

Solving Linear Boundary Value Problems by Approximating the Coefficients

By **Steven A. Pruess**

Abstract. A method for solving linear boundary value problems is described which consists of approximating the coefficients of the differential operator. Error estimates for the approximate solutions are established and improved results are given for the case of approximation by piecewise polynomial functions. For the latter approximations, the resulting problem can be solved by Taylor series techniques and several examples of this are given.

1. Introduction. In trying to solve the boundary value problem

$$(1a) \quad Lu = u^{(k+1)}(x) + \sum_{i=0}^k p_i(x)u^{(i)}(x) = f(x) \quad \text{in } (a, b),$$

$$(1b) \quad U_i u = \sum_{j=1}^{k+1} (a_{ij}u^{(j-1)}(a) + b_{ij}u^{(j-1)}(b)) = \gamma_i, \quad i = 1, 2, \dots, k + 1,$$

with $p_i, f \in C[a, b]$, we, instead, attempt to solve the "approximate problem"

$$(2a) \quad L_\pi u_\pi = u_\pi^{(k+1)}(x) + \sum_{i=0}^k \hat{p}_i(x)u_\pi^{(i)}(x) = \hat{f}(x) \quad \text{in } (a, b),$$

$$(2b) \quad U_i u_\pi = \gamma_i, \quad i = 1, 2, \dots, k + 1,$$

where \hat{p}_i, \hat{f} are approximations to the coefficients of (1). It is a classical result that, if (1) has a unique solution, then perturbations of $O(\epsilon)$ in the coefficients produce perturbations of at most $O(\epsilon)$ in the solution (this result is stated formally in Theorem 1 below). For computational purposes, piecewise polynomial functions are obvious choices for approximations and bounds for the errors incurred are given in Theorem 2. These results are for fairly arbitrary perturbations and, in fact, for specific piecewise polynomial approximations, the bounds can be sharpened considerably as demonstrated in Theorem 3.

Similar ideas have been used by Alexander and Gordon [1], Gordon [5], for boundary value problems and Canosa and Gomes de Oliveira [2] and Pruess [8] for eigenvalue problems. Gordon has explored some of the theoretical aspects using perturbation methods to approximate the error but no general rigorous development has been given. In [5], Gordon uses these techniques to compute high order bounds on the local error in certain quadratures related to the error in the solution at mesh points, however, the max-norm error is not studied. L. Ixaru, in as yet unpublished

Received March 20, 1972.

AMS (MOS) subject classifications (1970). Primary 65L10; Secondary 65B05, 34E10.

Key words and phrases. Linear boundary value problems, piecewise polynomial approximation, Richardson's extrapolation, Taylor series methods, shooting methods.

Copyright © 1973, American Mathematical Society

papers, has also established some results for piecewise polynomial approximations but these appear to be no stronger than what is predicted by Theorem 1.

Computationally, the most relevant result of this paper is the following: Given a partition $\pi = \{a = x_1 < x_2 < \dots < x_{N+1} = b\}$, if on each (x_n, x_{n+1}) the coefficients of (1) are interpolated at the roots of the $(m + 1)$ st degree Legendre polynomial transformed to (x_n, x_{n+1}) , then $|u(x_n) - u_\pi(x_n)| = O(|\pi|^{2m+2})$ and $\|u - u_\pi\| = O(|\pi|^s)$ where $s = \min(2m + 2, m + k + 2)$. Here and in what follows, $|\pi| = \max(x_{n+1} - x_n)$, m is the degree of the piecewise polynomial functions and $\|\cdot\|$ is the sup-norm. The paper concludes with a discussion of the numerical algorithm as well as several examples.

2. Error Bounds. For the remainder of the paper, we assume that (1) has a unique solution $u(x)$ in $C^{k+1}[a, b]$. The notation $D^k[a, b]$ is used for the space of k -times piecewise continuously differentiable functions with the understanding that, for any $f \in D[a, b]$,

$$(3) \quad f(x) = \frac{1}{2}(f(x^+) + f(x^-)).$$

For $\hat{p}_i \in D[a, b]$, $i = 0, 1, \dots, k$, $\hat{f} \in D[a, b]$, we are concerned with existence, uniqueness and error bounds for solutions to (2). Theorem 1 is found in various forms in the literature (see, for example, Nécas [7, Chapter 3]) but is presented here in some detail in order to yield the higher order results to follow.

THEOREM 1. *If $\hat{p}_i \in D[a, b]$, $i = 0, 1, \dots, k$, $\hat{f} \in D[a, b]$ and $\|p_i - \hat{p}_i\| \leq \epsilon$, $\forall i$, $\|f - \hat{f}\| \leq \epsilon$ for some constant ϵ , then, for ϵ sufficiently small, there exists a unique $u_\pi \in C^k[a, b] \cap D^{k+1}[a, b]$ satisfying (2) and $\|u^{(j)} - u_\pi^{(j)}\| = O(\epsilon)$, $j = 0, 1, \dots, k$.*

Proof. Since $Lu = f$ is assumed to have a unique solution, there is a Green's function $G(x, t)$ for L corresponding to homogeneous boundary conditions; the properties of G are detailed in Cole [3, Chapter 6]. Set $X = \{v \in C^k[a, b] / \|v - u\|_k \leq 1\}$ where $\|v\|_k = \sum_{i=0}^k \|v^{(i)}\|$. If z satisfies $Lz = 0$, $U_i z = \gamma_i$, $i = 1, 2, \dots, k + 1$, then, for $v \in X$, define

$$(Tv)(x) = z(x) + \int_a^b G(x, t)[(L - L_\pi)v + \hat{f}] dt.$$

Since $Lu = f$, this can be rewritten as

$$(Tv)(x) = u(x) + \int_a^b G(x, t)[(L - L_\pi)v + \hat{f} - f] dt.$$

From the assumptions concerning the coefficient functions, it is easily verified that, for ϵ sufficiently small, T is a contraction on X and thus has a unique fixed point, say $u_\pi(x)$. It then follows from the properties of the Green's function that $u_\pi \in D^{k+1}[a, b]$ satisfies (2) and

$$(4) \quad (u^{(j)} - u_\pi^{(j)})(x) = \int_a^b \frac{\partial^j G}{\partial x^j}(x, t)[(L_\pi - L)u_\pi + f - \hat{f}] dt.$$

3. Results from Approximation Theory. For computational purposes, piecewise polynomial functions are a natural choice for the coefficients of the approximate problem. For specific choices of piecewise polynomial approximations, the bounds of Theorem 1 can be improved.

The approximations considered are characterized in terms of linear projectors, i.e., linear idempotent maps, whose range is P_m , the space of polynomials of degree at most m . By Π is meant the set of partitions of the form $\pi = \{a = x_1 < x_2 < \dots < x_{N+1} = b\}$; for $\pi \in \Pi$, $|\pi| = \max_n (x_{n+1} - x_n)$.

Given a function $g \in C^{m+1}[a, b]$, a piecewise polynomial approximation \hat{g} to g is generated by the scheme:

$$(5) \quad \text{for some continuous linear projector } Q \text{ from } C[-1, 1] \text{ onto } P_m, \text{ given } \pi \in \Pi \text{ on each } (x_n, x_{n+1}), \text{ let } \hat{g}(x) = Q[g(x(t))] \text{ for } t \in (-1, 1) \text{ where } x(t) = \frac{1}{2}(x_{n+1} - x_n)(t + 1) + x_n; \hat{g}^{(j)}(a) = g^{(j)}(a), g^{(j)}(b) = \hat{g}^{(j)}(b), j = 0, 1, \dots, m, \text{ and convention (3) is observed at interior mesh points.}$$

Note that many piecewise polynomial approximations do not fit this context, e.g., higher order splines, because of the global continuity condition. For these approximations, Theorem 1 is applicable but succeeding results are not.

LEMMA 1. *If $g \in C^{m+1}[a, b]$ and \hat{g} is given by (5), then there is a constant K such that, for each $\pi \in \Pi$,*

$$\|g^{(j)} - \hat{g}^{(j)}\| \leq K_1 |\pi|^{m-j+1}, \quad j = 0, 1, \dots, m.$$

Proof. See [8].

If Q is chosen so that certain moments are integrated exactly for polynomials, then the error bounds are sharpened. The following result is needed for this.

LEMMA 2. *If, for some integer $M \geq m + 2$, $g \in C^M[a, b]$, \hat{g} is given by (5) and, for the projector Q , there is a constant K_2 such that, for $G \in C^M[-1, 1]$,*

$$\left| \int_{-1}^1 t^i (1 - Q)G dt \right| \leq K_2 \|G^{(M-1)}\|, \quad i = 0, 1, \dots, M - m - 2,$$

then, for $\pi \in \Pi$, $x_n \in \pi$ and $v \in C^{M-m-1}(x_n, x_{n+1})$ with bounded derivatives up to order $M - m - 1$,

$$(6) \quad \left| \int_{x_n}^{x_{n+1}} (g - \hat{g})v dx \right| \leq K \left\{ (x_{n+1} - x_n) \sum_{i=0}^{M-m-1} [\|v^{(i)}\| \cdot \|g^{(M-i)}\| / i!] \right\} |\pi|^M$$

where $K = \max\{K_1, K_2\}$, K_1 as in Lemma 1.

Proof. Expand v in a Taylor series about $z_n = \frac{1}{2}(x_n + x_{n+1})$, then for some $\xi_x \in (x_n, x_{n+1})$,

$$\begin{aligned} \int_{x_n}^{x_{n+1}} (g - \hat{g})v dx &= \sum_{i=0}^{M-m-2} v^{(i)}(z_n) \int_{x_n}^{x_{n+1}} (x - z_n)^i (g - \hat{g})(x) dx / i! \\ &+ \int_{x_n}^{x_{n+1}} (x - z_n)^{M-m-1} (g - \hat{g})(x) v^{(M-m-1)}(\xi_x) dx / (M - m - 1)!. \end{aligned}$$

Set $G(t) = g(x(t))$, $\hat{G}(t) = \hat{g}(x(t))$ where $x(t) = \frac{1}{2}(x_{n+1} - x_n)(t + 1) + x_n$ so that

$$\begin{aligned} \left| \int_{x_n}^{x_{n+1}} (x - z_n)^i (g - \hat{g}) dx \right| &= (x_{n+1} - x_n)^{i+1} \left| \int_{-1}^1 t^i (1 - Q)G dt \right| / 2^{i+1} \\ &\leq (x_{n+1} - x_n)^{i+1} K_2 \|G^{(M-i)}\| / 2^{i+1} \\ &\leq K_2 (x_{n+1} - x_n)^{M+1} \|g^{(M-i)}\| / 2^{M+1}. \end{aligned}$$

The remaining integral is bounded by

$$K_1(x_{n+1} - x_n)^{M-m} |\pi|^{m+1} \|g^{(m+1)}\| \cdot \|v^{(M-m-1)}\| / 2^{M-m}$$

as a consequence of Lemma 1. Q.E.D.

The restriction on Q is equivalent to

$$\int_{-1}^1 t^i (1 - Q)p \, dt = 0 \quad \forall p \in P_{M-i-1}, \quad i = 0, 1, \dots, M - m - 2,$$

and the largest M , for which such a Q (with range P_m) exists, is $M = 2m + 2$. Thus, for piecewise polynomial approximations generated as in (5), $M = 2m + 2$ is the highest attainable exponent in (6).

4. Applications and Examples. The above lemmas suffice to yield error bounds for the solutions of (2) when the coefficients have been approximated by piecewise polynomial functions generated as in (5). The first result is a simple consequence of Theorem 1.

THEOREM 2. *If $p_i \in C^{m+1}[a, b]$, $i = 0, \dots, k$, $f \in C^{m+1}[a, b]$ and \hat{p}_i, \hat{f} are generated from p_i, f , respectively, as in (5), then, for each $\pi \in \Pi$ as $|\pi| \rightarrow 0$,*

$$(7) \quad \|u^{(i)} - u_\pi^{(i)}\| = O(|\pi|^{m+1}), \quad j = 0, 1, \dots, k,$$

$$(8) \quad \|u^{(k+i)} - u_\pi^{(k+i)}\| = O(|\pi|^{m+2-i}), \quad j = 1, 2, \dots, m + 1.$$

A more significant result is obtained by imposing the restrictions of Lemma 2 on Q .

THEOREM 3. *If, for some integer $M \geq m + 2$, $p_i \in C^r[a, b]$, $i = 0, 1, \dots, k$, $f \in C^r[a, b]$, $r \geq \max(M, k + 1)$, \hat{p}_i, \hat{f} are generated as in (5) and, for the projector Q , there exists a constant K such that, for $G \in C^M[-1, 1]$,*

$$\left| \int_{-1}^1 t^i (1 - Q)G \, dt \right| \leq K \|G^{(M-i)}\|, \quad i = 0, 1, \dots, M - m - 2,$$

then, for $\pi \in \Pi$, $|\pi|$ sufficiently small, there exists a constant C independent of π such that

$$(9) \quad |u^{(i)}(x) - u_\pi^{(i)}(x)| \leq C |\pi|^M, \quad x \in \pi,$$

$$(10) \quad |u^{(i)}(x) - u_\pi^{(i)}(x)| \leq C |\pi|^s, \quad x \notin \pi,$$

where $s = \min(M, m + k + 2 - j)$ and $j = 0, 1, \dots, k$.

Proof. If $p_i \in C^r[a, b]$, then the coefficients of the adjoint operator are in $C^{r-k}[a, b]$, so $G(x, t)$ has continuous derivatives with respect to t of order $r - 1$ on $[a, x)$ and $(x, b]$ for each x , and continuous derivatives with respect to x of order $r + k - 1$ on $[a, t)$ and $(t, b]$ for each t . For $x \in \pi$, say $x = x_i$, (4) implies

$$u^{(i)}(x_i) - u_\pi^{(i)}(x_i) = \left(\int_a^{x_i} + \int_{x_i}^b \right) dt \left\{ \frac{\partial^i G}{\partial x^i} [(L_\pi - L)u_\pi + f - \hat{f}] \right\}.$$

The integrand is comprised of sums of the form $(p_i - \hat{p}_i)v$ or $(f - \hat{f})v$ where $v \in C^{m+1}(x_n, x_{n+1}) \forall n$ and v and its derivatives can be bounded over $[a, b]$, independently of π . Thus, application of Lemma 2 to each subinterval and summation over the

subintervals yield (9). If $x \notin \pi$, say $x \in (x_i, x_{i+1})$, the integral can be written as $\int_a^{x_i} + \int_{x_{i+1}}^b + \int_{x_i}^{x_{i+1}}$: Again Lemma 2 insures that the first two integrals are $O(|\pi|^M)$. However, the integrand is not smooth enough in (x_i, x_{i+1}) for Lemma 2 to be applicable to the remaining integral. Let $G_{j,l}$ denote $\partial^{j+l}G/\partial x^j \partial t^l$ and $z_i = \frac{1}{2}(x_i + x_{i+1})$.

$$\int_{x_i}^{x_{i+1}} G_{j,0}(x, t)(f - \hat{f}) dt = \sum_{l=0}^{k-1-j} G_{j,l}(x, z_i)/l! \int_{x_i}^{x_{i+1}} (t - z_i)^l (f - \hat{f}) dt + \int_{x_i}^{x_{i+1}} (t - z_i)^{k-j} G_{j,k-j}(x, \xi_i)(f - \hat{f}) dt / (k - j)!$$

The last integral is $O(|\pi|^{m+k+2-j})$ from Lemma 1; the others are $O(|\pi|^{M+1})$ by Lemmas 1 and 2. The remaining terms of the error have the same form and are treated similarly.

COROLLARY 1. *If $M = 2m + 2$, then*

$$(11) \quad |u^{(j)}(x) - u_\pi^{(j)}(x)| \leq C |\pi|^{2m+2}, \quad x \in \pi,$$

$$(12) \quad |u^{(j)}(x) - u_\pi^{(j)}(x)| \leq C |\pi|^s, \quad x \notin \pi,$$

where $s = \min(2m + 2, m + k + 2 - j)$.

COROLLARY 2. *If for a $(k + 1)$ st order differential equation $p_i, f \in C^{2k+2}[a, b]$, $M = 2k + 2, m = k$, i.e., we approximate by piecewise k th degree polynomials, then, as $|\pi| \rightarrow 0$,*

$$||u^{(j)} - u_\pi^{(j)}|| = O(|\pi|^{2k+2-j}), \quad j = 0, 1, \dots, k.$$

Note that in Corollary 2 the coefficient functions have been perturbed by $O(\epsilon)$, $\epsilon = |\pi|^{k+1}$, yet the solution is only perturbed by $O(\epsilon^2)$. As discussed above, for $\text{ran } Q = P_m$, the largest M for which the restriction on Q in Theorem 3 is satisfied is $M = 2m + 2$. Thus, for such approximations, (11)–(12) give best possible error bounds.

There are many examples of projectors for which Theorem 2 is applicable and several which satisfy the additional restriction of Theorem 3. Proofs of the following results and further examples are found in [8].

Example 1. Let Q be the map which takes $G \in C[-1, 1]$ into the m th degree polynomial which interpolates G on some arbitrary set $\{t_i\}_1^{m+1}$ in $[-1, 1]$, then Theorem 2 applies.

Example 2. If Q is chosen as in Example 1 with the restriction that the interpolating points are the roots of the $(m + 1)$ st degree Legendre polynomial, then Q satisfies the hypotheses of Theorem 3 with $M = 2m + 2$.

Example 3. If Q is the map which takes $G \in C[-1, 1]$ into its m th degree polynomial best approximation in the least squares sense, then Theorem 3 applies with $M = 2m + 2$.

5. Richardson's Extrapolation. If Q preserves evenness and oddness, then, for partitions consisting of equally spaced points, the error at mesh points can be expressed as a power series in h^2 : Thus, Richardson's extrapolation can be used to advantage. The representation of the error (4) has the disadvantage that it contains u_π on the right-hand side. This posed no difficulty in Theorem 3 since u_π and its derivatives could be bounded independently of π . To analyze the error more closely, however, requires a different characterization of the error in which the only dependence on π occurs in the coefficient functions.

Define $g_\pi(x) = \int_a^b G(x, t)(f - \hat{f}) dt$ where $G(x, t)$ is as in Theorem 1. Furthermore, for $v \in C^k[a, b]$, set

$$(I_\pi v)(x) = \int_a^b G(x, t)[L - L_\pi]v dt.$$

Then, by a Neumann series argument, (4) implies

$$u(x) - u_\pi(x) = g_\pi(x) + \sum_{i=0}^r I_\pi^i(g_\pi - u) - I_\pi^{r+1}u_\pi$$

for any positive integer r . For the case of equally spaced points, we have $x_n = a + (n - 1)h, n = 1, 2, \dots, N + 1, h = (b - a)/N$. Now, $\|I_\pi\| = O(h^{m+1})$ in the map norm subordinate to $\|\cdot\|_k$, so

$$(13) \quad u(x) - u_\pi(x) = g_\pi(x) + \sum_{i=0}^r I_\pi^i(g_\pi - u) + O(h^{r(m+1)})$$

which is the desired expression for the error.

THEOREM 4. *If $p, f \in C^{r+1}[a, b], r \geq m + 1$, and QF is odd (even) when F is odd (even), then for each n there exist constants α_i independent of h such that*

$$(14) \quad u(x_n) - u_\pi(x_n) = \sum_{i=0}^r \alpha_i h^{2i} + O(h^{m+r+2}).$$

Proof. From Cole [3, p. 165], the Green's operator $\int_a^b dtG(x, t)$ can be written as a sum of $(k + 1)^2$ terms of the form $v_1(x) \int_a^x dtv_2(t)$ and $v_3(x) \int_x^b dtv_4(t)$ where $v_i \in C^{r+1}$. Thus from (13), the desired formula (14) follows if it can be established that integrals of the form

$$\int_a^x v(t)(f - \hat{f}) dt \quad \text{or} \quad \int_x^b v(t)(f - \hat{f}) dt, \quad v \in C^{r+1},$$

have error expansions like the right-hand side of (14) when $x = x_n$, a mesh point. With the notation of Lemma 2,

$$\int_a^{x_n} (f - \hat{f})v dx = \sum_{k=1}^{n-1} \left\{ \sum_{i=0}^r v^{(i)}(z_k)/i!(\frac{1}{2}h)^{i+1} \int_{-1}^1 t^i(1 - Q)F dt \right\} + O(h^{m+r+2}).$$

Since $F \in C^{r+1}[-1, 1]$,

$$\begin{aligned} \int_{-1}^1 t^i(1 - Q)F dt &= \sum_{j=0}^r \left\{ F^{(j)}(0)/j! \int_{-1}^1 t^i(1 - Q)t^j dt \right\} \\ &\quad + \int_{-1}^1 t^{r+1}(1 - Q)(F^{(r+1)}(\xi_t)t^{r+1}) dt/(r + 1)!, \end{aligned}$$

so, with $\beta_{i,j} = \int_{-1}^1 t^i(1 - Q)t^j dt/(2^{i+j+1}i!j!)$, we have

$$(15) \quad \int_a^{x_n} (f - \hat{f})v dx = \sum_{i=0}^r \sum_{j=0}^r \beta_{i,j} h^{i+j} \left\{ h \sum_{k=1}^{n-1} v^{(i)}(z_k) f^{(j)}(z_k) \right\} + O(h^{m+r+2}).$$

The term in braces is just the composite midpoint rule applied to $\int_a^{x_n} v^{(i)}(x)f^{(j)}(x) dx$ which is known [4, p. 168] to have an error which is a power series in h^2 . Finally, from the assumption on $Q, \beta_{i,j} = 0$ for $i + j$ odd and thus only even powers of h occur in the low order part of (15).

In practice, Q is chosen to satisfy the restrictions of Theorem 3 as well, so that $\alpha_0 = \alpha_1 = \dots = \alpha_m = 0$. Examples 2 and 3 in Section 4 satisfy both the restrictions of Theorem 3 and those of Theorem 4.

6. The Numerical Method. With the exception of a special algorithm for the case $k = 1$ and $m = 0$, a shooting method is used to solve (2). The general scheme is described in Keller [6] and consists of computing a number of linearly independent solutions of (2a) with initial value conditions and then finding the appropriate linear combination which satisfies (2b). Thus, the problem is reduced to that of integrating a linear initial value problem where the coefficients are piecewise polynomial functions. This is easily done by Taylor series techniques as the required derivatives can be computed recursively from the differential equation. In particular, given initial values $u_\pi^{(j)}(x)$, $j = 0, 1, \dots, k$, from the differential equation

$$u_\pi^{(k+j+1)} = \hat{f}^{(j)} - \sum_{i=0}^k \sum_{l=0}^i \binom{j}{l} \hat{p}_i^{(l)} u_\pi^{(i+l-l)}, \quad j = 0, 1, \dots$$

But on each (x_n, x_{n+1}) , the coefficient functions are m th degree polynomials so all their derivatives of order $m + 1$ or higher are identically zero. Thus, we have

$$(16) \quad u_\pi^{(k+j+1)} = \hat{f}^{(j)} - \sum_{i=0}^k \sum_{l=0}^m \binom{j}{l} \hat{p}_i^{(l)} u_\pi^{(i+l-l)}, \quad j = 0, 1, \dots$$

For $m = 0, 1, 2$ these sums are not too complicated. If $T_p(h)$ is the operator $\sum_{i=0}^p (h^i/i!) d^i/dx^i$, then since the coefficient functions are analytic in (x_n, x_{n+1}) , for x, h such that $x, \dot{x} + h \in [x_n, x_{n+1}]$, we have

$$u_\pi(x + h) = T_p(h)u_\pi(x) + O(h^{p+1}).$$

Moreover, the quantity $(T_{p+1}(h) - T_p(h))u_\pi$ provides an approximation to the local error in computing u_π which can be used to monitor the accuracy and predict h for the next step. In most cases, π is fairly coarse so it is necessary to use a step size $h < x_{n+1} - x_n$. For the numerical data displayed below, $p = 12$ is used so the local error in computing u_π is $O(h^{13})$.

The computation is more efficient if (16) is rewritten in terms of the scaled derivatives $h^j u_\pi^{(j)}/j!$. For example, when $k = 1$ and $m = 2$, the formulas are

$$\begin{aligned} V_{-2} &= V_{-1} = 0, \\ V_0 &= u_\pi(x), \quad V'_0 = u'_\pi(x), \end{aligned}$$

and, for $j = 1, 2, \dots, p$,

$$\begin{aligned} V_j &= h V'_{j-1}/j, \\ V'_j &= h^j f^{(j-1)}/j! - \hat{p}_1 V_j \\ &\quad - h\{(\hat{p}_0 + (j-1)\hat{p}'_1) V_{j-1} + h[(\hat{p}'_0 + (j-2)\hat{p}''_1/2) V_{j-2} + h\hat{p}'''_0 V_{j-3}/2]\}/j. \end{aligned}$$

Then $u_\pi(x + h) \doteq \sum_{i=0}^p V_i$, $u'_\pi(x + h) \doteq \sum_{i=0}^p V'_i$ and $V_{p+1} = hV'_p/(p + 1)$ is an estimate of the local error in $u_\pi(x + h)$.

7. Numerical Examples. All computations were done in double precision on the IBM 360 using N subintervals of equally spaced points.

The first problem is a 4th order equation ($k = 3$), viz.,

$$u^{(4)}(x) = (x^4 + 14x^3 + 49x^2 + 32x - 12)e^x \quad \text{in } (0, 1),$$

$$u(0) = u'(0) = u(1) = u'(1) = 0,$$

which has the known solution $u(x) = x^2(x - 1)^2e^x$. It is easily seen that if $v(x)$ satisfies $v^{(4)} = f, v(0) = v'(0) = v''(0) = 0, v'''(0) = 1$, then

$$u(x) = v(x) + (v'(1) - 3v(1))x^2 + (2v(1) - v'(1))x^3.$$

Moreover, if \hat{f} is a piecewise m th degree polynomial, then the solution of $v^{(4)} = \hat{f}$ is a piecewise $(m + 4)$ th degree polynomial which can be computed in closed form. Table 1 displays the error for various types of approximation schemes and different mesh sizes. The following notation is used for the types of approximations:

G_m : On each (x_n, x_{n+1}) , \hat{f} is the m th degree polynomial which interpolates f at Gauss-Legendre points.

H_m (Hermite): On each $[x_n, x_{n+1}]$, $\hat{f}^{(j)}$ interpolates $f^{(j)}$, $j = 0, 1, \dots, (m - 1)/2$, at x_n, x_{n+1} .

S_2 (Spline): On each $[x_n, x_{n+1}]$, \hat{f} interpolates f at x_n, x_{n+1} and $\frac{1}{2}(x_n + x_{n+1})$.

S_3 : \hat{f} is the natural cubic interpolatory spline for f on π .

Theorem 3 predicts $\|u^{(j)} - u_\pi^{(j)}\| = O(|\pi|^s)$ where $s = \min(2m + 2, m + 5 - j)$ for G_m . The other schemes have errors at worst $O(|\pi|^{m+1-j})$, although there are some special cases which give slightly higher rates of convergence. The superiority of the methods based on interpolation at Gauss-Legendre points is evident and it is equally clear that demanding smoothness in \hat{f} , as in S_3 , contributes little to smaller errors in u_π .

The remaining examples are second-order problems and, for low degree polynomial approximations, it is not necessary to use shooting as the approximate problem can be solved exactly. For example, when $m = 0$, the equation

$$u''_\pi + p_n u'_\pi + q_n u_\pi = f_n, \quad x \in (x_n, x_{n+1}),$$

has the solution

TABLE 1

| Method | $\ u - u_\pi\ $ | | | $\ u' - u'_\pi\ $ | |
|--------|-----------------------|---------------|---------------|-------------------|---------------|
| | $ \pi = \frac{1}{2}$ | $\frac{1}{4}$ | $\frac{1}{8}$ | $\frac{1}{2}$ | $\frac{1}{4}$ |
| G_0 | 2.25E-2 | 4.69E-3 | 1.08E-3 | 8.44E-2 | 1.75E-2 |
| G_1 | 1.48E-3 | 2.51E-5 | 1.09E-6 | 8.51E-3 | 3.03E-4 |
| G_2 | 8.14E-5 | 1.24E-6 | 1.94E-8 | 4.83E-4 | 1.24E-5 |
| G_3 | 1.74E-6 | 8.25E-9 | 3.50E-11 | 1.30E-5 | 1.24E-7 |
| H_1 | 3.62E-2 | 8.50E-3 | 2.10E-3 | 1.25E-1 | 2.78E-2 |
| H_3 | 6.85E-4 | 4.09E-5 | 2.54E-6 | 2.38E-3 | 1.31E-4 |
| S_2 | 6.98E-4 | 3.41E-5 | 1.96E-6 | 2.80E-3 | 1.28E-4 |
| S_3 | 9.64E-3 | 1.38E-4 | 5.96E-6 | 4.10E-2 | 1.26E-3 |

$$(17) \quad u_{\pi}(x) = \exp[\alpha_n(x - x_n)]\{A_n \cos \beta_n(x - x_n) + B_n \sin \beta_n(x - x_n)\} + f_n/q_n$$

with $\alpha_n = -p_n/2$, $\beta_n = (q_n - \alpha_n^2)^{1/2}$. If β_n is imaginary, it is a simple matter to switch to hyperbolic functions of a real argument; if $q_n = 0$, adjustments can be made which lead to similar results. The continuity of u_{π} and u'_{π} , together with the boundary conditions, give rise to a linear system for $u_{\pi}(x_n)$, $u'_{\pi}(x_n)$, $n = 1, 2, \dots, N + 1$; moreover, the coefficient matrix has a simple band structure. For $m = 1$, when the u' term is not present, a similar procedure can be used with Airy functions, for $m = 2$ parabolic cylinder functions appear in the solutions. Gordon [5] has described ways of evaluating Airy functions but these can be awkward to use, particularly because of scaling and conditioning problems. No method (other than Taylor series) is available for computing parabolic cylinder functions, thus, in this paper, the results for $m > 0$ have been computed by the Taylor algorithm. For $m = 0$, the fact that u_{π} has the form (17) can be used to carry out exact integrations in the shooting method. Since the exponential and trig functions can be evaluated fairly rapidly, this is the method used here.

The following problem is a difficult one for methods based on polynomial approximation of solutions, particularly for conventional shooting techniques.

$$u'' + (3 \cot x + 2 \tan x)u' + .7u = 0, \quad x \in (30^{\circ}, 60^{\circ}),$$

$$u(30^{\circ}) = 0, \quad u(60^{\circ}) = 5.$$

This problem arose in considering stress distributions in a spherical membrane having normal and tangential loads [9]. The solution has a sharp peak near $x = 30.5$, so good local polynomial approximation requires fairly small step sizes. However, the coefficient functions are quite smooth and, in fact, the results in Table 2 have been generated with $N = 6$, i.e., $|\pi| = 5$. The "exact" answers are actually high accuracy results of other methods.

This data illustrates the power of this method for problems with smooth coefficients but ill-behaved solutions: The high order Taylor algorithm for piecewise polynomial functions is stable enough and accurate enough that u_{π} can be computed well by shooting even though u cannot.

TABLE 2

| x | $u(x)$ | $u_{\pi}(x)$ by G_0 | G_1 | G_2 |
|-----|----------|------------------------|----------|----------|
| 35 | 171.653 | 172.092 | 171.651 | 171.652 |
| 40 | 89.0707 | 89.2643 | 89.0694 | 89.0704 |
| 50 | 21.2680 | 21.2935 | 21.2676 | 21.2679 |
| | $u'(x)$ | $u'_{\pi}(x)$ by G_0 | G_1 | G_2 |
| 30 | 1896.44 | 1795.69 | 1889.92 | 1896.22 |
| 35 | -21.5363 | -20.5436 | -21.6044 | -21.5341 |
| 40 | -12.1522 | -11.7617 | -12.1821 | -12.1515 |
| 50 | -3.13100 | -3.11483 | -3.13746 | -3.13100 |

TABLE 3

| | $ \pi = \frac{1}{2}$ | $ (u - u_\pi)/u $ $\frac{1}{4}$ | $\frac{1}{8}$ | $\frac{1}{16}$ |
|-------|-----------------------|------------------------------------|---------------|----------------|
| G_0 | 6.47E-3 | 7.26E-4 | 1.60E-4 | 3.80E-5 |
| G_1 | 1.01E-3 | 4.42E-5 | 2.59E-6 | 1.60E-7 |

The final example is difficult to solve by conventional difference schemes:

$$u'' + \frac{4x}{1+x^2} u' + \frac{2}{1+x^2} u = 0 \quad \text{in } (0, \frac{1}{2}),$$

$$u'(0) = 0, \quad u(\frac{1}{2}) = 8000.$$

which has the known solution $u(x) = 10^4/(1+x^2)$. As in the previous example, the coefficients are more amenable to local polynomial approximation than is the solution. Table 3 contains results for methods based on interpolation at Gauss-Legendre points which yield uniform errors of $O(|\pi|^{2m+2})$ for $m = 0, 1$.

Theorem 4 states that for many approximation schemes, including those based on interpolation at Gauss-Legendre points, Richardson's h^2 -extrapolation can be used. The argument is sufficiently general to include initial value as well as boundary value problems so either local or global extrapolation can be used. The former is much more effective since the error can be monitored as the integration is carried out. For $m = 0$, the fact that u_π has the form (17) means the integration can be done exactly and much more efficiently than by the Taylor algorithm. For the preceding example, the missing initial condition $u(0)$ was computed by this approach and, in spite of the fact that u has very large derivatives, the extrapolation was very effective. With the extrapolation sequence $\{\frac{1}{2}, \frac{1}{4}, \frac{1}{6}, \frac{1}{8}, \frac{1}{12}, \dots\}$ to estimate $u(0)$ with a relative error of 5×10^{-8} required 19 evaluations of the coefficient functions and a relative error of 10^{-14} required only 55 function evaluations.

8. Conclusions and Generalizations. From the above examples, it can be seen that this method is particularly effective when the coefficients can be well-approximated locally by polynomials and the solutions cannot. In essence, this is due to the fact that the solutions are being approximated locally by transcendental functions which are far more able to reflect rapid growth or oscillatory behavior than are polynomials. Even if the coefficients do not lend themselves to accurate local polynomial approximation, the method presented here has one advantage over conventional methods in that one is dealing with known functions, viz., the coefficients. Thus, the degree of the polynomials and the mesh can be chosen to approximate the coefficients well without requiring a special knowledge about the solution.

The Taylor algorithm described above is also useful for problems where the original coefficients are polynomials. The method appears quite stable and the built-in error estimator for the high order method allows fairly large step sizes to be used in the integration.

All of the results in this paper pertain to problems in nonselfadjoint form. Analogous results can be derived by directly treating problems in selfadjoint form.

Department of Mathematics and Statistics
The University of New Mexico
Albuquerque, New Mexico 87106

1. M. ALEXANDER & R. GORDON, "A new method for constructing solutions to time dependent perturbation equations," *J. Chem. Phys.*, v. 55, 1971, pp. 4889–4898.
2. J. CANOSA & R. GOMES DE OLIVEIRA, "A new method for the solution of the Schrödinger equation," *J. Comput. Phys.*, v. 5, 1970, pp. 188–207.
3. R. COLE, *Theory of Ordinary Differential Equations*, Appleton-Century-Crofts, New York, 1968. MR 39 #512.
4. P. DAVIS & P. RABINOWITZ, *Numerical Integration*, Blaisdell, Waltham, Mass., 1967. MR 35 #2482.
5. R. GORDON, "Quantum scattering using piecewise analytic solutions," in *Computational Physics*. Vol. 10 (Edited by B. Adler), Academic Press, New York, 1971, pp. 81–109.
6. H. KELLER, *Numerical Methods for Two-Point Boundary Value Problems*, Blaisdell, Waltham, Mass., 1968. MR 37 #6038.
7. J. NÉCAS, *Les méthodes directes en théorie des équations elliptiques*, Masson, Paris; Academia, Prague, 1967. MR 37 #3168.
8. S. PRUESS, "Estimating the eigenvalues of Sturm-Liouville problems by approximating the differential equation," *SIAM J. Numer. Anal.*, v. 10, 1973, pp. 55–68.
9. A. ROARK & L. SHAMPINE, *On the Numerical Solution of a Linear Two-Point Boundary Value Problem*, Sandia Laboratory Technical Memorandum SC-TM-67-588, September 1967.