AN A POSTERIORI ERROR ESTIMATION OF APPROXIMATE SOLUTIONS OF TWO POINT BOUNDARY VALUE PROBLEMS FOR PIECEWISE SMOOTH SYSTEMS

By

Kumiko ITOH

(Received Oct. 22, 1975)

Abstract

In his paper [2], Urabe has given a method of a posteriori error estimation of approximate solutions for boundary value problems with nonlinear boundary conditions. In the present paper, the author applies Urabe's method to the boundary value problem with nonlinear boundary conditions derived, by means of the author's method [1], from two point boundary value problem concerning piecewise smooth systems in her previous paper [1], and she gives a method of a posteriori error estimation of approximate solutions to the two point boundary value problem in question.

§0. Introduction

In [1], the piecewise smooth system is defined as follows.

Let D be a region of tx -space (x : n -vector) intercepted by two hyperplanes $t=t_0$ and $t=t_r$; the cross sections of D at $t=t_0$ and $t=t_r$ are supposed to be an open set in each hyperplanes and these open sets are supposed to be included in D .

Let the region D be divided into a finite number of subregions E_λ , $\lambda \in A$, disjoint each other. Let $Y^\lambda(x, t)$, $\lambda \in A$ are smooth n -vector functions on D such that

$$(0.1) \quad Y^{\lambda_1}(x, t) = Y^{\lambda_2}(x, t), \quad \text{for } (t, x) \in \bar{E}_{\lambda_1} \cap \bar{E}_{\lambda_2},$$

where $\bar{E}$ denotes the closure set of E .

Then the piecewise smooth system is

$$(0.2) \quad \dot{x} = X(x, t): X(x, t) = Y^\lambda(x, t), \quad \text{for } (t, x) \in E_\lambda, \lambda \in A.$$

The present paper concerns with the solution $x(t)$ on $[t_0, t_r]$ of (0.2) satisfying the two point boundary condition

$$(0.3) \quad L_0 x(t_0) + L_1 x(t_r) = l,$$

where L_0 and L_1 are given $(n \times n)$ -matrices and l a given n -vector.

The above boundary value problem is then transformed to a boundary value problem with nonlinear boundary conditions by the method in Ref. [1] as follows.

Let the solution $x(t)$ passes through successively the subregions E_j and crosses the boundary hypersurfaces $f^j(t, x) = 0$ of subregions E_j at times t_j . Time t_j is called a switching time. Let the number of switching times is finite, that is, $r - 1$.

We treat switching times t_j as unknown parameters and the solution $x(t)$ restricted on subintervals divided by switching times as independent unknowns. The boundary value problem (0.2) with (0.3) is thus reduced to the boundary value problem of the smooth system of differential equations

$$(0.4) \quad \frac{dx^j}{d\tau} = Y^j(x^j(\tau), t_j(\tau)) \frac{t_j - t_{j-1}}{2}, \quad j = 1, \dots, r, \tau \in I = [-1, 1],$$

with nonlinear boundary conditions

$$(0.5) \quad \begin{cases} f^j(x^j(1), t_j) = 0 & j = 1, 2, \dots, r-1, \\ L_0 x^1(-1) + L_1 x^r(1) = l, \\ x^{j+1}(-1) = x^j(1) & j = 1, 2, \dots, r-1, \end{cases}$$

where

$$t_j(\tau) = \frac{t_j - t_{j-1}}{2} \tau + \frac{t_j + t_{j-1}}{2},$$

$$x^j(\tau) = x(t_j(\tau)), \quad \text{for } \tau \in I, j = 1, 2, \dots, r.$$

We assume that functions $Y^j(x, t)$, $j = 1, \dots, r$ and $f^j(x, t)$, $j = 1, \dots, r-1$ are continuously differentiable on D .

Let $(\bar{x}^j(\tau)$, $j = 1, 2, \dots, r$, $\bar{t}_j$, $j = 1, \dots, r-1$) be an approximate solution of (0.4) with (0.5).

For $j=1, 2, \dots, r$, let $\Phi_j(\tau)$ be a fundamental matrix of the first variation system

$$(0.6) \quad \dot{x}^j = \bar{Y}_x^j(\tau) \frac{\bar{t}_j - \bar{t}_{j-1}}{2} x^j, \quad j=1, \dots, r$$

satisfying

$$\Phi_j(-1) = E_{n,n},$$

where $E_{n,n}$ is a unit $(n \times n)$ -matrix, and

$$\bar{Y}_x^j(\tau) = Y_x^j(\bar{x}^j(\tau), \bar{t}_j(\tau)), \quad j=1, \dots, r.$$

Let G be the $\{nr + (r-1)\} \times \{nr + (r-1)\}$ -matrix

$$(0.7) \quad \begin{bmatrix} L_0 & L_1 \Phi_r(1) & L_1 \Phi_r(1) \bar{Y}^{r-1}(1) \\ \Phi_1(1), -E_{n,n} & \bar{Y}^1(1) & \\ & -\Phi_2(1) \bar{Y}^1(1), \bar{Y}^2(1) & \\ & \Phi_{r-1}(1), -E_{n,n} & \Phi_{r-1}(1) \bar{Y}^{r-2}(1), \bar{Y}^{r-1}(1) \\ \bar{f}_x^1 \Phi_1(1) & \bar{f}_x^1 \bar{Y}^1(1) + \bar{f}_t^1 & \\ & -\bar{f}_x^2 \Phi_2(1) \bar{Y}^1(1), \bar{f}_x^2 \bar{Y}^2(1) + \bar{f}_t^2 & \\ & & \ddots & \\ \bar{f}_x^{r-1} \Phi_{r-1}(1), 0 & -\bar{f}_x^{r-1} \Phi_{r-1}(1) \bar{Y}^{r-2}(1), \bar{f}_x^{r-1} \bar{Y}^{r-1}(1) + \bar{f}_t^{r-1} \end{bmatrix}$$

where

$$\bar{Y}^j(\tau) = Y^j(\bar{x}^j(\tau), \bar{t}_j(\tau)), \quad j=1, \dots, r,$$

$$\bar{f}^j = f^j(\bar{x}^j(1), \bar{t}_j)$$

$$\bar{f}_x^j = f_x^j(\bar{x}^j(1), \bar{t}_j),$$

$$\bar{f}_t^j = f_t^j(\bar{x}^j(1), \bar{t}_j), \quad j=1, \dots, r-1.$$

Further let A be the $nr \times \{nr + (r-1)\}$ matrix

$$(0.8) \quad \begin{bmatrix} \Phi_1(\tau) & \bar{Y}^1(\tau) \frac{\tau+1}{2} & & \\ \Phi_2(\tau) & \bar{Y}^2(\tau) \frac{1-\tau}{2} - \Phi_2(\tau) \bar{Y}^1(1), & \bar{Y}^2(\tau) \frac{\tau+1}{2} & \\ \vdots & \vdots & \vdots & \\ \Phi_r(\tau) & \bar{Y}^r(\tau) \frac{1-\tau}{2} - \Phi_r(\tau) \bar{Y}^{r-1}(1) & \bar{Y}^{r-1}(\tau) \frac{\tau+1}{2} & \end{bmatrix}$$

By $h_1, k_1, h_2^1, h_2^2, k_2^1$ and k_2^2 , we denote the following numbers:

$$h_1 = \sqrt{2} \|AG^{-1}\|_n k_3 + \sqrt{2} \left[\max_{j \in I} \int_{-1}^{\tau} \|\Phi_j(\tau) \Phi_j^{-1}(s)\|^2 ds \right]^{1/2},$$

$$k_1 = \sqrt{2} \|[O_{r-1, nr}, E_{r-1, r-1}]\| G^{-1} \|_n k_3,$$

$$h_2^1 = \left\| AG^{-1} \begin{bmatrix} O_{nr, r-1} \\ E_{r-1, r-1} \end{bmatrix} \right\|_n,$$

$$h_2^2 = \left\| AG^{-1} \begin{bmatrix} E_{nr, nr} \\ O_{r-1, nr} \end{bmatrix} \right\|_n,$$

$$k_2^1 = \left\| [O_{r-1, nr}, E_{r-1, r-1}] G^{-1} \begin{bmatrix} O_{nr, r-1} \\ E_{r-1, r-1} \end{bmatrix} \right\|,$$

$$k_2^2 = \left\| [O_{r-1, nr}, E_{r-1, r-1}] G^{-1} \begin{bmatrix} E_{nr, nr} \\ O_{r-1, nr} \end{bmatrix} \right\|,$$

where

$$k_3 = \max \left\{ \left[\int_{-1}^1 \|L_1 \Phi_r(1) \Phi_r^{-1}(s)\|^2 ds \right]^{1/2}, \left[\int_{-1}^1 \|\bar{f}_x^j \Phi_j(1) \Phi_j^{-1}(s)\|^2 ds \right]^{1/2}, \right. \\ \left. \left[\int_{-1}^1 \|\Phi_j(1) \Phi_j^{-1}(s)\|^2 ds \right]^{1/2}, \quad j=1, \dots, r-1 \right\},$$

$O_{m,k}$: the $m \times k$ zero matrix.

By symbol $\|\cdot\|$, we denote the Euclidean norm for a vector and $(\sum_{i,j} a_{ij}^2)^{1/2}$

for a matrix with elements a_{ij} . For a vector- or matrix- valued functions defined on I , we denote by $\|\cdot\|_n$ the supremum of the norm for these functions, defined in the above way.

The following Theorem is a result of the present paper.

THEOREM. *Suppose the boundary value problem (0.4) with (0.5) has an approximate solution $(\bar{x}^j(\tau), j=1, \dots, r, \bar{t}_j, j=1, \dots, r-1)$ satisfying*

$$(a) \quad \det(L_0 + L_1 \Phi_r(1)) \neq 0,$$

$$(b) \quad \bar{f}_x^j \bar{Y}^j(1) + \bar{f}_t^j \neq 0 \quad j=1, \dots, r-1.$$

In addition, suppose that there are a positive number δ and a nonnegative $\kappa < 1$ satisfying following conditions:

$$(c) \quad D_\delta = \{(t_j(\tau), x^j(\tau)): j=1, \dots, r, \tau \in I, (x^1 \dots x^r, t_1 \dots t_{r-1}) \in \Omega_\delta\} \subset D,$$

where

$$\Omega_\delta = \{(x^1 \dots x^r, t_1 \dots t_{r-1}): (\sum_{j=1}^r \|x^j - \bar{x}^j\|_n^2)^{1/2} + (\sum_{j=1}^{r-1} |t_j - \bar{t}_j|^2)^{1/2} \leq \delta, x^j(\tau) \text{ is a continuously differentiable function}\}.$$

$$(d) \quad \max \left[\left(\sum_{j=1}^r \left\| Y_x^j(x^j(\cdot), t_j(\cdot)) \frac{t_j - t_{j-1}}{2} - \bar{Y}_x^j \frac{\bar{t}_j - \bar{t}_{j-1}}{2} \right\|_n^2 \right)^{1/2} + \left(\sum_{j=1}^{r-1} \|f_x^j(x_{(1)}^j, t_j) - \bar{f}_x^j\|^2 \right)^{1/2}, \right. \\ \left. \left(2 \sum_{j=1}^r \left\{ \left\| Y_x^j(x^j(\cdot), t_j(\cdot)) \frac{t_j - t_{j-1}}{2} - \bar{Y}_x^j \frac{\bar{t}_j - \bar{t}_{j-1}}{2} \right\|_n + \frac{1}{2} \|Y^j(x_{(1)}^j, t_j(\tau)) - \bar{Y}^j(\tau)\|_n \right\}^2 \right)^{1/2} + \left(\sum_{j=1}^{r-1} |f_t^j(x_{(1)}^j, t_j) - \bar{f}_t^j|^2 \right)^{1/2} \right] \leq \frac{\kappa}{M_1},$$

for any $(x^1 \dots x^r, t_1, \dots, t_r) \in \Omega_\delta$,

where

$$\bar{Y}_t^j(\tau) = Y_t^j(\bar{x}^j(\tau), \bar{t}_j(\tau)).$$

$$(e) \quad \frac{M_1 r + M_2 \varepsilon}{1 - \kappa} \leq \delta.$$

Here r, ε, M_1 and M_2 are numbers such that

$$\begin{aligned}
 (f) \quad & \left[\sum_{j=1}^r \left\| \frac{d\bar{x}^j}{d\tau} - \bar{Y}^j \frac{\bar{t}_j - \bar{t}_{j-1}}{2} \right\|_n^2 \right]^{1/2} + \left[\sum_{j=1}^{r-1} f^{j^2} \right]^{1/2} \leq r, \\
 & [\|L_0 \bar{x}^1(-1) + L_1 \bar{x}^r(1) - l\|^2 + \sum_{j=1}^{r-1} \|\bar{x}^j(1) - \bar{x}^{j+1}(-1)\|^2]^{1/2} \leq \varepsilon. \\
 (g) \quad & M_1 = \max(h_1 + k_1, h_2^1 + k_2^1), \\
 & M_2 = h_2^2 + k_2^2.
 \end{aligned}$$

Then the boundary value problem (0.4) with (0.5) has the unique exact solution $(\hat{x}^j(\tau), j=1, \dots, r, \hat{t}_j, j=1, \dots, r-1)$ in Ω_δ such that

$$\left[\sum_{j=1}^{r-1} (\hat{t}_j - \bar{t}_j)^2 \right]^{1/2} + \left[\sum_{j=1}^r \|\hat{x}^j - \bar{x}^j\|_n^2 \right]^{1/2} \leq (M_1 r + M_2 \varepsilon) / (1 - \kappa).$$

REMARK 1. Nonsingularity of the matrix G under (a) and (b) will be shown in §4. From the nonsingularity it follows that the exact solution $(\hat{x}^j(\tau), \hat{t}_j, j=1, \dots, r-1)$ is the so-called isolated solution defined in [2].

In the present paper, we shall prove the above theorem by the use of the method of [2]. The method of [2] is based on the Newton method for operator equations in normed spaces. In the present, therefore, some remarks necessary for subsequent discussions will be made in §1 concerning the Newton method in normed spaces. In §2, as a preparation for the Newton method in normed space, we shall give the explicit form of weak Fréchet derivative of the operator, which appears in the operator equation derived from the boundary value problem with nonlinear boundary conditions. In §3 we shall give the inverse form of the weak Fréchet derivative. In §4, we evaluate the matrices G and A , which are necessary to express the inverse of the weak Fréchet derivative. Finally, in §5, we shall give a proof of our theorem.

§1. The Newton Method for Operator Equations

The boundary value problem (0.4) with (0.5) is reduced to the problem to seek the solution y satisfying the operator equation

$$(1.1) \quad F(y) = 0,$$

where

$$(1.2) \quad F(y) = \left[\frac{dx}{d\tau} - X(x, \tau, T), f(x, T), l_3 x - c \right]$$

with

$$\begin{aligned}
 y &= (x, T), \quad x = (x^1 \dots x^r), \quad T = (t_1 \dots t_{r-1}), \\
 X(x, \tau, T) &= \left(Y^1(\tau) \frac{t_1 - t_0}{2}, Y^2(\tau) \frac{t_2 - t_1}{2} \dots Y^r(\tau) \frac{t_r - t_{r-1}}{2} \right), \\
 Y^j(\tau) &= Y^j(x^j(\tau), t_j(\tau)), \quad j = 1, \dots, r. \\
 f(x, T) &= (f^1(x^1(1), t_1), f^2(x^2(1), t_2), \dots, f^{r-1}(x^{r-1}(1), t_{r-1})), \\
 l_3 x &= (L_0 x^1(-1) + L_1 x^r(1), x^1(1) - x^2(-1), \dots, x^{r-1}(1) - x^r(-1)), \\
 c &= \begin{pmatrix} I \\ 0 \end{pmatrix}.
 \end{aligned}$$

We shall define a normed space E and a Banach space B and show that $F(y)$ maps some open set of E into B .

We use the symbol R^m for m dimensional Euclidean space (m : a positive integer), $C[I]$ for the space consisting of nr -dimensional vector-functions continuous on I with norm $\|\cdot\|_n$, $C^1[I]$ for the subspace of $C[I]$ consisting of the functions continuously differentiable on I , B for the space $C[I] \times R^{(r-1)+nr}$ whose norm is defined by

$$\|z\|_b = \|\varphi\|_n + \|c\|, \quad \text{for } z = (\varphi, c) \in C[I] \times R^{(r-1)+nr},$$

B_1 for the subspace $C[I] \times R^{r-1}$ of B , B_2 for the subspace R^{nr} of B , and finally E for the space $C^1[I] \times R^{r-1}$ whose norm is defined by

$$\|y\|_e = \|x\|_n + \|T\|, \quad \text{for } y = (x, T) \in C^1[I] \times R^{r-1}.$$

Let Ω be a set of E such that

$$(1.3) \quad \Omega = \{(x(\cdot), T) \in E: (t_j(\tau), x^j(\tau)) \in D, \quad \text{for } \tau \in I, j = 1, \dots, r\}$$

By the definition of D , Ω is open in E .

From (c),

$$(1.4) \quad \Omega_\delta \subset \Omega.$$

Thus the operator $F(y)$ is defined for $y \in \Omega$ and maps into B : $\frac{dx}{d\tau} - X(x, \tau, T)$ maps Ω into $C[I]$, l_3 maps Ω into R^{nr} and f maps Ω into R^{r-1} .

The following Proposition 1 is concerned with error estimation of a solution of an operator equation by the Newton method in normed space. It is specialized in the form of $F(y)$ compared with Theorem 1 of [2], that is, in Proposition 1 $F(y)$ has two components, linear and nonlinear corresponding to boundary conditions, while in Theorem 1 of [2] $F(y)$ has only nonlinear component corresponding to boundary conditions.

PROPOSITION 1. *Let $F(y)$ be an operator, which maps an open set Ω of a normed space E with norm $\|\cdot\|_e$ into a Banach space B with norm $\|\cdot\|_b$. Let $F(y)$ is of the form*

$$F(y) = F_1(y) \oplus (F_2 y - c).$$

Here $F_1(y)$ is a nonlinear operator from Ω into B , $F_2 y$ is a linear operator from Ω into B , and c is a constant element in B . The symbol $\oplus$ denotes the direct sum in space B ; hence there exists subspaces B_1 and B_2 such that

$$B = B_1 \oplus B_2,$$

$$F_1(y) \in B_1,$$

$$F_2 y - c \in B_2.$$

Suppose $F_1(y)$ have the weak Fréchet derivative $J(y)$ continuous on Ω , and the system of the equations

$$(1.5) \quad \begin{cases} F_1(y) = 0, \\ F_2 y - c = 0, \end{cases}$$

has an approximate solution $y = \bar{y} \in \Omega$. In addition, suppose that there are an additive operator L mapping E into B_1 , a positive number δ , and a non-negative $\kappa < 1$ such that

$$(1.6) \quad L \oplus F_2 \text{ has a linear inverse operator } (L \oplus F_2)^{-1},$$

$$\Omega_\delta = \{y: \|y - \bar{y}\|_e \leq \delta, y \in E\} \subset \Omega,$$

$$(1.7) \quad \|J(y) - L\|_b \leq \kappa/M_1 \quad \text{on } \Omega_\delta,$$

$$(M_1 r + M_2 \varepsilon)/(1 - \kappa) \leq \delta.$$

Here $r \geq 0$, $\varepsilon \geq 0$, $M_1 > 0$ and $M_2 > 0$ are numbers such that

$$(1.8) \quad \begin{aligned} & \|F_1(\bar{y})\|_b \leq r, \quad \|F_2\bar{y} - c\|_b \leq \varepsilon, \\ & \sup_{\substack{\bar{z} \in B_1^B \\ \|\bar{z}\|_b = 1}} \|(L \oplus F_2)^{-1} \bar{z}\|_e \leq M_1, \quad \sup_{\substack{\bar{z} \in B_2^B \\ \|\bar{z}\|_b = 1}} \|(L \oplus F_2)^{-1} \bar{z}\|_e \leq M_2. \end{aligned}$$

Then the Newton iterative process

$$(1.9) \quad y_{n+1} = (L \oplus F_2)^{-1} \{(Ly_n - F_1(y_n)) \oplus c\}$$

with $y_0 = \bar{y}$ yields a fundamental sequence $\{y_n\}$, $n=0, 1, 2, \dots$ in Ω_δ and we have

$$(1.10) \quad \|y_n - \bar{y}\|_e \leq (M_1 r + M_2 \varepsilon) / (1 - \kappa), \quad n=0, 1, 2, \dots$$

If the above sequence y_n , $n=0, 1, 2, \dots$, converges to the point $\hat{y}$ in E , then $\hat{y}$ is a unique solution of (1.5) in Ω_δ and we have

$$(1.11) \quad \|\hat{y} - \bar{y}\|_e \leq (M_1 r + M_2 \varepsilon) / (1 - \kappa).$$

PROOF. From (1.9), follows

$$\begin{aligned} \|y_1 - \bar{y}\|_e &= \|(L \oplus F_2)^{-1} (L\bar{y} - F_1(\bar{y})) \oplus c\|_e - \|\bar{y}\|_e \\ &= \|(L \oplus F_2)^{-1} [\{(L\bar{y} - F_1(\bar{y})) \oplus c\} - (L \oplus F_2)\bar{y}]\|_e \\ &= \|(L \oplus F_2)^{-1} \{(-F_1(\bar{y})) \oplus (-F_2\bar{y} + c)\}\|_e \end{aligned}$$

and, by (1.8), follows

$$(1.12) \quad \|y_1 - \bar{y}\|_e \leq M_1 r + M_2 \varepsilon,$$

Further we have

$$y_{n+1} - y_n = (L \oplus F_2)^{-1} \left[\int_0^1 \{L - J(y_n + \theta(y_n - y_{n-1}))\} d\theta (y_n - y_{n-1}) \oplus 0 \right]$$

and, by (1.7),

$$(1.13) \quad \|y_{n+1} - y_n\|_e \leq \kappa \|y_n - y_{n-1}\|, \quad n=1, 2, \dots$$

From (1.12) and (1.13), Proposition 1 is proved similarly as Theorem 1 in [2].

REMARK 2. We call an additive operator $J(y)$, which maps E into B , the weak Fréchet derivative of $F(y)$ at $y \in \Omega$, if $J(y)$ satisfies

$$\|F(y+g) - F(y) - J(y)g\|_b / \|g\|_e \longrightarrow 0, \text{ as } \|g\|_e \longrightarrow 0, \quad (g \in E).$$

§2. Calculation of the weak Fréchet derivative of $F(y)$

In what follows, we assume that the boundary value problem (0.4) with (0.5) has an approximate solution $\bar{y}=(\bar{x}^j(\cdot), j=1, \dots, r, \bar{t}_j, j=1, \dots, r-1)$ and the assumptions (a)~(g) are satisfied.

We put

$$F_1(y) = \left[\frac{dx}{d\tau} - X(x, \tau, T), f(x, T) \right],$$

$$F_2 y = l_3 x.$$

It follows from the smoothness of $Y^j(x, t)$, $j=1, \dots, r$ and of $f^j(x, t)$, $j=1, \dots, r-1$ that $\frac{dx}{d\tau} - X(x, \tau, T)$ and $f(x, T)$ are continuously weak Fréchet differentiable at $[x(\cdot), T] \in \Omega$.

$F_1(y)$ has the weak Fréchet derivative $J(y)$, continuous on Ω , which is given by

$$(2.1) \quad J(y)g = \left[\frac{dh}{dt} - X_x(x, \tau, T)h - X_T(x, \tau, T)\mathcal{T}, f_x(x, T)h \right. \\ \left. + f_T(x, T)\mathcal{T} \right],$$

with $g=(h, T) \in E$. Here

$$(2.2) \quad X_x(x, \tau, T) = \begin{bmatrix} Y_x^1(\tau) \frac{t_1 - t_0}{2} & & \\ & \ddots & \\ & & Y_x^r(\tau) \frac{t_r - t_{r-1}}{2} \end{bmatrix},$$

$$(2.3) \quad X_T(x, \tau, T) = \begin{bmatrix} a^1 & & \\ b^2 & a^2 & \\ & \ddots & \ddots \\ b^3 & & a^{r-1} \\ & & \ddots & \ddots \\ & & & a^r & \\ & & & b^r & \end{bmatrix},$$

with

$$(2.4) \quad \begin{aligned} a^j &= \frac{t_j - t_{j-1}}{2} \frac{1 + \tau}{2} Y_t^j(\tau) + \frac{1}{2} Y^j(\tau), \quad j = 1, \dots, r-1, \\ b^j &= \frac{t_j - t_{j-1}}{2} \frac{1 - \tau}{2} Y_t^j(\tau) - \frac{1}{2} Y^j(\tau), \quad j = 2, \dots, r. \end{aligned}$$

$$(2.5) \quad f_x(x, T)h(\tau) = \begin{bmatrix} f_x^1 & & \\ & \ddots & \\ & & f_x^{r-1} & 0_{1,n} \end{bmatrix} h(1),$$

for $h(\tau) \in C[I]$,

$$(2.6) \quad f_T(x, T) = \begin{bmatrix} f_t^1 & 0 \\ & \ddots \\ 0 & f_t^{r-1} \end{bmatrix},$$

where

$$\begin{aligned} Y_x^j(\tau) &= Y_x(x^j(\tau), t_j(\tau)) \\ Y_t^j(\tau) &= Y_t(x^j(\tau), t_j(\tau)) \\ f_x^j &= f_x^j(x^j(1), t_j), f_t^j = f_t^j(x^j(1), t_j). \end{aligned}$$

It holds that, for $g = (h, \mathcal{T}) \in E$,

$$(2.7) \quad \begin{aligned} \|J(y)g - J(\bar{y})g\|_b &\leq \|X_x(x, \cdot, T) - X_x(\bar{x}, \cdot, \bar{T})\|_n \|h\|_n \\ &\quad + \|X_T(x, \cdot, T) - X_T(\bar{x}, \cdot, \bar{T})\|_n \|\mathcal{T}\| \\ &\quad + \|f_x(x, T) - f_x(\bar{x}, \bar{T})\| \|h\|_n + \|f_T(x, T) - f_T(\bar{x}, \bar{T})\| \|\mathcal{T}\| \\ &= \max(\|X_x(x, \tau, T) - X_x(\bar{x}, \tau, \bar{T})\|_n + \|f_x(x, T) - f_x(\bar{x}, \bar{T})\|, \\ &\quad \|X_T(x, T) - X_T(\bar{x}, \bar{T})\|_n + \|f_T(x, T) - f_T(\bar{x}, \bar{T})\|) \|g\|_e, \end{aligned}$$

where

$$\begin{aligned} \|X_x(x, \cdot, T) - X_x(\bar{x}, \cdot, \bar{T})\|_n &= \left(\sum_{j=1}^r \left\| Y_x^j(\cdot) \frac{t_j - t_{j-1}}{2} - \bar{Y}_x^j(\cdot) \frac{\bar{t}_j - \bar{t}_{j-1}}{2} \right\|_n^2 \right)^{1/2} \\ \|X_T(x, \cdot, T) - X_T(\bar{x}, \cdot, \bar{T})\|_n &\leq \sqrt{2} \left[\sum_{j=1}^r \left\{ \left\| Y_t^j(\cdot) \frac{t_j - t_{j-1}}{2} - \bar{Y}_t^j(\cdot) \frac{\bar{t}_j - \bar{t}_{j-1}}{2} \right\|_n^2 \right. \right. \\ &\quad \left. \left. + \frac{1}{2} \|Y^j(\tau) - \bar{Y}^j(\tau)\|_n^2 \right\} \right]^{1/2} \\ \|f_x(x, T) - f_x(\bar{x}, \bar{T})\| &= \left(\sum_{j=1}^{r-1} \|f_x^j - \bar{f}_x^j\|^2 \right)^{1/2}, \end{aligned}$$

$$\|f_T(x, T) - f_T(\bar{x}, \bar{T})\| = \left(\sum_{j=1}^{r-1} |f'_j - \bar{f}'_j|^2 \right)^{1/2}$$

§3. The inverse of the weak Fréchet derivative of $F(y)$

The weak Fréchet derivative of $F(y)$ at $\bar{y}$ is an additive operator as follows.

$$(3.1) \quad (L \oplus F_2)g = \left[\frac{dh}{dt} - E(\tau)h - F(\tau)\mathcal{T}, l_1 h + l_2 \mathcal{T} \right] \oplus [l_3 h],$$

for $g = (h, \mathcal{T}) \in E$.

Here

$$L = J(\bar{y}), \quad E(\tau) = X_x(\bar{x}, \tau, \bar{T}), \quad F(\tau) = X_T(\bar{x}, \tau, \bar{T}),$$

$$l_1 = f_x(\bar{x}, \bar{T}), \quad l_2 = f_T(\bar{x}, \bar{T}).$$

Let $\Phi(\tau)$ be a fundamental matrix of the differential system

$$(3.2) \quad \frac{dh}{dt} = E(\tau)h$$

satisfying $\Phi(-1) = E_{nr, nr}$.

The following lemma gives the existence of the inverse operator of $(L \oplus F_2)$.

LEMMA 1. *If the matrix*

$$(3.3) \quad G = \begin{bmatrix} l_3 \Phi(\tau) & l_3 \Phi(\tau) \int_{-1}^{\tau} \Phi^{-1}(s) F(s) ds \\ l_1 \Phi(\tau) & l_1 \Phi(\tau) \int_{-1}^{\tau} \Phi^{-1}(s) F(s) ds + l_2 \end{bmatrix}$$

is nonsingular, that is,

$$(3.4) \quad \det G \neq 0,$$

then the operator $(L \oplus F_2)$ has an inverse operator $(L \oplus F_2)^{-1}$ and it holds that

$$(3.5) \quad \sup_{\substack{z \in B_1 \\ \|z\|_b = 1}} \|(L \oplus F_2)^{-1} z\|_e \leq (\|H_1\|_n + \|K_1\|, \|H_2\|_n + \|K_2\|),$$

$$\sup_{\substack{z \in B_2 \\ \|z\|_b = 1}} \|(L \oplus F_2)^{-1} z\|_e \leq \|H_2\|_n + \|K_2\|,$$

where H_1 is a linear operator from $C[I]$ into $C^1[I]$ given by

$$(3.6) \quad H_1 \varphi = \Phi(\tau) \int_{-1}^{\tau} \Phi^{-1}(s) \varphi(s) ds - AG^{-1} \begin{bmatrix} l_3 \Phi(\tau) \int_{-1}^{\tau} \Phi^{-1}(s) \varphi(s) ds \\ l_1 \Phi(\tau) \int_{-1}^{\tau} \Phi^{-1}(s) \varphi(s) ds \end{bmatrix},$$

with

$$(3.7) \quad A = [\Phi(\tau), \Phi(\tau) \int_{-1}^{\tau} \Phi^{-1}(s) F(s) ds],$$

K_1 a linear operator from $C[I]$ into R^{r-1} by

$$(3.8) \quad K_1 = -[0_{r-1, nr}, E_{r-1, r-1}] G^{-1} \begin{bmatrix} l_3 \Phi(\tau) \int_{-1}^{\tau} \Phi^{-1}(s) \varphi(s) ds \\ l_1 \Phi(\tau) \int_{-1}^{\tau} \Phi^{-1}(s) \varphi(s) ds \end{bmatrix},$$

H_2^1 a linear operator from R^{r-1} into $C^1[I]$ by

$$(3.9) \quad H_2^1 c_1 = AG^{-1} \begin{bmatrix} O_{nr, r-1} \\ E_{r-1, r-1} \end{bmatrix} c_1,$$

H_2^2 a linear operator from R^{nr} into $C^1[I]$ by

$$(3.10) \quad H_2^2 c_2 = AG^{-1} \begin{bmatrix} E_{nr, nr} \\ O_{r-1, nr} \end{bmatrix} c_2,$$

and K_2^1, K_2^2 are given by

$$(3.11) \quad K_2^1 = [0_{r-1, nr}, E_{r-1, r-1}] G^{-1} \begin{bmatrix} O_{nr, r-1} \\ E_{r-1, r-1} \end{bmatrix},$$

$$(3.12) \quad K_2^2 = [0_{r-1, nr}, E_{r-1, r-1}] G^{-1} \begin{bmatrix} E_{nr, nr} \\ O_{r-1, nr} \end{bmatrix}.$$

PROOF. Put

$$(L \oplus F_2)g = z = (\varphi, c_1, c_2),$$

then, by (3.1), we have

$$(3.13) \quad \frac{dh}{d\tau} = E(\tau)h + F(\tau)\mathcal{T} + \varphi(\tau),$$

$$(3.14) \quad \begin{cases} l_1 h + l_2 \mathcal{T} = c_1, \\ l_3 h = c_2, \end{cases}$$

The general solution of (3.13) is, by (3.7),

$$(3.15) \quad \begin{aligned} h(\tau) &= \Phi(\tau)a + \Phi(\tau) \int_{-1}^{\tau} \Phi^{-1}(s)F(s)ds \mathcal{T} + \Phi(\tau) \int_{-1}^{\tau} \Phi^{-1}(s)\varphi(s)ds \\ &= A \begin{bmatrix} a \\ \mathcal{T} \end{bmatrix} + \Phi(\tau) \int_{-1}^{\tau} \Phi^{-1}(s)\varphi(s)ds \end{aligned}$$

where a is an arbitrary constant vector.

We substitute (3.15) into (3.14):

$$\begin{aligned} [l_3 \Phi(\tau)]a + l_3 \left[\Phi(\tau) \int_{-1}^{\tau} \Phi^{-1}(s)F(s)ds \right] \mathcal{T} + l_3 \left[\Phi(\tau) \int_{-1}^{\tau} \Phi^{-1}(s)\varphi(s)ds \right] &= c_2, \\ [l_1 \Phi(\tau)]a + \left\{ l_1 \left[\Phi(\tau) \int_{-1}^{\tau} \Phi^{-1}(s)F(s)ds \right] + l_2 \right\} \mathcal{T} + l_1 \left[\Phi(\tau) \int_{-1}^{\tau} \Phi^{-1}(s)\varphi(s)ds \right] &= c_1. \end{aligned}$$

It is readily seen from (3.3) that $\begin{bmatrix} a \\ \mathcal{T} \end{bmatrix}$ becomes as follows.

$$(3.16) \quad \begin{bmatrix} a \\ \mathcal{T} \end{bmatrix} = G^{-1} \begin{bmatrix} c_2 - l_3 \Phi(\tau) \int_{-1}^{\tau} \Phi^{-1}(s)\varphi(s)ds \\ c_1 - l_1 \Phi(\tau) \int_{-1}^{\tau} \Phi^{-1}(s)\varphi(s)ds \end{bmatrix}.$$

Thus

$$\mathcal{T} = [O_{r-1, nr}, E_{r-1, r-1}] G^{-1} \begin{bmatrix} c_2 - l_3 \Phi(\tau) \int_{-1}^{\tau} \Phi^{-1}(s)\varphi(s)ds \\ c_1 - l_1 \Phi(\tau) \int_{-1}^{\tau} \Phi^{-1}(s)\varphi(s)ds \end{bmatrix},$$

that is, by (3.8), (3.11) and (3.12),

$$(3.17) \quad \mathcal{T} = K_1 \varphi + K_1^2 c_1 + K_2^2 c_2.$$

We substitute (3.16) into (3.15):

$$h(\tau) = AG^{-1} \begin{bmatrix} c_2 - I_3 \Phi(\tau) \int_{-1}^{\tau} \Phi^{-1}(s) \varphi(s) ds \\ c^1 - I_1 \Phi(\tau) \int_{-1}^{\tau} \Phi^{-1}(s) \varphi(s) ds \end{bmatrix} + \Phi(\tau) \int_{-1}^{\tau} \Phi^{-1}(s) \varphi(s) ds ;$$

hence, by (3.6), (3.9) and (3.10),

$$(3.18) \quad h(\tau) = H_1 \varphi + H_2^1 c_1 + H_2^2 c_2.$$

Combining of (3.17) and (3.18) evidently implies the existence of the linear inverse operator $(L \oplus F_2)^{-1}$.

From (3.17) and (3.18) it follows that

$$\begin{aligned} \sup_{\substack{z \in B_1 \\ \|z\|_b = 1}} \| (L \oplus F_2)^{-1} z \|_e &= \| H_1 \varphi + H_2^1 c_1 \|_n + \| K_1 \varphi + K_2^1 c_1 \| \\ &\leq \max(\| H_1 \|_n + \| K_1 \|, \| H_2^1 \|_n + \| K_2^1 \|) \| z \|_b, \end{aligned}$$

where $z = (\varphi, c_1, 0) \in B_1$.

In addition, it follows that

$$\sup_{\substack{z \in B_2 \\ \|z\|_e = 1}} \| (L \oplus F_2)^{-1} z \|_e = \| H_2^2 c_2 \|_n + \| K_2^2 c_2 \| \leq (\| H_2^2 \|_n + \| K_2^2 \|) \| z \|_b,$$

where $z = (0, 0, c_2) \in B_2$.

§4. The matrices G and A

The submatrices which constitute the matrix A , are calculated as follows: by (3.2) and (0.6),

$$(4.1) \quad \Phi(\tau) = \begin{bmatrix} \Phi_1(\tau) & & 0 \\ & \ddots & \\ 0 & & \Phi_r(\tau) \end{bmatrix},$$

and, by (2.3),

$$(4.2) \quad \Phi(\tau) \int_{-1}^{\tau} \Phi^{-1}(s) F(s) ds = \begin{bmatrix} A^1(\tau) & & 0 \\ B^2(\tau) A^2(\tau) & & \\ & \ddots & \\ & & A^{r-1}(\tau) \\ 0 & & & B^r(\tau) \end{bmatrix},$$

where

$$(4.3) \quad \left\{ \begin{array}{l} A^j(\tau) = \Phi_j(\tau) \int_{-1}^{\tau} \Phi_j^{-1}(s) \bar{a}^j(s) ds, \\ \text{with } \bar{a}^j(s) = \frac{\bar{i}_j - \bar{i}_{j-1}}{2} \frac{1+s}{2} \bar{Y}_t^j(s) + \frac{1}{2} \bar{Y}^j(s), \\ \hspace{25em} j=1, \dots, r-1, \\ \\ B^j(\tau) = \Phi_j(\tau) \int_{-1}^{\tau} \Phi_j^{-1}(s) \bar{b}^j(s) ds, \\ \text{with } \bar{b}^j(s) = \frac{\bar{i}_j - \bar{i}_{j-1}}{2} \frac{1-s}{2} \bar{Y}_t^j(s) - \frac{1}{2} \bar{Y}^j(s), \\ \hspace{25em} j=2, \dots, r. \end{array} \right.$$

The submatrices which constitute G , are calculated as follows:

$$(4.4) \quad l_3 \Phi(\tau) = \begin{bmatrix} L_0 & L_1 \Phi_r(1) \\ \Phi_1(1), -E_{nn} & \ddots \\ & \Phi_{r-1}(1), -E_{nn} \end{bmatrix},$$

$$(4.5) \quad l_3 \Phi(\tau) \int_{-1}^{\tau} \Phi^{-1}(s) F(s) ds = \begin{bmatrix} A^1(1) & L_1 B^r(1) \\ B^2(1) & A^2(1) \\ & \ddots \\ B^{r-1}(1) & A^{r-1}(1) \end{bmatrix},$$

and, by (2.5) and (2.6),

$$(4.6) \quad l_1 \Phi(\tau) = \begin{bmatrix} \bar{f}_x^1 \Phi_1(1) & & 0 \\ & \ddots & \\ 0 & & \bar{f}_x^{r-1} \Phi_{r-1}(1) & 0_{1,n} \end{bmatrix},$$

$$(4.7) \quad l_1 \Phi(\tau) \int_{-1}^{\tau} \Phi^{-1}(s) F(s) ds + l_2 = \begin{bmatrix} \bar{f}_x^1 A^1(1) + \bar{f}_t^1 & & 0 \\ \bar{f}_x^2 B^2(1), \bar{f}_x^2 A^2(1) + \bar{f}_t^2 & & \\ & \ddots & \\ 0 & \bar{f}_x^{r-1} B^{r-1}(1), \bar{f}_x^{r-1} A^{r-1}(1) + \bar{f}_t^{r-1} \end{bmatrix}.$$

By the following lemma, we can simplify the matrices A and G .

LEMMA 2. For $A^j(\tau)$ and $B^j(\tau)$ in (4.3), we have

$$(4.8) \quad \left\{ \begin{array}{l} \left\| A^j(\tau) - \bar{Y}^j(\tau) \frac{\tau+1}{2} \right\|_n \leq K_1 r, \quad j=1, \dots, r-1 \\ \left\| B^j(\tau) - \left\{ \bar{Y}^j(\tau) \frac{1-\tau}{2} - \Phi_j(\tau) \bar{Y}^j(-1) \right\} \right\|_n \leq K_2 r, \quad j=2, \dots, r \end{array} \right.$$

where

$$K_1 = \sqrt{2} \left[\max_{\substack{\tau \in I \\ j=1, \dots, r-1}} \int_{-1}^{\tau} \left\| \Phi_j(\tau) \Phi_j^{-1}(s) \bar{Y}_x^j(s) \frac{s+1}{2} \right\|^2 ds \right]^{1/2},$$

$$K_2 = \sqrt{2} \left[\max_{\substack{\tau \in I \\ j=2, \dots, r}} \int_{-1}^{\tau} \left\| \Phi_j(\tau) \Phi_j^{-1}(s) \bar{Y}_x^j(s) \frac{1-s}{2} \right\|^2 ds \right]^{1/2}.$$

PROOF. It holds that

$$(4.9) \quad \begin{aligned} A^j(-1) &= 0, \\ \bar{Y}^j(-1) \frac{(-1)+1}{2} &= 0. \end{aligned}$$

$\bar{Y}^j(\tau) \frac{\tau+1}{2}$ is a solution of the system

$$(4.10) \quad \begin{aligned} \frac{dy}{d\tau} &= \bar{Y}_x^j(\tau) \frac{\bar{t}_j - \bar{t}_{j-1}}{2} y + \bar{a}^j(\tau) \\ &+ \bar{Y}_x^j(\tau) \left(\frac{d\bar{x}^j(\tau)}{d\tau} - \bar{Y}^j(\tau) \frac{\bar{t}_j - \bar{t}_{j-1}}{2} \right) \frac{\tau+1}{2} \end{aligned}$$

and $A^j(\tau)$ is a solution of the system

$$(4.11) \quad \frac{dy}{d\tau} = \bar{Y}_x^j(\tau) \frac{t_j - t_{j-1}}{2} y + \bar{a}_j(\tau).$$

Thus from (4.9), (4.10) and (4.11) it follows that

$$(4.12) \quad \begin{aligned} A^j(\tau) - \bar{Y}^j(\tau) \frac{\tau+1}{2} \\ = \Phi_j(\tau) \int_{-1}^{\tau} \Phi_j^{-1}(s) \bar{Y}_x^j(s) \left(\frac{d\bar{x}^j}{ds} - \bar{Y}^j(s) \frac{\bar{t}_j - \bar{t}_{j-1}}{2} \right) \frac{s+1}{2} ds. \end{aligned}$$

The first inequality of (4.8) follows from (4.12). The second of (4.8) can be proved similarly.

In addition to (4.8), $Y^{j-1}(1)$ is equal to $Y^j(-1)$ except for the error of $O(\varepsilon)$ for $j=2, \dots, r$. Since the residual errors r and ε we consider here are sufficiently small, A and G are rewritten as (0.8) and (0.9) respectively, neglecting those errors.

From (0.7) it follows that

$$(4.13) \quad \det G = \det(L_0 + L_1 \Phi_r(1)) \prod_{j=1}^{r-1} (\bar{f}_x^j \bar{Y}^j(1) + \bar{f}_t^j).$$

Thus conditions (a) and (b) imply that G is nonsingular and, by Lemma 1, there exists the inverse operator $(L \oplus F_2)^{-1}$.

For a general function $\varphi(s) \in C[I]$, it holds that

$$\begin{bmatrix} l_3 \Phi(\tau) \int_{-1}^{\tau} \Phi^{-1}(s) \varphi(s) ds \\ l_1 \Phi(\tau) \int_{-1}^{\tau} \Phi^{-1}(s) \varphi(s) ds \end{bmatrix} = \begin{bmatrix} L_1 \Phi_r(1) \int_{-1}^1 \Phi_r^{-1}(s) \varphi_r(s) ds \\ \Phi_1(1) \int_{-1}^1 \Phi_1^{-1}(s) \varphi_1(s) ds \\ \vdots \\ \Phi_{r-1}(1) \int_{-1}^1 \Phi_{r-1}^{-1}(s) \varphi_{r-1}(s) ds \\ \bar{f}_x^1 \Phi_1(1) \int_{-1}^1 \Phi_1^{-1}(s) \varphi_1(s) ds \\ \vdots \\ \bar{f}_x^{r-1} \Phi_{r-1}(1) \int_{-1}^1 \Phi_{r-1}^{-1}(s) \varphi_{r-1}(s) ds \end{bmatrix}$$

and

$$(4.14) \quad \left\| \begin{bmatrix} l_3 \Phi(\tau) \int_{-1}^{\tau} \Phi^{-1}(s) \varphi(s) ds \\ l_1 \Phi(\tau) \int_{-1}^{\tau} \Phi^{-1}(s) \varphi(s) ds \end{bmatrix} \right\|_n \leq \sqrt{2} k_3 \left(\sum_{j=1}^r \|\varphi_j\|_n^2 \right)^{1/2}.$$

Thus from (3.5), by (3.6)~(3.11) and (4.14), it follows that

$$(4.15) \quad \sup_{\|z\|_b^1=1} \|(L \oplus F_2)^{-1} z\|_e \leq M_1, \quad \sup_{\|z\|_b^2=1} \|(L \oplus F_2)^{-1} z\|_e \leq M_2.$$

From (2.7) and (d) it follows that

$$(4.16) \quad \|J(y) - L\|_b \leq \frac{\kappa}{M_1} \quad \text{on } \Omega_\delta.$$

§5. Proof of Theorem

For the operator equation

$$F(y)=0$$

and for given $\bar{y}$, L , F_1 , F_2 , it is easily seen from (1.4), (4.13), (4.15) and (4.16) that the conditions of Proposition 1 are all fulfilled.

Thus results of Proposition 1 hold.

Let $\{y_n\}$ be the fundamental sequence in (1.9). From the completeness of the space $C[I] \times R^{r-1}$, there exists $\hat{y}$ in $C[I] \times R^{r-1}$ such that

$$\|y_n - \hat{y}\|_e \longrightarrow 0,$$

where $\|\cdot\|_e$ is a norm of the space $C[I] \times R^{r-1}$ defined as the same way as E space.

Letting $n \rightarrow \infty$ in (1.9), we have

$$\hat{y} = (L \oplus F_2)^{-1} \{X(\hat{x}, \tau, \hat{T}) - E(\tau)\hat{x} - F(\tau)\hat{T}, l_2x - f(\hat{x}, \hat{T}) \oplus c\}.$$

Since

$$X(\hat{x}, \tau, \hat{T}) - E(\tau)\hat{x} - F(\tau)\hat{T} \in C[I],$$

$$l_2\hat{x} - f(\hat{x}, \hat{T}) \oplus c \in R^{r-1} \times R^{nr},$$

and $(L \oplus F_2)^{-1}$ a linear operator B into E , then

$$\hat{y} \in E.$$

Thus the fundamental sequence $\{y_n\}$ converges to $\hat{y}$ in E and $\hat{y}$ satisfies (1.11). From (1.11) follows (0.9) of Theorem.

The author greatly acknowledges suggestions and guidance of Dr. Minoru Urabe and the valuable discussions with Dr. Nagata Furukawa.

References

- [1] ИТОИ, К.: *On the numerical solvability of two point boundary value problems in a finite Chebyshev series for piecewise smooth differential system*, Memoirs of Numer. Math., **2**, (1975), 45-67.

- [2] URABE, M.: *The Newton method and its application to boundary value problems with non-linear boundary conditions*, Proceeding of U. S. Japan seminar on differential and functional equations, June, 1967, Benjamin, 383–410.