

Compact, Implicit Difference Schemes for a Differential Equation's Side Conditions

By Blair Swartz*

Abstract. Lynch and Rice have recently derived compact, implicit (averaged-operator) difference schemes for the approximate solution of an m th order linear ordinary differential equation under m separated side conditions. We construct here a simpler form for a compact, implicit difference scheme which approximates a more general side condition. We relax the order of polynomial exactness required for such approximate side conditions. We prove appropriate convergence rates of the approximate solution (and its first $m - 1$ divided differences) to (those of) the solution, even, of multi-interval differential equations. Appropriate, here, means k th order convergence for schemes whose interior equations are exact for polynomials of order $k + m$ and whose approximation of a side condition of order l is exact for polynomials of order $k + l$. We also prove the feasibility of shooting (and of multiple shooting) based on initial divided differences. The simplicity of the proofs is based upon the simplicity of form of the approximating side conditions, together with the crucial stability result of Lynch and Rice for their interior difference equations under divided-difference initial data.

0. Introduction. In an elegant and understandable expansion and completion of some work of Osborne [1967], Lynch and Rice [1980] derive compact, implicit, averaged-operator difference schemes for the approximate solution of a boundary value problem for a single m th order linear ordinary differential equation

$$(Mu)(t) = f(t), \quad t \text{ in } (A, B);$$
$$(\lambda_i \mu = c_i)_0^{m-1}; \quad M := D^m + \sum_{i < m} \alpha_i D^i.$$

The m side conditions $(\lambda_i)_0^{m-1}$ which they consider consist of separated endpoint functionals (of order less than m) which are assumed linearly independent over nullspace(M).

The interval $[A, B]$ is partitioned by mesh points

$$A = t_0 < t_1 < \cdots < t_n = B,$$

which are constrained by certain local mesh-ratio requirements together with the global restriction that the ratio of the maximum mesh interval H_n to the average mesh interval $(B - A)/n$ be bounded as n grows large. Their interior difference equations

$$(0.1) \quad (M_n U)_i = (A_n(Mu))_i, \quad 0 \leq i \leq n - m,$$

Received August 30, 1979.

1980 *Mathematics Subject Classification*. Primary 65L10, 65N05.

* Work supported by the U. S. Department of Energy.

are *compact* (Kreiss [1972, p. 612]) in that they involve values of the approximation $(U)_j \cong u(t_j)$ at only $m + 1$ successive mesh points. They are in *implicit* (Swartz and Wendroff [1974a, p. 156], [1974b, p. 992], Swartz [1974]), or *averaged-operator, form* in that $(A_n(Mu))_i$ consists of a linear combination of local values of $Mu (=f)$ (not necessarily mesh-point values) whose positive weights sum to one. The coefficients of the interior difference equations M_n and averages A_n are found (at solution time *via* the solution of n local linear systems) by requiring that the truncation error vanish for all polynomials of degree less than $k + m$; and enough local evaluation points for Mu (J of them) are chosen to accomplish this. The connection of J with $k + m$ depends upon whether one wants to specify the local evaluation points in advance of the calculation or to specify them by the solution of n local nonlinear systems depending upon M and upon the local mesh structure and sizes. From the convergence results and from their truncation error analysis, this can be put more memorably as follows: $O(H_n^k)$ convergence can be proven for k rather arbitrary evaluation points; and, if they are chosen more carefully, for down to $k/2$ evaluation points. One could well call this last a superconvergence result for difference schemes.

Lynch and Rice's convergence proof involves the following measures of size for a mesh function V :

$$\|V\|_{m-1,0} := \max_{0 \leq j < m} |\Delta_0^j V|, \quad \Delta_i^j V/j! := V[t_i, \dots, t_{i+j}] := j\text{th divided difference};$$

$$\|V\|_{m-1} := \max_{0 \leq j < m} \max_{0 \leq i \leq n-j} |\Delta_i^j V|, \quad (\text{this last a norm});$$

and begins with a crucial stability result (for their interior equations coupled with initial divided-difference data) which we paraphrase as follows:

STABILITY THEOREM (LYNCH AND RICE). *Suppose H_n is sufficiently small, and suppose that certain mesh-ratio restrictions hold. Then for any mesh function V*

$$\|V\|_{m-1} \leq \text{const}[\|V\|_{m-1,0} + \|M_n V\|_\infty].$$

They then turn to the derivation of a compact, implicit scheme approximating each functional λ_i (a scheme which is exact on the same polynomial space \mathbf{P}_{k+m}); conclude that the truncation error in the composite scheme is $O(H_n^k)$; and prove that $\|U - u\|_{m-1}$ converges at that rate, first for the initial value problem, then for their general boundary value problem.

In this note we express compact, implicit approximations to somewhat more general side conditions λ in a more concise form; a form which not only permits a structurally simple proof of convergence using the above stability theorem, but also points out how, for that portion of a side condition which is of order \hat{m} , the necessary order of polynomial exactness can be reduced from $k + m$ to $k + \hat{m}$.

Next, we observe that the crucial stability result (together with the form of the approximate side conditions) leads, immediately, not only to the feasibility of shooting (using ordinary divided differences as initial data) as a solution technique for the

composite implicit difference equations; but also to high-order convergence rates for multi-interval boundary value problems (e.g., problems where M is piecewise smooth); and finally to the feasibility of a multiple shooting technique.

Lastly, we wish to associate an acronym with our approximate side conditions; one which is more memorable (if not more meaningful) when preceded by the acronym HODIE promulgated by Lynch and Rice [1975], [1978a], [1978b], [1980] for their averaged-operator interior schemes. This acronym is based upon the phrase "Demonstrating Osborne's Obscure (but) Decent Implicit Expressions".

1. Compact, Implicit Difference Schemes for Lower Order Point Functionals.

The boundary value problem considered is

$$Mu := \left(D^m + \sum_{i=0}^{m-1} \alpha_i D^i \right) u = f \text{ in } (A, B), \quad (\lambda_q[u] = c_q)_{q=0}^{m-1},$$

where the linear functionals

$$\lambda_q = \mu_q + \nu_q, \quad \mu_q := \sum_{i=0}^{m_\mu} a_{i,q} (D^i)(A), \quad \nu_q = \sum_{i=0}^{m_\nu} b_{i,q} (D^i)(B)$$

have the property that $(\lambda_q)_0^{m-1}$ are linearly independent on the nullspace of M .

We concentrate our attention on approximating the q th functional and suppress dependence upon q . Furthermore, we approximate μ and ν separately; the sum of the approximations will approximate λ . And we shall approximate the condition $\mu[u] = c$ by the condition $\mu_h[U] = C[u]$ where, for any mesh function V and for v in $C^m[0, 1]$

$$\mu_h[V] := \sum_{i=0}^{m_\mu} a_i \Delta_0^i V + h \sum_{i=0}^{m-1} \gamma_i \Delta_0^i V h^{(i-m_\mu-1)_+}, \quad 0 \leq m_\mu < m; \tag{1.1a}$$

$$C[v] := \mu[v] + h^{m-m_\mu} \sum_{j=1}^J \beta_j (Mv)(t_0 + h\xi_j).$$

Here $\Delta_0^i V := i! V[t_0, \dots, t_i]$; $(j)_+$ is the positive part of j ; $h := t_{m-1} - t_0$; the $(\xi_j)_1^J$ are chosen, well separated, from a fixed interval $[0, a]$; and we propose, for h small enough, to determine the numbers $(\gamma_i)_0^{m-1}$ and $(\beta_j)_1^J$ uniquely by the condition that

$$\mu_h[p] = C[p], \quad p \text{ in } \mathbf{P}_{J+m} := \{\text{polynomials of degree less than } J + m\}.$$

Furthermore, we show that, as $h \rightarrow 0$, the γ 's and β 's are $O(1)$. To see this can be done, set

$$p_l(t) := (t - t_0)^l / l!.$$

Then, since

$$\Delta_0^i p_l = \begin{cases} (D^i p_l)(t_0), & i \geq l, \\ (D^i p_l)(t_0) + O(h^{l-i}), & i < l, \end{cases}$$

we see that

$$\begin{aligned} \mu_h[p_l] &= \mu[p_l] + \sum_{i < l}^{m_\mu} a_i \cdot O(h^{l-i}) + \sum_{i=0}^{m-1} \gamma_i h^{1+(l-m_\mu-1)+} (D^l p_l)(t_0) \\ &\quad + \sum_{i < l}^{m-1} \gamma_i O(h^{l+1-i+(l-m_\mu-1)+}) \\ &= \mu[p_l] + O(h^{\max(1, l-m_\mu)}) + \sum_{i=0}^{\min(l-1, m_\mu)} O(h^{l+1-i}) \gamma_i \\ &\quad + \sum_{i=m_\mu+1}^{\min(l-1, m-1)} O(h^{l-m_\mu}) \gamma_i \\ &\quad + \sum_{i=0}^{m-1} h^{\max(1, l-m_\mu)} \delta_{i,l} \gamma_i. \end{aligned}$$

On the other hand, with $\alpha_{l,j} := \alpha_j(t_0 + \xi_j h)$,

$$C[p_l] = \mu[p_l] + \begin{cases} h^{m-m_\mu} \sum_{j=1}^J [\alpha_{l,j} + O(h)] \beta_j, & l < m; \\ (l-m)! h^{l-m_\mu} \sum_{j=1}^J [\xi_j^{l-m} + O(h)] \beta_j, & l \geq m. \end{cases}$$

In order that $\mu_h = C$ on P_{J+m} we must satisfy the $J + m$ linear equations $(\mu_h[p_l] - C[p_l] = 0)_{l=0}^{J+m-1}$. The existence argument for the unknowns $(\gamma_i)_{i=0}^{m-1}$ and $(\beta_j)_{j=1}^J$ considers two cases. First we suppose that $m_\mu = m - 1$. The equations then take two forms:

(a) for $l < m$,

$$0 = O(h) + h \sum_{i=0}^{l-1} O(h^{l-i}) \gamma_i + h \gamma_l - h \sum_{j=1}^J [\alpha_{l,j} + O(h)] \beta_j;$$

(b) for $l \geq m$,

$$0 = O(h^{l+1-m}) + h \sum_{i=0}^{m-1} O(h^{l-i}) \gamma_i - (l-m)! h^{l+1-m} \sum_{j=1}^J [\xi_j^{l-m} + O(h)] \beta_j.$$

In this case one, therefore, solves a linear system (with $O(1)$ right-hand side) whose coefficients are an $O(h)$ perturbation of those in the following system:

$$\gamma_l - \sum_{j=1}^J \alpha_j(t_0) \beta_j = O(1), \quad 0 \leq l < m; \quad (l-m)! \sum_{j=1}^J \xi_j^{l-m} \beta_j = O(1), \quad l \geq m.$$

The last J equations involve $\text{diag}[(l-m)!]$ (a Vandermonde matrix). Hence the unperturbed (thus the perturbed) systems have uniformly bounded inverses (h small). Consequently, as $h \rightarrow 0$, unique and uniformly bounded $(\gamma_i)_{i=0}^{m-1}$ and $(\beta_j)_{j=1}^J$ exist for a given operator M and for the case $m_\mu = m - 1$.

The existence argument differs slightly when $m_\mu < m - 1$. Then the $m + J$ equations take three possible forms:

(a) for $l \leq \min(m_\mu + 1, m - 1)$,

$$0 = O(h) + h \sum_{i=0}^{l-1} O(h^{l-i})\gamma_i + h\gamma_l - O(h^2) \sum_{j=1}^J [\alpha_{l,j} + O(h)]\beta_j;$$

(b) if $m_\mu < m - 2$, then for $m_\mu + 1 < l < m$,

$$0 = O(h^{l-m_\mu}) + \sum_{i=0}^{m_\mu} O(h^{l+1-i})\gamma_i + \sum_{i=m_\mu+1}^{l-1} O(h^{l-m_\mu})\gamma_i + h^{l-m_\mu}\gamma_l - O(h^{l-m_\mu+1}) \sum [a_{l,j} + O(h)]\beta_j;$$

(c) for $l \geq m$,

$$0 = O(h^{l-m_\mu}) + \sum_{i=0}^{m_\mu} O(h^{l+1-i})\gamma_i + \sum_{i=m_\mu+1}^{m-1} O(h^{l-m_\mu})\gamma_i - (l-m)!h^{l-m_\mu} \sum_{j=1}^J [\xi_j^{l-m} + O(h)]\beta_j.$$

One, therefore, solves a linear system (with $O(1)$ right-hand side) whose coefficients are an $O(h)$ perturbation of those in the following system:

$$\begin{aligned} \gamma_l &= O(1), \quad l \leq \min(m_\mu + 1, m - 1); \\ \sum_{i=m_\mu+1}^{l-1} O(1)\gamma_i + \gamma_l &= O(1), \quad \left\{ \begin{array}{l} \text{(only if } m_\mu < m - 2; \\ \text{then for } m_\mu + 1 < l \leq m - 1); \end{array} \right. \\ \sum_{i=m_\mu+1}^{m-1} O(1)\gamma_i - (l-m)! \sum_{j=1}^J \xi_j^{l-m}\beta_j &= O(1), \quad l \geq m. \end{aligned}$$

The first m equations are lower triangular with 1's on the diagonal. We conclude, again, that as $h \rightarrow 0$, unique and uniformly bounded $(\gamma_i)_0^{m-1}$ and $(\beta_j)_1^J$ exist for given operator M and for $m_\mu < m - 1$.

For a general boundary condition $\lambda = \mu + \nu$ we approximate $\nu[u] = c$ separately by $\nu_h[U] = C_\nu[u]$ using $\nabla_n^i U := i!U[t_n, \dots, t_{n-i}]$, $0 \leq i < m$. Then the approximation of $\lambda[u] = c$ is given by

$$(1.1b) \quad \begin{aligned} \lambda_h[U] := \mu_h[U] + \nu_h[U] &= \lambda[u] + h_\mu^{m-m_\mu} \sum_{j=1}^{J_\mu} \beta_{j,\mu}(Mu)(t_0 + h_\mu \xi_{j\mu}) \\ &+ h_\nu^{m-m_\nu} \sum_{j=1}^{J_\nu} \beta_{j,\nu}(Mu)(t_n - h_\nu \xi_{j\nu}) \quad (= C_\mu[u] + C_\nu[u]). \end{aligned}$$

The truncation error $T_\lambda[u]$ is defined for u in $C^m[A, B]$ as

$$T_\lambda[u] := \lambda_h[u] - C_\mu[u] - C_\nu[u].$$

Approximate u by its Taylor polynomial of degree $J_\mu + m - 1$ on $[t_0, t_{m-1}]$, by a corresponding Taylor polynomial of degree $J_\nu + m - 1$ on $[t_{n-m+1}, t_n]$, by anything elsewhere. The truncation error is the truncation error in the remainder, and it satisfies (because the γ 's and the β 's are bounded)

$$T_\lambda[u] = O(H_n^{m+J_\mu-m\mu}) + O(H_n^{m+J_\nu-m\nu})$$

$$(u \in C^{m+J_\mu}[A, A + \epsilon_A] \cap C^{m+J_\nu}[B - \epsilon_B, B]).$$

To maintain *uniform* order of truncation error $O(H_n^k)$ (that of the interior equations) we set

$$(1.2) \quad J_\mu := k - (m - m_\mu), \quad J_\nu := k - (m - m_\nu);$$

i.e., $\mu_h - C_\mu$ is to vanish on \mathbf{P}_{k+m_μ} , $\nu_h - C_\nu$ on \mathbf{P}_{k+m_ν} . Thus, as far as truncation error goes, the order of polynomial approximation necessary for a boundary condition is appropriately smaller than that required for the differential operator.

Remark. The same argument yields compact, implicit difference schemes for multipoint functionals of the form

$$\lambda = \sum_{i=1}^k \sum_{j=0}^{m_i} a_{ij}(D^j)(z_i), \quad (z_i \text{ in } [A, B]_1^k, (m_i < m)_0^k),$$

as long as each mesh $(t_i)_0^n$ contains $(z_i)_0^k$.

Lynch has conjectured (personal communication) the existence of special evaluation points for Mu near t_0 so that J_μ can be less than (1.2) but the approximations are still exact on \mathbf{P}_{k+m_μ} . (For $M = D^2$, $\lambda = D|_A$, see Lynch and Rice [1980, Ex. (2.2')–(2.6')]. Each λ_h there is too accurate for its associated (M_n, A_n) .)

As an example of (1.1) we consider an approximation of $\lambda[u] := (Du)(t_0)$ using U at two values of t (namely t_0 and $t_\pm := t_0 \pm h, h > 0$) and two corresponding values of

$$(1.3) \quad Mu := a D^2u + b Du + cu.$$

It is

$$(1.4) \quad \lambda_\pm[U] = C_\pm[u];$$

where, with $D_\pm U := (U_\pm - U_0)/(t_\pm - t_0)$,

$$\lambda_\pm[U] := D_\pm U \pm hd^{-1} \{ [(4a_\pm \pm b_\pm h)c_0 + 2a_0c_\pm]U_0$$

$$+ [4a_\pm b_0 + 2a_0b_\pm \pm h(b_0b_\pm + 2a_0c_\pm)]D_\pm U \},$$

$$C_\pm[u] := (Du)(t_0) \pm hd^{-1} [(4a_\pm \pm b_\pm h)(Mu)_0 + 2a_0(Mu)_\pm];$$

and one assumes h small enough so that

$$d := 12a_0a_{\pm} \pm 4(a_0b_{\pm} - a_{\pm}b_0)h - b_0b_{\pm}h^2 \neq 0.$$

For $a \equiv 1, b \equiv c \equiv 0$, it reduces to $D_{\pm}U = (Du)(t_0) \pm h[2(D^2u)(t_0) + (D^2u)(t_{\pm})]/6$. Being exact for cubics, the $O(h^3)$ (but not $O(h^4)$) truncation error in (1.4) yields $O(H_n^3)$ convergence when used with the tridiagonal averaged-operator Lynch and Rice [1980] scheme associating $U_i, U_{i\pm 1}$ with $(Mu)_i, (Mu)_{i\pm 1}$ for a nonuniform mesh $(t_i)_0^n$. It is conceivable that an argument based on a maximum principle would show that, when the mesh is uniform, it yields $O(h^4)$ convergence when coupled with the tridiagonal, averaged-operator scheme for (1.3) found in Swartz [1974, pp. 304–305].

An examination of the form of $C_{\pm}[u]$ above shows that words like “identity expansion” or, even, “averaged-operator” are somewhat misleading when attempting to describe the structure of approximations like (1.1).

2. Convergence for the Initial, Then the Boundary, Value Problem. We now prove appropriate convergence rates for our implicit difference schemes. We first use the stability of the initial value difference scheme (divided-difference data) together with the high order truncation error of the implicit initial value difference scheme to prove high order convergence (discretization error) for the IVP. From this, from the initial divided-difference stability, and from the high order truncation error for the approximating (unseparated but linearly independent) boundary conditions, we then prove the same convergence rate holds for the general boundary value problem. A similar sequence of results is followed by Keller [1976, Chapter 2], except that he considers first order systems, general (noncompact) interior; difference equations, but exact boundary conditions. The result also constitutes an extension (to reasonably general nonuniform spatial mesh discretizations) of the easiest, but not the least significant, of Kreiss’ results, [1972], namely his convergence proof for what he calls compact difference schemes.

For any mesh function V we again set

$$\|V\|_{m-1,0} := \max_{0 \leq j < m} |\Delta_0^j V|, \quad \Delta_i^j V := j!V[t_i, \dots, t_{i+j}],$$

$$\|V\|_{m-1} := \max_{0 \leq j < m} \max_{0 \leq i \leq n-j} |\Delta_i^j V|.$$

First, we determine the convergence rate for the solution of the initial value problem

$$Mu = f; \quad \lambda_q[u] := (D^q u)(t_0) = c_q, \quad 0 \leq q < m,$$

as approximated by

$$M_n U = A_n f; \quad \lambda_{q,h}[U] = C_q[u], \quad 0 \leq q < m;$$

here $\lambda_{q,h}$ and C_q are given in Section 1. Recall, now, that the stability theorem states conditions under which $\|V\|_{m-1} \leq \text{const}(\|V\|_{m-1,0} + \|M_n V\|_{\infty})$, uniformly and for all mesh functions V . So consider, as usual, the mesh function $E := u - U$, which

satisfies the truncation error difference equations (Lynch and Rice [1980], Section 1 herein)

$$M_n E = T = O(H_n^k); \quad \lambda_{q,h}[E] = T_{\lambda_q}[u] = O(H_n^k), \quad 0 \leq q < m; k > 0.$$

Since, for $0 \leq q < m$,

$$\lambda_{q,h}[E] = \Delta_0^q E + h \sum_{j=0}^{m-1} \gamma_{j,q} \Delta_0^j E h^{(q-j-1)+}$$

with $(\gamma_{j,q})_0^{m-1}$ uniformly bounded, we may solve for the initial divided differences of E to find

$$\Delta_0^q E = O(H_n^k), \quad 0 \leq q < m.$$

Using the stability result, now, the convergence rate is $\|E\|_{m-1} = O(H_n^k)$ for the initial value problem with the implicit difference schemes (Section 1) as initial data.

We turn to the general boundary value problem

$$Mu = f, \quad (\lambda_q[u] = c_q)_{q=0}^{m-1},$$

as approximated by

$$M_n U = A_n f, \quad (\lambda_{q,h}[U] = C_{\mu,q}[u] + C_{\nu,q}[u])_{q=0}^{m-1},$$

using the approximate boundary conditions as given in Section 1. We first note that the mesh function $E := u - U$ satisfies the truncation error difference equations

$$M_n E = T = O(H_n^k), \quad (\lambda_{q,h}[E] = T_{\lambda_q}[u])_{q=0}^{m-1}.$$

Let U_{IN} be the above approximate solution to the initial value problem for u . Then

$$E = E_{\text{IN}} + (E - E_{\text{IN}}), \quad \|E_{\text{IN}}\|_{m-1} = O(H_n^k).$$

Furthermore,

$$M_n(E - E_{\text{IN}}) = 0, \quad (\lambda_{q,h}[E - E_{\text{IN}}] = T_{\lambda_q}[u] - \lambda_{q,h}[E_{\text{IN}}])_{q=0}^{m-1}.$$

But, from the uniform boundedness of the coefficients of the divided differences in $\lambda_{q,h}$ and from the convergence rate for the approximate initial value problem,

$$\lambda_{q,h}[E_{\text{IN}}] = O(H_n^k), \quad 0 \leq q < m,$$

as well. We would like to conclude from this that

$$\Delta_0^q(E - E_{\text{IN}}) = O(H_n^k), \quad 0 \leq q < m,$$

for then $\|E - E_{\text{IN}}\|_{m-1}$ (and hence $\|E\|_{m-1}$) would be $O(H_n^k)$ from the stability result.

Thus, it remains to show that $(\Delta_0^q V)_0^{m-1}$ is uniformly bounded by $(\lambda_{q,h}[V])_0^{m-1}$ on the subspace (which is m -dimensional by the stability result for H_n small)

$$V_{0,h} := \text{nullspace } M_n.$$

To see this we recall the assumption on the boundary conditions that $(\lambda_q)_0^{m-1}$ are

presumed linearly independent on

$$V_0 := \text{nullspace } M;$$

i.e., since the m initial value functionals are similarly independent,

$$(2.1) \quad (\lambda_0[v], \dots, \lambda_{m-1}[v])^T = \Lambda(v(t_0), \dots, D^{m-1}(t_0))^T, \quad v \in V_0;$$

where the $m \times m$ matrix Λ is invertible. We need a corresponding result to hold uniformly on $V_{0,h}$. We begin with

LEMMA 2.1. *Let $L: V_{0,h} \rightarrow V_0$ be defined by $(D^q(LV))(t_0) = \Delta_0^q V)_0^{m-1}$, $V \in V_{0,h}$, $LV \in V_0$. Then, uniformly for H_n small,*

$$\|D^i(LV)\|_\infty \leq \text{const} \|V\|_{m-1,0}, \quad 0 \leq i \leq k + m.$$

Proof. For H_n small the map is well defined. With $(v_q)_0^{m-1}$ satisfying $Mv_q = 0$, $(D^j v_q)(t_0) = \delta_{jq}$,

$$\|D^i(LV)\|_\infty = \left\| \sum_{q=0}^{m-1} \Delta_0^q V D^i(v_q) \right\|_\infty.$$

As a corollary, using the stability theorem, the truncation error estimates for the HODIE interior equations, and the usual error estimates for $\nabla_n^q r$ as approximations of $D^q r(t_n)$, we have

COROLLARY 2.2. *If $V \in V_{0,h}$, then for $0 \leq q < m$,*

$$|\nabla_n^q V - D^q(LV)(t_n)| \leq \text{const} \cdot H_n \|V\|_{m-1,0}.$$

COROLLARY 2.3. *If $V \in V_{0,h}$, then for $0 \leq q < m$,*

$$\left| \sum_{j=0}^{m_{\mu,q}} a_{j,q} \Delta_0^j V + \sum_{j=0}^{m_{\nu,q}} b_{j,q} \nabla_n^j V - \lambda_q [LV] \right| \leq \text{const} H_n \|V\|_{m-1,0}.$$

Since $(\gamma_{j,q})_0^{m-1} \ 0^{m-1}$ are uniformly bounded, we finally have

COROLLARY 2.4. *For $V \in V_{0,h}$,*

$$|(\lambda_{0,h}[V], \dots, \lambda_{m-1,h}[V])^T - \Lambda(\Delta_0^0 V, \dots, \Delta_0^{m-1} V)^T| \leq \text{const} H_n \|V\|_{m-1,0}.$$

That is to say, the linear map on $(\Delta_0^j V)_0^{m-1}$, V in $V_{0,h}$, defined by the first term satisfies the following desired analog of (2.1):

$$(\lambda_{0,h}[V], \dots, \lambda_{m-1,h}[V])^T = (\Lambda + O(H_n))(\Delta_0^0 V, \dots, \Delta_0^{m-1} V)^T, \quad V \text{ in } V_{0,h}.$$

Since Λ is invertible and independent of n ,

$$(\Delta_0^q V)_0^{m-1} \text{ is uniformly bounded (as } H_n \rightarrow 0) \text{ by } (\lambda_{q,h}[V])_0^{m-1}$$

on $V_{0,h}$, which (in addition to showing that the full set of difference equations has a unique solution) is what we needed to complete the proof of the following convergence

statement:

THEOREM 2.5. *The unique solution U (and its first $m - 1$ divided differences) of the compact, implicit difference scheme*

$$M_n U = A_n f \quad (\lambda_{q,h}[U] = C_q[u])_0^{m-1},$$

converges to (those of) the solution u of the m th order boundary value problem

$$Mu = f \quad (\lambda_q[u] = c_q)_0^{m-1}$$

$((\lambda_q)_0^{m-1}$ presumed independent on $\text{nullspace}(M)$) at a rate $O(H_n^k)$, i.e., $\|U - u\|_{m-1} = O(H_n^k)$, under the hypotheses that

- (a) the maximum mesh size H_n is small enough,
- (b) the local and global mesh-ratio assumptions for the Stability Theorem hold,
- (c) the interior difference equations (0.1) are exact for polynomials of degree less than $k + m$,
- (d) that portion of each side condition (1.1) which is of order m_μ is exact for polynomials of degree less than $k + m_\mu$, and
- (e) u is in $C^{k+m}[A, B]$.

A sufficient condition for (d) is given by (1.2).

COROLLARY 2.6 (Shooting using initial value divided differences). *Let $(U_i)_0^m$ satisfy the $m + 1$ difference equations*

$$(M_n U_i = 0, (\Delta_0^j U_i = \delta_{ij})_{j=0}^{m-1})_{i=0}^{m-1}; \quad M_n U_m = A_n f \quad (\Delta_0^j U_m = 0)_{j=0}^{m-1}.$$

Then

$$M_n U = A_n f \quad (\lambda_{i,h}[U] = C_{\mu_i}[u] + C_{\nu_i}[u])_{i=0}^{m-1};$$

where the linear system for the unknown coefficients in the expression $U = U_m - \sum_{j=0}^{m-1} c_j U_j$, namely,

$$\sum_{j=0}^{m-1} \lambda_{i,h}[U_j] c_j = \lambda_{i,h}[U_m] - C_{\mu_i}[u] - C_{\nu_i}[u], \quad 0 \leq i < m,$$

has a uniformly bounded inverse for H_n small.

Proof. $V := U_m - U = \sum_{j=0}^{m-1} c_j U_j \in V_{0,h}$ satisfies

$$\left(\Delta_0^i V = \sum_{j=0}^{m-1} c_j \Delta_0^i U_j = c_i \right)_0^{m-1}, \quad (\lambda_{i,h}[V] = \lambda_{i,h}[U_m] - C_{\mu_i}[u] - C_{\nu_i}[u])_0^{m-1}.$$

But $(\Delta_0^i)_0^{m-1}$ is uniformly bounded by $(\lambda_{i,h})_0^{m-1}$ on $V_{0,h}$ for H_n small.

Remark. The results of this section are easily extended to cover the multi-point side conditions mentioned at the end of the last section as long as those conditions are collectively linearly independent on $\text{nullspace}(M)$.

Remark. One presumes that the (at this point somewhat useless) stability result

$$\|V\|_{m-1} \leq \text{const}[\|\lambda_h[V]\|_\infty + \|M_n V\|_\infty], \text{ all mesh functions } V,$$

would follow from convergence (to smooth solutions) and seeming consistency via an appropriate application of uniform boundedness. We may have missed a simple trick, but we are unable to provide an elementary proof.

3. Multi-Interval Problems; Multiple Shooting. The derivation of compact, implicit difference schemes for multi-interval problems; the proof of their convergence rates; and a delineation of multiple shooting for single-interval problems are, at this point, mostly matters of appropriate definitions.

Thus, take the differential equations as

$$(M^\pm u)(t) = f(t), \quad t \in (A_\pm, B_\pm) =: I_\pm, \quad A_+ = B_-, \quad \text{order } M^\pm = m_\pm,$$

under side conditions $\lambda[u] = c$, where $\lambda = (\lambda_1, \dots, \lambda_{m_+ + m_-})^T$ consists of $m_+ + m_-$ linear functionals on $C^{m_-}[I_-] \times C^{m_+}[I_+]$. Each functional λ is presumed to be of point functional form

$$\lambda = \sum_{i=1}^{k_-} \mu_{z_-,i} + \sum_{i=1}^{k_+} \mu_{z_+,i}; \quad \text{each } z_{\pm,i} \in \bar{I}_\pm; \quad \text{each } \mu_{z_\pm} = \sum_{j=0}^{m_{z_\pm} < m_\pm} a_{z_\pm j} (D^j)(z_\pm);$$

and we suppose $(\lambda_j)_1^{m_- + m_+}$ to be linearly independent on

$$V_0 := \text{nullspace}(M^-) \times \text{nullspace}(M^+);$$

that is to say, that there is a nonsingular matrix Λ so that

$$\lambda[v] = \Lambda(v(A_-), \dots, D^{m_- - 1}v(A_-), v(A_+), \dots, D^{m_+ - 1}v(A_+))^T, \quad v \in V_0.$$

As to difference schemes, we suppose that the points $(z_{\pm,i})_1^{k_\pm}$ lie in every mesh $A_\pm = t_0^\pm < \dots < t_{n_\pm}^\pm = B_\pm$; and invoke, as ‘‘interior’’ equations, the compact, implicit schemes $M_{n_\pm}^\pm(U^\pm) = A_{n_\pm}^\pm f$, which utilize enough evaluation points for f so that they each have $O(H^k)$ truncation error. Here $H := \max_{+,-}(H_{n_\pm})$; and we now define norms on mesh functions as

$$\begin{aligned} \|V\|_{m-1,0} &:= \max_{+,-} \|V^\pm\|_{m_+ - 1,0}, \quad \|V\|_{m-1} := \max_{+,-} \|V^\pm\|_{m_\pm - 1}, \\ \|M_n V\| &:= \max_{+,-} \|M_{n_\pm}^\pm V^\pm\|_\infty; \end{aligned}$$

note that the last two expressions involve no differences across the common boundary of I_- and I_+ .

Given all these concepts, then, we conclude the following results:

LEMMA 3.1 (Initial value divided difference stability). *If H is small enough, then*

$$\|V\|_{m-1} \leq \text{const}(\|V\|_{m-1,0} + \|M_n V\|)$$

for all mesh functions V .

LEMMA 3.2. We may approximate each side condition $\lambda[u] = c$ with $\lambda_h[U] = C_\lambda[u]$ by summing approximations to its component functionals $\mu_{z_{\pm,i}}$, as in Section 1, utilizing forward or backward differences as appropriate together with enough auxiliary evaluations of $M^\pm u$ near each $z_{\pm,i}$ to attain $O(H^k)$ truncation error.

LEMMA 3.3. For $V \in V_{0,h} := \text{nullspace}(M_{n-}^-) \times \text{nullspace}(M_{n+}^+)$,

$$\|V\|_{m-1,0} \leq \text{const} \|\lambda_h[V]\|_\infty.$$

Proof. As in Lemma 2.1, with LV satisfying the pair of initial value problems

$$M^\pm(LV) = 0, \quad (D^j(LV))(t_0^\pm) = \Delta_0^j V^\pm)^{m_\pm-1},$$

we have

$$\|D^i(LV)\|_\infty \leq \text{const} \|V\|_{m-1,0}, \quad i \leq \text{const};$$

the remainder of the proof follows by tracing through Corollaries 2.2–2.4. And, as in Section 2, we conclude also

COROLLARY 3.4. (a) *Convergence:* $\|U - u\|_{m-1} = O(H^k)$; (b) *shooting, using the pair of initial value divided-difference problems approximating $M^\pm u = f$ on I_\pm , works.*

COROLLARY 3.5. *The version of multiple shooting for the problem*

$$Mu = f \quad \text{in } (A_-, B_+), \quad (\lambda_q[u] = c_q)_0^{m-1},$$

which is defined by shooting (using the pair of initial divided-difference problems) for the multi-interval problem

$$\begin{aligned} M^\pm u &= Mu \quad \text{in } I_\pm := (A_\pm, B_\pm), \quad B_- = A_+ = C \quad \text{in } (A_-, B_+); \\ \lambda_q[u] &= c_q, \quad 0 \leq q \leq m-1 \quad (\text{L.I.D. on } \text{nullspace}(M)); \\ \lambda_q[u] &= (D^{q-m}v)(B_-) - (D^{q-m}u)(A_+) = 0, \quad m \leq q \leq 2m-1; \end{aligned}$$

works and yields

$$\max_{+,-} (\|U^+ - u\|_{m-1}, \|U^- - u\|_{m-1}) = O(H^k)$$

(note we say nothing here about the error in divided differences which involve mesh points from both I_+ and I_-).

Proof. Because of Corollary 3.4 it suffices to show that $(\lambda_q)_0^{2m-1}$ are linearly independent on V_0 . But if $(\lambda_q[v] = 0)_m^{2m-1}$, v in V_0 , then $(D^j v)(B_-) = D_+^j v(A_+)^{m-1}$. By the continuity of the coefficients of M at $C = B_- = A_+$,

$$\bar{v}(t) := v_\pm(t), \quad t \text{ in } [A_\pm, B_\pm],$$

lies in $\text{nullspace}(M)$. Since $(\lambda_q[\bar{v}] = \lambda_q[v] = 0)_0^{m-1}$, $\bar{v} = 0$; i.e., $v = 0$.

4. Convergence for Systems via the Stability of Other Systems. Since there is nothing in Section 2 which assumes that u is a scalar function, that section provides proof of $O(H_n^k)$ convergence of difference approximations to the solution of a general

linear boundary value problem under the following hypotheses:

(a) the problem has the form

$$Mu = f, \quad \lambda[u] = c,$$

where M is of order $(m)_i$ in $(u)_i$, λ is of order less than $(m)_i$ in $(u)_i$, $\dim u = \dim m$, $\dim c = \Sigma(m)_i$;

(b) the initial value problem

$$Mu = f, \quad [u(A)^T, (Du)(A)^T, \dots, (D^{m-1}u)(A)^T]^T = c$$

is well posed;

(c) there is a nonsingular matrix Λ such that

$$\lambda[u] = \Lambda[u(A)^T, \dots, (D^{m-1}u)(A)^T]^T$$

for u in nullspace(M);

(d) with $\|V\|_{m-1,0}$ and $\|V\|_{m-1}$ involving the divided differences of $(V)_i$ through order $(m)_i - 1$, the initial value stability theorem

$$\|V\|_{m-1} \leq \text{const}[\|V\|_{m-1,0} + \|M_n V\|_\infty]$$

can be proven;

(e) each approximate side condition $(\lambda_n[U])_j = (C_n[u])_j$ can be expressed in a form analogous to (1.1) with bounded γ 's and β 's;

(f) the truncation error

$$\max[\|M_n v - A_n M v\|_\infty, \|\lambda_n[v] - C_n[v]\|_\infty]$$

can be shown to be $O(H_n^k \|v\|)$, $\|v\|$ the max norm of $v, \dots, D^{k+\sigma}v$, some σ . The hard parts, of course, are (d), (e), and (f); about which we conjecture nothing.

There may well be systems for which stability (perhaps uniformly in some parameter) is more easily proven for side conditions other than initial value conditions. In any event, one could also pursue convergence *via* assumptions like

(b') there is a problem $Mu = f, \mu[u] = c$ which is well posed;

(c') for some nonsingular $\Lambda, \lambda[u] = \Lambda\mu[u], u$ in nullspace(M);

(d') there exist approximate side conditions μ_n satisfying $\|(\mu_n - \mu)[v]\|_\infty = o(H_n)$, v smooth, such that it can be proven for all mesh functions V that

$$\|V\|_{m-1} \leq \text{const}[\|\mu_n[V]\|_\infty + \|M_n V\|_\infty].$$

The norm $\|V\|_{m-1}$, too, could be weakened just as long as one could go through the essential argument of Section 2, namely, that using the stability and truncation errors for (M_n, μ_n) and working *via* nullspace(M), one can show $\|\mu_n[V]\|_\infty \leq \text{const}\|\lambda_n[V]\|_\infty, V$ in nullspace(M_n).

Additional references, mostly concerning the interior implicit difference schemes (M_n, A_n) , are appended below; see especially Doedel's work, traceable from Doedel [1980], going back to 1976. Nonlinear problems are discussed in Stepleman [1976], two of Doedel's papers, and Keller and Pereyra [1979]. This last paper devised compact implicit schemes for certain classes of operators M , mentions compact implicit

boundary approximations, but (most importantly) considers deferred corrections. Osborne [1975] creates a nice connection between implicit schemes and collocation with $C^{(m-1)}$ piecewise polynomials, a connection which Doedel more thoroughly expresses and explores later.

T-Division
University of California
Los Alamos Scientific Laboratory
Los Alamos, New Mexico 87545

H. B. KELLER [1976], *Numerical Solution of Two Point Boundary Value Problems*, (esp. Chapter 2), Regional Conf. Series in Applied Math. #24, SIAM, Philadelphia, Pa.

H.-O. KREISS [1972], "Difference approximations for boundary and eigenvalue problems for ordinary differential equations," *Math. Comp.*, v. 26, pp. 605–624.

R. E. LYNCH & J. R. RICE [1975], *The HODIE Method: A Brief Introduction with Summary of Computational Properties*, Dept. of Comput. Sci. Report #170, Purdue Univ., West Lafayette, Ind.

R. E. LYNCH & J. R. RICE [1978a], "High accuracy finite difference approximation to solution of elliptic partial differential equations," *Proc. Nat. Acad. Sci. U.S.A.*, v. 75, pp. 2541–2544.

R. E. LYNCH & J. R. RICE [1978b], "The performance of the HODIE method for solving elliptic partial differential equations," in *Recent Developments in Numerical Analysis* (C. de Boor, Ed.), Proc. of an MRC Conf., Academic Press, New York.

R. E. LYNCH & J. R. RICE [1980], "A high-order difference method for differential equations," *Math. Comp.*, v. 34, pp. 333–372.

M. R. OSBORNE [1967], "Minimizing truncation error in finite difference approximations to ordinary differential equations," *Math. Comp.*, v. 21, pp. 133–145.

B. SWARTZ & B. WENDROFF [1974a], "The comparative efficiency of certain finite element and finite difference methods for a hyperbolic problem," in *Conf. on the Numerical Solution of Differential Equations* (G. A. Watson, Ed.), Lecture Notes in Math., Vol. 363, Springer-Verlag, New York, pp. 153–163.

B. SWARTZ & B. WENDROFF [1974b], "The relative efficiency of finite element methods. I. Hyperbolic problems and splines," *SIAM J. Numer. Anal.*, v. 11, pp. 979–993.

B. SWARTZ [1974], "The construction and comparison of finite difference analogs of some finite element schemes," in *Mathematical Aspects of Finite Elements in Partial Differential Equations* (C. de Boor, Ed.), Academic Press, New York, pp. 279–312.

E. J. DOEDEL [1980], "Some stability theorems for finite difference collocation methods on nonuniform meshes," *BIT*. (To appear.)

H. B. KELLER & V. PEREYRA [1979], "Difference methods and deferred corrections for ordinary boundary value problems," *SIAM J. Numer. Anal.*, v. 16, pp. 241–259.

M. R. OSBORNE [1975], "Collocation, difference equations, and stitched function representations," in *Topics in Numerical Analysis II* (John J. H. Miller, Ed.), Academic Press, New York, pp. 121–132.

R. S. STEPLEMAN [1976], "Tridiagonal fourth order approximations to general two-point nonlinear boundary value problems with mixed boundary conditions," *Math. Comp.*, v. 30, pp. 92–103.