MEMOIRS OF NUMERICAL MATHEMATICS

NUMBER 5

1978

The MEMOIRS OF NUMERICAL MATHEMATICS will publish original papers concerned with numerical mathematics, and henceforth will be published once a year.

All correspondence should be addressed to

Sin Hitotumatu
Research Institute for Mathematical Sciences
Kyoto University, Sakyo-ku, Kyoto
606, Japan

CONTENTS

Equivalence Class and Invariant Figures of Rational Iterations with special reference to the Global Convergence Properties of Newton's Method

Kohei Sato

A B S T R A C T

Basic concepts and theorems about global convergence features of rational iterations with one complex variable are summarized here. The concept of equivalence class of a rational function, which is implicit in the works of G. Julia and others, is reformulated in view of practical application. Properties of the "domain of direct convergence" presented by Julia are derived more systematically from those of "invariant figures" for a given rational iteration. Applying them to Newton's method for solving polynomial equations, we have made a scheme of sketching the shapes of its convergence regions mechanically.

1. Introduction

The iteration by a rational function with one complex variable is one of the processes most frequently appearing in numerical analysis. The essential part of its global convergence features was studied by Gaston Julia in his comprehensive work [4]. His work, being concerned mainly about the topological features of convergence areas, also contains many hints on the numerical methods for determining their boundaries. On the other hand, in the field of numerical analysis, we hardly find general treatments on global behavior of this kind of iteration.

The aim of this paper is to sum up such global properties of rational iterations that are considered to be useful in numerical analysis and to establish a practical method for determining the shape of its convergence areas. Some of the concepts and theorems presented here, including the concept of equivalence class and theorems about Newton's iteration, are either implicit in Juia's work [4] or readily derivable therefrom. However, they are reformulated in view of numerical analysis and their implications in the practically significant situations of applied mathematics are clarified. In particular, the idea of invariant figures will turn out to be useful for investigating the properties of convergence areas.

1.1. Definitions. By $\mathbb{C}$ we mean the closed complex plane or the Riemann sphere. A rational function Φ is considered as a mapping from $\mathbb{C}$ onto $\mathbb{C}$. The iteration by Φ, which will be denoted by $\mathrm{It}[\Phi]$, is a sequence of mappings: $\mathrm{It}[\Phi] \equiv \{\Phi_0, \Phi_1, \Phi_2, \cdots\}$ where $\Phi_0(z) \equiv z$, $\Phi_1 = \Phi$, $\Phi_2 = \Phi\Phi$, $\Phi_3 = \Phi\Phi\Phi$ etc. The set $\{w \mid \Phi_n(w) = z\}$ is represented by $\Phi_{-n}(z)$.

When two subsets A, B of $\mathbb{C}$ are related to each other by the relation $\Phi_n(A) = B$ $(n \geqq 1)$, we call A an antecedent of B and B a consequent of A in $\mathrm{It}[\Phi]$. Especially, if $\Phi_m(A) \neq \Phi_n(A) = B$ $(0 \leqq m < n)$, A is called an n-th antecedent of B and B the n-th consequent of A.

When $\Phi(A) = A$, we call A an *invariant figure* or simply a *figure* of $It[\Phi]$. $\mathbb{C}$, $\emptyset$, a set of fixed points of Φ, etc. are examples of the simplest (invariant) figures of $It[\Phi]$.

The set $\{z \mid \lim_{n\to\infty} \Phi_n(z) = x\}$ is denoted by $\mathbb{U}_\Phi(x)$ (if Φ or x is evident from the context, we shall denote it by $\mathbb{U}(x)$ or simply by $\mathbb{U}$). It is well known that, if $\mathbb{U}_\Phi(x) \neq \emptyset$, x is a fixed point of Φ. The interior, or the open kernel, of $\mathbb{U}(x)$ is denoted by $U(x)$, each of whose connected component is called a *convergence region* (or a *domain of convergence*) of $It[\Phi]$ toward x. $\mathbb{U}(x)$ and $U(x)$ are other examples of figures of $It[\Phi]$, while a convergence region is not always a figure of $It[\Phi]$.

1.2. Equivalence of rational functions. If Φ and Ψ are rational functions and there exists a linear fractional function T such that $T\Phi T_{-1} = \Psi$, Ψ is said to be equivalent to Φ and is denoted as $\Phi \simeq \Psi$ or $\Phi \xrightarrow{T} \Psi$. The following propositions follow readily from this definition.

(1) The relation $\simeq$ is an equivalence relation.

(2) If $\Phi \simeq \Psi$, then $\deg \Phi = \deg \Psi$.

(3) If $\Phi \xrightarrow{T} \Psi$ and A is an n-th antecedent of B in $It[\Phi]$, then $T(A)$ is an n-th antecedent of $T(B)$ in $It[\Psi]$.

(4) If $\Phi \xrightarrow{T} \Psi$ and A is an invariant figure of $It[\Phi]$, then $T(A)$ is an invariant figure of $It[\Psi]$.

(5) If $\Phi \xrightarrow{T} \Psi$, then $T[\mathbb{U}_\Phi(x)] = \mathbb{U}_\Psi[T(x)]$ and $T[U_\Phi(x)] = U_\Psi[T(x)]$.

1.3. Classification of fixed points. Let x be a fixed point of Φ. Then x is either a point of $U_\Phi(x)$ or a point of $\mathbb{U}_\Phi(x) - U_\Phi(x)$. In the former case we call x a *stable fixed point* and in the latter case an *unstable fixed* point. Now we define the "dissipation factor" of the fixed point x by

$$(1.1) \qquad \rho(x) = \begin{cases} \Phi'(x) & \text{if } x \neq \infty , \\ 1/\Phi'(x) & \text{if } x = \infty \end{cases}$$

which is invariant under an equivalence transformation, i.e., if $\Phi \xrightarrow{T} \Psi$ then $\rho_\Phi(x) = \rho_\Psi[T(x)]$. It can easily be verified that x is stable if $0 \leqq |\rho(x)| < 1$ and unstable if $1 \leqq |\rho(x)| < \infty$. Furthermore, it is known (see, e.g., [1], [4],[5]) that

(1) the convergence of $It[\Phi]$ is of the second or higher order in a vicinity of x if $\rho(x) = 0$;

(2) the convergence of $It[\Phi]$ is of the first order in a vicinity of x if $0 < |\rho(x)| < 1$;

(3) if $|\rho(x)| = 1$, either $U(x) \neq \emptyset$ or every vicinity of x contains infinitely many non-convergennt sequences of values generated by $It[\Phi]$;

(4) if $|\rho(x)| > 1$, then $U(x) = \emptyset$.

We call the case (1) strongly stable, the case (2) weakly stable, the case (3) semistable and the case (4) properly unstable. If $It[\Phi]$ shows the convergence of the n-th order in a vicinity of a strongly stable fixed point x, we may call x a stable fixed point of the n-th order, which is the case when x ($\neq \infty$) is an (n-1)-ple zero of $\Phi'(x)$ or $x = \infty$ is an n-ple pole of $\Phi(x)$.

It can elementarily be verified that there are $\deg \Phi + 1$ different fixed points of Φ if and only if all of them have dissipation factors other than 1. So we call x a multiple fixed point if $\rho(x) = 1$ and a simple fixed point otherwise.

Table 1. Classification of Fixed Points

<table>
<tr><td colspan="2">stable fixed point $(\mu)_0$</td><td colspan="3">unstable fixed point $(\lambda)_0$</td></tr>
<tr><td>strongly stable</td><td>weakly stable</td><td colspan="2">semistable</td><td>properly unstable</td></tr>
<tr><td colspan="2">simple</td><td>multiple</td><td colspan="2">simple</td></tr>
<tr><td>$\rho = 0$</td><td>$0 < |\rho| < 1$</td><td>$\rho = 1$</td><td>$|\rho| = 1$
$\rho \neq 1$</td><td>$1 < |\rho|$</td></tr>
<tr><td>second or higher order convergence</td><td>first order convergence</td><td colspan="2">domain of convergence exists</td><td>domain of convergence is absent</td></tr>
<tr><td colspan="2">the fixed point is contained in a domain of convergence</td><td colspan="3">the fixed point is not contained in any domain of convergence</td></tr>
</table>

The <u>multiplicity</u> of the fixed point x of Φ is the multiplicity of the root x of the equation $\Phi(x) = x$ when $x \neq \infty$, whereas it is the multiplicity of the root 0 of the equation $z\Phi(1/z) = 1$ when $x = \infty$.

<u>1.4. Julia's theory.</u> The results of Julia's study [4] on the global convergence properties of rational iterations may be summarized into the following four theorems.

[Theorem I] If $\deg \Phi \geq 2$ and Φ has no multiple fixed point, then Φ has at least one properly unstable fixed point.

[Theorem II] Let E be the set of all antecedents of all (1) properly unstable fixed points and (2) semistable fixed points whose dissipation factors are roots of 1, of Φ with $\deg \Phi \geq 2$ and E' be its derived set. Then every neighborhood of every point of E' contains antecedents of all the points on $\mathbb{C}$ except at most two points, and E' coincides with the boundary of $\mathbb{U}(x)$, where x is an arbitrary fixed point of Φ.

[Theorem III] Let x be a stable fixed point of Φ with $\deg \Phi \geqq 2$, and the convergence region of $It[\Phi]$ which contains x be denoted by $D_0(x)$. Then $D_0(x)$ contains at least one critical point of Φ, and every component of $U(x)$ other than $D_0(x)$ is an antecedent of $D_0(x)$ if $U(x) \neq D_0(x)$.

[Theorem IV] Let x be a semistable fixed point of Φ with $\deg \Phi \geqq 2$, and the convergence region of $It[\Phi]$ which has x on its boundary be $D_0(x)$. If x is a multiple fixed point with multiplicity p, there are $p - 1$ such regions, each of them containing at least one critical point of Φ. If $\rho(x)$ is an n-th primitive root of 1 with $n \geqq 2$, then the number of $D_0(x)$'s is a multiple of n and at least one of them contains at least one critical point of Φ. In both cases, every component of $U(x)$ is an antecedent of $D_0(x)$.

Theorem I is connected with the fact that, if Φ has $N + 1$ simple fixed points x_i $(i = 1, \cdots, N + 1 ; N = \deg \Phi)$ then we have

$$\sum_{i=1}^{N+1} \frac{1}{1 - \rho(x_i)} = 1 , \tag{1.2}$$

which will be used in the next section of this paper.

Theorem II implies that E' is a perfect set (i.e., every point of E' is its accumulation point) and that a convergence region is either simply connected or infinitely multiply connected.

Julia named the domain $D_0(x)$ in Theorem III and Theorem IV the <u>domain of direct convergence</u>. If $D_0(x) = U(x)$

(which is the case if x is a stable or a double fixed point and $\underset{0}{D}(x)$ contains $N-1$ critical points of Φ), $\underset{0}{D}(x)$ is naturally called a complete convergence region.

2. Special Classes of Rational Iterations

If $\Phi \simeq \Psi$ and the global convergence features of $It[\Phi]$ are known, those of $It[\Psi]$ can easily be derived therefrom. Hence in investigating the global features of the iteration by any given rational function, we may find a simplest function in the equivalence class to which it belongs. The following theorems are useful for this purpose.

[Theorem 2.1] A rational function Φ of degree N is equivalent to a fraction whose denominator is a polynomial of degree $N-1$.

Proof: If ∞ is a simple pole of Φ, the theorem obviously holds true. If ∞ is not a simple pole of Φ, it can be verified elementarily that there exists at least one fixed point x of Φ such that $\Phi'(x) \neq 0$ and $x \neq \infty$. Hence if we put $T = 1/(z-x)$, the point at infinity is a simple pole of $T\Phi T_{-1}$ which is equivalent to Φ.

[Theorem 2.2] The n.a.s.c. that a rational function Φ of degree N is equivalent to a polynomial is that it has a stable fixed point of the N-th order.

Proof: If ∞ is an N-th order stable fixed point of Φ, Φ is obviously a polynomial of degree N and vice versa. If $x \neq \infty$ is a stable fixed point of the N-th order, then the point at infinity is an N-th order stable fixed point of $T\Phi T_{-1}$ where $T = 1/(z-x)$, and vice versa.

[Theorem 2.3] If a rational function Φ of degree N has two stable fixed points of the N-th order, then Φ is equivalent to z^N and vice versa.

Proof: Let α and β be fixed points of the N-th order. Then putting $T = (z - \alpha)/(z - \beta)$ if $\alpha \neq \infty$ and $\beta \neq \infty$ and $T = z - \alpha$ if $\alpha \neq \infty$ and $\beta = \infty$, we have $T\Phi T_{-1} = z^N$ because z^N is the only polynomial that has two fixed points of the N-th order at 0 and ∞. The converse is obvious.

Now suppose that f is a rational function which is not a linear polynomial, and put

$$(2.1) \qquad \Phi(z) = z - f(z)/f'(z) .$$

It can easily be verified, by means of the Laurent expansions, that all the fixed points of Φ are simple, their dissipation factors being as shown in Table 2.

Table 2. Fixed points of $\Phi(z) \equiv z - f(z)/f'(z)$ where f is a rational function of z

	of f	order	fixed point of Φ	$\rho(x)$
$x \neq \infty$	zero	1	strongly stable	0
		m ($\geqq 2$)	weakly stable	$1 - 1/m$
	pole	m	properly unstable	$1 + 1/m$
$x = \infty$	neither a zero nor a pole		strongly stable	0
	zero	m	weakly stable	$m/(m + 1)$
	pole	m ($\geqq 2$)	properly unstable	$m/(m - 1)$

We call Φ defined by (2.1) the <u>Newton transform</u> of f, It[Φ] being Newton's method for solving the equation $f(z) = 0$. Hereafter in this paper, the "Newton transform" will be abbreviated as "N.T.".

[Theorem 2.4] The n.a.s.c. for a rational function Φ to be the N.T. of another rational function f is that the dissipation factor of every fixed point of Φ differs from 1 by the inverse of a natural number.

<u>Proof</u>: The necessity is evident from Table 2. Conversely, let λ_i ($i = 1, \cdots, L$) and μ_j ($j = 1, \cdots, M$) be the fixed points of Φ with

$$(2.2) \qquad \rho(\lambda_i) = 1 + 1/\ell_i \quad \text{and} \quad \rho(\mu_j) = 1 - 1/m_j ,$$

respectively, where ℓ_i and m_j are natural numbers. Then all these fixed points are simple so that from (1.2) we have

$$(2.3) \qquad \sum_{j=1}^{M} m_j - \sum_{i=1}^{L} \ell_i = 1 .$$

Hence, if we put

$$(2.4) \qquad \Psi(z) = z - f(z)/f'(z)$$

where

$$(2.5) \qquad f(z) = \begin{cases} \prod_j (z - \mu_j)/\prod_i (z - \lambda_i) & \text{if } \Phi(\infty) \neq \infty, \\ \prod_j (z - \mu_j)/\prod_{i \neq k} (z - \lambda_i) & \text{if } \lambda_k = \infty, \\ \prod_{j \neq k} (z - \mu_j)/\prod_i (z - \lambda_i) & \text{if } \mu_k = \infty, \end{cases}$$

then, referring to Table 2, we have the relation

$$(2.6) \qquad \Phi(z) = \Psi(z) , \qquad \Phi'(z) = \Psi'(z)$$

for $\deg \Phi + 1$ distinct points (fixed points of Φ). Hence $\Phi(z)$ and $\Psi(z)$ must coincide with each other for all values of z.

The next theorem is derived directly from Theorem 2.4.

[Theorem 2.5] A function equivalent to the N.T. of a rational function is also the N.T. of a rational function.

On the other hand, from (2.3) we have

[Theorem 2.6] The N.T. of a rational function has at least one stable fixed point.

[Theorem 2.7] If Φ is the N.T. of a rational function and has only one unstable fixed point, then Φ is equivalent to the N.T. of a polynomial.

Proof: If $\Phi(z) = z - f(z)/f'(z)$ and Φ has only one unstable fixed point at ∞, then ∞ is the only pole of f so that f is a polynomial. From this and the invariance of dissipation factors under equivalence transformation follows the theorem.

It must be noted that the equivalence of N.T.'s of two rational functions does not mean the equivalence of the rational functions themselves. For instance, the N.T. of $cf(z)$, where c is a constant other than 0 or 1, is the same as that of $f(z)$, whereas $cf \neq f$. On the other hand, if Φ is the N.T. of f and S is a linear polynomial, $S\Phi S_{-1}$ is (in general) not the N.T. of SfS_{-1} but that of fS_{-1}, which will be called a function similar to f.

Example 1. If a quadratic function has no multiple fixed point, it is equivalent to

$$(2.7) \qquad \Phi(z) \equiv z - \frac{z^2 - 1}{k(z - p)} \quad ,$$

where $k \neq 1$, $p \neq \pm 1$. Since the dissipation factors of the fixed points of the $\Phi(z)$ are

$$(2.8) \qquad \rho(\pm 1) = 1 - \frac{2}{k(1 \mp p)} \quad \text{and} \quad \rho(\infty) = \frac{k}{k - 1},$$

we may suppose $|\rho(\infty)| > 1$ or

$$(2.9) \qquad \mathrm{Re}(k) > 1/2$$

without loss of generality (cf. Theorem I in § 1.1). Within the restriction of (2.9), the real parts of both $k(1 + p)$ and $k(1 - p)$ may be less than 1, which shows that all fixed points of a rational function may be properly unstable.

From Theorem 2.2, $\Phi(z)$ is equivalent to a quadratic polynomial if either $k(1 - p)$ or $k(1 + p)$ equals 2.

From Theorem 2.4, $\Phi(z)$ is the N.T. of a rational function if k is a positive integer (greater than 1) and $k(1 + p)$ is an even integer. Especially, if both $k(1 - p)$ and $k(1 + p)$ are positive even integers, $\Phi(z)$ is the N.T. of a polynomial.

If $k = 2$ and $p = 0$, then Φ is the N.T. of $z^2 - 1$ which is similar to all quadratic polynomials with simple zeros. On the other hand, $\Phi \simeq z^2$ if $k = 2$ and $p = 0$. The well-known properties of Newton's method applied to quadratic equations with simple roots can therefore be deduced most naturally from the properties of $\mathrm{It}[z^2]$.

Note that z^2 itself is the N.T. of the linear function $z/(z-1)$. Furthermore, z^2 is equivalent to $2z - kz^2$ $(k \neq 0)$, the latter being the function iterated in the widely used method for calculating the inverse of k.

Example 2. If a quadratic function has a multiple fixed point, it is equivalent to either

$$z + 1 + a/z \quad (a \neq 0) \quad \text{or} \quad z + 1/z \; . \tag{2.10}$$

The former, having a double fixed point at ∞ and a simple fixed point at $-a$, is equivalent to polynomials if and only if $a = 1$. The latter has a triple fixed point at ∞ and is not equivalent to any polynomial.

Example 3. Bailey's iteration for root extraction is defined by

$$\Phi(z) \equiv \frac{z[(n-1)z^n + (n+1)a]}{(n+1)z^n + (n-1)a} \; , \tag{2.11}$$

where $n \geq 2$, $a \neq 0$ (cf. [7]). The dissipation factors of the fixed points of this function are

$$\rho(a^{\frac{1}{n}}) = 0 \; , \quad \rho(0) = \rho(\infty) = \frac{n+1}{n-1} \; . \tag{2.12}$$

Hence Φ must be the N.T. of a rational function if n is odd. In fact, we have $\Phi(z) = z - f(z)/f'(z)$ by putting

$$f(z) = z^{-\frac{n-1}{2}}(z^n - a) \; . \tag{2.13}$$

It is easy to verify that all the stable fixed points of Φ are of the third order. Especially, if $n = 2$, the function is equivalent to z^3 from Theorem 2.3.

3. Divergence Centers

If we cannot find a rational function Ψ in the equivalence class of Φ such that the E' of $It[\Psi]$ is known to be a circle or a closed line segment, then the E' of $It[\Phi]$ is not an analytic curve [3],[4]. In this most general case, the shape of the E' can be traced only numerically, for instance, by plotting as many points of E as possible. This method of numerical plotting has several theoretical advantages as will be shown later. For simplicity, the set of all unstable fixed points of Φ other than semistable ones with dissipation factors whose arguments are incommensurable multiples of π will be denoted by $\underset{0}{\Lambda} = \{\underset{0}{\lambda}\}$ and the n-th antecedent of $\underset{0}{\Lambda}$ by $\underset{n}{\Lambda} = \{\underset{n}{\lambda}\}$. If we write

$$E_n = \bigcup_{k=0}^{n} \underset{k}{\Lambda} , \tag{3.1}$$

then we have $\lim_{n\to\infty} E_n = E$. If the numerical values of $\underset{0}{\lambda}$'s $\in \underset{0}{\Lambda}$ are known, those of $\underset{1}{\lambda}$'s $\in \underset{1}{\Lambda}$, $\underset{2}{\lambda}$'s $\in \underset{2}{\Lambda}$ etc. can be calculated by solving $\Phi(\underset{k}{\lambda}) = \underset{k-1}{\lambda}$ successively. Hereafter $\underset{k}{\lambda}$ will be called a <u>divergence center</u> of the k-th order.

<u>Example 4.</u> It can easily be verified that every quadratic polynomial is equivalent to

$$\Phi(z) \equiv z^2 + p , \tag{3.2}$$

where p is a constant. Here we restrict ourselves to the case where p is real. The fixed points of (3.2) are

$$c_1 = (1 - \sqrt{1-4p})/2 , \quad c_2 = (1 + \sqrt{1-4p})/2 \tag{3.3}$$

and ∞. Their dissipation factors being $2c_1$, $2c_2$ and 0, we have

$$\underset{0}{\Lambda} = \begin{cases} \{c_2\} & \text{if } -3/4 < p \leqq 1/4 , \\ \{c_1 , c_2\} & \text{otherwise} . \end{cases} \tag{3.4}$$

The divergence centers are calculated successively by means of the formulae:

$$\lambda_1 = -\lambda_0 , \quad \lambda_k = \pm\sqrt{\lambda_{k-1} - p} \quad (k \geqq 2) . \tag{3.5}$$

Rough sketches of E' mechanically made from numerical values of the points of E_8 for several real values of p are shown in Fig. 1. (Cf. Example 5 in § 5.)

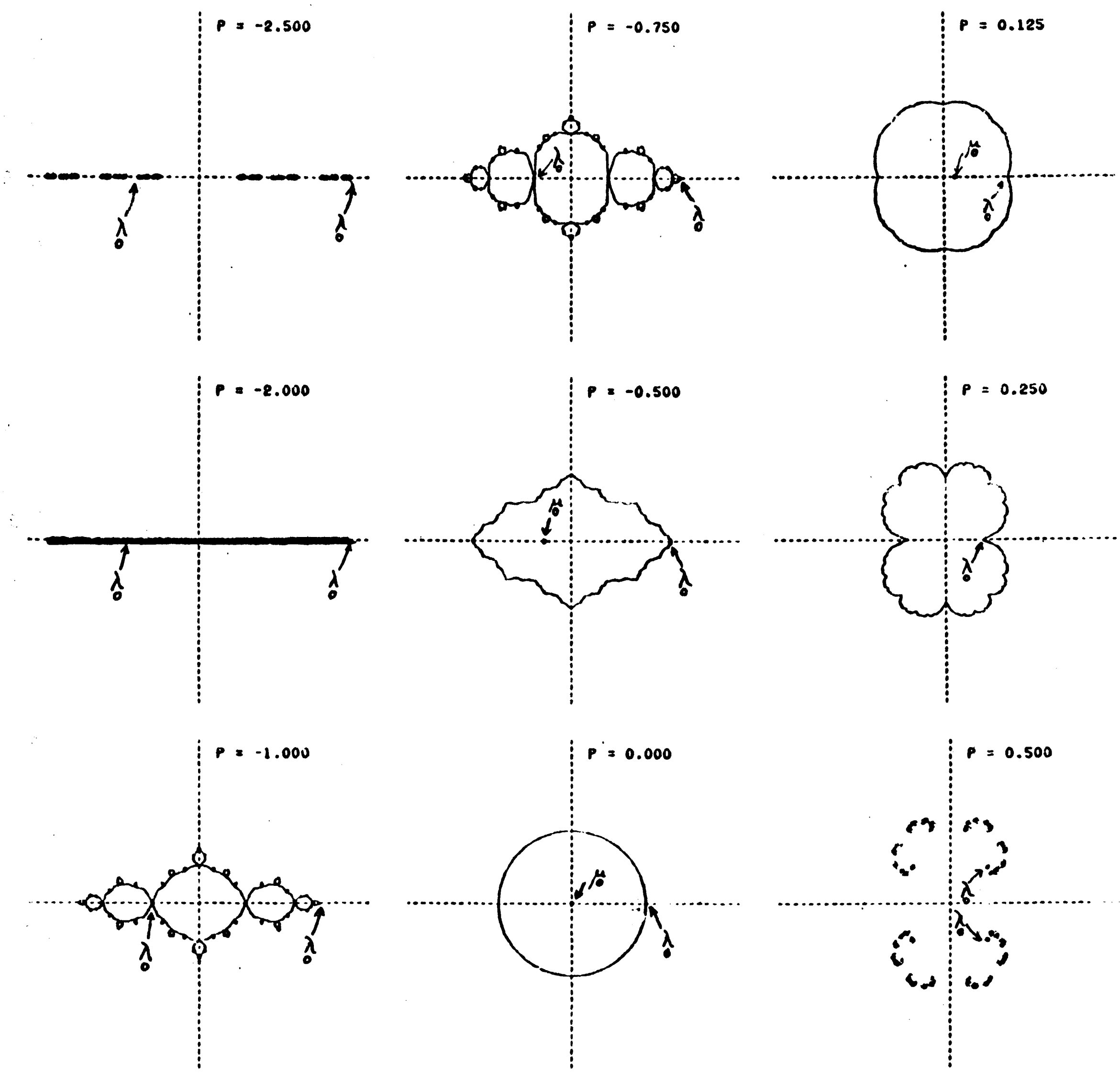


Fig. 1. Julia's singularity set E' for $It[z^2 + p]$ when p takes several real values (mechanically plotted from numerical values of the points of E_8)

4. Minimal Figurés

If $A \geq B$ and both A and B are figures of $It[\Phi]$. we call B a <u>subfigure</u> of A. If every subfigure B of a figure A such that

$$(4.1) \qquad \lim_{n\to\infty} \bigcup_{m=0}^{n} \Phi_{-m}(B) \equiv \lim_{n\to\infty} \Phi_{-n}(B) \geq A$$

has the same closure as A, we call A a <u>minimal figure</u> of $It[\Phi]$. A minimal subfigure of A which is not contained in another minimal subfigure of A will be called a <u>kernel</u> of A.

For example, either a fixed point of Φ or a set of fixed points of Φ is a minimal figure of $It[\Phi]$. If $\deg \Phi \geq 2$, then E, the set of all the divergence centers, is not a minimal figure of $It[\Phi]$, its kernel being $\underset{0}{\Lambda}(= E_0)$, i.e., the set of all unstable fixed points of Φ.

An n-cycle of $It[\Phi]$, i.e., the set of n distinct points $x_1, x_2, \cdots, x_n$ satisfying

$$(4.2) \qquad \Phi(x_k) = x_{k+1} \; (k = 1,\cdots,n-1), \quad \Phi(x_n) = x_1$$

is another example of minimal figure.

From the above definitions, the following propositions are easily derived.

(1) If a minimal figure is a finite set, each of its elements is either a fixed point or a member of some cycle.

(2) If a minimal figure consists of a finite number of connected components, each component coincides with one of its own consequents.

(3) The kernel of a figure which is perfect (i.e., which coincides with its derived set) is perfect.

(4) The kernel of a figure which is an open set is an open set.

5. Areas of Direct and Indirect Convergence

Hereafter we always suppose that

$$(5.1) \qquad N \equiv \deg \Phi \geqq 2 \quad \text{and} \quad U(x) \neq \emptyset \,.$$

The kernel of $U(x)$ will be denoted by $\underset{0}{U}(x)$ and called the <u>area of direct convergence</u> toward x. From the property (4) of the preceding section, $\underset{0}{U}(x)$ is an open set. A component domain of $\underset{0}{U}(x)$ is nothing but Julia's "domain of direct convergence" which is denoted by $\underset{0}{D}(x)$ in this paper. Especially, if x is either a stable or a double fixed point, $\underset{0}{U}(x)$ coincides with $\underset{0}{D}(x)$.

As a natural consequence of its definition, $\underset{0}{U}(x)$ is a minimal figure of $It[\Phi]$. If x is either a stable or a multiple fixed point, $\underset{0}{D}(x)$ is also a minimal figure of $It[\Phi]$.

If $U(x) \neq \underset{0}{U}(x)$, we call $U(x) - \underset{0}{U}(x)$ the <u>area of indirect convergence</u> toward x and its connected component a <u>domain of indirect convergence</u> toward x. A domain of indirect convergence that is an n-th antecedent of $\underset{0}{D}(x)$ will be denoted by $\underset{n}{D}(x)$. The union of all $\underset{n}{D}(x)$'s will be denoted by $\underset{n}{U}(x)$. Thus we have

$$(5.2) \qquad U(x) = \bigcup_{n=0}^{\infty} \underset{n}{U}(x) = \bigcup_{n=0}^{\infty} \bigcup \underset{n}{D}(x) \,.$$

<u>5.1. Properties of $\underset{0}{U}(x)$ and $\underset{0}{D}(x)$.</u> It has already been proved by Julia that $\underset{0}{U}(x)$ has a finite number of components if either $|\rho(x)| \leq 1$ or $\rho(x)$ is a root of a cyclotomic equation. The following theorems will hold

true even if $|\rho(x)| = 1$ and $\rho(x)$ is not a root of a cyclotomic equation. For simplicity, we may write $\underset{\circ}{U}$ for $\underset{\circ}{U}(x)$ and $\underset{\circ}{D}$ for $\underset{\circ}{D}(x)$ if x is either understood or indeterminate.

[Theorem 5.1] There is at least one unstable fixed point on $\partial\underset{\circ}{D}(x)$.

Proof: If x is a semistable fixed point, the theorem is obvious. If x is a stable fixed point, it is a simple fixed point so that at least one fixed point y distinct from x exists. The fixed point y must lie either on $\partial\underset{\circ}{D}(x)$ or in a region Δ bounded by a closed Jordan curve that is a subset of $\partial\underset{\circ}{D}(x)$. In the former case, y must be unstable and the theorem holds. In the latter case, both Δ and $\partial\Delta$ being invariant figures, we can prove the existence of at least one fixed point on $\partial\Delta$ by reductio ad absurdum, through applying Carathéodory's theorem [9] to Δ. This fixed point on $\partial\Delta$ is naturally unstable.

[Theorem 5.2] If $\underset{\circ}{U}(x)$ consists of a finite number of $\underset{\circ}{D}(x)$'s, then every $\underset{\circ}{D}(x)$ coincides with one of its own consequents, i.e., there exists n for each $\underset{\circ}{D}(x)$ such that $\underset{\circ}{D}(x)$ is a figure of $It[\Phi_n]$.

Proof: This theorem is a special case of the property (2) of minimal figures mentioned in the foregoing section.

[Theorem 5.3] If $\underset{\circ}{D}$ is a figure of $It[\Phi_n]$, it is multiply connected if and only if it contains a pair of critical points of Φ_n around which the same branches of Φ_{-n} exist.

Proof: Let Γ be a connected component of ∂D_0 on which a fixed point of Φ exists. Then Γ must be a figure of $It[\Phi_n]$. If D_0 is multiply connected, ∂D_0 contains at least one branch of $\Phi_{-n}(\Gamma)$ other than Γ itself. (For, every vicinity of a point on E' contains antecedents of all other points on E' due to Theorem II.) Hence D_0 must contain at least two branches of $\Phi_{-n}(z) \cap D_0$ divided by a closed curve on which a pair of critical points of Φ_n exist. Conversely, if D_0 contains such branches of $\Phi_{-n}(z) \cap D_0$, then $\Phi_{-n}(\Gamma) \cap \partial D_0$ has at least one branch other than Γ itself so that D_0 cannot be simply connected.

As was already pointed out by Julia, a D_0 that is not simply connected is infinitely multiply connected.

The next theorem can readily be derived from Theorem 5.3.

[Theorem 5.4] If D_0 is a figure of $It[\Phi_n]$, it is

(1) simply connected if there is only one unstable fixed point of Φ_n on ∂D_0 ;

(2) simply connected if there is only one critical point of Φ_n in D_0 ; and

(3) multiply connected if all the critical points of Φ_n are contained in D_0 .

5.2. Divergence centers on ∂D_n . On the other hand, from Theorem 5.1 and the definition of D_n, we have

[Theorem 5.5] ∂D_n contains at least one λ_n and no λ_m such that $m < n$.

Furthermore, since $\Phi_{-1}(z) \cap D_0(x)$ has two or more branches if x is either a stable or a multiple fixed point, there must be at least one λ_1 on $\partial D_0(x)$. Hence the following theorem is verified by mathematical induction.

[Theorem 5.6] If x is either a stable or a multiple fixed point, there exists on $\partial D_n(x)$ a divergence center of any order not less than n.

The following theorem can be proved in a similar way.

[Theorem 5.7] If $\rho(x)$ is a root of a cyclotomic equation, there exist on $\partial U_0(x)$ divergence centers of all orders.

These theorems are useful for numerically tracing the shapes of convergence regions by the method described in § 3.

[Theorem 5.8] If it is known that D_0 is a simply connected region which is a figure of $\mathrm{It}[\Phi_n]$ and that $\Phi_{-n}(z) \cap D_0$ has M branches, then ∂D_0 is a closed Jordan curve or not according as the number of fixed points of Φ_n on ∂D_0 equals $M - 1$ or not.

Proof: If D_0 is a simply connected domain, there exists a univalent analytic function g which maps D_0 onto the interior of the unit circle C [9]. Since $\Psi \overset{\text{def}}{=} g\Phi_n g_{-1}$ maps the interior of C onto itself and its inverse Ψ_{-1} has M branches on C, Ψ has $M - 1$ different fixed points on C. In order that ∂D_0 may be a closed Jordan curve, these fixed points of Ψ on C must correspond one-to-one to the fixed points of Φ_n on ∂D_0. Conversely, if ∂D_0 contains $M - 1$ different fixed points, it can be proved that the mapping $\partial D_0 \xrightarrow{g} C$ is a bijection and hence ∂D_0 must be a closed Jordan curve.

5.3. Polynomial case. If $\Phi(z)$ is a polynomial with $N = \deg \Phi \geqq 2$, the point at infinity is a stable fixed point of order N so that we have

$$(5.3) \qquad \partial \underset{\circ}{D}(\infty) = \partial \underset{\circ}{U}(\infty) = \partial U(\infty) = E'$$

and all branches of $\Phi_{-1}(z)$ are around ∞ which is a critical point of Φ. Therefore, applying Theorem 5.3 and Theorem 5.8 to $\underset{\circ}{D}(\infty)$, we have

(1) $\underset{\circ}{D}(\infty)$ is simply connected if and only if it contains no finite critical point of Φ ;

(2) E' is a closed Jordan curve if and only if no finite critical point is in $\underset{\circ}{D}(\infty)$ and Φ has $N - 1$ different unstable fixed points.

Example 5. In case $\Phi(z) = z^2 + p$ where p is a real constant, $\underset{\circ}{D}(\infty)$ is simply connected if and only if $0 \notin \underset{\circ}{D}(\infty)$, which is equivalent to $-2 \leqq p \leqq 1/4$ as is verified elementarily. On the other hand, Φ has only one unstable fixed point if $-3/4 < p \leqq 1/4$ and two unstable fixed points otherwise (cf. Example 4). So E' is a closed Jordan curve if and only if $-3/4 < p \leqq 1/4$. (See Fig. 1.)

Example 6. If $N \geqq 3$, it sometimes happens that a multiply connected convergence region coexists with a simply connected one. For instance, $\underset{\circ}{D}(\infty)$ of $It[z^3 - 3z + 3]$ is multiply connected because it contains a finite critical point -1, while another finite critical point 1 coincides with a stable fixed point and hence $\underset{\circ}{D}(1) = \underset{\circ}{U}(1)$ is simply connected. Moreover, $\partial \underset{\circ}{D}(1)$ is a closed Jordan curve because it contains only one single critical point of Φ.

Fixed points other than 1 and ∞ are properly unstable. Therefore E' consists of a countably infinite number of discrete closed Jordan curves which separate $U(1)$ from $U(\infty)$.

6. Newton's Method Applied to Polynomial Equations

Let $f(z)$ be a polynomial with N (≥ 2) different zeros and $\Phi(z)$ be its Newton transform. We have

$$N^* \overset{\text{def}}{=} \deg f \geq N = \deg \Phi \ . \tag{6.1}$$

Zeros of f are stable fixed points of Φ, while the point at infinity is the only unstable fixed point of Φ, with the dissipation factor

$$\rho(\infty) = N^*/(N^* - 1) \ . \tag{6.2}$$

Hence, the following properties of $It[\Phi]$ are readily derived from the theorems in the preceding section.

(1) Every D_0 is a simply connected infinite domain, while D_n $(n \geq 1)$ is a simply connected finite domain.

(2) ∂D_0 is a closed Jordan curve if and only if D_0 contains only one critical point of Φ which is a simple critical point.

(3) ∂D_n $(n \geq 0)$ contains divergence centers of all orders equal to or greater than n, and never contains those of orders less than n.

For simplicity, a D_0 which is bounded by a closed Jordan curve will be called a <u>normal</u> D_0. Since there are N different stable fixed points and no more than $2N - 2$ different critical points of Φ, the following two propositions are derived from this definition and (2).

(4) At least two $\underset{0}{D}$'s are normal.

(5) If one $\underset{0}{D}$ is complete, all the other $\underset{0}{D}$'s are normal.

On the other hand, it can be proved from (2) and Caratheodory's theorem that $\Phi_{-1}(z)$ has two branches on the boundary of a normal $\underset{0}{D}$, which implies the following fact.

(6) If $\underset{0}{D}$ is normal, then $\partial\underset{0}{D}$ contains only one $\underset{1}{\lambda}$ and 2^{n-1} $\underset{n}{\lambda}$'s $(n \geqq 2)$, the arrangement of $\{\underset{k}{\lambda} \mid 0 \leqq k \leqq n\}$ on $\partial\underset{0}{D}$ being the same as that on the unit circle in the case of $\mathrm{It}[z^2]$.

6.1. Center of gravity of roots. Now we define symbols $\underset{k}{\mu}$, m_i and G as follows.

$\underset{0}{\mu}$: a stable fixed point of Φ (i.e., a root of $f = 0$); (another index i is attached as the subscript if necessary).

$\underset{k}{\mu}$: a root of $\Phi_k(x) = \underset{0}{\mu} \neq \Phi_{k-1}(x)$ $(k \geqq 1)$.

m_i : the multiplicity of $\underset{0}{\mu}_i$ as a root of $f = 0$.

G : the center of gravity of the roots of $f = 0$, i.e.,

$$G = \sum_{i=1}^{N} m_i \underset{0}{\mu}_i / N^* . \tag{6.3}$$

Note that, if $|z - G|$ is sufficiently large, we have

$$\Phi(z) = G + \frac{z - G}{\rho(\infty)} + O\left(\frac{1}{z - G}\right) . \tag{6.4}$$

This implies that, if a figure of $\mathrm{It}[\Phi]$ contains a point in every neighborhood of ∞, it is a repetition of similar configurations which are arranged asymptotically like a geometrical progression with G as the center of similitude.

For example, if $\underset{0}{D}$ is normal and we define $X(R)$ as the exterior of the circle with center G and radius R, then

$\partial \underset{0}{D} \cap X(R)$ is approximated by

$$(6.5) \qquad \bigcup_{n=0}^{\infty}{}' \{\rho(\infty)^n (z - G) + G \mid z \in A\} ,$$

where

$$(6.6) \qquad A = \partial \underset{0}{D} \cap X(R) - \partial \underset{0}{D} \cap X[\rho(\infty)R]$$

usually consists of two half-closed finite arcs of $\partial \underset{0}{D}'$. Hence it is inferred that $\partial \underset{0}{D}$ is an analytic curve only when it coincides with a straight line which passes G. In the other (general) case, $\partial \underset{0}{D}$ must be a curve somewhat like a branch of hyperbola --- more exactly, $X(R) \cap \partial \underset{0}{D}$ has enumerable common tangents with a hyperbola-like analytic curve whose asymptotes are two straight lines passing G, and all the divergence centers on $\partial \underset{0}{D}$ are nodes of $\partial \underset{0}{D}$.

6.2. Quadratic case. If $N = 2$, both $\underset{0}{D}$'s are normal and complete (irrespective of the values of N^*) and hence

$$(6.7) \qquad \partial \underset{0}{D}(\underset{0}{\mu}_i) = \partial U(\underset{0}{\mu}_i) = E' \qquad (i = 1, 2)$$

is reduced to an infinite closed Jordan curve. It can easily be verified that $G = \underset{1}{\lambda}$ if $m_1 = m_2$ while $G \in \underset{0}{D}(\underset{0}{\mu}_1)$ if $m_1 > m_2$. Therefore, E' is a straight line passing G if $m_1 = m_2$ while it is a non-analytic curve as was described in § 6.1 if $m_1 \neq m_2$.

6.3. Main part of $\partial \underset{0}{D}$. If $M \geqq 3$, every neighborhood of an arbitrary point on $\partial \underset{0}{D}$ contains an infinite number of closed Jordan curves. We shall call each of these closed Jordan curves a Jordan component of $\partial \underset{0}{D}$. From Julia's Theorem II, the exterior (with respect to $\underset{0}{D}$) of every Jordan component contains at least one convergence region toward another stable fixed point. Hence the closed curve is an invariant figure of It[Φ] if and only if its exterior contains another domain of direct convergence, i.e., if and only if it is an infinite Jordan curve. Such a Jordan component of $\partial \mathbf{D}_0$ will be called a <u>main component</u> of $\partial \mathbf{D}_0$ and the union of all the main components of $\partial \mathbf{D}_0$ will be called the <u>main part</u> of $\partial \mathbf{D}_0$.

(7) Both a main component and the main part of $\partial \mathbf{D}_0$ are minimal figures of It[Φ], the main part being the kernel of the union of all the Jordan components of $\partial \mathbf{D}_0$.

Note that $\partial \mathbf{D}_0$ is a tree-like chain of infinitely many closed curves and the tops of its infinitely many "twigs" do not belong to any Jordan component.

On the other hand, the following propositions are proved by the same reasoning as was used in the proof of Theorem 5.8.

(8) If $M \geqq 3$, $\partial \mathbf{D}_0$ has $M - 1$ main components.

(9) Each of the main components of $\partial \mathbf{D}_0$ contains only one λ_1 and 2^{n-1} λ_n's ($n \geqq 2$), their arrangement being the same as that on the boundary of a normal $\mathbf{D}_0$.

Furthermore, the foregoing description on the shape of the boundary of a normal $\mathbf{D}_0$ applies also to the shape of each main component of $\partial \mathbf{D}_0$ in this case. Hence a normal $\mathbf{D}_0$ can be considered as a special $\mathbf{D}_0$ whose boundary consists of only one main component.

<u>6.4. Mechanical sketching.</u> Based upon the above properties of Newton's method for polynomial equations, we have

schemed a process of making a sketch of the convergence regions mechanically. First, we obtain numerically all the zeros of f, all the critical points of Φ and all the divergence centers of a few lowest orders. Then, we draw smooth curves each of which contains divergence centers of lowest orders on a main component of ∂D_0 and is asymptotic to a pair of half-straight lines starting from G. Let the union of all such curves be denoted by Γ_0 and $\Phi_{-n}(\Gamma_0)$ by Γ_n. Then Γ_n gives a sketch of $\{\partial D_k \mid 0 \leqq k \leqq n\}$ and, among the parts of Γ_n approximating ∂D_k's, those for lower values of k will approach their exact form as n increases --- theoretically, at least. In practice, we cannot draw too microscopic details of the convergence regions. We have omitted them in the process of numerical calculation through neglecting such divergence centers that are too close to those of lower orders.

If we calculate all μ_k's parallel to Γ_k and plot their locations (except those which are too close to Γ_k) with appropriate symbols, most remarkable convergence regions toward each zero of f will be discriminated from those toward other zeros. (See Figs. 2--5, where the convergence regions toward a fixed point identified by a numeral are distinguished by the same small numerals at μ's in them.)

6.5. Equations whose roots are all simple. If $N = N^*$, the critical points of Φ are zeros of ff'' and λ_1's are zeros of f'. On the other hand, $\Phi(z) = \mu_0 \neq z$ is equivalent to $[f(z)/(z - \mu_0)]' = 0$ if μ_0 is a simple zero of f.

Hence follows the following property of $It[\Phi]$.

(10) If all the zeros of f are simple, then

1° $\underset{0}{D}$ is normal if no zero of f'' is in it,

2° $\underset{0}{D}$ is complete if all the zeros of f'' are in it,

3° the gravity center of all the critical points of Φ is G,

4° the gravity center of all the $\underset{1}{\lambda}$'s is G, and

5° the gravity center of all the $\underset{1}{\mu}$'s contained in $U(\underset{0}{\mu}_i)$ is

(6.7) $$\sum_{j \neq i} \underset{0}{\mu}_j/(N - 1) \overset{\text{def}}{=} G_i .$$

We may restrict ourselves to the cases where $G = 0$ without loss in generality, since the other cases are obtained simply by parallel displacements therefrom.

Example 7. In case $f(z) = z^3 + pz + q$, where p and q are real constants satisfying $\Delta \equiv 4p^3 + 27q^2 \neq 0$, the origin coincides with G and the only zero of f''. So $\underset{0}{D}$ is complete or normal according as it contains 0 or not. Since all the coefficients of Φ is real, $\underset{0}{\mu}$ must be real if $\underset{0}{D}(\underset{0}{\mu}) \ni 0$, which happens only when $p\Delta > 0$. The boundary of the complete $\underset{0}{D}$ is E' so that each of its two main components must coincide with the boundary of a normal $\underset{0}{D}$.

If $p\Delta < 0$, then we have $p < 0$ so that there are two real λ's, one of which must be a common boundary point of two $\underset{0}{D}$'s toward conjugate complex zeros of f.

If $p\Delta = 0$, the origin is the only $\underset{1}{\lambda}$ which must be common to the boundaries of all the three $\underset{0}{D}$'s.

Thus in this case, there are four types of global convergence features of It[Φ] as shown in Fig. 2.

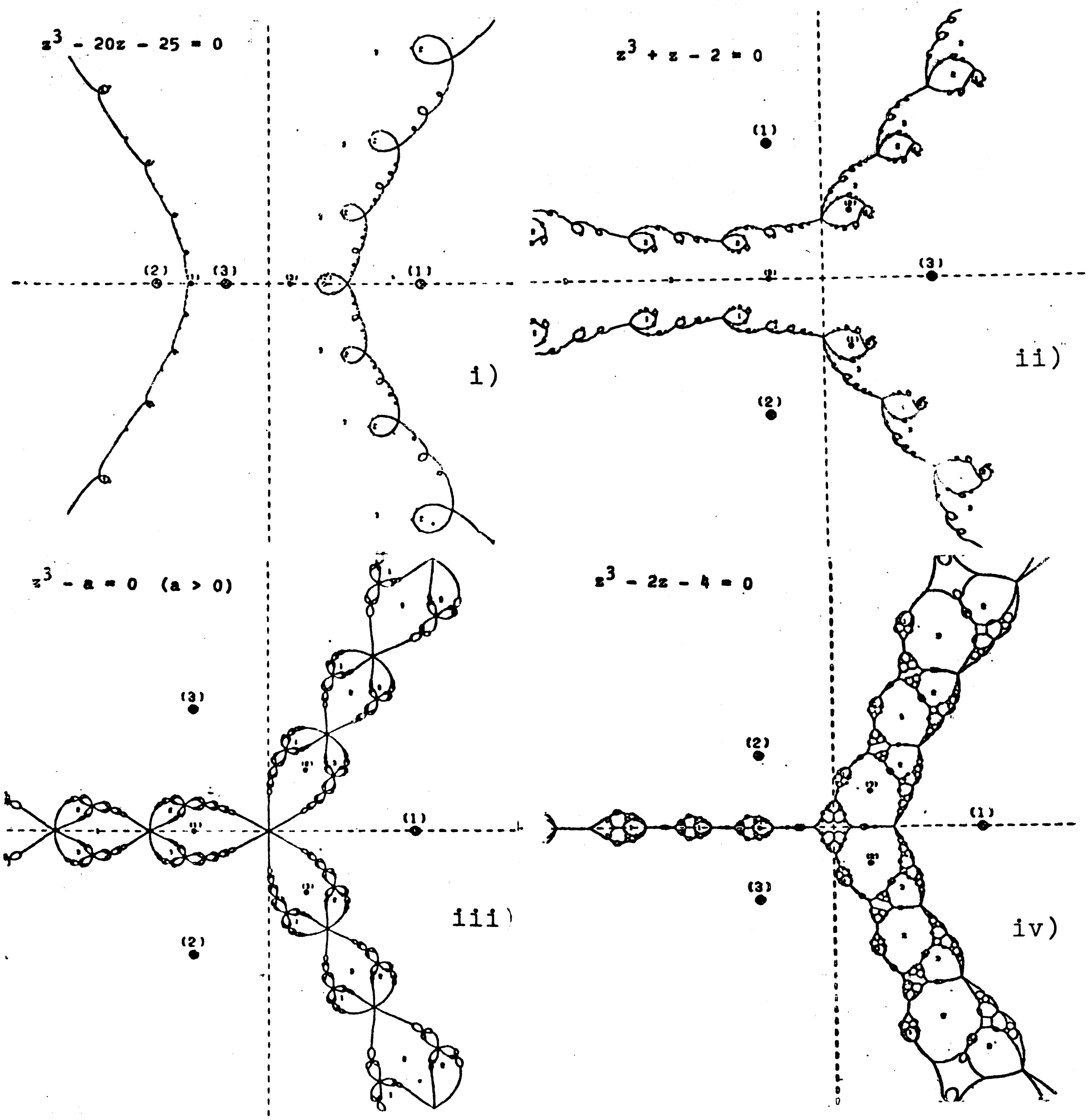


Fig. 2. Four types of global convergence features of Newton's method applied to the equation $z^3 + pz + q = 0$, where p and q are real and satisfy $\Delta \equiv 4p^3 + 27q^2 \neq 0$: i) $p < 0$, $\Delta < 0$; ii) $p > 0$, $\Delta > 0$; iii) $p = 0$, $\Delta > 0$; iv) $p < 0$, $\Delta > 0$.

Example 8. In case $f(z) = z^N - a$ $(a \neq 0, N \geqq 3)$, the origin is the only zero of f'' and the only λ_1 at the same time. Hence all the D_0's are normal and, from Theorem 5.5, the origin is common to the boundaries of all the D_0's and all the D_1's. Every U has one D_1 if $N = 3$, and $N - 2$ D_1's if $N \geqq 4$. Every finite divergence center is therefore common to the boundaries of

$$N + N(N - 2) = N(N - 1)$$

convergence regions. (See Fig. 2 iii) and Fig. 3.)

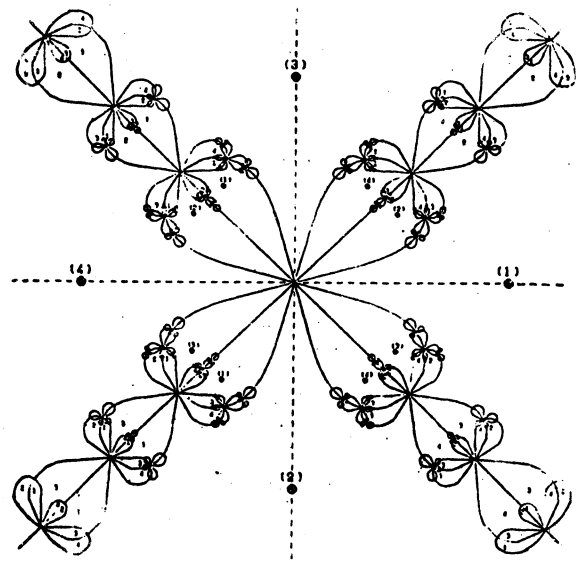
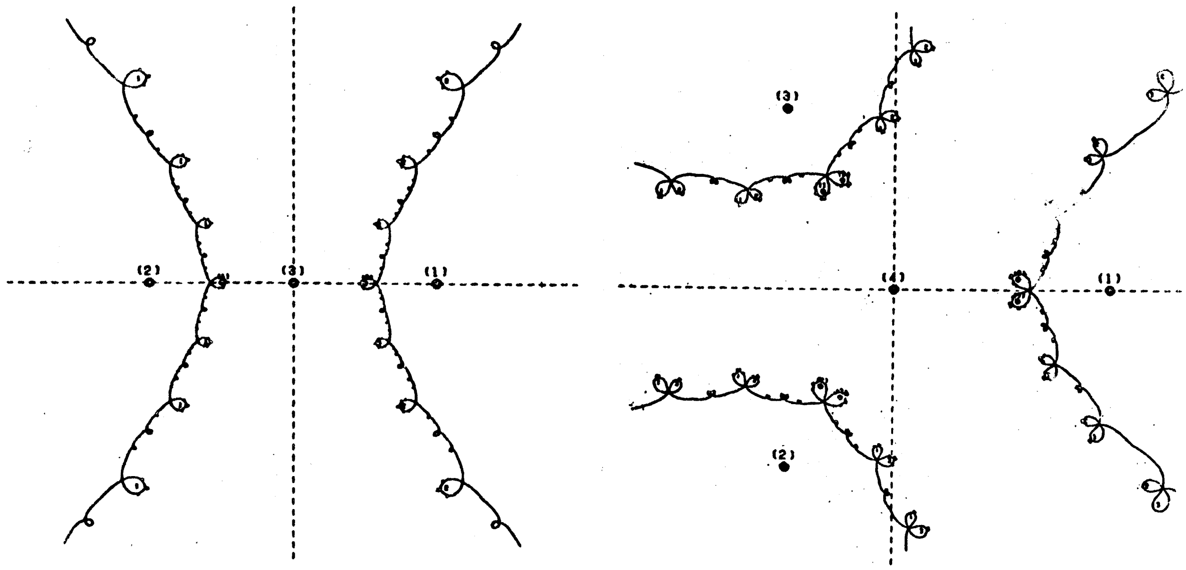


Fig. 3

Convergence regions of Newton's method applied to $z^4 - a = 0$, where a is a positive real constant.

Example 9. In case $f(z) = z^N - az$ $(a \neq 0, N \geqq 3)$, the origin is the only zero of f'' and a root of $f = 0$ at the same time. Therefore, $D_0(0)$ is complete and the other D_0's are all normal. From (8), the main part of $\partial D_0(0)$ consists of $N - 1$ components, each of which must coincide with the boundary of another D_0 because $\partial D_0(0) = E'$. The total number of D_1's being $(N - 1)(N - 2)$, every λ_1 is common to the boundaries of two D_0's and $N - 2$ D_1's. Therefore, every finite divergence center is common to $N - 1$ Jordan components of $\partial D_0(0) = E'$. (See Fig. 4.) Note that, in this case, Φ is equivalent to a polynomial.



i) ii)

Fig. 4. Convergence regions of Newton's method applied to i) $z^3 - az = 0$ and ii) $z^4 - az = 0$ $(a > 0)$.

<u>Example 10.</u> Let us take up the case where f has N real zeros $\underset{0}{\mu}_i$ $(i = 1, \cdots, N)$ where $N \geqq 3$ and $\underset{0}{\mu}_i < \underset{0}{\mu}_{i+1}$. For simplicity, $\underset{0}{D}(\underset{0}{\mu}_i)$ will be meant by $\underset{0}{D}_i$. It is known that $\underset{0}{D}_i$ contains only one zero of f'' if $2 \leqq i \leqq N - 1$. (Cf. [7],[8].) Therefore, $M = 2$ in $\underset{0}{D}_1$ and $\underset{0}{D}_N$, whereas $M = 3$ in the other $\underset{0}{D}$'s. In other words, $\underset{0}{D}_1$ and $\underset{0}{D}_N$ are normal, whereas the boundary of each of the other $\underset{0}{D}$'s has two main components. Each of the main components crosses the real axis at $\underset{1}{\lambda}$ which is a simple zero of f', wherefrom it can be proved that $\partial\underset{0}{D}_i$ shares a main component with $\partial\underset{0}{D}_{i+1}$. Every divergence center on $\partial\underset{0}{D}_i$ $(2 \leqq i \leqq N - 1)$ is common to a pair of Jordan components of $\partial\underset{0}{D}_i$. (See Fig. 2 i), Fig. 4 i) and Fig. 5.)

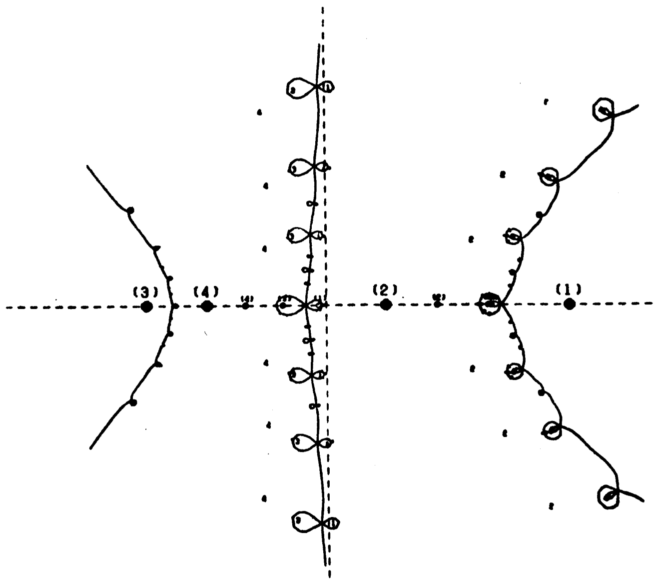


Fig. 5. Convergence regions of Newton's method applied to $z^4 - 15z^2 - 10z + 24 = 0$ whose roots are all real.

A C K N O W L E D G E M E N T S

This study started from preliminary experimental research about the convergence regions of Newton's method applied to cubic equations, which was performed by Mr. Toshiyuki Iida, Mr. Masato takeichi and the author as a generalization of Rall's "convergence map" [12] and was reported at the meeting of the Union of the Japanese Scientists and Engineers (Nikka-giren) two years ago. The author would like to express his sincere thanks to Prof. Sigeiti Moriguti and Prof. Masao Iri for many theoretical suggestions and continuing encouragement. He is also indebted to Prof. Toshio Uno, Dr. Tsutomu Date and Mr. Takahide Ogawa whose comments on several theoretical issues were inspiring to the author.

Numerical calculations of the examples contained in this

paper were conducted by the HITAC 8800-8700 computer system in the Computer Centre of the University of Tokyo, and the mechanical plotting of convergence regions was made with the flat-bed type X-Y plotter connected to the system.

REFERENCES

[1] G. Königs, Sur les intégrals de certaines équations fonctionelles. Ann. de l'Ecole Normale Supérieure, Suppl. n° 2, nos 24, 25 (1884), pp. 3 - 41.

[2] M. Fatou, Sur les solutions uniformes de certaines équations fonctionelles. C. R. Acad. Sci. Paris, 15 Oct. (1906), pp. 546 - 548.

[3] M. Fatou, Sur les substitutions rationelles. C. R. Acad. Sci. Paris, 21 Mai (1917), pp. 806 - 808.

[4] G. Julia, Memoires sur l'itération des fonctions rationelles. J. de Mathématique, 7e série, Tome IV, n° 1 (1918), pp. 2 - 245.

[5] J. F. Ritt, On the Iteration of Rational Functions. Trans. Amer. Math. Soc., 21 (1920), pp. 348 - 356.

[6] N. Bourbaki, Les structures fondamentales de l'analyse III, Topologie générale. Actualité Sci. Ind., Hermann & Cie, Paris (1948).

[7] A. S. Householder, Principles of Numerical Analysis. McGraw-Hill, New York (1953).

[8] S. Gorn, Maximal Convergence Interval and a Gibbs-Type Phenomenon for Newton's Approximation Procedure. Ann. of Math., 59-3 (1954), pp. 463 - 491.

[9] J. A. Jenkins, Univalent Functions and Conformal Mappings. Springer, New York (1958).

[10] Y. Yoshida, Tenshûgôron Nyûmon (Introduction to the Theory of Point-Sets). Baifûkan, Tokyo (1960).

[11] Y. Yoshida, Kansûron (Functions of a Complex Variable). Iwanami Shoten, Tokyo (1965).

[12] B. Rall, Computational Solution of Nonlinear Operational Equations. John Wiley & Sons, New York (1969).

Kohei Sato
Department of Mathematical Engineering
and Instrumentation Physics
Faculty of Engineering
University of Tokyo
Tokyo, Japan

On Approximate Solutions of Nonlinear Volterra Integrodifferential Equations in Chebyshev Series

Mitsunobu Kurihara

(University of Yamanashi, Japan)

0. Introduction

The present paper is concerned with the multi-point boundary value problem for nonlinear Volterra integrodifferential equations. Let

$$(0.\ 1)\qquad \frac{d}{dt}x(t)=f(t,x(t),\int_{-1}^{t}g(t,s,x(s))ds)$$

be a given system of nonlinear integrodifferential equations subject to a given multi-point boundary condition

$$(0.\ 2)\qquad \sum_{i=0}^{N}L_i x(t_i)=\ell.$$

Here $x(t)$ is the unknown vector valued function of t defined on the interval $J=\{t\,|\,-1\leqq t\leqq 1\}$. Denote that $S=\{(t,s)\,|\,-1\leqq s\leqq t\leqq 1\}$. Assume that $f(t,x,y)$ and $g(t,s,x)$ are vector valued twice continuously differentiable functions of (t,x,y) and (t,s,x) on the domains $J\times D\times E$ and $S\times D$ respectively, where D and E are bounded open domains in the Euclidean space with the Euclidean norm $\|\ \|$ which satisfy the relation

$$(0.\ 3)\qquad \{x\,|\,\|x\|\leqq 2\max\{\|g(t,s,x)\|\,|\,(t,s,x)\in S\times D\}\}\subset E.$$

Moreover assume that $L_i\,(i=0,1,\cdots\cdots,N)$ are square matrices, ℓ is a vector and $t_i\,(i=0,1,\cdots\cdots,N)$ are given points belonging to the interval J such that

$$-1=t_0<t_1<\cdots\cdots<t_N=1.$$

It is clear that our boundary value problem includes as the special cases Cauchy problem, two-point boundary value problem and Hukuhara's problem.

The system of linear integrodifferential equations which corresponds to the system (0. 1) is written in the form

$$(0.\ 4) \qquad \frac{d}{dt}x(t)=A(t)x(t)+\int_{-1}^{t} B(t,s)x(s)ds+h(t),$$

where A(t) and B(t,s) are square matrix valued continuous functions of t on the interval J and of (t,s) on the domain S respectively and h(t) is a vector valued continuous function of t on the interval J.

In the present paper we shall construct the Green function of the multi-point boundary value problem (0. 4) and (0. 2). The Green function plays an impotant role in studying approximate solutions of the nonlinear integrodifferential equations (0. 1) and the boundary condition (0. 2). We shall prove three main theorems on approximate solutions of the boundary value problem (0. 1) and (0. 2). The first theorem, Theorem 4.1, says that for any isolated solution there exists an approximate solution accurately as it is desired by computing finite Chebyshev polinomial series. The second theorem, Theorem 4.2, says that the obtained Chebyshev approximate solution corresponds one to one to the isolated solution. The third theorem, Theorem 4.3, says that one can always assure the existence of an exact solution by checking several conditions on the obtained Chebyshev approximate solution and further it gives a method to obtain an error bound of the obtained approximate solution.

The analogous theorems were originally proved by M. Urabe [5],

[6] concerning the multi-point boundary value problem for nonlinear differential equations. Later concerning nonlinear integral equations, the theorem analogous to Theorem 4.3 was proved by M. Shimasaki and K. Kiyono [2] and the theorems analogous to Theorem 4.1 and Theorem 4.2 were proved by the author [7] especially for nonlinear integral equations of Fredholm type. On the other hand K. Tsuruta and K. Ohmori [3] proved the theorem analogous to Theorem 4.3 concerning the Cauchy problem for nonlinear integro-differential equations and gave some examples of the numerical solutions with their aposteori error bounds.

Throughout the present paper we denote Euclidean norms for vectors and matrices by the symbol $\| \ \|$. Moreover for any vector valued function $x=x(t)$ continuous of t on the interval J we use two kinds of norms $\|x\|_C$ and $\|x\|_Q$ which are defined as follows

$$(0.\ 5)\qquad \|x\|_C=\max\{\|x(t)\| \mid t\in J\}$$

and

$$(0.\ 6)\qquad \|x\|_Q=[\frac{1}{\pi}\int_{-1}^{1}\|x(t)\|^2(1-t^2)^{-1/2}dt]^{1/2}.$$

In order to prove Theorem 4.1, we use the following lemma proved by M. Urabe [4] based on Newton-Raphson's procedure for nonlinear algebraic equations.

Lemma 0.1 Let

$$(0.\ 7)\qquad F(\alpha)=0$$

be a given real system of equations, where α and $F(\alpha)$ are voctors of the same dimension and $F(\alpha)$ is a continuously differentiable function of α defined in some region Ω of the α-space. Assume that the system (0. 7) has an approximate solution $\alpha=\hat{\alpha}$ for which the determinant of the Jacobian matrix $J(\alpha)$ of $F(\alpha)$ with respect to α does not vanish and that there are a positive constant δ and a nonnegative constant $\kappa<1$ such that

$$\text{(0. 8)} \qquad \begin{aligned} &(1) \quad \Omega_\delta=\{\alpha \mid \|\alpha-\hat{\alpha}\|\leqq\delta\}\subset\Omega, \\ &(2) \quad \|J(\alpha)-J(\hat{\alpha})\|\leqq\kappa/M \quad \text{for any } \alpha\in\Omega_\delta, \\ &(3) \quad Mr/(1-\kappa)\leqq\delta, \end{aligned}$$

where r and M are numbers such that

$$\text{(0. 9)} \qquad \|F(\hat{\alpha})\|\leqq r \qquad \text{and} \qquad \|J^{-1}(\hat{\alpha})\|\leqq M.$$

Then the system (0. 7) has one and only one solution $\alpha=\bar{\alpha}$ in Ω_δ and for $\alpha=\bar{\alpha}$ it holds that

$$\text{(0.10)} \qquad \det J(\bar{\alpha})\neq 0 \qquad \text{and} \qquad \|\alpha-\bar{\alpha}\|\leqq Mr/(1-\kappa).$$

1. Some Properties of Chebyshev Series

Denote by $T_n(t)$ Chebyshev polynomial of degree n, that is,

$$\text{(1. 1)} \qquad T_n(t)=\cos n\theta \qquad \text{as } t=\cos\theta$$

for $n=0,1,2,\cdots\cdots$. Then it is well known that for any continuous function $x(t)$ of t on the interval J we have Chebyshev polynomial series expansion of the form

$$\text{(1. 2)} \qquad x(t)\simeq\sum_{n=0}^{\infty} e_n a_n T_n(t),$$

where

(1. 3) $\qquad e_0=1 \qquad$ and $\qquad e_n=\sqrt{2}$ for $n=1,2,\cdots\cdots$

and

(1. 4) $$a_n=\frac{1}{\pi}e_n\int_{-1}^{1}x(t)T_n(t)(1-t^2)^{-1/2}dt.$$

For the expansion (1. 2) we obtain the Parseval's equality of the form

(1. 5) $$\|x\|_Q^2=\sum_{n=0}^{\infty}\|a_n\|^2.$$

In particular for finite Chebyshev series of the form

(1. 6) $$x_m(t)=\sum_{n=0}^{m}e_na_nT_n(t)$$

we have

(1. 7) $\qquad \|x_m\|_Q=\|\alpha\| \qquad$ and $\qquad \|x_m\|_C\leqq\sqrt{2m+1}\|\alpha\|,$

where $\alpha=(a_0,a_1,\cdots\cdots,a_m)$. In fact the inequalities (1. 7) are proved making use of the Parseval's equality (1. 5) and Schwarz's inequality.

Suppose that $x(t)$ is a continuously differetiable function of t on the interval J. Let the Chebyshev polynomial series of the derivative of $x(t)$ be

$$\frac{d}{dt}x(t)\simeq\sum_{n=0}^{\infty}e_na_n'T_n(t).$$

Then it is easily proved that

(1. 8) $\qquad (\sqrt{2}/e_{n-1})a_{n-1}'-a_{n+1}'=2na_n \qquad (n=1,2,\cdots\cdots),$

where a_n $(n=1,2,\cdots\cdots)$ are the coefficients of the expansion (1. 2) of $x(t)$. Moreover making use of the relation (1. 8), we obtain

$$(1.\ 9)\qquad a_n'=\sqrt{2}e_n\sum_{p=1}^{\infty}(n+2p-1)a_{n+2p-1}.$$

We define the operator P_m which expresses the truncation of the Chebyshev polynomial series (1. 2) of the operand discarding the terms of the order higher that m, that is, for any continuous function $x(t)$ of t on the interval J expanded in the form (1. 2)

$$(1.10)\qquad (P_m x)(t)=\sum_{n=0}^{m}e_n a_n T_n(t).$$

If $x=x(t)$ is a continuously differentiable function of t on the interval J, it is proved that for $m=0,1,\cdots\cdots$

$$(1.11)\qquad \|(I-P_m)x\|_C\leq\sigma(m)\|(I-P_{m-1})\frac{dx}{dt}\|_Q\leq\sigma(m)\|\frac{dx}{dt}\|_Q,$$

$$(1.12)\qquad \|(I-P_m)x\|_Q\leq\sigma_1(m)\|(I-P_{m-1})\frac{dx}{dt}\|_Q\leq\sigma_1(m)\|\frac{dx}{dt}\|_Q,$$

$$(1.13)\qquad \|\frac{d}{dt}(I-P_m)x\|_C\leq(m+2)\|(I-P_{m-1})\frac{dx}{dt}\|_Q+\|(I-P_{m+1})\frac{dx}{dt}\|_Q$$

and

$$(1.14)\qquad \|\frac{d}{dt}(I-P_m)x\|_Q\leq\sqrt{(m+2)/2}\|(I-P_{m-1})\frac{dx}{dt}\|_Q+\|(I-P_{m+1})\frac{dx}{dt}\|_Q.$$

Here and hereafter I is the identity operator, $P_{-1}=0$ and $\sigma(m)$ and $\sigma_1(m)$ are monotone decreasing functions of m satisfying

$$(1.15)\qquad \sqrt{2}/(m+1)<\sigma(m)<\sqrt{2/m}\ \text{and}\ \sigma_1(m)=1/(m+1).$$

These properties of Chebyshev polynomial series are proved in detail in the paper by M. Urabe [6].

2. Linear Integro-Differential Equations

In the present section we study a system of linear integro-differential equations (0. 4). We put

$$(2.\ 1)\qquad Q(t,s)=A(t)+\int_s^t B(t,s)ds$$

and consider the matrix equation

$$(2.\ 2)\qquad R(t,s)=I+\int_s^t R(t,u)Q(u,s)du$$

on the domain S. It is well known that there exists uniquely the continuous function R(t,s) satisfying the equation (2. 2). R(t,s) is also the unique solution of the adjoint equation

$$(2.\ 3)\qquad \frac{\partial}{\partial s}R(t,s)=-R(t,s)A(s)-\int_s^t R(t,u)B(u,s)du$$

satisfying the condition

(2. 4) $R(t,t)=I$ (identity matrix)

on the domain S. The function R(t,s) is called Resolvent matrix with respect to the matrices A(t) and B(t,s).

We introduce the following lemma proved by Tsuruta K. and K. Ohmori [3].

Lemma 2.1 The resolvent matrix R(t,s), which is the unique solution of the equation (2. 3) satisfying the condition (2. 4) on the domain S, is differentiable with respect to t and satisfies the equation

$$(2.\ 5)\qquad \frac{\partial}{\partial t}R(t,s)=A(t)R(t,s)+\int_s^t B(t,u)R(u,s)du.$$

Let us consider the Cauchy condition

(2. 6) $x(-1)=x_0$.

Then we shall prove the following lemma.

Lemma 2.2 For any vector x_0 and any continuous function $h(t)$ the system (0. 4) subject to the condition (2. 6) is equivalent to the system

(2. 7) $$x(t)=R(t,-1)x_0+\int_{-1}^{t}R(t,s)h(s)ds$$

on the interval J.

In fact, similarly to the proof by Grossman S. I. and R. K. Miller [1], for any solution $x=x(t)$ of the system (0. 4) satisfying the condition (2. 6) we have

$$\int_{-1}^{t}\{R(t,s)\frac{d}{ds}x(s)+[\frac{\partial}{\partial s}R(t,s)]x(s)\}ds$$

$$=[R(t,s)x(s)]_{s=-1}^{s=t}=x(t)-R(t,-1)x_0.$$

Then it follows from the equation (2. 3) that

$$x(t)-R(t,-1)x_0-\int_{-1}^{t}R(t,s)h(s)ds$$

$$=\int_{-1}^{t}R(t,s)[A(s)x(s)+\int_{-1}^{s}B(s,u)x(u)du+h(s)]ds$$

$$+\int_{-1}^{t}[\frac{\partial}{\partial s}R(t,s)]x(s)ds-\int_{-1}^{t}R(t,s)h(s)ds$$

$$=\int_{-1}^{t}[R(t,s)A(s)+\int_{s}^{t}R(t,u)B(u,s)du+\frac{\partial}{\partial s}R(t,s)]x(s)ds$$

$$=0.$$

Conversely if x(t) solves the system (2. 7), it follows from the equation (2. 5) and the condition (2. 4) that

$$\frac{d}{dt}x(t)-A(t)x(t)-\int_{-1}^{t}B(t,u)x(u)du-h(t)$$

$$=\frac{\partial}{\partial t}R(t,-1)x_0+\int_{-1}^{t}\frac{\partial}{\partial t}R(t,s)h(s)ds+h(t)$$

$$-A(t)[R(t,-1)x_0+\int_{-1}^{t}R(t,s)h(s)ds]$$

$$-\int_{-1}^{t}B(t,u)[R(u,-1)x_0+\int_{-1}^{u}R(u,s)h(s)ds]du-h(t)$$

$$=[\frac{\partial}{\partial t}R(t,-1)-A(t)R(t,-1)-\int_{-1}^{t}B(t,u)R(u,-1)du]x_0$$

$$+\int_{-1}^{t}[\frac{\partial}{\partial t}R(t,s)-A(t)R(t,s)-\int_{s}^{t}B(t,u)R(u,s)du]h(s)ds$$

$$=0.$$

It is clear that x(t) satisfies the condition (2. 6). This completes the proof of Lemma 2.2.

We shall construct the Green function H(t,s) for the system (0. 4) and the multi-point boundary value condition (0. 2).

Lemma 2.3 Let R(t,s) be the resolvent matrix with respect to the matrices A(t) and B(t,s). If the matrix

$$(2.\ 8)\qquad G=\sum_{i=0}^{N}L_iR(t_i,-1)$$

is nonsingular, then for any constant vector ℓ and any continuous function h(t) the system (0. 4) subject to the condition (0. 2) is equivalent to the system

(2. 9) $$x(t)=R(t,-1)G^{-1}\ell+\int_{-1}^{1}H(t,s)h(s)ds,$$

where for $t_{k-1}\leqq t<t_k$ $(k=1,2,\cdots\cdots,N)$

(2.10) $$H(t,s)=\begin{cases} -R(t,-1)G^{-1}\sum_{i=p}^{N}L_iR(t_i,s)+R(t,s); & t_{p-1}\leqq s<t_p \\ & (p=1,2,\cdots\cdots,k-1) \\ -R(t,-1)G^{-1}\sum_{i=k}^{N}L_iR(t_i,s)+R(t,s); & t_{k-1}\leqq s<t \\ -R(t,-1)G^{-1}\sum_{i=k}^{N}L_iR(t_i,s) & ;t\leqq s<t_k \\ -R(t,-1)G^{-1}\sum_{i=p}^{N}L_iR(t_i,s) & ;t_{p-1}\leqq s<t_p \\ & (p=k+1,\cdots\cdots,N). \end{cases}$$

In fact, by Lemma 2.2, for any constant vector x_0 any solution of the system (0. 4) satisfying the condition (2. 6) is expressed in the form (2. 7). The solution (2. 7) satisfies the boundary condition (0. 2) if and only if

$$\ell=\sum_{i=0}^{N}L_ix(t_i)=Gx_0+\sum_{i=0}^{N}L_i\int_{-1}^{t_i}R(t_i,s)h(s)ds.$$

Since G is nonsingular by the assumption, we obtain

(2.11) $$x_0=G^{-1}\ell-G^{-1}\sum_{i=0}^{N}L_i\int_{-1}^{t_i}R(t_i,s)h(s)ds.$$

Substituting (2.11) into (2. 7), we have the desired equality (2. 9) with the function (2.10). This completes the proof of Lemma 2.3.

The function H(t,s) in (2.10) is called Green function for the multi-point boundary value problem (0. 4) and (0. 2) or Green function with respect to the matrices $A(t)$, $B(t,s)$ and L_i $(i=0,1,\cdots\cdots,N)$. If we put

$$(2.12) \qquad x(t)=\int_{-1}^{1} H(t,s)h(s)ds,$$

then $x(t)$ satisfies the equations (0. 4) and the homogeneous boundary condition

$$\sum_{i=0}^{N} L_i x(t_i)=0$$

by Lemma 2.3. Therefore the expression (2.12) defines a bounded linear mapping in the normed space $C(J)$ which is defined to consist of all continuous vector valued functions of t on the interval J. For brevity we express (2.12) in the form

$$(2.13) \qquad x=Hh, \qquad \text{where } H;C(J)\to C(J).$$

Hereafter the mapping H defined in (2.12) or (2.13) is called H-mapping with respect to the matrices $A(t)$, $B(t,s)$ and L_i $(i=0,1,\cdots\cdots,N)$. It it noted by Lemma 2.3 that the H-mapping can be always defined so far as the matrix (2. 8) is nonsingular for the resolvent matrix $R(t,s)$ with respect to $A(t)$ and $B(t,s)$ and the matrices L_i $(i=0,1,\cdots\cdots,N)$. The norms of the H-mapping are defined corresponding to the norms (0. 5) and (0. 6) of vector valued functions belonging to the space $C(J)$. Hence we have two kinds of norms $\|H\|_C$ and $\|H\|_Q$, which are defined in the usual ways in normed spaces.

3. Isolated Solutions

We return to our multi-point boundary value problem (0. 1) and (0. 2) under all assumptions written in the section 0. Denote by $C(J;D)$ the space of all continuous vector valued functions of t on the interval J lying in the domain D for any $t\in J$. Let

(3. 1) $\qquad f_x(t,x,y),\ f_y(t,x,y)$ and $g_x(t,s,x)$

be Jacobian matrices of the functions $f(t,x,y)$ and $g(t,s,x)$ with respect to the variables x, y and x respectively. We define for any function $x=x(t)\in C(J;D)$ the functions

$$\Phi(t;x)=f_x(t,x(t),\int_{-1}^{t} g(t,s,x(s))ds),$$

$$\Psi(t,s;x)=f_y(t,x(t),\int_{-1}^{t} g(t,s,x(s))ds)g_x(t,s,x(s)),$$

where we note that for any function $x=x(t)\in C(J;D)$ and for any $t\in J$ the vector $\int_{-1}^{t} g(t,s,x(s)ds\in E$ by the relation (0. 3).

Let $x=\hat{x}(t)$ be any solution of the system (0. 1) satisfying the condition (0. 2) lying in the domain D for any $t\in J$. The system of linear integrodifferential equations

(3. 3) $$\frac{d}{dt}y(t)=\Phi(t;\hat{x})y(t)+\int_{-1}^{t}\Psi(t,s;\hat{x})y(s)ds$$

is called the first variation equations of (0. 1) with respect to the solution $x=\hat{x}(t)$. When we denote by $\hat{R}(t,s)$ the resolvent matrix with respect to the matrices $\Phi(t;\hat{x})$ and $\Psi(t,s;\hat{x})$, the solution $x=\hat{x}(t)$ is called isolated solution of the boundary value problem

(0. 1) and (0. 2) if the matrix

$$\hat{G}=\sum_{i=0}^{N} L_i\hat{R}(t_i,-1) \tag{3. 4}$$

is nonsingular. The word "isolated" comes from the following fact.

Lemma 3.1 If the matrix (3. 4) is nonsingular, then, besides the solution $x=\hat{x}(t)$, there is no other solution of the boundary value problem (0. 1) and (0. 2) in a sufficiently small neighborhood of $x=\hat{x}(t)$.

The above lemma can be proved in the following way. Since $\hat{G}$ is nonsingular, by Lemma 2.3, there is the H-mapping $\hat{H}$ with respect to the matrices $\Phi(t;\hat{x})$, $\Psi(t,s;\hat{x})$ and $L_i(i=0,1,\cdots\cdots,N)$. Let us take a positive number ε so that

$$\varepsilon<1/3\|H\|_C. \tag{3. 5}$$

For such ε, by the uniform continuities of the derivatives of the functions $f(t,x,y)$ and $g(t,s,x)$ and by the definitions of $\Phi(t;\hat{x})$ and $\Psi(t,s;\hat{x})$ in (3. 2), there exists a positive constant δ such that

$$U=\{x\,|\,\|x-\hat{x}(t)\|\leqq\delta \text{ for some } t\in J\}\subset D \tag{3. 6}$$

and

$$\|\Phi(t;\hat{x}+z)-\Phi(t;\hat{x})\|<\varepsilon \text{ for any } t\in J \tag{3. 7}$$

and

$$\|\Psi(t,s,\hat{x}+z)-\Psi(t,s;\hat{x})\|<\varepsilon \text{ for any } (t,s)\in S \tag{3. 8}$$

for any $z=z(t)\in C(J;D)$ satisfying $\|z\|_C$.

Suppose that, besides $x=\hat{x}(t)$, there is a solution $x=x(t)$ of the boundary value problem (0. 1) and (0. 2) satisfying $\|x-\hat{x}\|_C\leqq\delta$.

Then if we put

(3.9) $\quad y(t)=x(t)-\hat{x}(t),$

we have

(3.10) $\quad \|y\|_C \leq \delta,$

(3.11) $\quad \sum_{i=0}^{N} L_i y(t_i)=0$

and

(3.12) $$\frac{d}{dt}y(t)=f(t,x(t),\int_{-1}^{t} g(t,s,x(s))ds) -f(t,\hat{x}(t),\int_{-1}^{t} g(t,s,\hat{x}(s))ds).$$

Using a mean value theorem, we rewrite (3.12) in the form

(3.13) $$\frac{d}{dt}y(t)=\int_0^1 [\Phi(t;\hat{x}+\theta y)y(t)+\int_{-1}^{t} \Psi(t,s;\hat{x}+\theta y)y(s)ds]d\theta$$

$$=\Phi(t;\hat{x})y(t)+\int_{-1}^{t} \Psi(t,s;\hat{x})y(s)ds+h(t),$$

where

$$h(t)=-\int_0^1 [\Phi(t;\hat{x})-\Phi(t;\hat{x}+\theta y)]y(t)d\theta$$

$$-\int_0^1\int_{-1}^{t} [\Psi(t,s;\hat{x})-\Psi(t,s;\hat{x}+\theta y)]y(s)dsd\theta.$$

Noting that for any $\theta\in[0,1]$ $\hat{x}+\theta y=\hat{x}+\theta(x-\hat{x})\in C(J;D)$, that is, $\|\theta y\|_C\leq\delta$, we have by (3.7) and (3.8)

$$\|h\|_C \leq \varepsilon\|y\|_C+2\varepsilon\|y\|_C=3\varepsilon\|y\|_C.$$

On the other hand, by Lemma 2.3, for the equations (3.13) and the boundary condition (3.11) we have

$$y(t)=\int_{-1}^{1}\hat{H}(t,s)h(s)ds.$$

Thus it follows that

$$\|y\|_C \leq \|\hat{H}\|_C \|h\|_C \leq 3\varepsilon \|\hat{H}\|_C \|y\|_C.$$

By the inequality (3. 5) this implies that

$$\|y\|_C=0,$$

that is,

$$y(t)=0 \qquad \text{for any } t\in J.$$

This completes the proof of Lemma 3.1.

4. Main Theorems

In order to obtain an approximate solution of the multi-point boundary value problem (0. 1) and (0. 2), we consider finite Chebyshev polynomial series with unknown coefficients $a_0, a_1, \cdots\cdots, a_m$, that are

$$(4.\ 1) \qquad x_m(t)=\sum_{n=0}^{m} e_n a_n T_n(t).$$

Here $T_n(t)$ $(n=0,1,\cdots\cdots)$ are Chebyshev polynomials defined in (1. 1) and e_n $(n=0,1,\cdots\cdots)$ are constants defined in (1. 3). It is reasonable to determine the m+1 coefficients $a_0, a_1, \cdots\cdots, a_m$ in the finite Chebyshev polynomial series (4. 1) so as to satisfy the conditions that $x_m(t)$ lies in D for any $t\in J$ and that

$$(4.\ 2) \qquad \frac{d}{dt}x_m(t)=P_{m-1}[f(t,x_m(t),\int_{-1}^{t} g(t,s,x_m(s))ds)]$$

and

(4. 3) $$\sum_{i=0}^{N} L_i x_m(t_i)=\ell,$$

where P_m is the truncation operator defined in (1.10). In what follows the finite Chebyshev polynomial series (4. 1) satisfying (4. 2) and (4. 3) is called Chebyshev approximation of order m. In the present paper we shall prove the following three main theorems.

Theorem 4.1 Suppose that there exists an isolated solution $x=\hat{x}(t)$ of the boundary value problem (0. 1) and (0. 2) lying in D for any $t\in J$. Then for some sufficiently large m_0 there exists a Chebyshev approximation $x=\bar{x}_m(t)$ of any order $m\geq m_0$ such that

$$\bar{x}_m(t)\to\hat{x}(t) \text{ and } \frac{d}{dt}\bar{x}_m(t)\to\frac{d}{dt}\hat{x}(t)$$

uniformly on the interval J as $m\to\infty$.

Theorem 4.2 The Chebyshev approximation $x=\bar{x}_m(t)$ stated in Theorem 4.1 is determined uniquely in a sufficiently small neighborhood of the solution $x=\hat{x}(t)$ provided that the order m of the Chebyshev approximation $x=\bar{x}_m(t)$ is sufficiently high.

Theorem 4.3 Assume that the boundary value problem (0. 1) and (0. 2) has an approximate solution $x=\bar{x}(t)$, for which there are a positive constant δ, a nonnegative constant $\kappa<1$ and the matrices A(t) and B(t,s) continuous of t on the interval J and of (t,s) on the domain S respectively such that

$$
(4.\ 4)\quad \begin{cases} (1) & G=\sum_{i=0}^{N} L_i R(t_i,-1) \text{ is nonsingular} \\ (2) & U=\{x \mid \|x-\bar{x}(t)\| \leqq \delta \text{ for some } t\in J\} \subset D \\ (3) & \|\Phi(t;x)-A(t)\| \leqq \kappa/2M_1 \text{ for any } t\in J \text{ and any } x\in C(J,U) \\ (4) & \|\Psi(t,s;x)-B(t,s)\| \leqq \kappa/4M_1 \text{ for any } (t,s)\in S \text{ and for any } x\in C(J,U) \\ (5) & (M_1\varepsilon+M_2 r)/(1-\kappa) \leqq \delta. \end{cases}
$$

Here $R(t,s)$ is the resolvent matrix with respect to the matrices $A(t)$ and $B(t,s)$. $\Phi(t;x)$ and $\Psi(t,s;x)$ are the functions defined in (3. 2) by use of the Jacobian matrices of the functions $f(t,x,y)$ and $g(t,s,x)$. M_1 and M_2 are constants satisfying

$$
(4.\ 5)\qquad \|H\|_C \leqq M_1 \text{ and } \|R(t,-1)G^{-1}\| \leqq M_2 \text{ for any } t\in J
$$

respectively, where H is the H-mapping with respect to $A(t)$, $B(t,s)$ and L_i $(i=0,1,\cdots\cdots,N)$. r and ε are constants satisfying the inequalities

$$
(4.\ 6)\qquad \left\|\frac{d}{dt}\bar{x}(t)-f\left(t,\bar{x}(t),\int_{-1}^{t} g(t,s,\bar{x}(s))ds\right)\right\| \leqq r \text{ for any } t\in J
$$

and

$$
(4.\ 7)\qquad \left\|\sum_{i=0}^{N} L_i\bar{x}(t_i)-\ell\right\| \leqq \varepsilon.
$$

Then there exists uniquely an exact solution $x=\hat{x}(t)$ of the boundary value problem (0. 1) and (0. 2) lying in U for any $t\in J$. Moreover this is an isolated solution and it holds that

$$
(4.\ 8)\qquad \|x(t)-\bar{x}(t)\| \leqq (M_1\varepsilon+M_2 r)/(1-\kappa) \text{ for any } t\in J.
$$

The proof of these theorems are given later in the present paper.

The coefficients $\alpha=(a_0,a_1,\cdots\cdots,a_m)$ of our desired Chebyshev approximations $x=x_m(t)$ in (4. 1) of order m are determined by the equations (4. 2) and (4. 3) if $x_m(t)$ lies in D for any $t\in J$. The equations (4. 2) and (4. 3) are equivalent to the system of nonlinear algebraic equations

$$F^{(m)}(\alpha)=(F_0(\alpha),F_1(\alpha),\cdots\cdots,F_m(\alpha))=0, \tag{4. 9}$$

where

$$F_0(\alpha)=\sum_{i=0}^{N} L_i x_m(t_i)-\ell \tag{4.10}$$

and

$$P_{m-1}[f(t,x_m(t),\int_{-1}^{t} g(t,s,x_m(s))ds]-\frac{d}{dt}x_m(t) \tag{4.11}$$

$$=\sum_{n=0}^{m-1} e_n F_{n+1}(\alpha)T_n(t),$$

which implies by the expressions (1. 4) and (1. 9)

$$F_n(\alpha)=\frac{1}{\pi}e_{n-1}\int_{-1}^{1}[f(t,x_m(t),\int_{-1}^{t} g(t,s,x_m(s))ds)T_{n-1}(t) \tag{4.12}$$

$$\times(1-t^2)^{-1/2}]dt-\sqrt{2}e_{n-1}\sum_{p=1}^{\infty}(n+2p-2)a_{n+2p-2}$$

for $n=1,2,\cdots\cdots,m$. The system (4. 9) is called determining equations of Chebyshev approximations. For a solution $\bar{\alpha}=(\bar{a}_0,\bar{a}_1,\cdots\cdots,\bar{a}_m)$ of the system (4. 9) the finite Chebyshev polynomial series

$$\bar{x}_m(t)=\sum_{n=0}^{m} e_n\bar{a}_n T_n(t)$$

is a Chebyshev approximation of order m.

5. Inequalities for Determining Equations

Suppose that $x=\hat{x}(t)$ is an isolated solution of the boundary value problem (0. 1) and (0. 2) lying in D for any $t\in J$. There exists a positive constant δ such that

(5. 1) $\quad U=\{x\,|\,\|x-\hat{x}(t)\|\leqq\delta$ for some $t\in J\}\subset D$.

Denote that $\hat{x}_m=P_m\hat{x}$. It follows from the inequalities (1.11)-(1.14) that

(5. 2) $\quad \|\hat{x}_m-\hat{x}\|_C\leqq K_1\sigma(m)/m(m-1),$

(5. 3) $\quad \|\hat{x}_m-\hat{x}\|_Q\leqq K_1/(m+1)m(m-1)$

(5. 4) $\quad \|\frac{d}{dt}\hat{x}_m-\frac{d}{dt}\hat{x}\|_C\leqq K_1[(m+2)/m(m-1)+\sigma(m+1)/(m+1)]$

(5. 5) $\quad \|\frac{d}{dt}\hat{x}_m-\frac{d}{dt}\hat{x}\|_Q\leqq K_1[\sqrt{m+2}/\sqrt{2}m(m-1)+1/(m+2)(m+1)],$

where K_1 is a constant satisfying

(5. 6) $\quad \|\frac{d^3}{dt^3}x\|_Q=\|\frac{d^2}{dt^2}f(t,\hat{x}(t),\int_{-1}^{t}g(t,s,\hat{x}(s))ds)\|_Q\leqq K_1.$

K_1 may depend only on the structure of the given system (0. 1). Hereafter we denote by K's the constants depending only on the structure of the given system (0. 1).

In order to determine a domain where the function $F^{(m)}(\alpha)$ of the system (4. 9) of determining equations is well defined, we choose the number m_1 sufficiently large such that

(5. 7) $\quad K_1\sigma(m)/m(m-1)<\delta \quad$ for any $m\geqq m_1$.

This is possible by the properties (1.15) of the constant $\sigma(m)$. Then it follows from the inequalities (5. 2) and (5. 7) that

$$\hat{x}_m(t)\in U\subset D \quad \text{for any } t\in J.$$

Let us put

$$(5.\ 8)\qquad \hat{x}_m(t)=(P_m\hat{x})(t)=\sum_{n=0}^{m} e_n\hat{a}_nT_n(t)$$

and

$$(5.\ 9)\qquad \hat{\alpha}=(\hat{a}_0,\hat{a}_1,\cdots\cdots,\hat{a}_m).$$

Define the domain

$$(5.10)\qquad \Omega_m=\{\alpha\,|\,\|\alpha-\hat{\alpha}\|\leqq[\delta-K_1\sigma(m)/m(m-1)]/\sqrt{2m+1}\}.$$

The domain Ω_m is the desired domain where the function $F^{(m)}(\alpha)$ is well defined. In fact, if we put for any vector $\alpha=(a_0,a_1,\cdots\cdots,a_m)$ belonging to the domain Ω_m

$$x_m(t)=\sum_{n=0}^{m} e_na_nT_n(t),$$

then we obtain

$$x_m(t)\in U\subset D \qquad \text{for any } t\in J$$

since for any $t\in J$

$$\|x_m(t)-\hat{x}(t)\|\leqq\|x_m(t)-\hat{x}_m(t)\|+\|\hat{x}_m(t)-\hat{x}(t)\|$$

$$\leqq\sqrt{2m+1}\,\|\alpha-\hat{\alpha}\|+\|\hat{x}_m-\hat{x}\|_C$$

$$\leqq\sqrt{2m+1}[\delta-K_1\sigma(m)/m(m-1)]/\sqrt{2m+1}+K_1\sigma(m)/m(m-1)=\delta$$

by the inequalities (5. 2) and (1. 7). Therefore it follows from the expression (4.10) and (4.12) that the function $F^{(m)}(\alpha)$ is continuously differentiable of α on the domain Ω_m.

Let $J_m(\alpha)$ be the Jacobian matrix of the function $F^{(m)}(\alpha)$. To investigate the properties of the matrix $J_m(\alpha)$, we consider a system of linear equations of the form

$$(5.11)\qquad J_m(\alpha)\xi+\gamma=0,$$

where $\quad \alpha=(a_0,a_1,\cdots\cdots,a_m)\in\Omega_m,$

$\xi=(u_0,u_1,\cdots\cdots,u_m)$ and $\gamma=(c_0,c_1,\cdots\cdots,c_m)$.

Let us put that

$$x_m(t)=\sum_{n=0}^{m} e_n a_n T_n(t)\in U\subset D,$$

$$y(t)=\sum_{n=0}^{m} e_n u_n T_n(t) \text{ and } h(t)=\sum_{n=0}^{m-1} e_n c_{n+1} T_n(t).$$

Then we can prove that the system (5.11) is equivalent to the boundary value problem

$$(5.12)\qquad \sum_{i=0}^{N} L_i y(t_i)=-c_0$$

and

$$(5.13)\qquad \frac{d}{dt}y(t)=P_{m-1}[\Phi(t;x_m)y(t)+\int_{-1}^{t}\Psi(t,s;x_m)y(s)ds]+h(t),$$

where $\Phi(t;x)$ and $\Psi(t,s;x)$ are the functions defined in (3. 1) and (3. 2).

Substituting $\hat{x}_m(t)$ into $x_m(t)$ in the equations (5.13), equivalently doing $\hat{\alpha}$ into α in the system (5.11), we obtain

$$(5.14)\qquad \frac{d}{dt}y(t)=\Phi(t;\hat{x})y(t)+\int_{-1}^{t}\Psi(t,s;\hat{x})y(s)ds+h(t)+p(t),$$

where

$$(5.15)\qquad p(t)=-(I-P_{m-1})[\Phi(t;\hat{x})y(t)+\int_{-1}^{t}\Psi(t,s;\hat{x})y(s)ds]$$

$$-P_{m-1}\{[\Phi(t;\hat{x})-\Phi(t;\hat{x}_m)]y(t)+\int_{-1}^{t}[\Psi(t,s;\hat{x})-\Psi(t,s;\hat{x}_m)]y(s)ds\}.$$

Let $\hat{R}(t,s)$ be the resolvent matrix with respect to the matrices $\Phi(t;\hat{x})$ and $\Psi(t,s;\hat{x})$. The matrix

$$\hat{G}=\sum_{i=0}^{N} L_i\hat{R}(t_i,-1)$$

is nonsingular from the isolatedness of the solution $x=\hat{x}(t)$. Then, applying Lemma 2.3 for the boundary value problem (5.12) and (5.14), we have

$$y(t)=-\hat{R}(t,-1)\hat{G}^{-1}c_0+\int_{-1}^{1}\hat{H}(t,s)[h(s)+p(s)]ds,$$

where $\hat{H}(t,s)$ is the Green function with respect to the matrices $\Phi(t;\hat{x})$, $\Psi(t,s;\hat{x})$ and $L_i(i=0,1,\cdots\cdots,N)$. Denote by H the H-mapping eith respect to the above matrices. This implies that

$$\|y\|_Q\leq M_1\|c_0\|+\|\hat{H}\|_Q(\|h\|_Q+\|p\|_Q), \tag{5.16}$$

where

$$M_1=\max\{\|\hat{R}(t,-1)\hat{G}^{-1}\|\,|\,t\in J\}.$$

On the other hand for the function (5.15) we have

$$\begin{aligned}\|p\|_Q\leq\sigma_1(m-1)\|\frac{d}{dt}[\Phi(t;\hat{x})y(t)+\int_{-1}^{t}\Psi(t,s;\hat{x})y(s)ds]\|_Q\\ +\|[\Phi(t;\hat{x})-\Phi(t;\hat{x}_m)]y(t)\\ +\int_{-1}^{t}[\Psi(t,s;\hat{x})-\Psi(t,s;\hat{x}_m)]y(s)ds\|_Q\end{aligned}$$

by the inequality (1.12) and the Parseval's equality (1. 5). Moreover using Schwartz's inequality and a mean value theorem, we obtain

$$\left\|\frac{d}{dt}[\Phi(t;\hat{x})y(t)+\int_{-1}^{t}\Psi(t,s;\hat{x})y(s)ds]\right\|_Q$$

$$\leqq\left\|\frac{d}{dt}\Phi(t;\hat{x})y(t)\right\|_Q$$

$$+\left\|\Phi(t;\hat{x})\{P_{m-1}[\Phi(t;\hat{x}_m)y(t)+\int_{-1}^{t}\Psi(t,s;\hat{x}_m)y(s)ds]+h(t)\}\right\|_Q$$

$$+\left\|\Psi(t,t;\hat{x})y(t)+\int_{-1}^{t}\frac{\partial}{\partial t}\Psi(t,s;\hat{x})y(s)ds\right\|_Q$$

$$\leqq K_2\|y\|_Q+K_3\|h\|_Q$$

and

$$\left\|[\Phi(t;\hat{x})-\Phi(t;\hat{x}_m)]y(t)+\int_{-1}^{t}[\Psi(t,s;\hat{x})-\Psi(t,s;\hat{x}_m)]y(s)ds\right\|_Q$$

$$\leqq K_4\|\hat{x}-\hat{x}_m\|_C\|y\|_Q.$$

Then for the inequality (5.16) we have

$$\|y\|_Q\leqq M_1\|c_0\|+\|\hat{H}\|_Q(1+\sigma_1(m-1)K_3)\|h\|_Q$$
$$+\|\hat{H}\|_Q[K_2\sigma_1(m-1)+K_4\sigma(m)/m(m-1)]\|y\|_Q$$

using the inequality (5. 2). If we choose a number $m_2\geqq m_1$ sufficiently large, then we have for any $m\geqq m_2$

$$\|y\|_Q\leqq\frac{\sqrt{M_1^2+\|\hat{H}\|_Q^2(1+\sigma_1(m-1)K_3)^2}\sqrt{\|c_0\|^2+\|h\|_Q^2}}{1-\|\hat{H}\|_Q[K_2\sigma_1(m-1)+K_4\sigma(m)/m(m-1)]} \tag{5.17}$$

$$\leqq M\sqrt{\|c_0\|^2+\|h\|_Q^2},$$

where M is a constant independent of m. The inequality (5.17) is equivalent to the inequality

(5.18) $\|\xi\| \leq M\|\gamma\|$

by the Parseval's eqiality (1. 7) for the finite Chebyshev polynomial series y(t) and h(t). From the inequality (5.18) and the relation (5.11) it readily follows that for any $m \geq m_2$

(5.19) $\det J_m(\hat{\alpha}) \neq 0$

and

(5.20) $\|J_m^{-1}(\hat{\alpha})\| \leq M.$

The inequality (5.20) will play an impotant role in the proof of Theorem 4.1.

Let

$$\alpha' = (a_0', a_1', \cdots\cdots, a_m') \text{ and } \alpha'' = (a_0'', a_1'', \cdots\cdots, a_m'')$$

be arbitrary vectors belonging to the domain Ω_m. For any vector $\xi = (u_0, u_1, \cdots\cdots, u_m)$ we consider the systems of linear equations

(5.21) $J_m(\alpha')\xi = \gamma'$ and $J_m(\alpha'')\xi = \gamma''$,

where

$$\gamma' = (c_0', c_1', \cdots\cdots, c_m') \text{ and } \gamma'' = (c_0'', c_1'', \cdots\cdots, c_m'').$$

Let us put

$$x_m'(t) = \sum_{n=0}^{m} e_n a_n' T_n(t), \quad x_m''(t) = \sum_{n=0}^{m} e_n a_n'' T_n(t),$$

$$y(t) = \sum_{n=0}^{m} e_n u_n T_n(t),$$

$$h'(t) = \sum_{n=0}^{m} e_n c_{n+1}' T_n(t) \text{ and } h''(t) = \sum_{n=0}^{m} e_n c_{n+1}'' T_n(t).$$

Then, corresponding to the systems (5.21), we have

$$(5.22)\quad \begin{cases} \sum_{i=0}^{N} L_i y(t_i) = -c_0' \\ \frac{d}{dt}y(t) = P_{m-1}[\Phi(t;x_m')y(t) + \int_{-1}^{t} \Psi(t,s;x_m')y(s)ds] + h'(t) \end{cases}$$

and

$$(5.23)\quad \begin{cases} \sum_{i=0}^{N} L_i y(t_i) = -c_0'' \\ \frac{d}{dt}y(t) = P_{m-1}[\Phi(t;x_m'')y(t) + \int_{-1}^{t} \Psi(t,s;x_m'')y(s)ds] + h''(t) \end{cases}$$

respectively by the definition of $J_m(\alpha)$. From (5.22) and (5.23) it readily follows that

$$(5.24)\qquad c_0' = c_0''$$

and

$$(5.25)\qquad h'(t) - h''(t) = -P_{m-1}\{[\Phi(t;x_m') - \Phi(t;x_m'')]y(t) + \int_{-1}^{t} [\Psi(t,s;x_m') - \Psi(t,s;x_m'')]y(s)ds\}.$$

The relation above implies that

$$\|h' - h''\|_Q \leqq K_5 \|x_m' - x_m''\|_C \|y\|_Q.$$

Then by the relation (1. 7) and (5.24) we obtain

$$\|\gamma' - \gamma''\| \leqq K_4 \sqrt{2m+1}\, \|\alpha' - \alpha''\|\, \|\xi\|,$$

which implies from the systems (5,21) that for any $m \geqq m_2$ and any α', $\alpha'' \in \Omega_m$

$$(5.26)\qquad \|J_m(\alpha') - J_m(\alpha'')\| \leqq K_4 \sqrt{2m+1}\, \|\alpha' - \alpha''\|.$$

The inequality (5.26) will also play an impotant role in the proof of Theorem 4.1.

6. Proof of Theorem 4.1

Suppose that there exists an isolated solution $x=\hat{x}(t)$ of the boundary value problem (0. 1) and (0. 2) lying in D for any $t\in J$. It is concluded that in the previous section 5 that there exists a positive constant δ satisfying the relation (5. 1) and a number m_2 sufficiently large such that for any $m\geqq m_2$ the inequalities (5. 2)-(5. 5) hold and such that the function $F^{(m)}(\alpha)$ of the determining equations (4. 9) is continuously differentiable of α in the domain Ω_m and its Jacobian matrix $J_m(\alpha)$ has the inverse $J_m^{-1}(\alpha)$ at $\alpha=\hat{\alpha}$ satisfying the inequality (5.20) and satisfies the inequality (5.26), where $\hat{x}_m$, $\hat{\alpha}$ and Ω_m are defined in the formula (5. 8), (5. 9) and (5.10) respectively.

Let us put

$$(6.\ 1)\qquad \sum_{i=0}^{N} L_i\hat{x}_m(t_i)-\ell=r_0$$

and

$$(6.\ 2)\qquad \frac{d}{dt}\hat{x}_m(t)-P_{m-1}[f(t,\hat{x}_m(t),\int_{-1}^{t} g(t,s,\hat{x}_m(s))ds]=h(t).$$

Using the inequality (5. 2), we have for the relation (6. 1)

$$(6.\ 3)\qquad \|r_0\|=\|\sum_{i=0}^{N} L_i\hat{x}_m(t_i)-\sum_{i=0}^{N} L_i\hat{x}(t_i)\|\leqq(\sum_{i=0}^{N}\|L_i\|)K_1\sigma(m)/m(m-1).$$

We rewrite the equation (6. 2) in the form

$$h(t)=\{\frac{d}{dt}\hat{x}_m(t)-\frac{d}{dt}\hat{x}(t)\}$$

$$+(I-P_{m-1})[f(t,\hat{x}(t),\int_{-1}^{t}g(t,s,\hat{x}(s))ds)]$$

$$-P_{m-1}[f(t,\hat{x}_m(t),\int_{-1}^{t}g(t,s,\hat{x}_m(s))ds)$$

$$-f(t,\hat{x}(t),\int_{-1}^{t}g(t,s,\hat{x}(s))ds)]$$

using the fact that $x=\hat{x}(t)$ is a solution of the equations (0. 1). Then by the inequalities (5. 5), (1.12), (1. 5) and (5. 3) and a mean value theorem used in (3.12) we obtain

(6. 4) $$\|h\|_Q \leqq K_1[\sqrt{m+2}/\sqrt{2}m(m-1)+1/(m+2)(m+1)]$$

$$+K_1/m(m-1)+K_6K_1/(m+1)m(m-1).$$

It follows from the inequalities (6. 3) and (6. 4) that

(6. 5) $$\sqrt{\|r_0\|^2+\|h\|_Q^2}=O(m^{-3/2}) \text{ as } m\to\infty.$$

By the definition of the function $F^{(m)}(\alpha)$ of the determining equations (4. 9) the boundary value problem (6. 1) and (6. 2) is equivalent to a system

$$F^{(m)}(\hat{\alpha})=\rho^{(m)}.$$

This implies that

$$\|\rho^{(m)}\|=\sqrt{\|r_0\|^2+\|h\|_Q^2}.$$

Then there exists s number $m_3 \geqq m_2$ such that for any $m \geqq m_3$

(6. 6) $\quad \|\rho^{(m)}\| \leqq K_7 m^{-3/2}$

for some constant K_7 by the asymptotic behavior (6. 5).

We shall apply Lemma 0.1 in the section 0 to the determining equations (4. 9) to complete the proof of Theorem 4.1. In order to check the conditions (0. 8) in Lemma 0.1 we choose an arbitrary nonnegative constant $\kappa<1$ and put

$$\delta_1 = \min\{\kappa/K_5 M,\ \delta - K_1\sigma(m_3)/m_3(m_3-1)\},$$

where constants K_5, M and K_1 are defined in (5.26), (5.20) and (5. 6) respectively. There exists a number $m_4 \geqq m_3$ so that

$$[M/(1-\kappa)]K_7 m^{-3/2} < \delta_1/\sqrt{2m+1}$$

for any $m \geqq m_4$ since

$$m^{-3/2}\sqrt{2m+1} = O(m^{-1/2}) \text{ as } m \to \infty.$$

If we choose a number δ_m such that

$$[M/(1-\kappa)]K_7 m^{-1/2} < \delta_m < \delta_1/\sqrt{2m+1}, \tag{6. 7}$$

then we obtain

$$\Omega_{\delta_m} = \{\alpha \mid \|\alpha - \hat{\alpha}\| \leqq \delta_m\} \subset \Omega_m. \tag{6. 8}$$

In fact, for any $\alpha \in \Omega_{\delta_m}$ and $m \geqq m_4 \geqq m_3$

$$\begin{aligned} \|\alpha - \hat{\alpha}\| &\leqq \delta_m < \delta_1/\sqrt{2m+1} \\ &\leqq [\delta - K_1\sigma(m_3)/m_3(m_3-1)]/\sqrt{2m+1} \\ &\leqq [\delta - K_1\sigma(m)/m(m-1)]/\sqrt{2m+1}, \end{aligned}$$

which implies $\alpha \in \Omega_m$

Moreover it follows from the inequality (5.26) that

(6. 9) $$\|J_m(\alpha)-J_m(\hat{\alpha})\| \leqq K_5\sqrt{2m+1}\|\alpha-\hat{\alpha}\|$$

$$\leqq K_5\sqrt{2m+1}\delta_m \leqq K_5\delta_1 \leqq K_5(\kappa/K_5 M)=\kappa/M$$

for any $\alpha\in\Omega_{\delta_m}$ and any $m \geqq m_4$. Finally by the inequalities (6. 6) and (6. 7) we have

(6.10) $$[M/(1-\kappa)]\|\rho^{(m)}\| \leqq [M/(1-\kappa)]K_7 m^{-3/2} \leqq \delta_m.$$

Thus the inequalities (6. 8), (6. 9) and (6.10) show that the conditions (0. 8) in Lemma 0.1 are fulfilled. Hence by Lemma 0.1 we see that the determining equations (4. 9) has one and only one solution $\alpha=\bar{\alpha}$ in the domain Ω_{δ_m} satisfying

$$\det J_m(\bar{\alpha})\neq 0$$

and

(6.11) $$\|\bar{\alpha}-\hat{\alpha}\| \leqq [M/(1-\kappa)]\|\rho^{(m)}\| \leqq [M/(1-\kappa)]K_7 m^{-3/2}.$$

If we put

$$\bar{\alpha}=(\bar{a}_0,\bar{a}_1,\cdots\cdots,\bar{a}_m)$$

and

$$\bar{x}_m(t)=\sum_{n=0}^{m} e_n\bar{a}_n T_n(t),$$

then $x=\bar{x}_m(t)$ is a Chebyshev approximation and satisfies

(6.12) $$\|\bar{x}_m-\hat{x}\|_C \leqq \|\bar{x}_m-\hat{x}_m\|_C+\|\hat{x}_m-\hat{x}\|_C \leqq \sqrt{2m+1}\|\bar{\alpha}-\hat{\alpha}\|+\|\hat{x}_m-\hat{x}\|_C$$

$$\leqq [M/(1-\kappa)]K_7 m^{-3/2}\sqrt{2m+1}+K_1\sigma(m)/m(m-1)$$

for any $m \geqq m_4$ from the inequalities (1. 7) and (6.11). The inequality (6.12) implies the uniform convergence of the Chebyshev approximation $\bar{x}_m(t)$ to the solution $\hat{x}(t)$.

In order to prove the uniform convergence of the derivatives of the Chebyshev approximations $\bar{x}_m(t)$, we have

$$(6.13)\qquad \frac{d}{dt}\bar{x}_m(t)-\frac{d}{dt}\hat{x}(t)$$

$$=-(I-P_{m-1})[f(t,\hat{x}(t),\int_{-1}^{t} g(t,s,\hat{x}(s))ds)]$$

$$+P_{m-1}[f(t,\bar{x}_m(t),\int_{-1}^{t} g(t,s,\bar{x}_m(s))ds)$$

$$-f(t,\hat{x}(t),\int_{-1}^{t} g(t,s,\hat{x}(s))ds)].$$

On the other hand, by the inequalities (1.11) and (1.12), we have

$$\|(I-P_{m-1})[f(t,\hat{x}(t),\int_{-1}^{t} g(t,s,\hat{x}(s))ds)]\|_C$$

$$\leqq K_1\sigma(m-1)/(m-1)$$

and by the inequalities (1. 5), (1. 7), (5. 3) and (6.11) we have

$$\|P_{m-1}[f(t,\bar{x}_m(t),\int_{-1}^{t} g(t,s,\bar{x}_m(s))ds)$$

$$-f(t,\hat{x}(t),\int_{-1}^{t} g(t,s,\hat{x}(s))ds)]\|_C$$

$$\leqq\sqrt{2(m-1)+1}K_6\|\bar{x}_m-\hat{x}\|_Q\leqq K_6\sqrt{2m-1}\{\|\bar{x}_m-\hat{x}_m\|_Q+\|\hat{x}_m-\hat{x}\|_Q\}$$

$$\leqq K_6\sqrt{2m-1}\{[M/(1-\kappa)]K_7m^{-3/2}+K_1/(m+1)m(m-1)\}.$$

Then for the relation (6.13) we have

$$\left\|\frac{d}{dt}\bar{x}_m-\frac{d}{dt}\hat{x}\right\|_C=O(m^{-1}) \text{ as } m\to\infty,$$

which implies the uniform convergence of the derivatives $\frac{d}{dt}\bar{x}_m(t)$ to $\frac{d}{dt}x(t)$.

This completes the proof of Theorem 4.1.

7. Proof of Theorem 4.2

Let $x=\hat{x}(t)$ be an isolated solution of the boundary value problem (0. 1) and (0. 2) lying in D for any $t\in J$. There exists a positive constant δ satisfying (5. 1). We choose an arbitrary constant ε satisfying $0<\varepsilon\leqq\delta$. Then there exists a number m_0 such that for any $m\geqq m_0$

(7. 1) $\quad \sigma_1(m-1)=1/m<\varepsilon.$

Suppose that for any $m\geqq m_0$ there are two Chebyshev approximations

$$x=\bar{x}_m(t) \text{ and } x=\bar{x}'_m(t)$$

satisfying

(7. 2) $\quad \|\bar{x}_m-\hat{x}\|_C\leqq\varepsilon$ and $\|\bar{x}'_m-\hat{x}\|_C\leqq\varepsilon.$

We denote that

$$U_\varepsilon=\{x\,|\,\|x-\hat{x}(t)\|\leqq\varepsilon \text{ for some } t\in J\}\subset U\subset D.$$

Then it follows that for any $m\geqq m_0$ $\bar{x}_m(t)$ and $\bar{x}'_m(t)$ belong to the domain U_ε for any $t\in J$. Let us put

$$y(t)=\bar{x}_m(t)-\bar{x}'_m(t).$$

By the definition of Chebyshev approximations we have

$$(7.\ 3)\qquad \sum_{i=0}^{N} L_i y(t_i)=0$$

and

$$(7.\ 4)\qquad \frac{d}{dt}y(t)=P_{m-1}[f(t,\bar{x}_m(t),\int_{-1}^{t} g(t,s,\bar{x}_m(s))ds)$$

$$-f(t,\bar{x}'_m(t),\int_{-1}^{t} g(t,s,\bar{x}'_m(s))ds)].$$

We rewrite the above equations (7. 4) in the form

$$(7.\ 5)\qquad \frac{d}{dt}y(t)=\Phi(t;\hat{x})y(t)+\int_{-1}^{t}\Psi(t,s;\hat{x})y(s)ds+h(t),$$

where

$$h(t)=-(I-P_{m-1})[\Phi(t;\hat{x})y(t)+\int_{-1}^{t}\Psi(t,s;\hat{x})y(s)ds]$$

$$+P_{m-1}\int_{0}^{1}\{[\Phi(t;\bar{x}_m^{\theta})-\Phi(t;\hat{x})]y(t)$$

$$+\int_{-1}^{t}[\Phi(t,s;\bar{x}_m^{\theta})-\Psi(t,s;\hat{x})]y(s)ds\}d\theta,$$

and

$$\bar{x}_m^{\theta}(t)=\bar{x}'_m(t)+\theta[\bar{x}_m(t)-\bar{x}'_m(t)].$$

Noting that $\bar{x}_m^{\theta}(t)\in U_\varepsilon$ for any $t\in J$ and $\theta\in[0,1]$, we obtain, by the same argument as that used in proceeding from (5.16) to (5.17),

$$(7.\ 6)\qquad \|h\|_Q\leq\sigma_1(m-1)K_8\|y\|_Q+K_4\|\bar{x}_m^{\theta}-\hat{x}\|_C\|y\|_Q\leq(K_8+K_4)\varepsilon\|y\|_Q$$

from (7. 1) and (7. 2).

On the other hand, applying Lemma 2.3 to the boundary value problem (7. 5) and (7. 3), we obtain

$$y(t)=\int_{-1}^{1}\hat{H}(t,s)h(s)ds,$$

where $\hat{H}(t,s)$ is the Green function with respect to the matrices $\Phi(t;\hat{x})$, $\Psi(t,s;\hat{x})$ and $L_i(i=0,1,\cdots\cdots,N)$. Then it follows from (7. 6) that

$$\|y\|_Q\leq\|\hat{H}\|_Q\|h\|_Q\leq(K_8+K_4)\|\hat{H}\|_Q\varepsilon\|y\|_Q. \tag{7. 7}$$

Since ε is arbitrary, the above inequality (7. 7) implies that

$$\|y\|_Q=0,$$

that is

$$y(t)=0 \text{ for any } t\in J$$

by the Parseval's equality for the finite Chebyshev polynomial series $y(t)$.

This proves the uniqueness of Chebyshev approximations and hence completes the proof of Theorem 4.2.

8. Proof of Theorem 4.3

For the given approximate solution $x=\bar{x}(t)$ of the boundary value problem (0. 1) and (0. 2) we put

$$\frac{d}{dt}\bar{x}(t)=f(t,\bar{x}(t),\int_{-1}^{t}g(t,s,\bar{x}(s))ds)+q(t) \tag{8. 1}$$

and

$$(8.\ 2)\qquad \sum_{i=0}^{N} L_i\bar{x}(t_i)=\ell'.$$

Introducing the matrices $A(t)$ and $B(t,s)$, we rewrite the equations (8 .1) in the form

$$(8.\ 3)\qquad \frac{d}{dt}\bar{x}(t)=A(t)\bar{x}(t)+\int_{-1}^{t} B(t,s)\bar{x}(s)ds+h(t;\bar{x})+q(t).$$

Here we denote that

$$h(t;x)=f(t,x(t),\int_{-1}^{t} g(t,s,x(s))ds)$$

$$-A(t)x(t)-\int_{-1}^{t} B(t,s)x(s)ds$$

for any $x=x(t)\in C(J;D)$. Applying Lemma 2.3 to the boundary value problem (8. 3) and (8. 2), we have

$$(8.\ 4)\qquad \bar{x}(t)=R(t,-1)G^{-1}\ell'+\int_{-1}^{1} H(t,s)[h(s;\bar{x})+q(s)]ds,$$

where $H(t,s)$ is the Green function with respect to the matrices $A(t)$, $B(t,s)$ and $L_i (i=0,1,\cdots\cdots,N)$.

To seek an exact solution of the system (0. 1) satisfying the boundary condition (0. 2), we consider the iterative process

$$(8.\ 5)\qquad \begin{cases} x_{n+1}(t)=R(t,-1)G^{-1}\ell+\int_{-1}^{1} H(t,s)h(s;x_n)ds \\ \qquad\qquad\qquad\qquad (n=0,1,\cdots\cdots) \\ x_0(t)=\bar{x}(t). \end{cases}$$

For the iterative process (8. 5) we shall prove that it can be continued infinitely in the space $C(J;D)$ and that

(8. 6) $\|x_{n+1}-x_n\|_C \leq \kappa^n \|x_1-x_0\|_C$

and

(8. 7) $\|x_{n+1}-x_0\|_C \leq \delta$,

for n=0,1,·····. In fact, for n=0 the inequality (8. 6) is evident. Since

$$x_1(t)-x_0(t)=R(t,-1)G^{-1}(\ell-\ell')-\int_{-1}^{1} H(t,s)q(s)ds,$$

then by the assumptions of the theorem we obtain

$$\|x_1-x_0\|_C \leq M_2\varepsilon+M_1 r,$$

which implies by the condition (5) in (4. 4)

$$\|x_1-x_0\|_C \leq (1-\kappa)\delta<\delta.$$

This proves (8. 7) for n=0. To prove our statement by induction, let us suppose that the iterative process (8. 5) has been continued up to n-1 and we have obtained (8. 6) and (8. 7) up to n-1. Then by the inequality (8. 7) for n-1 we can make $x_{n+1}(t)$ and from (8. 5) we have

(8. 8) $$x_{n+1}(t)-x_n(t)=\int_{-1}^{1} H(t,s)[h(s;x_n)-h(s;x_{n-1})]ds,$$

where

$$h(t;x_n)-h(t;x_{n-1})$$

$$=f(t,x_n(t),\int_{-1}^{t} g(t,s,x_n(s))ds)$$

$$-f(t,x_{n-1}(t),\int_{-1}^{t} g(t,s,x_{n-1}(s))ds)$$

$$-A(t)[x_n(t)-X_{n-1}(t)]-\int_{-1}^{t} B(t,s)[x_n(s)-X_{n-1}(s)]ds.$$

Moreover, by a mean value theorem we obtain

$$h(t;x_n)-h(t;x_{n-1})$$

$$=\int_0^1\{\Phi(t;x_n^\theta)[x_n(t)-x_{n-1}(t)]+\int_{-1}^t \Psi(t,s;x_n^\theta)[x_n(s)-x_{n-1}(s)ds\}d\theta$$

$$-A(t)[x_n(t)-x_{n-1}(t)]-\int_{-1}^t B(t,s)[x_n(s)-x_{n-1}(s)]ds,$$

where

$$x_n^\theta(t)=x_{n-1}(t)+\theta[x_n(t)-x_{n-1}(t)]\in C(J;D).$$

It follows that

$$\|h(t;x_n)-h(t;x_{n-1})\|$$

$$\leqq\int_0^1\|\Phi(t;x_n^\theta)-A(t)\|\|x_n(t)-x_{n-1}(t)\|d\theta$$

$$+\int_{-1}^t\int_0^1\|\Phi(t,s;x_n^\theta)-B(t,s)\|\|x_n(s)-x_{n-1}(s)\|d\theta ds.$$

Hence for the relation (8. 8) we have

$$\|x_n-x_{n-1}\|_C\leqq\|H\|_C\|h(t;x_n)-h(t;x_{n-1})\|_C$$

$$\leqq M_1(\kappa/2M_1+2\kappa/4M_1)\|x_n-x_{n-1}\|_C=\kappa\|x_n-x_{n-1}\|_C$$

by the assumption (3) and (4) in (4. 4) of the theorem. This implies (8. 6) by the assumption of the induction and

$$\|x_{n+1}-x_0\|_C\leqq(\kappa^n+\kappa^{n-1}+\cdots\cdots+\kappa+1)\|x_1-x_0\|_C \tag{8. 9}$$

$$\leqq(M_2\varepsilon+M_1r)/(1-\kappa)\leqq\delta.$$

This completes the induction and hence we see that the iterative process (8. 5) can be continued infinitely and satisfies inequalities (8. 6) and (8. 7) for every n.

By the inequalities (8. 6) and (8. 7) we see that the sequence $\{x_n(t)\}\subset C(J;D)$ obtained by the iterative process (8. 5) converges uniformly to a function $\hat{x}(t)\in C(J;D)$. It readily follows from (8. 9) and (8. 5) that

$$\|\hat{x}-\bar{x}\|_C\leqq(M_2\varepsilon+M_1 r)/(1-\kappa)\leqq\delta$$

and

$$\hat{x}(t)=R(t,-1)G^{-1}\ell+\int_{-1}^{1}H(t,s)h(s;\hat{x})ds. \tag{8.10}$$

The equation (8.10) implies by Lemma 2.3 that

$$\sum_{i=0}^{N}L_i\hat{x}(t_i)=\ell$$

and

$$\frac{d}{dt}\hat{x}(t)=A(t)\hat{x}(t)+\int_{-1}^{t}B(t,s)\hat{x}(s)ds+h(t;\hat{x})$$

$$=f(t,\hat{x}(t),\int_{-1}^{t}g(t,s,\hat{x}(s))ds).$$

Therefore the function $\hat{x}(t)$ is a solution of the boundary value problem (0. 1) and (0. 2) belonging to the space $C(J;U)$.

In order to prove the uniqueness of the solution of our boundary value problem, we consider another solution $x=\hat{x}'(t)$ of the problem (0. 1) and (0. 2) belonging to the space $C(J;U)$. Then

$$\sum_{i=0}^{N}L_i\hat{x}'(t_i)=\ell \tag{8.11}$$

and

$$(8.12)\qquad \frac{d}{dt}\hat{x}'(t)=f(t,\hat{x}'(t),\int_{-1}^{t} g(t,s,\hat{x}'(s))ds)$$

$$=A(t)\hat{x}'(t)+\int_{-1}^{t} B(t,s)\hat{x}'(s)ds+h(t;\hat{x}').$$

Equivalently the relations (8.11) and (8.12) imply that

$$(8.13)\qquad \hat{x}'(t)=R(t,-1)G^{-1}\ell+\int_{-1}^{1} H(t,s)h(s;\hat{x}')ds.$$

Subtracting (8.13) from (8.10), we obtain

$$\|x-x'\|_C\leq\kappa\|x-x'\|_C,$$

which implies

$$\|x-x'\|_C=0.$$

This proves the uniqueness of the solution of the problem (0. 1) and (0. 2) lying in U for any $t\in J$.

In order to prove the isolatedness of the solution $x=\hat{x}(t)$, it is enough to see that the matrix

$$\hat{G}=\sum_{i=0}^{N} L_i\hat{R}(t_i,-1)$$

is nonsingular, where $\hat{R}(t,s)$ is the resolvent matrix with respect to the matrices $\Phi(t;\hat{x})$ and $\Psi(t,s;\hat{x})$. Suppose that $\hat{G}$ is singular. Then there exists a nontrivial vector c such that

$$\hat{G}c=0.$$

For such c let us put

$$y(t)=\hat{R}(t,-1)c.$$

Then $y=y(t)$ is a nontrivial solution of the boundary value problem of the form

$$(8.14)\qquad \sum_{i=0}^{N} L_i y(t_i)=\sum_{i=0}^{N} L_i R(t_i,-1)c=Gc=0$$

and

$$(8.15)\qquad \frac{d}{dt}y(t)=\Phi(t;\hat{x})y(t)+\int_{-1}^{t}\Psi(t,s;\hat{x})y(s)ds$$

$$=A(t)y(t)+\int_{-1}^{t}B(t,s)y(s)ds$$

$$+[\Phi(t;\hat{x})-A(t)]y(t)+\int_{-1}^{t}[\Psi(t,s;\hat{x})-B(t,s)]y(s)ds.$$

Applying Lemma 2.3 to the problem (8.14) and (8.15), we obtain

$$y(t)=\int_{-1}^{1}H(t,s)\{[\Phi(s;\hat{x})-A(s)]y(s)$$

$$+\int_{-1}^{s}[\Psi(s,u;\hat{x})-B(s,u)]y(u)du\}ds.$$

Then, by the assumptions (3) and (4) in (4. 4), we have

$$\|y\|_C\leq\|H\|_C(\kappa/2M_1+2\kappa/M_1)\|y\|_C=\kappa\|y\|_C,$$

which implies that

$$\|y\|_C=0.$$

This is contradiction. Hence the matrix $\hat{G}$ is nonsingular, that is, the solution $x=\hat{x}(t)$ is isolated. This completes the proof of Theorem 4.3.

References

[1] Grossman S. I. and R. K. Miller; Perturbation theory for Volterra integro-differential systems. J. Diff. Eqs., 8, 457-474 (1970).

[2] Shimasaki M. and T. Kiyono; Numerical solution of integral equations in Chebyshev series. Numer. Math., 21, 373-380 (1973).

[3] Tsuruta K. and K. Ohmori; A posteriori error estimation for Volterra integro-differential equations. Memoirs of Numer. Math., No.3, 33-47 (1976).

[4] Urabe M.; Galerkin's procedure for nonlinear periodic systems. Arch. Rat. Mech. Anal., 20, 120-152 (1965).

[5] Urabe M.; An existence theorem for multipoint boundary value problems. Funk. Ekvac., 9, 43-60 (1966).

[6] Urabe M.; Numerical solution of multipoint boundary value problems in Chebyshev series Theory of the method. Numer. Math., 9, 341-366 (1967).

[7] Kurihara M.; On the existence of an approximate solution in Chebyshev series of a nonlinear integral equation of Fredholm type. Memoirs of Numer. Math., No.3, 49-67 (1976).

The Buckling of Plates by the Mixed Finite Element Method

By

Kazuo ISHIHARA*

1. Introduction

We shall consider the mixed finite element approximation applied to the buckling problem of the thin plate with the clamped boundary condition:

$$(1)\qquad \left\| \begin{aligned} &\Delta\Delta u + \lambda \sum_{i,j=1}^{2} \tau_{ij} \frac{\partial^2 u}{\partial x_i \partial x_j} = 0 && \text{in } \Omega, \\ &u = \partial u/\partial \nu = 0 && \text{on } \partial\Omega. \end{aligned} \right.$$

Here Ω is a bounded convex domain in the x_1x_2-plane with boundary $\partial\Omega$ consisting of a finite number of smooth arcs, Δ is the Laplace operator, $\partial/\partial\nu$ is the outward normal derivative along $\partial\Omega$, and τ_{ij} $(i,j=1,2)$ are given smooth functions such that

$$\tau_{12} = \tau_{21},$$

$$\frac{\partial\tau_{11}}{\partial x_1} + \frac{\partial\tau_{12}}{\partial x_2} = 0, \qquad \frac{\partial\tau_{21}}{\partial x_1} + \frac{\partial\tau_{22}}{\partial x_2} = 0.$$

The buckling of the plate is possible only for certain definite values of λ. The minimum of these values determines the critical buckling load. The associated buckling configuration is the function u corresponding to the buckling load λ. A simple case is buckling under pure

*Department of Mathematics, Faculity of Science, Ehime University, Matsuyama 790, Japan.

compression

$$\Delta\Delta u + \lambda\Delta u = 0.$$

The aim of this paper is to give the rates of convergence for the finite element approximate solutions of the eigenvalues and the eigenfunctions, by applying the mixed method with piecewise linear polynomials proposed by Miyoshi [5],[6].

2. Notations and weak formulation

Let $L_2(\Omega)$ be the real space of square integrable functions on Ω, whose inner product and norm are denoted by $(\ ,\ )$ and $\|\cdot\|$, respectively. Let $H^m(\Omega)$ be the real m-th order Sobolev space(m=1,2, 3,$\cdots$) provided with the norm

$$\|u\|_m = \Big(\sum_{|\alpha|\leq m} \|D^\alpha u\|^2\Big)^{1/2}$$

Here $\alpha=(\beta,\gamma)$ is a two-component index with non-negative integers,

$$|\alpha| = \beta + \gamma \quad \text{and} \quad D^\alpha = \partial^{|\alpha|} \big/ \partial x_1^\beta\, \partial x_2^\gamma.$$

The spaces $H_0^1(\Omega)$ and $H_0^2(\Omega)$ are given by

$$H_0^1(\Omega) = \{f;\ f\in H^1(\Omega), \quad f = 0 \quad \text{on} \quad \partial\Omega\},$$

$$H_0^2(\Omega) = \{f;\ f\in H^2(\Omega), \quad f = \partial f/\partial\nu = 0 \quad \text{on} \quad \partial\Omega\}.$$

In order to deal with the buckling problem (1) in a weak form, let us define bilinear forms $<\ ,\ >$ and $[\ ,\ ]$ by

$$<u,v> = \sum_{|\alpha|=2} (D^\alpha u, D^\alpha v),$$

$$[u,v] = \sum_{i,j=1}^{2} (\tau_{ij}\partial u/\partial x_i, \partial v/\partial x_j).$$

We assume that

$$[u,u] \geq 0 \qquad \text{for each} \quad u \in H^1(\Omega),$$

and that the norm $|||\cdot|||$ induced by $[\ ,\ ]$ is equivalent to the norm $\|\cdot\|_1$ in $H_0^1(\Omega)$. The standard weak form of (1) consists of finding a real eigenvalue λ and a non-zero eigenfunction $u \in H_0^2(\Omega)$ defined by

$$\langle u,\phi\rangle = \lambda[u,\phi] \qquad \text{for each} \quad \phi \in H_0^2(\Omega).$$

It is well known that all the eigenvalues $\{\lambda_i\}$ are arranged as

$$0 < \lambda_1 \leq \lambda_2 \leq \cdots < \infty.$$

The multiplicity of each eigenvalue is always finite. The corresponding eigenfunctions are denoted by $\{u_i\}$ with the normalization condition

$$[u_i,u_j] = \delta_{ij}$$

where δ_{ij} is Kronecker's delta. It is also well known that the eigenfunctions $\{u_i\}_{i=1}^{\infty}$ belong to the space $H_0^2(\Omega) \cap H^4(\Omega)$ when $\partial\Omega$ is sufficiently smooth and that $\{u_i\}_{i=1}^{\infty}$ belong to $H_0^2(\Omega) \cap H^3(\Omega)$ when Ω is a convex polygon. From the Rayleigh principle the eigenvalues are characterized by

$$\lambda_i = \min_{\substack{u \in H_0^2(\Omega) \\ u \neq 0 \\ [u,u_j]=0 \\ j=1,\cdots,i-1}} \frac{\langle u,u\rangle}{[u,u]}, \qquad i=1,2,\cdots,$$

and the minimum is attained by u_i.

In order to construct the mixed finite element scheme, we introduce another formulation: Find $(u,U_{11},U_{12},U_{21},U_{22}) \in H_0^1(\Omega) \times H^1(\Omega) \times H^1(\Omega) \times H^1(\Omega) \times H^1(\Omega)(U_{12}=U_{21})$ such that

$$
(2)\quad \left\|\begin{aligned}
&\sum_{i,j=1}^{2}[(\partial u/\partial x_i,\partial\phi_{ij}/\partial x_j)+(U_{ij},\phi_{ij})]=0\\
&\qquad\qquad\qquad\qquad \text{for each}\quad \phi_{ij}\in H^1(\Omega)\quad(\phi_{12}=\phi_{21}),\\
&\sum_{i,j=1}^{2}(\partial U_{ij}/\partial x_j,\partial\phi/\partial x_i)-\lambda\Big(\sum_{i,j=1}^{2}\tau_{ij}U_{ij},\phi\Big)=0\\
&\qquad\qquad\qquad\qquad \text{for each}\quad \phi\in H_0^1(\Omega).
\end{aligned}\right.
$$

We assume that $u\in H_0^2(\Omega)\cap H^3(\Omega)$. The solution $(u,U_{11},U_{12},U_{21},U_{22})$ of (2) is related by

$$
U_{ij}=\partial^2 u/\partial x_i\partial x_j\quad,\quad 1\le i,j\le 2.
$$

We define a space V by

$$
\begin{aligned}
V=\{(u,U_{11},U_{12},U_{21},U_{22})\in H_0^1(\Omega)\times H^1(\Omega)\times H^1(\Omega)\times H^1(\Omega)\times H^1(\Omega);\\
U_{12}=U_{21},\quad \sum_{i,j}[(\partial u/\partial x_i,\partial\phi_{ij}/\partial x_j)+(U_{ij},\phi_{ij})]=0\\
\text{for each}\quad \phi_{ij}\in H^1(\Omega)\quad(\phi_{12}=\phi_{21})\},
\end{aligned}
$$

and a bilinear form $E(U,W)$ by

$$
E(U,W)=(U_{11},W_{11})+(U_{12},W_{12})+(U_{21},W_{21})+(U_{22},W_{22})
$$

where $U=(U_{11},U_{12},U_{21},U_{22})$, $W=(W_{11},W_{12},W_{21},W_{22})$. The eigenvalues are characterized by

$$
\lambda_i=\min_{\substack{\{u,U\}\in V\\ u\neq 0\\ [u,u_j]=0\\ j=1,\cdots,i-1}}\frac{E(U,U)}{|||u|||^2},\qquad i=1,2,\cdots,
$$

and the minimum is attained by $(u_i,U^i)=(u_i,U^i_{11},U^i_{12},U^i_{21},U^i_{22})$.

3. Convergence of the finite element scheme

For simplicity, we assume that the domain Ω is a convex polygon. Then the eigenfunction u belongs to $H_0^2(\Omega)\cap H^3(\Omega)$. The domain is decomposed into disjoint triangular elements in the usual manner. Let $P_i, 1 \le i \le n$ (or $P_i, n+1 \le i \le n+J$) be the nodal points of the triangulation T^h which belong to Ω (or $\partial\Omega$). Here h is the largest side length of all the triangular elements. It is also assumed that the triangulation T^h is uniform in the interior of Ω in the sense of Miyoshi ([5],[6]). Let $\{\hat{\phi}_i\}(i=1,2,\cdots,n+J)$ be the piecewise linear functions such that

$$\hat{\phi}_i(P_j) = \delta_{ij}, \qquad 1 \le i,j \le n+J.$$

Let Y^h be the subspace of $H^1(\Omega)$ spanned by $\{\hat{\phi}_1,\cdots,\hat{\phi}_{n+J}\}$ and Y_0^h be the subspace of $H_0^1(\Omega)$ spanned by $\{\hat{\phi}_1,\cdots,\hat{\phi}_n\}$.

We now define the mixed finite element solution $(\hat{\lambda},\hat{u},\hat{U}_{11},\hat{U}_{12},\hat{U}_{21},\hat{U}_{22})(\hat{u}\in Y_0^h,\ \hat{U}_{ij}\in Y^h,\ \hat{U}_{12}=\hat{U}_{21})$ of the consistent mass scheme for the problem (2) by

$$(3)\qquad \left\|\ \begin{aligned} &\sum_{i,j=1}^{2}[(\partial\hat{u}/\partial x_i,\partial\hat{\phi}_{ij}/\partial x_j) + (\hat{U}_{ij},\hat{\phi}_{ij})] = 0 \\ &\qquad\qquad\qquad \text{for each } \hat{\phi}_{ij}\in Y^h\ (\hat{\phi}_{12}=\hat{\phi}_{21}), \\ &\sum_{i,j=1}^{2}(\partial\hat{U}_{ij}/\partial x_j,\partial\hat{\phi}/\partial x_i) - \hat{\lambda}(\sum_{i,j=1}^{2}\tau_{ij}\hat{U}_{ij},\hat{\phi}) = 0 \\ &\qquad\qquad\qquad \text{for each } \hat{\phi}\in Y_0^h. \end{aligned}\right.$$

We define a space V^h by

$$\begin{aligned} V^h = \{(\hat{u},\hat{U}_{11},\hat{U}_{12},\hat{U}_{21},\hat{U}_{22}) &\in Y_0^h\times Y^h\times Y^h\times Y^h\times Y^h;\ \ \hat{U}_{12}=\hat{U}_{21}, \\ &\sum_{i,j}[(\partial\hat{u}/\partial x_i,\partial\hat{\phi}_{ij}/\partial x_j) + (\hat{U}_{ij},\hat{\phi}_{ij})] = 0 \\ &\qquad\qquad \text{for each } \hat{\phi}_{ij}\in Y^h\ (\hat{\phi}_{12}=\hat{\phi}_{21})\}. \end{aligned}$$

It is also assumed that for $(\hat{u},\hat{U}_{11},\hat{U}_{12},\hat{U}_{21},\hat{U}_{22})\in V^h$ and $\hat{w}\in Y_0^h$,

$$(4)\qquad (\sum_{i,j=1}^{2}\tau_{ij}\hat{U}_{ij},\hat{w}) = -[\hat{u},\hat{w}].$$

By (3) and (4), we can obtain a set of the matrix equations

$$Kx + My = 0, \qquad \tilde{K}y + \hat{\lambda}Gx = 0,$$

where K and $\tilde{K}$ are elastic stiffness matrices, G is the geometric stiffness matrix, and M is the consistent mass matrix. The eigenvalues $\{\hat{\lambda}_i\}_{i=1}^{n}$ of (3) are arranged as

$$0 < \hat{\lambda}_1 \leq \hat{\lambda}_2 \leq \cdots \leq \hat{\lambda}_n .$$

The corresponding eigenfunctions $(\hat{u}_i, \hat{U}^i) = (\hat{u}_i, \hat{U}^i_{11}, \hat{U}^i_{12}, \hat{U}^i_{21}, \hat{U}^i_{22})$ $(i=1,2,\cdots,n)$ can be normalized by

$$[\hat{u}_i, \hat{u}_j] = \delta_{ij}, \qquad [\hat{u}_i, u_i] \geq 0.$$

Then the eigenvalues are characterized by

$$\hat{\lambda}_i = \min_{\substack{(\hat{u},\hat{U}) \in V^h \\ u \neq 0 \\ [\hat{u},\hat{u}_j]=0 \\ j=1,\cdots,i-1}} \frac{E(\hat{U},\hat{U})}{|||\hat{u}|||^2}, \qquad i=1,2,\cdots,n,$$

and the minimum is attained by $(\hat{u}_i, \hat{U}^i)$.

Consider the static plate bending problem with the clamped boundary condition:

$$(5) \qquad \left\| \begin{array}{ll} \Delta\Delta w = f & \text{in } \Omega, \\ w = \partial w/\partial\nu = 0 & \text{on } \partial\Omega. \end{array} \right.$$

Here f is a given function belonging to $L_2(\Omega)$. Miyoshi proved the following proposition ([5],[6]).

<u>Proposition 1.</u> Let w be the solution of (5) and $(\hat{w}, \hat{W}_{11}, \hat{W}_{12}, \hat{W}_{21}, \hat{W}_{22}) \in Y^h_0 \times Y^h \times Y^h \times Y^h \times Y^h$ $(\hat{W}_{12} = \hat{W}_{21})$ be the finite element solution defined by

$$
(6)\quad \left\|\begin{aligned}
&\sum_{i,j=1}^{2}[(\partial\hat{w}/\partial x_i,\partial\hat{\phi}_{ij}/\partial x_j)+(\hat{W}_{ij},\hat{\phi}_{ij})]=0\\
&\qquad\qquad\qquad \text{for each } \hat{\phi}_{ij}\in Y_0^h\ (\hat{\phi}_{12}=\hat{\phi}_{21}),\\
&\sum_{i,j=1}^{2}(\partial\hat{W}_{ij}/\partial x_j,\partial\hat{\phi}/\partial x_i)+(f,\hat{\phi})=0\\
&\qquad\qquad\qquad \text{for each } \hat{\phi}\in Y^h.
\end{aligned}\right.
$$

Then

$$\|w-\hat{w}\|_1 \leq C\,\|f\|\,h^{1/2},$$

where C is a constant independent of h.

We can now obtain the following results using Proposition 1. The proof will be published elsewhere.

<u>Theorem 1.</u> Let $\hat{\lambda}_i$ be the approximate eigenvalue of λ_i. Then for sufficiently small h, there exists a constant C_1 which is independent of h such that

$$|\lambda_i-\hat{\lambda}_i| \leq C_1 h^{1/2}.$$

<u>Theorem 2.</u> Let λ_i be of multiplicity $p+1$ $(p \geq 0,\ \lambda_{i-1}<\lambda_i=\cdots=\lambda_{i+p}<\lambda_{i+p+1})$, and $u_i,\cdots,u_{i+p}$ be the corresponding eigenfunctions. Let $\hat{u}_k$ be the approximate eigenfunction corresponding to $\hat{\lambda}_k (k=1,2,\cdots,n)$. Then for sufficiently small h, there exists a constant C_2 which is independent of h such that

$$\operatorname{dist}\{u_j,\ \operatorname{span}[\hat{u}_i,\cdots,\hat{u}_{i+p}]\} \leq C_2 h^{1/2},\quad j=i,\cdots,i+p,$$

where

$$\operatorname{dist}\{u,\ B\} = \inf_{b\in B} \|\!|\!| u-b |\!|\!|.$$

As a corollary to Theorem 2, we have

<u>Corollary 1.</u> If λ_1 is a simple eigenvalue, then for sufficiently small h there exists a constant C_3 which is independent of h such that

$$|||u_1 - \hat{u}_1||| \leq C_3 h^{1/2}.$$

Remark. Furthermore, we can propose the generalized mixed mass scheme with a parameter θ, $0 \leq \theta \leq 1$([2])

$$Kx + \{\theta M + (1-\theta)M_1\}y = 0, \qquad \tilde{K}y + \tilde{\lambda}Gx = 0,$$

where M_1 is the diagonal lumped mass matrix. The error estimate of the above scheme is similar to the one of the consistent mass scheme.

4. Numerical example

To show the validity of the theoretical results, we deal with the following example of the square plate, which is the same as the one given by Weinstein and Stenger ([7], p193).

Example.

$$\Omega: \quad -\pi/2 < x_1, x_2 < \pi/2,$$
$$\Delta\Delta u + \lambda\Delta u = 0 \qquad \text{in } \Omega,$$
$$u = \partial u/\partial\nu = 0 \qquad \text{on } \partial\Omega.$$

Although the exact first eigenvalue (buckling load) λ_1 is not known, they obtained the inequality

$$5.30362 \leq \lambda_1 \leq 5.31173.$$

We divide Ω into $n \times n$ uniform mesh as shown in Figure 1. Our choices for the parameter θ are 0, 0.5, 1. Table 1 shows the numerical results, from which we can see that the approximate first eigenvalues converge as h tends to zero. All numerical computations were performed on the FACOM 230-28 computer at Ehime University, and the FACOM 230-75 computer at Kyushu University.

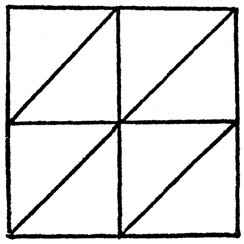


Figure 1. Mesh pattern (3×3)

Table 1. The results for Example

mesh($n \times n$)		7×7	8×8	9×9	10×10
h		$\sqrt{2}\pi/6$	$\sqrt{2}\pi/7$	$\sqrt{2}\pi/8$	$\sqrt{2}\pi/9$
θ	0	4.82007	4.94137	5.02223	5.07883
	0.5	5.45497	5.42827	5.40559	5.38766
	1	6.31138	6.03642	5.86073	5.74163

Acknowledgements

The author would like to express his thanks to Professor H. Fujita of the University of Tokyo and Dr. M. Tabata of Kyoto University for their advices.

References

[1] Brezzi, F. : On the existence, uniqueness and approximation of saddle-point problem arising from Lagrangian multipliers. R. A. I. R. O., 8-R2 (1974) 129-151.

[2] Ishihara, K. : Convergence of the finite element method applied to the eigenvalue problem $\Delta u+\lambda u=0$. Publ. RIMS, Kyoto Univ., 13 (1977) 47-60.

[3] Kikuchi, F. : Convergence of the ACM finite element scheme for plate bending problems. Publ. RIMS, Kyoto Univ., 11 (1975) 247-265.

[4] Kikuchi, F. : On a finite element scheme based on the discrete Kirchhoff assumption. Numer. Math., 24 (1975) 211-231.

[5] Miyoshi, T. : A finite element method for the solutions of fourth order partial differential equations. Kumamoto J. Sci. (Math.), 9 (1972) 87-116.

[6] Miyoshi, T. : A mixed finite element method for the solution of the von Karman equations. Numer. Math., 26 (1976) 255-269.

[7] Weinstein, A. and Stenger, W. : Methods of intermediate problems for eigenvalue problems, theory and ramifications. Academic Press 1972.