1)}[k]$	[1,1]	[-4.234,-4.206]	[10.267,10.446]	

Table-3

Calculate $f(y^{(1)})$ & $f'(y^{(1)})$ by Horer's method:

	$k=0$	$k=1$	$k=2$	$k=3$
$b^{(-1)}[k]$	1	0	-3	3
$b^{(0)}[k-1]y$		-2.1035	4.4247	-2.9968
$b^{(0)}[k]$	1	-2.1035	1.4247	0.0032
$b^{(1)}[k-1]y$		-2.1035	8.8494	
$b^{(1)}[k]$	1	-4.2070	10.2741	

Table-4

Calculate $F'(X^{(1)})$ by Horer's method:

	$k=0$	$k=1$	$k=2$	$k=3$
$b^{(-1)}[k]$	[1,1]	0	[-3.-3]	[3,3]
$b^{(0)}[k-1]y$		[-2.104,-2.103]	[4.4226,4.4269]	[-3.0022,-2.9917]
$b^{(0)}[k]$	[1,1]	[-2.104,-2.103]	[1.4226,1.4269]	[-0.0022,0.0083]
$b^{(1)}[k-1]y$		[-2.104,-2.103]	[8.8452,8.8537]	
$b^{(1)}[k]$	[1,1]	[-4.208,-4.206]	[10.2678,10.2806]	

4.3 Numerical Computations for a system of nonlinear equations:

We consider the system of nonlinear equations

$$f_1(x_1,x_2) = x_1^2 + x_2^2 - 1 = 0$$
$$f_2(x_1,x_2) = x_1^2 - x_2 = 0$$

here

$$x^* = (0.7861513377\ ,\ 0.618033988)$$

be the exact solution .

Let $z = y^0 = (.80,.62)$ be the approximate solution.

The Jacobian matrix for the system is given by

$$f'(x) = \begin{pmatrix} 2x_1 & 2x_2 \\ 2x_1 & -1 \end{pmatrix} = J(x)\,.$$

Here $$J(z) = \begin{pmatrix} 2\times 0.80 & 2\times 0.62 \\ 2\times 0.80 & -1 \end{pmatrix} = \begin{pmatrix} 1.6 & 1.24 \\ 1.6 & -1 \end{pmatrix}.$$

<u>Application of Urabe's theorem:</u>

$$r = \| f(z) \| = 0.0244$$

and

$$\| [J(z)]^{-1} \| = \| \begin{pmatrix} 1.6 & 1.24 \\ 1.6 & -1 \end{pmatrix}^{-1} \| = 0.892857142 \le 0.90 = M \text{ (say).}$$

In the closed domain $\Omega_\delta = \{x \in R^2 : \| x - z \| \le \delta\}$

Let $\| x-z \|_a \le \delta \;\Rightarrow\; \| \begin{pmatrix} x_1 - z_1 \\ x_2 - z_2 \end{pmatrix} \|_a \le \delta \;\Rightarrow\; \| \begin{pmatrix} x_1 - 0.80 \\ x_2 - 0.62 \end{pmatrix} \|_a \le \delta$,

i.e., $| x_1 - 0.80 | + | x_2 - 0.62 | \le \delta$.

Again
$$\| J(x) - J(z) \|_a = \| \begin{pmatrix} 2x_1 & 2x_2 \\ 2x_1 & -1 \end{pmatrix} - \begin{pmatrix} 2\times 0.80 & 2\times 0.62 \\ 2\times 0.80 & -1 \end{pmatrix} \|_a$$

$$= \| \begin{pmatrix} 2(x_1 - 0.80) & 2(x_2 - 0.62) \\ 2(x_1 - 0.80) & 0 \end{pmatrix} \|_a ,$$

where $\| . \|_a =$ sum of all entries.

Therefore

$$\| J(x) - J(z) \|_a = 4 | x_1 - 0.80 | + 2 | x_2 - 0.62 | \le 4 | x_1 - 0.80 | + 4 | x_2 - 0.62 |$$
$$= 4 \{ | x_1 - 0.80 | + | x_2 - 0.62 | \} \le 4\delta$$

But $\frac{Mr}{1-\kappa} \le \delta$ is satisfied if $\delta > \delta(1-\kappa) \ge Mr = 0.90 \times 0.0244 = 0.02196$ and

$\kappa \ge 4 \times 0.90\delta = 3.6\delta$.

If we choose $\delta = 0.10$, then

$$\Omega_\delta = \{ | x - z | \le \delta \} = \{ | x - z | \le 0.10 \} = [\, 0.70 , 0.90 \,] \times [\, 0.52 ,\; 0.72 \,] = D$$

and all the conditions of Urabe's theorem are satisfied.

In the closed domain Ω_δ, $x^* = (0.7861513377, 0.618033988)$ is the unique solution of $f(x) = 0$.

The error estimation gives

$$\| z - x^* \| \le \frac{Mr}{1-\kappa} = \frac{0.90 \times 0.0244}{1 - 0.36} = 0.0343125 \le 0.035 \text{ (say)}$$

$$i,e\ \ 0.80-0.035 < x_1 < 0.80+0.035 \ \Rightarrow\ x_1 = [0.765\,,\,0.835]$$

and

$$0.62-0.035 < x_2 < 0.62+0.035 \ \Rightarrow\ x_2 = [0.585\,,\,0.655]$$

We now to check whether there is a solution in the region

$$X^{(0)} = \begin{pmatrix} [0.765\,,\,0.835] \\ [0.585\,,\,0.655] \end{pmatrix}$$

At the same time we may obtain improved error bounds.

We compute that

$$f(y^{(0)}) = \begin{pmatrix} 0.0244 \\ 0.02 \end{pmatrix}$$

and

$$f'(y^{(0)}) = \begin{pmatrix} 1.6 & 1.24 \\ 1.6 & -1 \end{pmatrix}$$

And also

$$F'(X^{(0)}) = \begin{pmatrix} [1.53\,,1.67] & [1.17\,,1.31] \\ [1.53\,,1.67] & [-1,-1] \end{pmatrix}$$

For Y, we take the following approximation to $[f'(y)]^{-1}$:

$$Y^{(0)} = [f'(y^{(0)})]^{-1} = \begin{pmatrix} 0.279017857 & 0.345982142 \\ 0.446428571 & -0.446428571 \end{pmatrix}$$

Using rounded interval arithmetic, we compute that for Krawczyk transformation

$$K(X^{(0)}) = y^{(0)} - Y^{(0)}\left(f(y^{(0)}) + \{I - Y^{(0)}F'(X^{(0)})\}(X^{(0)} - y^{(0)}\right),$$

$$K(X^{(0)}) = \begin{pmatrix} [0.784\,,0.788] \\ [0.615\,,0.621] \end{pmatrix} \subseteq X^{(0)}$$

We have for r_0,

$$r_0 = \| I - Y^{(0)}F'(X^{(0)}) \| = 0.9462$$

Now for $X^{(0)}$, we have

$$X^{(1)} = K(X^{(0)}) \cap X^{(0)} = \begin{pmatrix} [0.784\,,0.788] \\ [0.615\,,0.621] \end{pmatrix}.$$

Again we can find

$$y^{(1)} = m(X^{(1)}) = \begin{pmatrix} 0.786 \\ 0.618 \end{pmatrix}.$$

Now we can compute that

$$f(y^{(1)}) = \begin{pmatrix} -0.00028 \\ -0.000204 \end{pmatrix} \text{ and } f'(y^{(1)}) = \begin{pmatrix} 1.572 & 1.236 \\ 1.572 & [-1,-1] \end{pmatrix}$$

and

$$F'(x^{(1)}) = \begin{pmatrix} [1.568\,,1.576] & [1.23\,,1.242] \\ [1.568\,,1.576] & [-1,-1] \end{pmatrix}.$$

Also we have

$$Y^{(1)} = [f'(y^{(1)}]^{-1} = \begin{pmatrix} 0.284 & 0.352 \\ 0.447 & -0.447 \end{pmatrix}$$

We using rounded interval arithmetic for $K(X^{(1)})$, we compute that for Krawczyk transformation

$$K(X^{(1)}) = y^{(1)} - Y^{(1)}\left(f(y^{(1)}) + \{I - Y^{(1)}F'(X^{(1)})\}(X^{(1)} - y^{(1)}\right),$$

$$K(X^{(1)}) = \begin{pmatrix} [0.786\,,0.786] \\ [0.618\,,0.618] \end{pmatrix} \subseteq X^{(1)}.$$

then

$$X^{(2)} = K(X^{(1)}) \cap X^{(1)} = \begin{pmatrix} [0.786\,,0.786\,] \\ [\,0.618\,,0.618\,] \end{pmatrix}.$$

Therefore

$$X^{(0)} \supset X^{(1)} \supset X^{(2)} \text{ is satisfied.}$$

For r_1,

$$r_1 = \| I - Y^{(1)}F'(X^{(1)}) \| = 0.006766.$$

In this example, we find $r_0 = 0.09462 < 1$, $r_1 = 0.006766 < 1$.

As
$$w(X^{(1)}) = w\begin{pmatrix} [0.784\,,0.788] \\ [0.615\,,0.621] \end{pmatrix} = 0.006$$

and

$$w(X^{(0)}) = w\begin{pmatrix} [0.765\,,0.835] \\ [0.585\,,0.655] \end{pmatrix} = 0.07,$$

it is seen that

$$w(X^{(k)}) \le r_0^k\, w(X^{(0)}).$$

Hence we can conclude that both the conditions of theorem 4.1.1 are satisfied. Here we also conclude that the degenerate intervals $[\,0.786\,,0.786\,]$ and $[\,0.618\,,0.618\,]$ indicate that the approximate values of x_1 and x_2 are correct to three decimal places.

Chapter-5

Solution to Nonlinear Differential Operator Equation

5.1 Introduction. In the previous chapter, we have applied Urabe's theorem (1965) to approximate the exact solution of algebraic as well as of nonlinear system of equations in interval number system. Basically, Urabe gave his theorem while studding the quasiperiodic differential operator equations for approximating numerical solutions using Galerkin's method. Galerkin's method produces system of equations that needs Newton's method to solve. We here study the error bound for approximating the exact quasiperiodic solution to Van der Pol type equations.

Here, It is to be noted that Mitsui T. (1977) gave an error estimation for the quasiperiodic solution of van der Pol type equation while studying Galerkin's procedure. In this chapter, we estimate a new error bound that we claim to be more general compared to that given by Mitsui.

Firstly, we have an attempt to discuss the existence and uniqueness of quasiperiodic solution to Van der Pol type equations.

Let us consider the following Van der Pol type equation

$$\frac{d^2x}{dt^2} - 2\lambda(1-x^2)\frac{dx}{dt} + x = \sum_{k=1}^{m}(a_k \cos \vartheta_k t + b_k \sin \vartheta_k t) \qquad (5.1)$$

with quasi-periodic forcing term.

Now we need to introduce some basic notations and terminologies.

5.2 Almost Periodic and Quasi-periodic Functions:

Let d denotes some positive integer and R the set of real numbers.

Definition 5.2.1. A function $f(t) \in C(R;R^d)$ is said to be quasi-periodic with periods $\omega_1, \omega_2, \cdots\cdots, \omega_m$ if there exists some continuous periodic function $f_0(u_1, u_2, \cdots\cdots, u_m) \in C(R^m, R^d)$ such that

$$f(t) \in f_0(t, t, \cdots, t) \qquad \text{for all } t \in R, \tag{5.2}$$

and $f_0(u_1, u_2, \cdots, u_m)$ is periodic with periods ω_i in which $u_i \quad (i = 1, 2, \cdots, m)$.

We assume here that $\omega_1, \omega_2, \cdots\cdots, \omega_m$ are all positive and also the reciprocals of $\omega_1, \omega_2, \cdots\cdots, \omega_m$ are rationally linearly independent. In fact, assume ω_m^{-1} is rationally linearly dependent on $\omega_1^{-1}, \omega_2^{-1}, \cdots\cdots, \omega_m^{-1}$, that is

$$\omega_m^{-1} = \sum_{k=1}^{m-1} a_k \omega_k^{-1},$$

where $a_k = \dfrac{b_k}{c_k}$ for b_k ,$c_k > 0$ and $k = 1, 2, \cdots, m-1$ are all integers. Here if we take ω_k in place of c_k $(k = 1, 2, \cdots, m-1)$ then we assume

$$\frac{1}{\omega_m} = \sum_{k=1}^{m-1} b_k \frac{1}{\omega_k}. \tag{5.3}$$

Now consider the following function

$$f_1(u_1, u_2, \cdots, u_{m-1}) = f_0\left(u_1, u_2, \cdots, u_{m-1}, \omega_m \sum_{k-1}^{m-1} b_k \frac{u_k}{\omega_k}\right) \tag{5.4}$$

which is periodic with periods ω_i in each u_i $(i = 1, 2, \cdots, m-1)$ and

$$f_1(\underbrace{t, t, \cdots, t}_{m-1}) = f_0(\underbrace{t, t, \cdots, t}_{m}),$$

because we have

$$\omega_m \sum_{k=1}^{m-1} b_k \frac{u_k + \omega_k}{\omega_k} = \omega_m \sum_{k=1}^{m-1} b_k \frac{u_k}{\omega_k} + \omega_m \sum_{k=1}^{m-1} b_k$$

$$t = \omega_m \sum_{k=1}^{m-1} b_k \frac{t}{\omega_k} \qquad \text{from (5.3).}$$

So we can reduce the case of periods $\omega_1, \omega_2, \cdots\cdots, \omega_m$.

Definition 5.2.2. A function $f(t)$ is said to be almost periodic if, from any sequence $\{\alpha_n'\} \subset R$, we can extract subsequence $\{\alpha_n\}$ such that $\{f(t + \alpha_n)\}$ is uniformly convergent.

Note that an almost periodic function f is bounded on R.

We show here a basic fact that a quasi-periodic function f is almost periodic.

Lemma 5.2.1. If a function $f(t)$ is quasi-periodic with periods $\omega_1, \omega_2, \cdots\cdots, \omega_m$ then $f(t)$ is almost periodic.

Proof. If f is quasi-periodic with periods $\omega_1, \omega_2, \cdots\cdots, \omega_m$, then there exists a continuous function $f_0(u_1, u_2, \cdots\cdots, u_m)$ which is periodic with periods ω_i in each u_i $(i=1,2,\cdots,m)$ such that $f(t)=f_0(t,t,\cdots,t)$. Then from any sequence $\{\alpha_n'\}\subset R$, we can extract a subsequence $\{\alpha_n^{(1)}\}$ such that $\{\alpha_n^{(1)}(\bmod\,\omega_1)\}$ converges to $a_0^{(1)}$. In the same way, we can extract subsequence $\{\alpha_n^{(2)}\}\subset\{\alpha_n^{(1)}\}$ such that $\{\alpha_n^{(2)}(\bmod\,\omega_2)\}$ converges to $a_0^{(2)}$. Continuing this process, we obtain subsequence $\{\alpha_n^{(k)}\}\subset\{\alpha_n^{(k-1)}\}$ such that $\{\alpha_n^{(k)}(\bmod\,\omega_k)\}$ converges to $a_0^{(k)}$ for $k=1,2,\cdots,m$. So if we select subsequence

$$\{a_n\}=\{a_n^{(m)}\}\subset\{a_n'\},$$

then $\{\alpha_n\}$ satisfies

$$\lim_{n\to\infty} f(t+a_n)=\lim_{n\to\infty} f_0(t+a_n,\cdots,t+a_n)$$

$$=\lim_{n\to\infty} f_0(t+a_n^{(1)},\cdots,t+a_n^{(m)})$$

$$=f_0(t+a_0^{(1)},\cdots,t+a_0^{(m)})$$

uniformly.

This shows that f is almost periodic.

Definition 5.2.3. A subset $S\subset R$ is called relatively dense if there exists a positive number L such that

$$[a,a+L]\cap S\neq\phi \text{ for all } a\in R. \tag{5.5}$$

The number L is called the inclusion length.

Definition 5.2.4. For any bounded function f and $\varepsilon > 0$, we define

$$T(f,\varepsilon) \equiv \{\tau; | f(t+\tau) - f(t) | < \varepsilon\} \text{ for all } t \in R. \tag{5.6}$$

$T(f,\varepsilon)$ is called the ε – translation set of f.

Lemma 5.2.2. For any almost periodic function $f(t)$ and any real number σ, there exists

$$a(f,\sigma) = \lim_{T\to\infty} \frac{1}{T} \int_0^T f(t) e^{-i\sigma t} dt. \tag{5.7}$$

Proof. Note first that it is enough to show the existence of $a(f,0)$. In fact, because $e^{-i\sigma t} f(t)$ is almost periodic and $a(f,\sigma) = a(e^{-i\sigma t} f,0)$, we obtain the existence of $a(f,\sigma)$ for any σ from the existence of $a(f,0)$.

In order to prove the existence of $a(f,0)$, let M be a bound for $|f|$ and $T(f,\varepsilon)$ be the ε – translation set of f. If $\varepsilon > 0$ is given, let ℓ be the inclusion length for $T(f,\frac{\varepsilon}{8})$ and A a real number so that $A > \frac{16M\ell}{\varepsilon}$. Then we have

$$\frac{1}{nA}\int_0^{nA} f(t)dt = \sum_{k=0}^{n-1} \frac{1}{nA}\int_{kA}^{(k+1)A} f(t)dt.$$

Now it is easy to see

$$\int_{kA}^{(k+1)A} f(t)\,dt = \int_{kA-\tau}^{(k+1)A-\tau} f(t+\tau)\,dt$$

$$= \int_0^A f(t)\,dt + \int_0^A (f(t+\tau) - f(t))\,dt + \int_{kA-\tau}^{0} f(t+\tau)\,dt + \int_A^{(k+1)A-\tau} f(t+\tau)\,dt$$

If

$$\tau \in T\left(f, \frac{\varepsilon}{8}\right) \cap (kA, kA+\ell)$$

then we have the estimate

$$\left| \int_{kA}^{(k+1)A} f(t)\,dt - \int_0^A f(t)\,dt \right| \le \frac{A\varepsilon}{8} + 2M\ell$$

and consequently,

$$\left| \frac{1}{nA} \int_0^{nA} f(t)\,dt - \frac{1}{A} \int_0^A f(t)\,dt \right| \le \frac{\varepsilon}{8} + \frac{2M\ell}{A} < \frac{\varepsilon}{4}. \tag{5.8}$$

Now if T is a real number, there is a unique n so that $nA \le T < (n+1)A$ and thus from the following

$$\lim_{T\to\infty} \left| \frac{1}{T} \int_0^T f(t)\,dt - \frac{1}{nA} \int_0^{nA} f(t)\,dt \right| = 0$$

we have

$$\left| \frac{1}{T} \int_0^T f(t)\,dt - \frac{1}{nA} \int_0^{nA} f(t)\,dt \right| < \frac{\varepsilon}{4} \tag{5.9}$$

if $T \ge T_0$ for some T_0 where T_0 depends on M.

Now if $T_1, T_2 \ge T_0$ and n_1, n_2 are chosen so that $n_i A \le T_i < (n_i+1)A$ for $i=1,2$ then from (5.8) and (5.9)

$$\left| \frac{1}{T_1} \int_0^{T_1} f(t)dt - \frac{1}{T_2} \int_0^{T_2} f(t)dt \right| < \left| \frac{1}{T_1} \int_0^{T_1} f(t)dt - \frac{1}{n_1 A} \int_0^{n_1 A} f(t)dt \right|$$

$$+ \left| \frac{1}{n_1 A} \int_0^{n_1 A} f(t)dt - \frac{1}{A} \int_0^{A} f(t)dt \right|$$

$$+\left|\frac{1}{A}\int_0^A f(t)dt-\frac{1}{n_2A}\int_0^{n_2A} f(t)dt\right|$$

$$+\left|\frac{1}{n_2A}\int_0^{n_2A} f(t)dt-\frac{1}{T_2}\int_0^{T_2} f(t)dt\right|<\varepsilon$$

so the limit value

$$a(f,\sigma)=\lim_{T\to\infty}\frac{1}{T}\int_0^T f(t)e^{-i\sigma t}\,dt$$

exists and the lemma is thus proved.

Further there is a countable set $\sum$ of real numbers, which is called the set of exponents of f, such that $a(f,\sigma)=0$ if $\sigma\notin\sum$.

This is shown from Bessel's types inequality

$$\sum_{n=1}^{N}|a(f,\sigma_n)|^2\le a(|f|^2,0)$$

for any finite set of distinct real numbers $\sigma_1,\sigma_2,\sigma_3,\cdots,\sigma_N$.

Definition 5.2.4. For an almost periodic function f, the module of f, $\mathrm{Mod}(f)$, is defined to be the smallest additive group of real numbers that contains the set $\sum$ for which $a(f,\sigma)\neq 0$ if $\sigma\in\sum$.

5.3 Linear Differential Operators. Let x be a d-dimensional vector and $A(t)$ a $d\times d$ square matrix. Then define the differential operator L as follows

$$Lx=\frac{dx}{dt}-A(t)x. \tag{5.10}$$

Let $\Phi(t)$ be the fundamental matrix of the linear homogeneous equation

$$Lx = 0 \tag{5.11}$$

satisfying the initial condition $\Phi(0) = E$, where E is the $d \times d$ unit matrix. Here we call $\Phi(t)$ the fundamental matrix of the linear homogeneous equation (5.11) and $\Phi(t)$ satisfies the following matrix differential equation

$$\frac{d\Phi(t)}{dt} = A(t)\Phi(t). \tag{5.12}$$

Definition 5.3.1. We say that the differential equation (5.11) has bounded growth on R if, for some fixed $h > 0$, there exists a constant $C \geq 1$ such that every solution $x = x(t)$ of (5.11) satisfies

$$|x(t)| \leq C|x(s)| \text{ for } s, t \in R \text{ and } s \leq t \leq s + h. \tag{5.13}$$

It is easy to see that (5.11) has bounded growth if and only if there exist real constants K, σ such that its fundamental matrix $\Phi(t)$ satisfies

$$|\Phi(t)\Phi^{-1}(s)| \leq Ke^{\sigma(t-s)} \text{ for } t \geq s.$$

Moreover, (5.11) has bounded growth if its coefficient matrix $A(t)$ is bounded. In fact, from Gronwall's inequality, we have

$$|\Phi(t)\Phi^{-1}(s)| \leq \exp\left|\int_s^t |A(u)|\,du\right|.$$

If (5.13) holds only for $s, t \in R^+$, then we say that (5.11) has bounded growth on R^+ where R^+ denotes the interval $[0, \infty)$.

Definition 5.3.2. A linear differential operator (5.10) is said to be quasi-periodic if $A(t)$ is a quasi-periodic matrix.

Analogously we use the following definition.

Definition 5.3.3. A linear differential operator (5.10) is said to be almost periodic if $A(t)$ is almost periodic.

We have the following definitions about an almost periodic (or a quasi-periodic) operator.

Definition 5.3.4. An almost periodic operator L is said to be regular if for any almost periodic function $f(t)$, the equation

$$Lx = f(t) \tag{IH}$$

has at least one solution bounded for all $t \in R$.

Definition 5.3.5. A quasi-periodic operator is said to be regular if it is regular as an almost periodic operator.

5.4 Conditions for bounded solutions. Here we assume that all functions considered are continuous on R. Let us consider a linear homogeneous differential equation

$$Lx = 0 \tag{5.14}$$

with the operator L given by

$$Lx = \frac{dx}{dt} - A(t)x. \tag{5.15}$$

First we need some basic definitions and lemmas. We say that (5.14) satisfies the exponential dichotomy on R if there exist a projection P and positive constants C, σ such that

$$|\Phi(t) P \Phi^{-1}(s)| \leq C e^{-\sigma(t-s)} \text{ for } s \leq t, \tag{5.16}$$

$$|\Phi(t)(E-P)\Phi^{-1}(s)| \leq C e^{-\sigma(s-t)} \text{ for } t \leq s, \tag{5.17}$$

where $\Phi(t)$ is the fundamental matrix of the linear homogeneous equation (5.14) with $\Phi(0) = E$. If (5.16) holds only for $0 \leq s \leq t$ and (5.17) only for $0 \leq s \leq t$ then we say that (5.14) satisfies the exponential dichotomy on R^+. Then we have the following.

Lemma 5.4.1. Suppose (5.14) has bounded growth R^+. Then the inhomogeneous equation $Lx = f(t)$ has at least on bounded solution for every f bounded on R^+ if and only if (5.14) satisfies exponential dichotomy on R^+.

Note that the similar state for $R^- = (-\infty, o]$ also holds.

Lemma 5.4.2. The inhomogeneous differential equation $Lx = f(t)$ has at least one solution bounded on R for every bounded $f(t)$ if and only if the following three conditions are satisfied.

(i) The equation $Lx = f(t)$ has at least one solution bounded on R^+ for every $f(t)$ bounded on R^+.

(ii) The equation $Lx = f(t)$ has at least one solution bounded on R^- for every $f(t)$ bounded on R^-.

(iii) Every solution of the homogeneous equation $Lx = 0$ is the sum of a solution which is bounded on R^+ and a solution which is bounded on R^-.

5.5 Regular Differential operators. By the results in the previous subsections, we can show some useful results in this subsection. First let L be almost periodic then we have the following propositions.

Proposition 5.5.1. A linear differential operator L is regular if and only if there is a $d \times d$ square matrix P such that

(i) $P^2 = P$,

(ii) $\| \Phi(t) P \Phi^{-1}(s) \| \leq Ce^{-\sigma(t-s)}$ for $t \geq s$

(iii) $\| \Phi(t)(E-P)\Phi^{-1}(s) \| \leq Ce^{-\sigma(s-t)}$ for $t < s$

where C and σ are positive numbers and $\Phi(t)$ is the fundamental matrix of (5.14).

For quasi-periodic L, we have the following.

Proposition 5.5.2. If a quasi-periodic operator L with periods $\omega_1, \omega_2, \cdots\cdots, \omega_m$ defined by (5.15) is regular, then for any quasi-periodic function $f(t)$ with periods $\omega_1, \omega_2, \cdots\cdots, \omega_m$ the differential equation (5.14) possesses a unique quasi-periodic solution $x = x(t)$ with the same periods given by

$$x(t)=\int_{-\infty}^{\infty} G(t,s)f(s)ds, \tag{5.18}$$

where

$$G(t,s)=\begin{cases} \Phi(t)p\Phi^{-1}(s) & for\ t\geq s, \\ -\Phi(t)(E-p)\Phi^{-1}(s) & for\ t<s. \end{cases} \tag{5.19}$$

$G(t,s)$ is called a Green function for L, and satisfies the inequality

$$\| G(t,s) \| \leq Ce^{-\sigma|t-s|} \tag{5.20}$$

5.6 Generalized Exponential Dichotomy. We slightly extend the conditions in proposition 5.5.1 as follows. Consider a linear differential operator L in (5.15) and let $\Phi(t)$ be the fundamental matrix of $Lx=0$ satisfying the condition $\Phi(0)=E$ as before.

The linear homogeneous equation $Lx=0$ is called to satisfy a generalized exponential dichotomy if there exist a $d\times d$ projection P, positive constant σ_1,σ_2 and nonnegative functions $C_1(t,s),C_2(t,s)$ such that

(i) $\| \Phi(t)P\Phi^{-1}(s) \| \leq C_1(t,s)e^{-\sigma_1(t-s)}$ for $t\geq s$,

(ii) $\| \Phi(t)(E-P)\Phi^{-1}(s) \| \leq C_2(t,s)e^{-\sigma_2(s-t)}$ for $t<s$,

(iv) the integral

$$\int_{-\infty}^{t} C_1(t,s)e^{-\sigma_1(t-s)}ds+\int_{t}^{\infty} C_2(t,s)e^{-\sigma_2(s-t)}ds \tag{5.21}$$

is bounded on R by a positive number M. Then we have the following propositions.

Proposition 5.6.1. Let L be a linear differential operator. Given in (5.15) and $A(t)$ a quasi-periodic $d \times d$ matrix with periods $\omega_1, \omega_2, \cdots\cdots, \omega_m$. Suppose that the equation $Lx = 0$ satisfies the generalized exponential dichotomy. Then for any quasi-periodic function $f(t)$ with periods $\omega_1, \omega_2, \cdots\cdots, \omega_m$ the inhomogeneous equation (IH) has a unique quasi-periodic solution $x(t)$ with the same periods given by

$$x(t) = \int_{-\infty}^{\infty} G(t,s) f(s) ds \tag{5.22}$$

where

$$G(t,s) = \begin{cases} \Phi(t) P \Phi^{-1}(s) & for\ t \geq s, \\ -\Phi(t)(E-P)\Phi^{-1}(s) & for\ t < s. \end{cases}$$

Moreover, the solution $x(t)$ satisfies the relation

$$\| x \| \leq M \| f \|. \tag{5.23}$$

The study is based on the following theorem:

Theorem 5.6.1. [Mitsui T. (1977)]

Given a nonlinear differential equation

$$\frac{dx}{dt} = X(t,x) \tag{5.24}$$

where x and $X(t,x)$ are vectors and $X(t,x)$ is quasi-periodic in t with periods $\omega_1, \omega_2, \cdots\cdots, \omega_m$ and is continuously differentiable with respect to x and x belongs to a region D of the $x-$space.

Suppose that there is a quasi-periodic function $x_0(t)$ with periods $\omega_1,\omega_2,\cdots\cdots,\omega_m$ such that

$$\begin{cases} x_0(t)\in D, \\ \| \dfrac{dx_0(t)}{dt} - X[t,x_0(t)] \| \le r \end{cases}, \tag{5.25}$$

for all t. Further suppose that there are a positive number δ, a nonnegative number $\kappa<1$ and a quasi-periodic matrix $A(t)$ with periods $\omega_1,\omega_2,\cdots\cdots,\omega_m$ such that

(i) the linear differential equation operator L define by

$$Ly=\frac{dy}{dt}-A(t)\,y \tag{5.26}$$

where $A(t)$ is quasi-periodic matrix ,

(ii) $$\begin{cases} D_\delta=\{x;\| x-x_0(t)\|\le\delta \quad for\ some\ t\}\subset D, \\ \| \psi(t,x)-A(t)\| \le \dfrac{\kappa}{M} \quad whenever \| x-x_0(t)\| \le\delta, \\ \dfrac{Mr}{1-\kappa}\le\delta \end{cases} \tag{5.27}$$

where $\psi(t,s)$ is the Jacobian matrix of $X(t,x)$ with respect to x and

$$M=\frac{2C}{\sigma}, \tag{5.28}$$

where C and σ are positive numbers such that Green function $G(t,s)$ for L satisfies

$$\| G(t,s)\| \le Ce^{-\sigma|t-s|} \tag{5.29}$$

Then the given equation (5.24) possesses a solution $x=\tilde{x}(t)$ quasi-periodic in t with periods $\omega_1, \omega_2, \cdots\cdots, \omega_m$ such that

$$\| x_0(t) - \tilde{x}(t) \| \leq \frac{Mr}{1-\kappa} \tag{5.30}$$

for all t. For the solution $\tilde{x}(t)$, a quasi-periodic differential operator $\hat{L}$ defined by

$$\hat{L}y = \frac{dy}{dt} - \psi[t, \hat{x}(t)]y.$$

is regular as an almost periodic differential operator. Furthermore, to the equation (5.24) there is no other quasi-periodic solution belonging to D_δ besides $x=\tilde{x}(t)$.

5.7 Quasi-periodic solution to second order differential equation with constant coefficients.

Consider the second order differential equation.

$$\frac{d^2x}{dt^2} + 2\mu\frac{dx}{dt} + \vartheta^2 x = a\cos\vartheta_1 t + b\sin\vartheta_2 t \tag{5.31}$$

where ϑ and ϑ_i $(i=1,2)$ are all positive numbers such that $\vartheta > 0$, $0 < |\mu| < \vartheta$ and $\omega_k = \dfrac{2\pi}{\vartheta_k}$ $(k=1,2)$. It is clear that $\dfrac{\omega_1}{\omega_2}$ is irrational.

We intend to get quasi-periodic solutions with periods ω_1 and ω_2 for the equation (5.24). The equation (5.31) written as the vector form

$$\frac{d\bar{x}}{dt} - A\bar{x} = \varphi(t), \tag{5.32}$$

where

$$\bar{x} = \begin{pmatrix} x \\ y \end{pmatrix}, \qquad A = \begin{pmatrix} 0 & 1 \\ -\vartheta^2 & -2\mu \end{pmatrix}, \quad \varphi(t) = \begin{pmatrix} 0 \\ a\cos\vartheta_1 t + b\sin\vartheta_2 t \end{pmatrix}$$

and $y = \frac{dx}{dt}$

Let L be the operator defined by

$$L\bar{x} = \frac{d\bar{x}}{dt} - A\bar{x} \tag{5.33}$$

The fundamental metric $\phi(t)$ of

$$L\bar{x} = 0 \tag{5.34}$$

can be given by

$$\phi(t) = e^{At} = e^{-\mu t}\begin{pmatrix} \cos\theta t + \frac{\mu}{\theta}\sin\theta t & \frac{1}{\theta}\sin\theta t \\ -\frac{\vartheta^2}{\theta}\sin\theta t & \cos\theta t - \frac{\mu}{\theta}\sin\theta t \end{pmatrix} \tag{5.35}$$

where

$$\theta = \sqrt{\vartheta^2 - \mu^2} \tag{5.36}$$

Here we introduce the ℓ_∞ – norm $\|\cdot\|$ of vectors and matrices.

From (5.35) we then have an estimation to $\phi(t)$ as follows:

$$\| \phi(t) \| \le Ce^{-\mu t} \tag{5.37}$$

where

$$C = \frac{\vartheta + 1}{\theta}\max(1, \vartheta).$$

Depending on the sign of μ, we have two cases.

Case (I)

If $0 < \mu < \vartheta$ holds, we can take E as the matrix P in proposition 5.5.1, and the Green function for the operator L is given by

$$G(t,s) = \begin{cases} \phi(t-s) & for\ t \geq s, \\ 0 & for\ t < s, \end{cases}$$

$$= \begin{cases} e^{-\mu(t-s)}\begin{pmatrix} \cos\theta(t-s) + \frac{\mu}{\theta}\sin\theta(t-s) & \frac{1}{\theta}\sin\theta(t-s) \\ -\frac{\vartheta^2}{\theta}\sin\theta(t-s) & \cos\theta(t-s) - \frac{\mu}{\theta}\sin\theta(t-s) \end{pmatrix} & for\ t \geq s, \\ 0 \quad for\ t < s. \end{cases}$$

(5.38)

Case (II)

If $0 > \mu > -\vartheta$ holds, we can take 0 as the matrix P in Proposition 5.5.1 and the Green function for the operator L is given by

$$G(t,s) = \begin{cases} 0 & for\ t \geq s, \\ -\phi(t-s) & for\ t < s, \end{cases}$$

$$= \begin{cases} 0 \quad for\ t \geq s. \\ -e^{-\mu(t-s)}\begin{pmatrix} \cos\theta(t-s) + \frac{\mu}{\theta}\sin\theta(t-s) & \frac{1}{\theta}\sin\theta(t-s) \\ -\frac{\vartheta^2}{\theta}\sin\theta(t-s) & \cos\theta(t-s) - \frac{\mu}{\theta}\sin\theta(t-s) \end{pmatrix} & for\ t < s, \end{cases}$$

(5.39)

In both cases, L is regular as the quasi-periodic differential operator with periods ω_1 and ω_2. From (5.18), quasi-periodic solution for (5.32) is given by

$$x(t)=\int_{-\infty}\phi(t-s)\phi(s)ds \tag{5.40}$$

for $\mu>0$. Substituting (5.38) into (5.40) and integrating , we have

$$x(t)=\frac{a}{(\vartheta^2-\vartheta_1^{\,2})^2+4\mu^2\vartheta_1^{\,2}}\left[(\vartheta^2-\vartheta_1^{\,2})\cos\vartheta_1 t+2\mu\vartheta_1\sin\vartheta_1 t\right]$$

$$+\frac{b}{(\vartheta^2-\vartheta_2^{\,2})^2+4\mu^2\vartheta_2^{\,2}}\left[(\vartheta^2-\vartheta_2^{\,2})\cos\vartheta_2 t+2\mu\vartheta_2\sin\vartheta_2 t\right]. \tag{5.41}$$

For $\mu<0$, the result coincides with (5.41). With above considerations, we have

Proposition 5.7.1. If $0<|\mu|<\vartheta$, the operator L defined by (5.33) is regular as the quasi-periodic differential operator with periods ω_1 and ω_2, and its Green function is given by (5.38) (for $\mu>o$) or (5.39) (for $\mu<o$). The unique quasi-periodic solution of equation (5.32) (or (5.31)) with periods ω_1 and ω_2 is given by (5.41).

5.8 Van der Pol Type Equation. Consider a Van dar Pol type equation with two frequencies quasi-periodic forcing term in the following form:

$$\frac{d^2x}{dt^2}-2\lambda(1-x^2)\frac{dx}{dt}+x=a\cos\vartheta_1 t+b\sin\vartheta_2 t, \tag{5.42}$$

where λ is a positive parameter, $\vartheta_1 = \frac{2\pi}{\omega_1}$, $\vartheta_2 = \frac{2\pi}{\omega_2}$ and $\frac{\omega_2}{\omega_1}$ is irrational.

Introducing $\bar{x} = \begin{pmatrix} x \\ y \end{pmatrix}$, $A(\lambda) = \begin{pmatrix} 0 & 1 \\ -1 & 2\lambda \end{pmatrix}$, $\varphi(t) = \begin{pmatrix} 0 \\ a\cos\vartheta_1 t + b\sin\vartheta_2 t \end{pmatrix}$,

$\eta(x) = \begin{pmatrix} 0 \\ -2x^2 y \end{pmatrix}$, the equation (5.42) can be rewritten as

$$\frac{d\bar{x}}{dt} = A(\lambda)\bar{x} + \varphi(t) + \lambda\eta(\bar{x}). \tag{5.43}$$

Let $L(\lambda)$ be the differential operator depending on λ such that

$$L(\lambda)z = \frac{dz}{dt} - A(\lambda)z. \tag{5.44}$$

We apply Proposition 5.7.1 to the case $\mu = -\lambda$, $\vartheta = 1$. Then we have that $L(\lambda)$ is regular as a quasi-periodic operator and that the Green Function for $L(\lambda)$ is given by

$$G_\lambda(t,s) = \begin{cases} 0 \quad for\ t \geq s. \\ -e^{\lambda(t-s)} \begin{pmatrix} \cos\theta(t-s) - \frac{\lambda}{\theta}\sin\theta(t-s) & \frac{1}{\theta}\sin\theta(t-s) \\ -\frac{1}{\theta}\sin\theta(t-s) & \cos\theta(t-s) + \frac{\lambda}{\theta}\sin\theta(t-s) \end{pmatrix} \quad for\ t < s, \end{cases} \tag{5.45}$$

where $\theta = \sqrt{1-\lambda^2}$. Therefore, the Green function $G_\lambda(t,s)$ satisfies

$$\| G_\lambda(t,s) \| \leq \frac{2}{\theta} e^{-\lambda} \, |t-s| \tag{5.46}$$

and quasi-periodic solution of the nonlinear equation

$$L(\lambda)\omega = \varphi(t) \tag{5.47}$$

is given by $\omega = \omega_0(t,\lambda) = \begin{pmatrix} x_0(t,\lambda) \\ y_0(t,\lambda) \end{pmatrix}$,

where

$$x_0(t,\lambda) = \frac{a}{(1-\vartheta_1^{\,2})^2 + 4\lambda^2\vartheta_1^{\,2}}\left\{(1-\vartheta_1^{\,2})\cos\vartheta_1 t - 2\lambda\vartheta_1 \sin\vartheta_1 t\right\}$$

$$+\frac{b}{(1-\vartheta_2^{\,2})+4\lambda^2\vartheta_2^{\,2}}\left\{(1-\vartheta_2^{\,2})^2\cos\vartheta_2 t - 2\lambda\vartheta_2\sin\vartheta_2 t\right\}, \quad (5.48)$$

$$y_0(t,\lambda) = \frac{a\vartheta_1}{(1-\vartheta_1^{\,2})^2+4\lambda^2\vartheta_1^{\,2}}\left\{-(1-\vartheta_1^{\,2})\sin\vartheta_1 t - 2\lambda\vartheta_1\cos\vartheta_1 t\right\}$$

$$+\frac{b\vartheta_2}{(1-\vartheta_2^{\,2})^2+4\lambda^2\vartheta_2^{\,2}}\left\{-(1-\vartheta_2^{\,2})^2\sin\vartheta_2 t - 2\lambda\vartheta_2\cos\vartheta_2 t\right\}.$$

Introducing $\tan\alpha = -\frac{1-\vartheta_1^{\,2}}{2\lambda\vartheta_1}$ and $\tan\beta = -\frac{1-\vartheta_2^{\,2}}{2\lambda\vartheta_2}$, we can write (5.48) as

$$x_0(t,\lambda) = \frac{a\sin(\vartheta_1 t+\alpha)}{\left\{(1-\vartheta_1^{\,2})^2+4\lambda^2\vartheta_1^{\,2}\right\}^{\frac{1}{2}}} + \frac{b\sin(\vartheta_2 t+\beta)}{\left\{(1-\vartheta_2^{\,2})^2+4\lambda^2\vartheta_2^{\,2}\right\}^{\frac{1}{2}}}, \quad (5.49)$$

Then, by differentiating (5.49), we have

$$y_0(t,\lambda) = \frac{a\vartheta_1\cos(\vartheta_1 t+\alpha)}{\left\{(1-\vartheta_1^{\,2})^2+4\lambda^2\vartheta_1^{\,2}\right\}^{\frac{1}{2}}} + \frac{b\vartheta_2\cos(\vartheta_2 t+\beta)}{\left\{(1-\vartheta_2^{\,2})^2+4\lambda^2\vartheta_2^{\,2}\right\}^{\frac{1}{2}}}. \quad (5.50)$$

For all t, from (5.49) and (5.50) we obtain

$$|x_0(t,\lambda)| \le \frac{|a|}{\left\{(1-\vartheta_1^{\,2})^2+4\lambda^2\vartheta_1^{\,2}\right\}^{\frac{1}{2}}} + \frac{|b|}{\left\{(1-\vartheta_2^{\,2})^2+4\lambda^2\vartheta_2^{\,2}\right\}^{\frac{1}{2}}}$$

$$\le \frac{|a|}{|1-\vartheta_1^{\,2}|} + \frac{|b|}{|1-\vartheta_2^{\,2}|}$$

and

$$|y_0(t,\lambda)| \le \frac{|a|\,\vartheta_1}{|1-\vartheta_1^{\,2}|} + \frac{|b|\,\vartheta_2}{|1-\vartheta_2^{\,2}|}.$$

Defining the constant K by

$$K = \max\left[\frac{|a|}{|1-\vartheta_1^{\,2}|} + \frac{|b|}{|1-\vartheta_2^{\,2}|}, \frac{|a|\,\vartheta_1}{|1-\vartheta_1^{\,2}|} + \frac{|b|\,\vartheta_2}{|1-\vartheta_2^{\,2}|}\right]$$

we get

$$|x_0(t,\lambda)|, |y_0(t,\lambda)| < K \tag{5.51}$$

for all t and $0 < \lambda < 1$. Using estimate (5.51), we can estimate the residual function for $\omega_0(t,\lambda)$ as follows

$$\| \frac{d\,\omega_0(t,\lambda)}{dt} - A(t)\omega_0(t,\lambda\} - \varphi(t) - \lambda\eta(\omega_0(t,\lambda)) \|$$

$$= \| -\lambda\eta(\omega_0(t,\lambda)) \|$$

$$= \lambda \left\| \begin{pmatrix} 0 \\ -2x_0^{\,2}y_0 \end{pmatrix} \right\|$$

$$\le \lambda\,|-2|\,|x_0|^2\,|y_0|$$

$$\le 2\lambda K^2.K$$

$$= 2\lambda K^3$$

Accordingly we can choose

$$r = 2\lambda K^3 \tag{5.52}$$

Let $D_k = \{x : \|x\| \le 2K\}$, $D' = \bigcup_{t\in R}\{x : \| -\omega_0(t,\lambda)\| \le K\}$.

It is clear that $\omega_0(t,\lambda) \in D_k$ for any t and $D' \subset D_k$.

Let us denote the Jacobian matrix of the right-hand side of (5.43) with respect to x by $\psi(x,\lambda)$. Then we have the inequality

$$\|\psi(x,\lambda) - A(\lambda)\| = \left\|\begin{pmatrix} 0 & 1 \\ -4\lambda xy - 1 & -2\lambda x^2 + 2\lambda \end{pmatrix} - \begin{pmatrix} 0 & 1 \\ -1 & 2\lambda \end{pmatrix}\right\|$$

$$= \left\|\begin{pmatrix} 0 & 0 \\ -4\lambda xy & -2\lambda x^2 \end{pmatrix}\right\|$$

$$\leq 4\lambda |x||y| + 2\lambda |x|^2$$

$$\leq 2\lambda |x|(2|y| + |x|)$$

$$\leq 2\lambda .2K(2.2K + 2K)$$

$$= 4\lambda K(6K)$$

$$= 24\lambda K^2.$$

for $x \in D'$. The inequality (5.46) yields

$$M = \frac{2.2}{\lambda\sqrt{1-\lambda^2}} = \frac{4}{\lambda\sqrt{1-\lambda^2}}.$$

In order to apply the theorem 5.6.1 to the present case, we have to cheek with the inequalities (5.27). The question is "Is it possible to take a nonnegative number $\kappa < 1$ satisfying the both of inequalities

$$24\lambda K^2 \leq \frac{\lambda\sqrt{1-\lambda^2}}{4}\kappa, \tag{5.53}$$

$$\frac{4}{\lambda\sqrt{1-\lambda^2}} \cdot 2\lambda K^3 \leq (1-\kappa)K \tag{5.54}$$

under the assumption $0<\lambda<1$?"

Answer is affirmative, if, when the inequality

$$K \le \sqrt{\frac{\sqrt{1-\lambda^2}}{52(2-\lambda^2)}} = \frac{1}{8}\sqrt{\frac{\sqrt{1-\lambda^2}}{3\eta}}, \text{ where } 0<\lambda<1 \tag{5.55}$$

holds, because we have

$$\left.\begin{aligned} \frac{4}{\lambda\sqrt{1-\lambda^2}}\cdot 24\lambda K^2 &= \frac{4}{\lambda\sqrt{1-\lambda^2}}\cdot 24\lambda\cdot\frac{1}{64}\left(\frac{\sqrt{1-\lambda}}{3\eta}\right) = \frac{1}{2\eta}, \\ \frac{4}{\lambda\sqrt{1-\lambda^2}}\cdot 2\lambda K^2 &= \frac{8}{\sqrt{1-\lambda^2}}\cdot\frac{1}{64}\left(\frac{\sqrt{1-\lambda^2}}{3\eta}\right) = \frac{1}{24\eta} \end{aligned}\right\} \tag{5.56}$$

So, we can choose such a nonnegative number $\kappa<1$ such that both the inequalities (5.55) and (5.56) may hold.

Summing up the considerations, we have the following

Proposition 5.8.1. If $0<\lambda<1$ holds, and if the constant

$$K = \max\left[\frac{|a|}{|1-\vartheta_1^{\ 2}|}+\frac{|b|}{|1-\vartheta_2^{\ 2}|}, \frac{|a|\,\vartheta_1}{|1-\vartheta_1^{\ 2}|}+\frac{|b|\,\vartheta_2}{|1-\vartheta_2^{\ 2}|}\right]$$

satisfies the inequality

$$K \le \frac{1}{8}\sqrt{\frac{\sqrt{1-\lambda^2}}{3\eta}}, \text{ where } 0<\eta<1$$

then the equation (5.42) possesses a quasi-periodic solution $x=\hat{x}(t)$ with periods ω_1 and ω_2 such that

$$\|\hat{x}(t)-x_0(t,\lambda)\| \le K \text{ for all } t .$$

Here, it is our error bound.

Reference

[1] Alefeld G. and Herzberger J. (1983), "Introduction to interval computations" Acad. Press, New York.

[2] Antosievice H. A. (1962), "Numerical analysis and functional analysis" In a survey of Numerical analysis, McGraw-Hill, New York, pp.485-517.

[3] Assem S. Deif. (1991), "Advanced Matrix Theory for Scientists and Engineers." Gordon and Breach Science Publishers.

[4] Bennet A. (1916), "Newton's method in general analysis." Proc. Nat. Acad. Sci. Vol. **2** pp. 592-598.

[5] Chan H. S. Y., K. W. Chung and Z. Xu (1996), "A perturbation-incremental method for strongly non-linear oscillators." Int. J. non-linear Mechanics, **31** pp.59-72.

[6] Charles W. Groetsck (1980), "Elements of Applicable Functional Analysis" Dakker, New York.

[7] Copson E.T (1968), "Matric space" Cambridge University Press. Cambridge.

[8] Datta K. B. (2002), "Matrix and Linear Algebra." Prenlice-Hall of India Private. Limited. New Delhi.

[9] Dayanand Reddy B. (1986), "Functional Analysis and Boundary-Value Problems: an Introductory Treatment" Longman Science and Technical. Copublished in the United States with John Wiley and Sons. Inc. New York.

[10] Dennis J. (1969), "On the Kantorovich Hypothesis for Newton's method." SIAM J. Numer. Anal. Vol. **11** pp.493-507.

[11] Fenyo I. (1954), "Uber die Losung der in Banachschen Raum definierten nichtlinearen Gleichungen." Act Math. Acad. Sci. Hungar. **5** pp.85-93.

[12] Fink A. M. (1974), "Almost periodic differential equations." Lecture notes in Math., 377, Springer-Verlag.

[13] Fine H. B. (1916), "On Newton's method of approximation." Pro. Nat'l. Acad. Sci. Vol. **2** pp.546-552.

[14] Gragg W. B. and R. A. Tapia. (1974), "Optimal error bounds for the Newton-Kantorovich theorem" SIAM J.Numer. Anal,Vol. **11** pp.10-13.

[15] Hansen E. (1969), "Topics in interval analysis." Clarendon Press. Oxford.

[16] Jean-Pierre Aubin. (1977), "Applid abstract analysis" John Wiley Sons Inc.

[17] Kantorovich L. V. and Akilov G. P. (1964), "Functional analysis in normed spaces." Fizmaliz, Moscow, English transl., Pergamoon Press, Oxford,

[18] Kantorovich L. V. (1949), "On Newton's method." Trudy Mat. Inst. Steklov. Vol. **28** pp.104-144. MR 12 (1951) 419.

[19] Kantorovich L. V. (1948), "Functional analysis and applied mathematics."Uspehi. Mat. Nauk. Vol. **3** pp.89-185, Tramsl. C.D. Benster, Nart. Bur. stand. 1948. MR 10, 380.

[20] Kantorovich L.V. (1949), "On Newton's method for functional equations (Russian)." Dokl. Akad. Nauk. sssR, Vol. **59** pp.1237-1240.

[21] Kohda A. and Y. Shinohara. (1993), "Numerical analysis of the quasiperiodic solutions to Duffing type equations." Japan J. Indust. Appl. Math. **10** pp.367-378.

[22] Kohda A. and Y. Shinohara. (1989), "On uniform limit of quasiperiodic functions." J. Math. Tokushima Univ. Vol. **23** pp.39-40.

[23] Kolmogorov V. & Fomin S. V. (1975), "Introductory Real Analysis" Dover publications Inc. New York.

[24] Krawczyk R. (1969), "Newton-Algorithmen zur Bestimmung von Nullstellen mit Fehlerschranken" Computing, Vol. **4** pp.187-201.

[25] Majedur Rarman M. & Zulfikar Ali Md. (2001) "Remark on A Note on Convergence of Newton's Method" Ann. Bull. of the Res. Ins. of Man. and Inf. Sci., Shikoku Univ., Japan. No. **7** pp. 83-86.

[26] Mitsui T. (1977), "Investigation of numerical solutions of some nonlinear quasiperiodic differential equations." Publ. RIMS, Kyoto Univ. **13** pp.793-820.

[27] Mitsui T. and Y. Shinohara. (1995), "Numerical analysis of ordinary differential equations and its applications." World Scientific.

[28] Moore R. E. (1978), "A computational test for Convergence of Iterative Methods for Nonlinear systems" SIAM J. Numer. Anal. Vol. **15** pp.1194-1196

[29] Moore R. E. (1977), "A test for existence of solutions to nonlinear systems" SIAM J. Numer. Anal. Vol. **14** pp.611-615.

[30] Moore R. E. (1966), "Interval Analysis" Prentice-Hall, Inc. Englewood Cliffs, N, J.

[31] Moser J. (1961), "A new technique for the construction of solution of nonlinear differential equations." Proc. Nat. Acad. Sci. USA Vol. **47** pp.1824-1831.

[32] Nakajima F. (1972), "Existence of quasiperiodic solutions of quasiperiodic systems." Funkcial Ekvac., **15** pp.61-73.

[33] Oetti W. & Prager W. (1964), "Compatibility of approximate solution of linear equations with given error bounds for coefficients and right-hand sides."Numer. Math. Vol. **6** pp.405-409.

[34] Ortega J. M. (1968), "The Newton-Kantorovich theorem", Amer. Math. Monthly. Vol. **75** pp. 658 – 660.

[35] Ostrowski A. M. (1973), "Solution of Equations in Euclidean and Banach spaces." 3rd Ed. Academic Prees, New York.

[36] Rall L. B. (1974), "A note on the convergence of Newton's method" SIAM J. Numer. Anal. Vol. **11** pp.34-36.

[37] Rall L. B. (1980), "A Comparison of the Existence Theorems of Kantorovich and Moore" SIAMJ.Numer. Anal. Vol. **17** pp.148-161

[38] Rall L. B. (1969), "Computational solution of Nonlinear operator Equations." John Wiley. New York.

[39] Shinohara Y. (1972), "A geometric method of numerical solution of nonlinear equations and its application to nonlinear oscillations." Publ. RIMS, Kyoto Univ., **8** pp.13-42.

[40] Shinohara Y. (1977), "Numerical analysis of periodic solutions and their periods to autonomous differential system." J. Math. Tokushima Univ., **11** pp.11-32.

[41] Shinohara Y., A Kohda and T. Mitsui. (1984), "On quasiperiodic solutions to van der Pol equation." J. Math. Tokushima Univ. **18** pp.1-9.

[42] Shinohara Y., M. Kurihara and A Kohda. (1986), "Numerical analysis of quasiperiodic solutions to nonlinear differential equations." Japan J. Appl. Math. **3** pp.315-330.

[43] Tapia R. A. (1971), "The Kantorovich theorem for Newton's method", Amer. Math. Monthly, Vol. **78** pp.389-392.

[44] Urabe. M. (1972), "On the existence of quasiperiodic solutions to nonlinear quasiperiodic differential equations." Proc. 6th ICNO, Polish Akad. Sci., Poznan, Marsaw. pp.1-38.

[45] Urabe. M. (1965), "Galerkin's procedure for nonlinear periodic systems." Arch. Rational Mech. Anal. **20** pp.120-152.

[46] Urabe. M. (1971), "Green Functions of pseudoperiodic differential operators."Japan-US Seminar on Ordinary Differential and Functional Equations, Lecture Notes in Math., 243 Springer Verlag.

[47] Urabe. M. (1972), "Existence theorem of quasiperiodic solutions to nonlinear differential Systems." Funkcial. Ekvac. **15** pp.75-100.

[48] Zulfikar Ali Md. & Majedur Rarman M. (2002), "On Convergence of Newton's Method" Ann. Bull. of the Res. Ins. of Man. and Inf. Sci., Shikoku Univ., Japan. No. **8** pp. 99-103.

[49] Zulfikar Ali Md. & Majedur Rarman M. (2002), "Remark on Moore's Theorem", Acta Ciencia Indica, India, (Submitted)

[50] Zulfikar Ali, Y Shinohara, H Imai, A Kohda, K Okamoto, H Sakaguchi, H Miyamoto. (1997), "Existence and Uniqueness of Quasi-periodic solutions to Van der Pol type equations" J Math. Tokushima University Vol. **31** pp. 69-80.