COLLOCATION FOR SYSTEMS OF ORDINARY DIFFERENTIAL EQUATIONS

by

John H. Cerutti

Computer Sciences Technical Report #230

December 1974

THE UNIVERSITY OF WISCONSIN
Computer Sciences Department
1210 West Dayton Street
Madison, Wisconsin 53706

Received: December 4, 1974

COLLOCATION FOR SYSTEMS OF ORDINARY DIFFERENTIAL EQUATIONS

by

John H. Cerutti

ABSTRACT

C. de Boor and B. Swartz have shown that by using C^{m-1} piecewise polynomials of order $k + m$ the C^{m+2k} solution of an m-th order non-linear differential equation can be approximated to within $O(|\Delta|^{k+m})$ globally and $O(|\Delta|^{2k})$ at the knots by collocation at Gaussian points. An extension of their results to systems of nonlinear equations is given.

COLLOCATION FOR SYSTEMS OF ORDINARY DIFFERENTIAL EQUATIONS*

by

John H. Cerutti

I. In [1], de Boor and Swartz discuss the use of collocation with piecewise polynomials to approximate the solution of an m-th order ordinary differential equation. Their results place collocation methods on an equal theoretical footing with other projection methods (e.g. Galerkin, least squares). They show that by careful choice of the collocation points the same order of convergence can be achieved with collocation methods as with Galerkin's or the least squares method. More precisely, given a partition Δ, they obtain the convergence rate, $O(|\Delta|^{k+m})$, when collocating at Gaussian points with splines of order $k+m$ that are in C^{m-1}. Furthermore, at the knots of Δ, the approximation to the solution and its first $m-1$ derivatives is $O(|\Delta|^{2k})$.

Wittenbrink [10] extends the results of [1] to include nonlinear boundary conditions. He treats simultaneously moment methods and collocation methods using the same techniques as de Boor and Swartz [1]. R. Weiss [9] studies implicit Runge-Kutta methods for first order systems and (i) showed their equivalence to appropriate collocation schemes, and (ii) obtained the corresponding high order convergence.

*Supported by the Office of Naval Research under Contract N00014-67-A-0128-0004

R. D. Russell [7] extended the work of de Boor and Swartz to first order systems. He also shows that Richardson's extrapolation can be applied at the knots when the mesh is uniform. Of course the system of differential equations treated in this paper can be written as a first order system, however, this reduction is not necessary in order to apply the collocation procedure described below. Moreover, the reduction to a first order system leads to a different algebraic problem, and hence, a different approximation procedure. In many instances there are natural reasons to deal with the higher order system rather than make the reduction to a first order system.

This note is a direct extension of the results of de Boor and Swartz [1] to systems of ordinary differential equations. In section 3 we follow Wittenbrink [10] and use a general theorem of Vainikko [8] to obtain $O(|\Delta|^k)$ error estimates. Section 4 follows de Boor and Swartz [1] to obtain higher order point-wise and global error estimates.

ACKNOWLEDGEMENT

We are indebted to Seymour Parter and Carl de Boor for many helpful discussions during the preparation of this report.

II. NOTATION

Let $\Delta = \{x_i\}_{i=0}^{N}$ be a partition of the interval $[a,b]$,

$$a = x_0 < x_1 < \cdots < x_N = b, \tag{2.1}$$

with

$$\Delta x_j = x_j - x_{j-1},\ j = 1,\ldots,N,$$

and

$$|\Delta| = \max_j \Delta x_j .$$

Let $m = (m_1,\ldots,m_d)$ be a d-multi-index and denote by $C^m[a,b]$ the space of d-dimensional real vector valued functions on $[a,b]$, where $u = (u^1,\ldots,u^d)^T \varepsilon\, C^m[a,b]$ means $u^j \varepsilon\, C^{m_j}[a,b]$, $j = 1,\ldots,d$. That is, each u^j has m_j continuous derivatives on $[a,b]$. More generally,

$$u \,\varepsilon\, C_\Delta^m = C^m[x_0,x_1] \times \cdots \times C^m[x_{N-1},x_N]$$

is a function of N pieces where $u = (u^1,\ldots,u^d)$ and each u^j is in $C^{m_j}[x_0,x_1] \times \cdots \times C^{m_j}[x_{N-1},x_N]$. We think of each u^j, $j = 1,\ldots,d$, as having two values at each of the interior breakpoints $\{x_j\}_{j=1}^{N-1}$, and on the k'th interval we denote u^j by u_k^j. $C_\Delta^0[a,b]$ is a Banach space with respect to the norm

$$\|u\| \equiv \max_{1\le j\le d}\ \max_{1\le i\le N}\ \max_{x\varepsilon[x_{i-1},x_i]} |u_i^j(x)|. \tag{2.2}$$

Let S_i, $i = 1,\ldots,N$, be defined by

$$(S_i u)(x) = u\left(\frac{x_i+x_{i-1}}{2} + x\,\frac{x_i-x_{i-1}}{2}\right),\ -1 \le x \le 1. \tag{2.3}$$

Hence S_i is a linear map from $C^0[x_{i-1},x_i]$ to $C^0[-1,1]$. Thus if P is a map on $C^0[-1,1]$ we may use S_i to obtain a map P_Δ on $C^0_\Delta[a,b]$ by setting

$$\left.\begin{aligned} (P_\Delta u)(x) = (S_i^{-1} P S_i f)(x),\ x \in [x_{i-1},x_i], \\ i = 1,\ldots,N. \end{aligned}\right\} \tag{2.4}$$

When P is a bounded linear map then so is P_Δ and

$$\|P_\Delta\| = \|P\| . \tag{2.5}$$

We denote by $\mathbb{P}_{m,\Delta}$ the linear subspace of $C^0_\Delta[a,b]$ consisting of those $u \in C^0_\Delta$ whose j'th component is a pieceiwse polynomial of order m_j (degree strictly less than m_j), i.e. $u = (u^1,\ldots,u^d) \in \mathbb{P}_{m,\Delta}$ means that u^j is a polynomial of order m_j on $[x_{i-1},x_i]$, $i = 1,\ldots,N$; $j = 1,\ldots,d$. We are particularly interested in P_Δ when P is a linear projector from $C^0[-1,1]$ onto $\mathbb{P}_m$. In this case we obtain a bound on $\|u-Pu\|$ since (see e.g. [4] pg. 338)

$$\left.\begin{aligned} \|u-Pu\| &\le \|I-P\| \inf_{p\in\mathbb{P}_m} \|u-p\| \\ &\le \text{const } \|I-P\|\,\omega_u\left(\frac{1}{(\underline{m}-1)}\right) \end{aligned}\right\} \tag{2.6}$$

where ω_u , the modulus of continuity, is defined by

$$\omega_u(\delta) = \max_{1\le j\le d} \{\sup |u^j(s)-u^j(t)| : s,\ t\in[a,b]\ |s-t|\le\delta\}$$

and

$$\underline{m} \equiv \min_{1\le j\le d} m_j .$$

Hence for all $u \in C^0_\Delta[a,b]$

$$\left.\begin{aligned} \|u-P_\Delta u\| &\le \|I-P\| \inf_{p\in\mathbb{P}_{m,\Delta}} \|u-p\| \\ &= \|I-P\| \operatorname{dist}(u,\mathbb{P}_{m,\Delta}) \end{aligned}\right\} \tag{2.7}$$

and thus for $u \in C^0[a,b]$

$$\|u-P_\Delta u\| \le \text{const } \|I-P\|\, \omega_u\left(\frac{|\Delta|}{(\underline{m}-1)}\right). \tag{2.8}$$

We will use the following conventions: if m is a d-multi-index and i is an integer then $m - i \equiv (m_1-i,\ldots,m_d-i)$, $m' \le m$ iff $m_i' \le m_i$ $i = 1,\ldots,d$; if an integer i appears where a multi-index is necessary interpret it as a multi-index each of whose terms is i (e.g. $D^i u \equiv \operatorname{diag}(D^i,\ldots,D^i)u$, when u is a vector). Also let $\bar{m} \equiv \max_i m_i$ and $\underline{m} \equiv \min_i m_i$. We will assume throughout that the equations have been ordered so that $m_1 \le m_2 \le \cdots \le m_d$.

III. COLLOCATION AND PRELIMINARY ESTIMATES

Consider the system of nonlinear differential equations

$$\left.\begin{array}{l} D^m u(x) = (Fu)(x) \quad x \in [a,b] \\ \text{subject to the linear boundary conditions} \\ B_i u = c_i \quad i = 1,\ldots,|m| \ . \end{array}\right\} \tag{3.1}$$

Here $D^m = \operatorname{diag}(D^{m_1},\ldots,D^{m_d})$ is a $d \times d$ matrix, $\{B_i\}_1^{|m|}$ is a known sequence of continuous linear maps from $C^{m-1}[a,b]$ to $\mathbb{R}^1$, and c_i is a known sequence of constants. For example, the B_i could be of the form $\sum_j C_j(u^1,\ldots,D^{m_d-1}u^d)^T(s_j)$ where each C_j is a $|m| \times |m|$ matrix and $\{s_j\}$ is a finite set of points of $[a,b]$. More general boundary conditions are possible and we discuss this at the end of section III.

We assume $F = (F^1,\ldots,F^d)^T$ is of the form

$$\left.\begin{array}{l} F^j(u(x)) = F^j(x,u^1,D^1u^1,\ldots,D^{m_1-1}u^1,\ldots,u^d,\ldots,D^{m_d-1}u^d) \\ \qquad\qquad\qquad\qquad j = 1,\ldots,d \end{array}\right\} \tag{3.2}$$

(i.e. F^j depends on u^k and at most $m_k - 1$ derivatives of u^k, $k = 1,\ldots,d$). Thus F is a d-vector valued map on a suitable domain of $\mathbb{R}^{|m|+1}$. We assume that (3.1) has a solution $u(x) \in C^m[a,b]$.

If $v \in W^{m,1}[a,b] \equiv \{v \in C^{(m-1)}[a,b],\ D^{m-1}v \text{ abs. cont.}, D^m v \in L^1[a,b]\}$ then each component v^j of v can be written in exactly one way in terms of $D^{m_j} v^j$ and $(D^i v^j)(a)$ $i = 0,\ldots,m_{j-1}$. That is

$$\left.\begin{aligned} D^i v^j(x) &= \int_a^x \frac{(x-s)^{m_j-i-1}}{(m_j-i-1)!} D^{m_j} v^j(s)\,ds + \sum_{k=i}^{m_j-1} a_k^i \frac{(x-a)^{k-1}}{(k-i)!} \\ &\qquad j = 1,\dots,d,\ i = 0,\dots,m_j-1, \\ \text{where}\quad a_k^j &= (D^k v^j)(a) \quad j = 1,\dots,d,\ k = 0,\dots,m_j-1. \end{aligned}\right\} \quad (3.3)$$

Let $H_i^j v(x)$ denote the right side of (3.3). Hence if $w(x) = D^m v$ then v is uniquely determined by the $|m| + d$ dimensional vector

$$y(x) \equiv (w(x), a_0^1, \dots, a_{m_1-1}^1, \dots, a_0^d, \dots, a_{m_d-1}^d). \quad (3.4)$$

Conversely, given a function $w \in L^1[a,b]$ and $|m|$ constants $\{a^i\}_{i=1}^{|m|}$ the $|m| + d$ vector $\{w, (a^i)_1^{|m|}\}$ uniquely determines an element v of $W^{m,1}[a,b]$ via (3.3). With $w = D^m u$ and the representations (3.3), (3.4) we rewrite the differential equation (3.1) as

$$w(x) = F(x, H_0^1 y, H_1^1 y, \dots, H_{m_1-1}^1 y, \dots, H_1^d y, \dots, H_{m_d-1}^d y) \quad (3.5)$$

(i.e. $w(x) = F(y(x))$, and the boundary conditions as

$$\left.\begin{aligned} b_i = b_i + B_i\{(H_0^1 y, H_1^1 y, \dots, H_{m_i-1}^1 y), \dots, \\ (H_1^d y, \dots, H_{m_d-1}^d y)\}^T - c_i, \\ i = 1,\dots,|m|, \end{aligned}\right\} \quad (3.6)$$

(i.e. $b_i = b_i + B_i(y) - c_i$). If $y = (w, b_1, \dots, b_{|m|})$ we define T by

$$Ty = \{Fy, (b_1 + B_1(y) - c_1), \dots, (b_{|m|} + B_{|m|}(y) - c_{|m|})\}. \quad (3.7)$$

Hence (3.5) and (3.6) can be written in the form

$$y = Ty. \tag{3.8}$$

Notice that $C^0_\Delta[a,b] \times \mathbb{R}^{|m|}$ is a Banach space with the norm

$$\|y\|_m = \|w\| + \max_{1 \le i \le |m|} |e^i| \tag{3.9}$$

where $w \in C^0_\Delta[a,b]$, $e = (e^1,\ldots,e^{|m|}) \in \mathbb{R}^{|m|}$, and $y = (w,e)$.

We have $C^0[a,b] \times \mathbb{R}^{|m|} \subset C^0_\Delta[a,b] \times \mathbb{R}^{|m|} \subset W^{m,1}[a,b] \times \mathbb{R}^{|m|}$. T is a non-linear operator mapping each of these three linear spaces into itself. From the definition of T, solving (3.1) is equivalent to finding $y^* \in C^0[a,b] \times \mathbb{R}^{|m|}$ so that y^* is a fixed point of T (i.e. y^* satisfies (3.8)).

In order to find an approximation to $u(x)$, the solution of (3.1), we employ the collocation method.

Let $\Delta = \{x_i\}_{i=0}^N$ be a partition of [a,b] as in (2.1). Let k be a fixed positive integer and let $\{\rho_i\}_{i=1}^k$ be a fixed set of distinct points of [-1,1] with $-1 \le \rho_1 < \rho_2 < \ldots < \rho_k \le 1$. Set

$$\tau_{kj+i} = \begin{cases} x_j + & \text{if } -1 = \rho_1 \\ [(x_j+x_{j+1}) + \rho_i \Delta x_{j+1}]/2 & \\ x_{j+1}- & \text{if } +1 = \rho_k. \end{cases} \tag{3.10}$$

We seek a function

$$u_\Delta \in S_\Delta \equiv \mathbb{P}_{m+k,\Delta} \cap C^{m-1}[a,b]$$

which satisfies the collocation equations

$$\begin{aligned} (D^m u_\Delta)(\tau_i) &= F(u_\Delta)(\tau_i) \quad i = 1,\ldots,kN \\ B_i u_\Delta &= c_i \qquad\qquad i = 1,\ldots,|m| . \end{aligned} \tag{3.11}$$

Using the notation of section 2 we see that the points $\{\rho_i\}_{i=1}^{k}$ uniquely determine a bounded linear projector P from $C^0[-1,1]$ onto $\mathbb{P}_k[-1,1]$ described by the interpolation conditions

$$(Pu)(\rho_i) = u(\rho_i), \qquad i = 1,\ldots,k.$$

Furthermore, the points $\{\tau_i\}_{i=1}^{kN}$ and the partition Δ determine a bounded linear projector P_Δ from $C_\Delta^0[a,b]$ onto $\mathbb{P}_{k,\Delta}[a,b]$ described by the interpolation conditions

$$(P_\Delta u)^j(\tau_i) = u^j(\tau_i) \quad j = 1,\ldots,d; \; i = 1,\ldots,kN. \tag{3.12}$$

Finding a solution to the collocation equations (3.11) is equivalent to finding $y_\Delta \in \mathbb{P}_{k,\Delta}[a,b]$ which satisfies

$$y_\Delta = \tilde{P}_\Delta T y_\Delta \tag{3.13}$$

where $\tilde{P}_\Delta(w,a_0^1,\ldots,a_{m_d-1}^d) \equiv (P_\Delta w,a_0^1,\ldots,a_{m_d-1}^d)$.

As in Wittenbrink we use a slightly modified form of theorem 3 of Vainikko [8] to obtain existence and uniqueness of y_Δ in a neighborhood of y^*. Additionally, this theorem also gives error estimates that will allow us to obtain convergence rates for $\|u-u_\Delta\|$.

<u>Theorem 1 (Vainikko)</u>.

Let Ω be an open set contained in a Banach space E. Suppose $T:\Omega \to \Omega$ is a continuous, possibly nonlinear, map and $\{\tilde{P}_\Delta\}$ is a collection of bounded linear projectors defined on E with the following properties:

(1) there exists $y^* \in \Omega$ such that $y^* = Ty^*$,

(2) $\|(I-\tilde{P}_\Delta)y^*\| \to 0$ as $|\Delta| \to 0$,

(3) T is continuously Fréchet differentiable at y^* and $[I-T'(y^*)]^{-1}$ is bounded,

(4) $\tilde{P}_\Delta T$ is Fréchet differentiable in a neighborhood of y^* and for every $\varepsilon > 0$ there exists a δ_ε, $0 < \delta_\varepsilon < \delta$, and $\eta_\varepsilon > 0$ such that $\|(I-\tilde{P}_\Delta)T'(y)\| \le \varepsilon$ when $\|y-y^*\| \le \delta_\varepsilon$ and $|\Delta| \le \eta_\varepsilon$.

Then one can find a δ_0 and a η_0 such that the fixed point y^* of T is unique in the sphere $\{y| \, \|y-y^*\| \le \delta_0\}$, and when $|\Delta| \le \eta_0$ the equation

$$y_\Delta = \tilde{P}_\Delta T y_\Delta \tag{3.14}$$

has in this same sphere a unique solution y_Δ. Furthermore there exist $c_1, c_2 > 0$ such that

$$c_1 \|(I-\tilde{P}_\Delta)y^*\| \le \|y^*-y_\Delta\| \le c_2 \|(I-\tilde{P}_\Delta)y^*\| . \tag{3.15}$$

Remark: T is said to be continuously Fréchet differetiable at y^* if T is Fréchet differentiable in an open neighborhood $N(y^*)$ of y^* and for every $\varepsilon > 0$ there exists a $\delta > 0$ such that $\|T'(y) - T'(y^*)\| < \varepsilon$ when $y \in \{y| \, \|y-y^*\| < \delta\} \cap N(y^*)$.

Proof: Theorem 1 is a direct consequence of theorem 3 of Vainikko [8]. □

It follows from (2.5) that the family of continuous linear projectors with which we are concerned is bounded, (i.e. $\|\tilde{P}_\Delta y\| \le M \|y\|$ for some M independent of Δ). Since $(\tilde{P}_\Delta T)' = \tilde{P}_\Delta T'$ in our application, hypothesis (4) of theorem 1 follows from the assumption

$$\|(I-\tilde{P}_\Delta)T'(y^*)\| \le \varepsilon \quad \text{when} \quad |\Delta| \le \eta(\varepsilon). \tag{3.16}$$

To see this observe

$$\|(I-\tilde{P}_\Delta)T'(y)\| \le \|(I-\tilde{P}_\Delta)(T'(y)-T'(y^*))\| + \|(I-\tilde{P}_\Delta)T'(y^*)\| . \tag{3.17}$$

Since $\|\tilde{P}_\Delta\|$ is bounded so is $\|I-\tilde{P}_\Delta\|$, and the first term on the right side of (3.17) may be made small since T is continuously Frechet differentiable in a neighborhood of y^* (hypothesis (3)). That the second term may be made small follows from (3.16).

Corollary. In addition to the hypotheses of theorem 1 suppose that in some neighborhood $N(y^*)$ of y^* that

$$\|Ty_1-Ty_2-T'(y_2)(y_1-y_2)\| \leq \|y_1-y_2\|\,\omega(\|y_1-y_2\|)$$

uniformly for $y_1, y_2 \in N(y^*)$ where ω is some modulus of continuity.

Then there exists $\eta > 0$ so that the Newton iteration defined by

$$\mathcal{N}(y) \equiv y - [I-\tilde{P}_\Delta T'(y)]^{-1}(y-\tilde{P}_\Delta Ty) \tag{3.18}$$

is well defined for $|\Delta| < \eta$, $y \in N(y^*)$. Since

$$y_\Delta - \mathcal{N}(y) = [I-\tilde{P}_\Delta T'(y)]^{-1}\tilde{P}_\Delta[Ty_\Delta-Ty-T'(y)(y_\Delta-y)] \tag{3.19}$$

there is an $\varepsilon > 0$ so that $\mathcal{N}$ maps $\{y|\ \|y-y_\Delta\| \leq \varepsilon\} \cap N(y^*)$ into itself and

$$\|y_\Delta-y\| \leq C\|y_\Delta-y\|\,\omega(\|y_\Delta-y\|). \tag{3.20}$$

Proof.

Obviously

$$I-\tilde{P}_\Delta T'(y) = I-T'(y^*) + (I-\tilde{P}_\Delta)T'(y) + [T'(y^*)-T'(y)]. \tag{3.21}$$

Under hypotheses (3) and (4) of theorem 1 we choose a small neighborhood of y^*, $N(y^*)$ and a $\eta_0 > 0$ so that $y \in N(y^*)$, $|\Delta| < \eta_0$ imply

$$\|[I-T'(y^*)]^{-1}\|\,(\|I-\tilde{P}_\Delta T'(y)\| + \|T'(y^*)-T'(y)\|) < \frac{1}{2}.$$

Hence the Banach lemma (see e.g. [ ] pg.) shows that $I-\tilde{P}'_\Delta T(y)$ is uniformly boundedly invertible in a neighborhood of y^* for $|\Delta|$ sufficiently small. Thus $\mathcal{N}$ is well defined. We have, since y_Δ is a fixed point of $\tilde{P}_\Delta T$,

$$\left.\begin{aligned} y_\Delta - \mathcal{N}(y) &= y_\Delta - y + [I-\tilde{P}_\Delta T'(y)]^{-1}(y-\tilde{P}_\Delta Ty) \\ &= [I-\tilde{P}_\Delta T'(y)]^{-1}\{y-\tilde{P}_\Delta Ty + (I-\tilde{P}_\Delta T'(y))(y_\Delta-y)\} \\ &= [I-\tilde{P}_\Delta T'(y)]^{-1}\tilde{P}_\Delta\{Ty_\Delta-Ty-T'(y)(y_\Delta-y)\} \end{aligned}\right\} \tag{3.22}$$

which shows (3.19).

By (3.15) we may choose $\eta\,(\le\eta_0)$ so small that $y_\Delta \in N(y^*)$ (i.e. $\|y_\Delta - y\| < \eta$ for η sufficiently small).
We choose $\varepsilon > 0$ so that

$$\|[I-\tilde{P}_\Delta T'(y)]^{-1}\|.\ \|\tilde{P}_\Delta\|\,\omega(\varepsilon) < \min[1, \frac{\eta}{\|y_\Delta - y^*\|} - 1],$$

hence N maps $\{y|\ \|y-y^*\| \le \min(\eta,\varepsilon)\}$ into itself. Estimate (3.20) follows directly from (3.22) and the hypothesis of the corollary. □

This corollary is useful since it shows that in a neighborhood of y_Δ, the fixed point of $\tilde{P}_\Delta T$, Newton's method converges superlinearly to y_Δ. When we can show (as below) that the modulus of continuity, ω, of the corollary is $O(\|y_\Delta - y\|)$ then we obtain quadratic convergence.

Before we restate theorem 1 in terms of problem (3.1) we discuss conditions under which we may verify the hypotheses of the theorem and corollary.

We regard F as a d-vector valued map on some domain of $\mathbb{R}^{|m|+1}$ and write

$$F(x,z) = F(x,z_0^1,z_1^1,\ldots,z_{m_1-1}^1,\ldots,z_0^d,\ldots,z_{m_d-1}^d) \qquad (3.23)$$

where $z \in \mathbb{R}^{|m|}$. So if $D^i u^j$ is an argument of F (see (3.3)) then it corresponds to z_i^j. Suppose $F = (F^1,\ldots,F^d)$ and $F^j \in C^1(\Omega)$, $\Omega \subset \mathbb{R}^{|m|+1}$, $j = 1,\ldots,d$. Let $F_n(x,v)$, $n = 0,1,\ldots,\bar{m}-1 \equiv \max_j m_j - 1$, denote the $d \times d$ matrix

$$\left.\begin{aligned} [F_n(x,v)]_{kj} \equiv \frac{\partial F^k(x,z)}{\partial z_n^j}\Bigg|_{z_i^j = D^i v^j} \quad & k=1,\ldots,d; \\ & j=1,\ldots,d; \\ & n=0,\ldots,\bar{m}-1 \end{aligned}\right\} \qquad (3.24)$$

Let L_v denote the linear matrix differential operator

$$L_v \equiv D^m - \sum_{i=0}^{\bar{m}-1} F_i(x,v)D^i . \tag{3.25}$$

When $i \geq m_j$ and $k \leq i$ then $[F_i(x,v)]_{\ell k} = 0$ for $\ell = 1,\ldots,d$, so that (3.25) is of the same form as (3.1) (i.e. (3.3) is satisfied). L_v is the Fréchet derivative at v of the nonlinear operator defined by (3.1). The linear differential operator L_v also gives us the form of the Fréchet derivative of T. That is if $y(x)$ is of the form

$$y(x) = (w(x),e), \qquad e \in \mathbb{R}^{|m|}$$

then the derivative of T at y_1 acting on y is given by

$$T'(y_1)(y) = \Big(\sum_{i=0}^{\bar{m}-1} F_i(x,H_0^1 y_1,\ldots,H_{m_d-1}^d y_1)(H_i^1 y,\ldots,H_i^d y)^T, \tag{3.26}$$
$$e_1 + B_1(y),\ldots,e_{|m|} + B_{|m|}(y)\Big).$$

Hence, finding a fixed point of $T'(y)$ is equivalent to solving

$$L_w\, v = 0$$
$$B_i(v) = 0 \qquad\qquad i = 1,\ldots,|m|. \tag{3.27}$$

Furthermore, if $y = (D^m v,e)$ then solving the inhomogeneous equation

$$(I-T'(y_1))y = (g,a_1,\ldots,a_{|m|}) \in C_\Delta^0 \times \mathbb{R}^{|m|} \tag{3.28}$$

is equivalent to solving

$$\left.\begin{aligned} L_v v &= g \\ B_i(v) &= a_i \qquad i = 1,\ldots,|m| \end{aligned}\right\} \tag{3.29}$$

for some $v \in C^{m-1}[a,b]$.

We now reformulate thoerem 1 in terms of problem (3.1) rather than in terms of problem (3.8). Following de Boor and Swartz [1] we have

Theorem 2.

Let $u \in C^{m+n}[a,b]$ $n \geq 0$ be a solution of (3.1) (hence $(D^m u,c)$ is a solution of (3.8)) and suppose

(1) that F is sufficiently smooth near u for the hypothesis of the corollary to hold (e.g. Let $\Omega = \{(x,z) \mid a \leq x \leq b, z \in \mathbb{R}^{|m|}, |z_j^i - D^i u^j(x)| < \delta, j=1,\ldots,d, i=1,\ldots,m_j, \delta > 0\}$ and let $F \in C^2(\Omega)$)

(2) the linear problem

$$\left.\begin{aligned} L_u\, v &= g \\ B_i(v) &= 0 \end{aligned}\right\} \qquad (3.30)$$

has a unique solution;

(3) associated with the linear problem (3.30) is a $d \times d$ Green's matrix, $G(x,t)$, and

$$D^j v^k(x) = \int_a^b \sum_{i=1}^{d} \left(\frac{\partial}{\partial x}\right)^j [G(x,t)]_{ki}\, g^i(t)\,dt,$$

$$j = 0,\ldots,m_k-1, k=1,\ldots,d.$$

(See appendix 2 for a further discussion of this Green's matrix.)

Then there exist $\gamma > 0$, $\eta > 0$ so that

(1) in the sphere $\{(w,b) \mid (w,b) \in C_\Delta^0 \times \mathbb{R}^{|m|}$ $\|(D^m u,a) - (w,b)\| < \varepsilon\}$ u is unique;

(2) the collocation equations (3.11) have a unique solution $u_\Delta \in S_\Delta$ in this same sphere for $|\Delta| \leq \eta$;

(3) Newton's method for approximately solving the collocation equations converges quadratically in some neighborhood of u_Δ for $|\Delta| \le \eta$;

(4) u_Δ satisfies

$$\|D^i(u-u_\Delta)\| \le c_1|\Delta|^{\min(n,k)} \quad i \le m; \tag{3.31}$$

(5) the collocation approximation v_Δ to the solution of the linear problem

$$\begin{aligned} L_u v &= L_u u \\ B_i v &= B_i u \qquad i = 1,\dots,|m| \end{aligned} \tag{3.32}$$

satisfies

$$\|D^i(u-v_\Delta)\| \le C_2|\Delta|^{\min(n,k)} \qquad i \le m \tag{3.33}$$

and

$$D^i(u-u_\Delta) = D^i(u-v_\Delta) + 0(|\Delta|^{2\min(n,k)}) \quad i \le m. \tag{3.34}$$

Proof.

The hypotheses of this theorem are precisely those needed to apply theorem 1. Since $u \in C^{m-1}[a,b]$ is a solution of (3.1) $y^* \equiv (D^m u, c)$ is a solution of (3.8) where $c = (D^i u^j(a))$ $i = 0,\dots,m_j-1, j=1,\dots,d$. Since $\tilde{P}_\Delta(D^m u,c) = (P_\Delta D^m u,c)$ we have $\|y^*-\tilde{P}_\Delta y^*\| = \|D^m u-P_\Delta D^m u\| \to 0$ as $|\Delta| \to 0$ by (2.7). That T is Fréchet continuously differentiable at y^* follows immediately from the differentiability assumption on F.

We use the Green's matrix $G(x,t)$ associated with the linear problem (3.30) to show that $[I-T'(y^*)]^{-1}$ is bounded. Let $v \in C^m[a,b]$ satisfy the boundary conditions of (3.30) and let $y = (D^m v, a) = (w,a)$ be its representation in $C^0[a,b] \times \mathbb{R}^{|m|}$ via (3.3). With $y_1 = (v_1, a_1)$, (3.26) shows that

$$y - T'(y_1)y = y - \Big(\sum_{i=0}^{\bar{m}-1} F_i(y_1)\vec{H}_i y, e_1 + B_1(y), \ldots, e_{|m|} + B_{|m|}(y)\Big)$$
$$= (L_{v_1} v, B_1(y), \ldots, B_{|m|}(y)). \tag{3.35}$$

Thus

$$y = y - T'(y_1)y + \Big(\sum_{i=0}^{\bar{m}-1} F_i(y_1) D^i v, e_1 + B_1(y), \ldots, e_{|m|} + B_{|m|}(y)\Big) \tag{3.36}$$

or

$$y = y - T'(y_1)y +$$
$$\Big(\sum_{i=0}^{\bar{m}-1} F_i(y_1) \int_a^b \Big(\frac{\partial}{\partial x}\Big)^i G(x,t)\,[w - \sum_{j=0}^{\bar{m}-1} F_i(y_1)\vec{H}_i(y)],$$
$$e_1 + B_1(y), \ldots, e_m + B_m(y)\Big) \tag{3.37}$$

Thus with $y_1 \equiv y^*$, (3.37) implies

$$\|y\| \le \|y - T'(y^*)y\|$$

and hypothesis 3 of theorem 1 is satisfied.

Hypotheses (4) of theorem 1 follows from the differentiability of F, the uniform bound on $\|P_\Delta\|$ and the remark immediately following theorem 1. Conclusions (1) and (2) follow immediately from the conclusions of theorem 1.

Newton's method for approximately solving the collocation equations may be interpreted from (3.18) as: find $v_{\Delta,n+1}$ from $v_{\Delta,n}$ by solving the linear problem

$$\begin{aligned} L_{v_{\Delta,n}} v &= F(x,v_{\Delta,n}) - \sum_{i=0}^{\bar{m}-1} F_i(x,v_{\Delta,n}) D^i v_{\Delta,n} \\ B_i v &= B_i u = c_i \qquad i = 1,\ldots,|m| \end{aligned} \tag{3.38}$$

by collocation.

That is, $v_{\Delta,n+1}$ satisfies

$$\begin{aligned} P_\Delta L_{v_{\Delta,n}} v_{\Delta,n+1} &= P_\Delta\{F(x,v_{\Delta,n}) - \sum_{i=0}^{\bar{m}-1} F_i(x,v_{\Delta,n}) D^i v_{\Delta,n}\} \\ B_i(v_{\Delta,n+1}) &= B_i(u) \qquad i = 1,\ldots,|m| \ . \end{aligned} \tag{3.39}$$

But the differentiability of F implies the hypothesis of the corollary with the modulus of continuity $\omega(x) \le \lambda x$. (i.e. quadratic convergence occurs in some neighborhood of u).

Conclusion (4) is an immediate consequence of (3.15) together with our smoothness assumption on u and the approximability result (2.7) of the previous section.

Since $y^* = Ty^*$, $y_\Delta = \tilde{P}_\Delta T y_\Delta$ we have

$$y_\Delta = \tilde{P}_\Delta (y^* + T'(y)(y_\Delta - y^*) + Ty_\Delta - Ty^* - T'(y^*)(y_\Delta - y^*))$$

or

$$[I - \tilde{P}_\Delta T'(y^*)]y_\Delta = \tilde{P}_\Delta (y^* - T'(y^*)y^* + Ty_\Delta - Ty^* - T'(y^*)(y_\Delta - y^*)).$$

Hence

$$y_\Delta = [I - \tilde{P}_\Delta T'(y^*)]^{-1} \tilde{P}_\Delta (I - T'(y^*))y^* + [I - \tilde{P}_\Delta T'(y^*)]^{-1} \tilde{P}_\Delta (Ty_\Delta - Ty^* - T'(y^*)(y_\Delta - y^*)).$$

But $[I - \tilde{P}_\Delta T'(y^*)]^{-1} \tilde{P}_\Delta (I - T'(y^*))y^* = v_\Delta$,
$\| [I - \tilde{P}_\Delta T'(y^*)]^{-1} \tilde{P}_\Delta \|$ is bounded and $(Ty_\Delta - Ty^* - T'(y^*)(y_\Delta - y^*))$ is $O(\|y_\Delta - y^*\|^2)$ by the hypothesis on F.
Hence

$$y_\Delta - y = v_\Delta - y + O(\|y_\Delta - y\|^2), \text{ and } \quad (3.34)$$

follows immediately by using our representation of u in terms of y^*. □

The boundary conditions of (3.1) may be nonlinear. The treatment given in this section does not require that the boundary conditions be linear, but only that the linearized problem (3.30) possess a unique solution and an associated Green's function.

IV. HIGHER ORDER POINTWISE AND GLOBAL ERROR ESTIMATES

As in theorem 2 let u be the solution of (3.1), u_Δ the solution of (3.11) and v_Δ the solution of (3.32). From (3.34) we see that in order to estimate $D^i(u-u_\Delta)$ $(i \le m)$ to within terms of order $|\Delta|^{2\min(k,n)}$ it is sufficient to estimate $D^i(u-v_\Delta)(i \le m)$. Again let $G(x,s)$ denote the $d \times d$ Green's matrix for the problem (3.29) with homogeneous boundary conditions, and let

$$[G_i(x,s)]_{kj} = [(\frac{\partial}{\partial x})^i G(x,s)]_{kj} \qquad k,j = 1,\ldots d. \qquad (4.1)$$

We assume there exists a constant C_1 such that when $G_i(x,\cdot)$ is considered as an element of $C^{(n)}[a,x] \times C^{(n)}[x,b]$ we have

$$\|D_s^j G_i(x,s)\| \le C_1 \qquad i = 0,\ldots,\underline{m}-1 \quad j = 0,\ldots,n. \qquad (4.2)$$

Recall $\underline{m} = \min\limits_{1 \le i \le d} m_i$.

We discuss this assumption in Appendix 2. Let

$$r(x) = L_u(u-u_\Delta) \qquad (4.3)$$

and note that $r(x)$ vanishes at the collocation points τ_i.

<u>Lemma 1</u>

If L_u is smooth enough (e.g. if $[F_i(x,v)]_{\ell j} \in C^{k+n}[a,b]$ $\ell,j = 1,\ldots,d$, $i = 0,\ldots,\bar{m}-1$, and if $u \in C^{m+k+n}[a,b]$) then there exists a constant C_2 such that

$$\max_{x \in [x_j,x_{j+1}]} \|(D^{k+i}r)(x)\| \le C_2(\frac{|\Delta|}{|\Delta x_j|})^k \qquad \begin{array}{l} j = 0,\ldots,N-1, \\ i = 0,\ldots,n . \end{array} \qquad (4.4)$$

Proof.

This lemma is an immediate consequence of lemma 4.1 of de Boor and Swartz [1] applied to each component of r.

Theorem 3

Assume the hypotheses of theorem 2, lemma 1 and the estimate (4.2). With $n \le k$ (in the hypothesis of lemma 1) choose the k points $\{\rho_i\}_1^k$ in $[-1,1]$ so that

$$\int_{-1}^{1} p(s) \prod_{i=1}^{k} (s-\rho_i)ds = 0 \qquad (4.5)$$

for every polynomial $p \in \mathbb{P}_n$. (e.g. when $n = k$ the ρ_i are the Gaussian points). These $\{\rho_i\}_1^k$ are used along with the partition $\Delta = \{x_i\}_{i=0}^N$ to determine the collocation points $\{\tau_i\}_{i=1}^{kN}$ in $[a,b]$ (see (3.10)).

Then the solution of the collocation equations (3.11) $u_\Delta \in \mathbb{P}_{m+k,\Delta}[a,b] \cap C^{m-1}[a,b]$ satisfies

$$\left.\begin{aligned} \|D^i(u-u_\Delta)(x_j)\| \le C|\Delta|^{k+n} \quad & x_j \in \Delta = \{x_\ell\}_{\ell=0}^N \\ & i = 0,\dots,\underline{m}-1, \end{aligned}\right\} \qquad (4.6)$$

and

$$\|D^i(u-u_\Delta)\| \le C|\Delta|^{k+\min(n,\underline{m}-i)} \quad i = 0,\dots,\underline{m}, \qquad (4.7)$$

where C is a constant independent of the partition Δ.

Proof.

This thoerem follows immediately from theorem 4.1 of de Boor and Swartz [1]. We let their E_j be a vector $\{E_j^\ell\}_{\ell=1}^d$ and work with each component separately. That is

$$E_j^{\ell}(x) \equiv \int_{x_j}^{x_{j+1}} \sum_{q=1}^{d} [G_i(x,s)]_{\ell,q} r^q(s)ds \qquad j = 0,\ldots,N-1.$$

We give the complete details of this proof in appendix 1. □

Again consider $G_i(x,\cdot) \in C^n[a,x] \times C^n[x,b]$ and suppose

$$\|D_s^j[G_i(x,s)]_{k\ell}\| \le C_1 \quad i = 0,\ldots,m_k-1, \quad j = 0,\ldots,n, \quad \ell = 1,\ldots,d, \tag{4.8}$$

which is stronger than (4.2). Then estimate (4.6) at the breakpoints of Δ holds for all multi-indicies i such that $i < m$. That is,

$$\|D^i(u-u_\Delta)(x_j)\| \le C|\Delta|^{k+m}, \quad x_j \in \Delta, \quad i < m. \tag{4.9}$$

Estimate (4.7), under assumption (4.8) becomes

$$\|D^i(u-u_\Delta)\| \le C|\Delta|^{k+\min(n,\underline{m-i})}, \quad i \le m \,. \tag{4.10}$$

In general, of course, we cannot expect to obtain better results than (4.9) and (4.10) since if some index were greater than m_j we would be approximating $D^{m_j+1}u^j$ to order $|\Delta|^k$ by $D^{m_j+1}u_\Delta^j$, a polynomial of order k-1.

NOTE: When all d equations are of the same order (i.e. m such that $m_1 = \ldots = m_d$) then (4.6) and (4.7) are the same as (4.8) and (4.9).

APPENDIX

In this appendix we give the details of the proof of theorem 3.

Recall from (3.34)

$$D^i(u-u_\Delta) = D^i(u-v_\Delta) + O(|\Delta|^{2\min(n,k)}) \quad i \le m.$$

Thus it is sufficient to prove (4.6) and (4.7) with u_Δ (the solution of the collocation equations) replaced by v_Δ (the solution of the collocation equations associated with the linearized problem at u). Since (3.35) has a unique solution and there is an associated Green's function $G(x,t)$ we have

$$(u-v_\Delta)(x) = \sum_{j=1}^{N} E_j(x), \qquad a \le x \le b,$$

where

$$E_j(x) \equiv \int_{x_{j-1}}^{x_j} G(x,t)\ L_u(u-v_\Delta)(t)\ dt$$

and L_u is the linearization of problem (3.1) at its solution u.

Let

$$[G_i(x,t)]_{\ell,r} = \left(\frac{\partial}{\partial x}\right)^i [G(x,t)]_{\ell,r} \quad \ell,r = 1,\ldots,d.$$

Then (for $0 \le i < \underline{m}$)

$$D^i(u-v_\Delta)(x) = \sum_{j=1}^{N} \int_{x_{j-1}}^{x_j} G_i(x,t) L_u(u-v_\Delta)(t)\,dt. \qquad \text{(A.1)}$$

Let $E_j^{\ell}(x)$ denote the ℓ'th component of the integral on the right side of (A.1). Fix i, the interval j, and ℓ. On the interval $[x_{j-1}, x_j]$ all components of $L_u(u-v_\Delta)$ vanish at the k collocation points

$$\tau_{kj+s} = [x_{j-1}+x_j+\rho_s \Delta x_j]/2 \qquad s = 1,\ldots,k.$$

With $r(x) = (r^1(x), r^2(x), \ldots, r^d(x))^T = L_u(v-v_\Delta)(x)$ we have

$$[G_i(x,t)r(x)]^{\ell} = \sum_{s=1}^{d} [G_i(x,t)]_{\ell,s} r^s(x)$$

which vanishes at the k collocation points τ_{kj+s}, $s = 1,\ldots,k$. Let $T_n(t)$ be the Taylor's series expansion for

$$f_x(t) \equiv \sum_{s=1}^{d} [G_i(x,t)]_{\ell,s}\; r^s[\tau_{kj+1},\ldots,\tau_{kj+k},t] \qquad \text{(A.2)}$$

around $t = x_{j-1}$ up to terms of degree $< n$.

Here $r^s[\tau_{kj+1},\ldots,\tau_{kj+k},t]$ is the k'th divided difference of r^s. Hence

$$[G_i(x,t)r(t)]^{\ell} = p_k(t)[T_n(t)+(t-x_{j-1})^n (D^n f_x)(\theta_t)/n!]$$

where $\theta_t \in (x_{j-1}, x_j)$, and

$$p_k(t) \equiv \prod_{s=1}^{k} (t-\tau_{k_{j+s}}).$$

We chose the collocation points to make $p_k(x)$ orthogonal to all polynomials of order k, hence

$$|E_j^{\ell}(x)| = \left|\int_{x_{j-1}}^{x_j} p_k(t)(t-x_{j-1})^n (D^n f_x)(\theta_t)/n!\,dt\right|$$

$$\leq K\,|\Delta x_j|^{k+n+1}\,\|D^n f_t\|_{(j)}\;,$$

where $\|D^n f_t(\cdot)\|_j = \|D^n f_t(\cdot)\|_{L^\infty[x_{j-1},x_j]}$.

and K is a constant independent of Δ.
From (A.2) we see that

$$(D_t^n f_x)(t) = \sum_{s=1}^{d} \frac{1}{(n+k)!} \sum_{j=0}^{n} \binom{n+k}{k+j} D_t^{n-j}[G_i(x,t)]_{\ell,s}\cdot (D^{k+j} r^s)(\theta_{t,j,d}) \qquad (A.3)$$

where $\theta_{t,j,d} \in (x_{j-1},x_j)$.

We now consider two cases:

(1) If $x \notin (x_{j-1},x_j)$ then, from (4.2),

$$\|D_t^{n-p}[G_i(x,t)_{\ell,s}]\|_{(j)} \leq C_1 \quad p = 0,\ldots,n$$

and by lemma 1

$$\|D^{k+p} r^s\|_{(j)} \leq C_2\left(\frac{|\Delta|}{\Delta x_j}\right)^k, \; p = 0,\ldots,n;\; s = 1,\ldots,d;$$

hence,

$$\|(D_t^n f_x)(t)\|_{(j)} \leq C_3\left(\frac{|\Delta|}{\Delta x_j}\right)^k .$$

Thus, if $x \notin (x_{j-1},x_j)$

$$|E_j^{\ell}(x)| \leq C_4 |\Delta|^k\,|\Delta x_j|^{n+1} . \qquad (A.4)$$

(2) If $x \in (x_{j-1},x_j)$, then

$$[G_i(x,t)]_{\ell,s} \in H_\infty^{m_\ell - i-1}[x_{j-1},x_j]$$

so that

$$||D^p f_x(t)||_{(j)} \le C_s \left(\frac{|\Delta|}{\Delta x_j}\right)^k$$

when $p \le \min(n,m_\ell - i-1)$. Hence,

$$|E_j^\ell(x)| \le C_6 \; |\Delta|^k \cdot |\Delta x_j|^{\min(n+1,m_\ell - i)} \tag{A.5}$$

when $x \in (x_{j-1},x_j)$.

Since we have restricted i, $0 \le i \le \underline{m}$, (A.5) is well defined for all $\ell = 1,\ldots,d$.

If $x \in \Delta \equiv \{x_\ell\}_0^N$ then summing (A.4) over j and maximizing over ℓ we obtain (4.6). If $x \in (x_{j-1},x_j)$ for some j we combine (A.4) and (A.5) to obtain

$$||D^i(u-v_\Delta)(x)|| \le |\Delta|^k \left(C_4|\Delta|^n + C_s|\Delta x_j|^{\min(n+1,\underline{m}-i)}\right)$$

when $x_{j-1} < x < x_j$. Thus (4.7) follows directly from this estimate.

From (A.3) we see that the constant, C, in theorem 3 depends on bounds on the coefficients of the linearized problem (see 3.24) and $k + n + m_\ell$ derivatives of the ℓ'th component of the solution u of (3.1).

APPENDIX 2

In this appendix we discuss the construction of a Green's function, $G(x,t)$, associated with a linear problem of the form (3.30). That is, using the notation of section 3,

$$D^m u = \sum_{i=0}^{\bar{m}-1} C_i(x) D^i u + g(x) ,$$

$$B_i(u) = c_i , \qquad i = 1,\dots,|m| , \tag{A.1}$$

where u, g are d-vectors, and C_j is a $d \times d$ matrix. Our assumptions on F in (3.1) show that

$$[C_j(x)]_{\ell k} = 0 , \qquad j \geq m_i , k \leq i , \ell = 1,\dots,d . \tag{A.2}$$

This natural assumption on the form of F assures us that (A.1) may be written as a first order system

$$Dv(x) = C(x)v(x) + \tilde{g}(x) , \tag{A.3}$$

where v, $\tilde{g}$ are $|m|$-vectors and $C(x)$ is an $|m| \times |m|$ matrix. We will use this fact to advantage in a moment, but first we consider an important special case.

Let L_i, $i = 1,\dots,d$, be defined by

$$(L_i u)(x) \equiv \left(\frac{d}{dx}\right)^2 u(x) + b_i(x)\frac{d}{dx}u(x) + q_i(x)u(x) . \tag{A.4}$$

Suppose (A.1) is of the form

$$\begin{aligned} L_1 u^1 &= u^2 + f_1(x) & u^1(a) &= u^1(b) = 0 \\ L_2 u^2 &= u^3 + f_2(x) & u^2(a) &= u^2(b) = 0 \\ &\ \vdots \\ L_d u^d &= f_d(x) & u^d(a) &= u^d(b) = 0 . \end{aligned} \tag{A.5}$$

Each L_i is of the form (A.4) where $b_i \in C^{p-1}[a,b]$, $q_i \in C^{p-2}[a,b]$. We assume that the homogeneous system associated with (A.5) is incompatible, that is when $f_i \equiv 0, i=1,\dots,d$, the system admits only the zero solution. Under this condition we can construct the classical Green's function $K_i(x,t)$ for each L_i with zero boundary conditions (see e.g. Courant and Hilbert [3]).

With these $K_i(x,t)$ it is easy to construct the $d \times d$ Green's matrix associated with (A.5). Since

$$u^d(x) = \int_a^b K_d(x,t)f_d(t)dx$$

let

$$[G(x,t)]_{dd} = K_d(x,t) \quad . \tag{A.6}$$

The equation for u^{d-1} shows that

$$u^{d-1}(x) = \int_a^b K_{d-1}(x,t)[u^d(t)+f_{d-1}(t)]dt$$

$$= \int_a^b K_{d-1}(x,t)[\int_a^b K_d(t,s)\, f_d(s)ds+f_{d-1}(t)]dt$$

and, upon interchange of the order of integration,

$$= \int_a^b K_{d-1}(x,t)f_{d-1}(t)dt +$$

$$+ \int_a^b \int_a^b K_{d-1}(x,t)K_d(t,s)dt\;\; f_d(s)ds.$$

Thus,

$$\begin{aligned} [G(x,t)]_{d-1,d-1} &= K_{d-1}(x,t), \\ [G(x,t)]_{d-1,d} &= \int_a^b K_{d-1}(x,s)K_d(s,t)ds \quad . \end{aligned} \tag{A.7}$$

Continuing in this manner we see that $G(x,t)$ is upper triangular with diagonal formed by the $K_i(x,t)$. The off-diagonal elements will be iterated integrals of the form (A.7). In particular, for $j > k$,

$$[G(x,t)]_{kj} = \int_a^b \cdots \int_a^b K_k(x,x_1)K_{k+1}(x_1,x_2)\cdots K_j(x_{j-k},t)\,dx_1\cdots dx_{j-k} \quad (A.8)$$

The continuity properties of the $K_i(x,t)$ are well known, and the equation (A.8) shows

$$\begin{aligned} G(\cdot,t) &\in C^{p+2}[a,t] \times C^{p+2}[t,b], \\ G(x,\cdot) &\in C^{p+2}[a,x] \times C^{p+2}[x,b] . \end{aligned} \quad (A.9)$$

This method of building the Green's matrix may be used also when the L_i are of different orders provided the problem (A.1) is of the form (A.5). That is, associated with an L_i of order m_i are m_i boundary conditions, λ_j, of the form

$$\{\lambda_j\}_1^{m_i} \subset \text{span}\,\{\delta_a,\ldots,\delta_a^{m_i-1},\delta_b,\ldots,\delta_b^{m_i-1}\}$$

which are applied only to u^i .

In order to treat more general problems we turn to the representation (A.3). We write the homogeneous first order system associated with (A.1) as

$$Dv(x) = C(x)v(x) \quad (A.10)$$

Here $v(x) = (u^1(x),Du^1(x),\ldots,D^{m_1-1}u^1(x),u^2(x),\ldots,D^{m_d-1}u^d(x))^T$. The corresponding homogeneous boundary conditions are

$$\lambda_i(v) = 0 \qquad i = 1,\ldots,|m| . \quad (A.11)$$

We assume $C(x) \in C^p[a,b]$, $p \geq 0$, although the following discussion holds if the elements of C have a finite number of jump discontinuities on $[a,b]$.

Let $\Phi(x)$ be a fundamental matrix for the system (A.10). Then the unique solution, $v(x)$, of (A.3) satisfying zero initial conditions ($v(a)\equiv 0$) is given by

$$v(x) = \Phi(x)\int_a^x \Phi^{-1}(t)\tilde{g}(t)dt, \quad (A.12)$$

(see e.g. Coddington and Levinson [2], p. 74).
With $H(x,t) = \Phi(x)\Phi^{-1}(t)(x-t)_+^0$ we may rewrite (A.12) as

$$v(x) = \int_a^b H(x,t)\tilde{g}(t)dt. \tag{A.13}$$

Thus $H(x,t)$ is the Green's matrix for the problem (A.3) with zero initial conditions. Note that

$$\begin{aligned} H(\cdot,t) &\in C^{p+1}[a,t] \\ H(x,\cdot) &\in C^{p+1}[a,x] \, . \end{aligned} \tag{A.14}$$

A necessary and sufficient condition for the problem (A.3), together with boundary conditions (A.11), to have a solution for every $\tilde{g}$ is that the set $\{\lambda_i\}_{i=1}^m$ be linearly independent over the kernel of L, where

$$Lv(x) \equiv Dv(x) - C(x)v(x). \tag{A.15}$$

That is, the boundary conditions are linearly independent over the columns of $\Phi(x)$.

If A is a matrix let $\vec{A}_j$ denote the jth column of A. An explicit representation of the Green's matrix $G(x,t)$ for the problem (A.3) with boundary (A.11) can be obtained by modifying the Green's matrix for the initial value problem. Under the assumption that the boundary conditions are linearly independent over the columns of $\Phi(x)$ we have that

$$G(x,t) = H(x,t) - \Phi(x)QA(t), \tag{A.16}$$

where $Q_{ij} = [\lambda_i \vec{\Phi}_j]^{-1}$, $A_{ij}(t) = \lambda_i \vec{H}_j(x,t)$.

Here it is understood that $\vec{H}_j(x,t)$ is thought of as a function of x when the functional λ_i is applied.

Since $G(x,t)$ is merely a modification of $H(x,t)$ by a linear combination of elements of the kernel of L, equation (A.3) is satisfied. It remains only to verify that the boundary conditions are satisfied. The computation goes as follows:

$$\lambda_k(v) = \lambda_k[\int_a^b G(x,t)\tilde{g}(t)dt]$$

$$= \int_a^b \sum_j \lambda_k[\vec{H}_j(x,t)]\tilde{g}^j(t)dt$$

$$- \lambda_k[\Phi(x)Q\int_a^b A(t)\tilde{g}(t)dt] \ .$$

However, with $\lambda_k(v) = \sum_\ell \lambda_k^\ell v^\ell$, we have that

$$\lambda_k[\Phi(x)Q\int_a^b A(t)g(t)dt]$$

$$= \sum_\ell \lambda_k^\ell[\sum_j \sum_m \sum_r \Phi_{\ell r}(x)Q_{rm} \int_a^b A_{mj}(t)\tilde{g}^j(t)dt]$$

$$= \sum_m \sum_r \sum_\ell \lambda_k^\ell(\Phi_{\ell r}(x))Q_{rm}\int_a^b \sum_j A_{mj}(t)\tilde{g}^j(t)dt$$

$$= \sum_m \delta_{km} \int_a^b \sum_j (\lambda_m \vec{H}_j(x,t))\tilde{g}^j(t)dt$$

$$= \int_a^b \sum_j \lambda_k[\vec{H}_j(x,t)]\tilde{g}^j(t)dt \ .$$

Hence, we conclude that

$$\lambda_k(v) = 0 \ , \qquad k = 1,\ldots,|m| \ .$$

Combining (A.16) and (A.14) allows us to make continuity statements about $G(x,t)$. For example, suppose the boundary conditions are of the form

$$\sum_{n=1}^{N} B_n v(s_n) = c. \tag{A.17}$$

Here $B_1,\ldots,B_N$ are $|m| \times |m|$ matrices, c is a constant $|m|$-vector, and the $N(\geq 1)$ distinct points $\{s_n\}$ lie in $[a,b]$. H. B. Keller [5] discusses this type of system and shows that for $A(x)$, $\tilde{g}(x) \in C_\Gamma^p[a,b]$ there exists a solution $v(x) \in C_\Gamma^{p+1}[a,b]$ satisfying (A.3) and (A.17) iff

$$\sum_{n=1}^{N} B_n \Phi(s_n)$$

is nonsingular. Here, as above, Φ is a fundamental matrix for equation (A.10), and Γ is a partition of [a,b]. R. D. Russell [7] gives an explicit representation for the Green's matrix for a first order system in this case. The less detailed representation given by (A.16) is sufficient to determine the continuity properties of $G(x,t)$. If ${}_jB_n$ is the j-th row of the n-th matrix then

$$\lambda_j(v) = \sum_{n=1}^{N} {}_jB_n v(t_j) .$$

Hence, $G(x,t)$ may have discontinuities introduced at the points $\{s_i\}$. Thus, if Γ is a partition of [a,b] such that $\Gamma \supset \{s_i\}$ then

$$\begin{aligned} G(\cdot,t) &\in C_{\Gamma}^{p+1}[a,t] \times C_{\Gamma}^{p+1}[t,b] \\ G(x,\cdot) &\in C_{\Gamma}^{p+1}[a,x] \times C_{\Gamma}^{p+1}[x,b] \ . \end{aligned} \tag{A.18}$$

This result is pertinent to the estimate (4.2) on the Green's matrix. Recall that theorems 2 and 3 also hold in the more general case when the functions involved are merely piecewise smooth, provided that the partition Δ contains such points of discontinuity, i.e. we would require $\Delta \supset \Gamma \supset \{s_i\}$. In the case that $N = 2$ and $s_1 = a$, $s_2 = b$ we see that (A.18) shows that (4.2) holds for first order systems provided p is sufficiently large.

In the case of functionals of the form

$$\lambda_j(v) = \int_a^b g_i(t)v(t)dt$$

where g_i is an $|m|$-vector whose components are continuous on [a,b] we again see that (A.18) holds. In this case no additional discontinuities are introduced.

Of course the Green's matrix for the original system (A.1) is "contained" in the Green's matrix for the first order system. When (A.1) is written in the form (A.3) with

$$v(x) = (u^1(x),\ldots,D^{m_1-1}u^1(x),\ldots,D^{m_d-1}u^d(x))$$

we see that $\tilde{g}$ has only d possibly nonzero components. These occur in positions $C \equiv \{m_1, m_1+m_2,\ldots,|m|\}$. Furthermore we only need rows $R \equiv \{1, 1+m_1,\ldots,1+|m|-m_d\}$ of the $|m|$-vector v to determine the d-vector u. Hence we use the columns of the set C and the rows of the set R of the $|m| \times |m|$ Green's matrix $G(x,t)$ to form the $d \times d$ Green's matrix $\mathcal{G}(x,t)$ for the system (A.1).

Consider row d of $\mathcal{G}(x,t)$ and let $\ell = m_d-1$. Obviously

$$D^\ell u^d(x) = D^\ell \int_a^b \sum_{j=1}^{d} [\mathcal{G}(x,t)]_{dj} g^j(t)dt$$

$$= \int_a^b \sum_{j=1}^{|m|} [G(x,t)]_{|m|j} \tilde{g}^j(t)dt \ .$$

But this equality holds for all continuous $g(t)$, and we have seen (A.18) hold for all rows of $G(x,t)$. Furthermore, a similar discussion holds for each row of $\mathcal{G}(x,t)$. Thus, if the coefficient matrices $C_i(x)$ of (A.1) are in $C^p[a,b]$, we have, for $j = 1,\ldots,d$,

$$\begin{aligned} &[\mathcal{G}(\cdot,t)]_{kj} \in C^{p+m_k}[a,t] \times C^{p+m_k}[t,b] \\ &[\mathcal{G}(x,\cdot)]_{kj} \in C^{p+1}[a,x] \times C^{p+1}[x,b]. \end{aligned} \tag{A.19}$$

Of course when (A.1) is self-adjoint we have additional smoothness on $\mathcal{G}(x,\cdot)$. However, with the appropriate p, (A.19) justifies assumptions of the form (4.2).

REFERENCES

1. C. de Boor and B. K. Swartz, Collocation at Gaussian Points, SIAM J. Numer. Anal., 10 (1973), 582-606.

2. E. A. Coddington and N. Levinson, Theory of Ordinary Differential Equations, McGraw-Hill, New York, 1955.

3. R. Courant and D. Hilbert, Methods of Mathematical Physics, Vol. 1, Interscience, New York, 1953.

4. Ph. J. Davis, Interpolation and Approximation, Blaisdell, New York, 1963.

5. H. B. Keller, Accurate Difference Methods for Linear Ordinary Differential Systems Subject to Linear Constraints, SIAM J. Numer. Anal., 6 (1969), 8-30.

6. M. A. Krasnosel'skii et al., Approximate Solution of Operator Equations, Wolters-Noordhoff, Groningen, 1972.

7. R. D. Russell, Collocation for Systems of Boundary Value Problems, Numer. Math., 23 (1974), 119-133.

8. G. M. Vainikko, Galerkin's Perturbation Method and the General Theory of Approximate Methods for Nonlinear Equations, USSR Comp. Math. Physics, 7(1967), 1-41.

9. R. Weiss, The Application of Implicit Runge-Kutta and Collocation Methods to Boundary Value Problems, Math. Comp. 28 (1974), 449-464.

10. K. A. Wittenbrink, High Order Projection Methods of Moment and Collocation Type for Nonlinear Boundary Value Problems, Comp., 11 (1973), 255-274.