

## Approximation Methods for Nonlinear Problems with Application to Two-Point Boundary Value Problems

By H. B. Keller\*

**Abstract.** General nonlinear problems in the abstract form  $F(x) = 0$  and corresponding families of approximating problems in the form  $F_h(x_h) = 0$  are considered (in an appropriate Banach space setting). The relation between "isolation" and "stability" of solutions is briefly studied. The main result shows, essentially, that, if the nonlinear problem has an isolated solution and the approximating family has stable Lipschitz continuous linearizations, then the approximating problem has a stable solution which is close to the exact solution. Error estimates are obtained and Newton's method is shown to converge quadratically. These results are then used to justify a broad class of difference schemes (resembling linear multistep methods) for general nonlinear two-point boundary value problems.

**1. Introduction.** We present a general abstract study of methods for approximating the solution of nonlinear problems formulated in a Banach space setting. Our basic results are of the following kind: *If the nonlinear problem has a solution, and a consistent approximating problem has a stable Lipschitz continuous linearization (i.e., Fréchet derivative), then the approximating problem has a stable solution which is close to the exact solution.* Estimates of the error are given in terms of the order of consistency, and Newton's method is shown to converge quadratically for computing the approximate solution. Asymptotic error expansions can also be derived under appropriate assumptions. We illustrate the theory by studying difference methods for approximating the *isolated* solutions of nonlinear two-point boundary value problems. Of course, all of these results are local in that they are confined to some sphere about the exact solution. The phenomenon of nonlinear instability does not occur here since, as has been shown by Stetter [10], this requires departing from the sphere. Indeed, parts of our theory are closely related to that of Stetter whose interest, however, was confined to the question of nonlinear instabilities; he therefore assumed the existence of solutions of the approximating problems. Similar existence results have been obtained by Pereyra [13] but his proofs are not constructive.

---

Received by the editors March 18, 1974.

AMS (MOS) subject classifications (1970). Primary 65J05, 65L10.

Key words and phrases. Nonlinear stability, Newton's method, two point boundary problems, finite difference schemes.

\*This work was supported under A.E.C. Contract AT(04-3)-767.

For general nonlinear problems, the isolated nature of the solution replaces or is equivalent to the well posedness required of Cauchy problems in the Lax theory [7]. In particular, we show in Section 2 that “stability” and “isolation” are essentially equivalent. In Section 3, the family of approximating problems are introduced, and the main theorem is proven. This shows that a stable family of approximate solutions exists, that they can be obtained by Newton’s method, and error estimates are given. Asymptotic error expansions are not discussed in Section 3 as they are easily obtained from our results by employing the techniques indicated in [3] or [9].

The present basic theory was initially developed and applied to study a specific difference method for nonlinear two-point boundary value problems [3]. However, the general simplicity and applicability of the theory to a variety of approximation problems prompted the more general treatment. In addition to the above cited use, the theory has already been applied by R. K. Weiss in [11] to study implicit Runge-Kutta and collocation schemes for nonlinear two-point boundary value problems. It can also be used for nonlinear Fredholm problems, for mildly nonlinear elliptic problems, and to justify the Box scheme applied to nonlinear parabolic problems, etc. Such applications will be presented elsewhere. However, we do show, in Section 4, how the theory can be applied to justify very general difference schemes for approximating isolated solutions of very general nonlinear two-point boundary value problems. A crucial step in this demonstration is supplied by a powerful stability result in [4] for linear problems. Indeed, we essentially show that any difference scheme which is stable and consistent for the initial-value problem is so for isolated solutions of the boundary value problem. Our theory is also used in [4] to study the most general difference methods for nonlinear boundary value problems in ordinary differential equations.

**2. Stable and Isolated Solutions.** For a mapping  $F: \mathcal{B}_1 \rightarrow \mathcal{B}_2$ , where the  $\mathcal{B}_\nu$  are appropriate Banach spaces, we consider the problem

$$(2.1) \quad F(x) = 0.$$

With the sphere  $S_\rho(u) \equiv \{x: x \in \mathcal{B}_1, \|x - u\| \leq \rho\}$ , we introduce the

(2.2) *Definitions.* (a) *The mapping  $F(\cdot)$  is stable on  $S_\rho(u)$  iff there exists a constant  $K_\rho > 0$  such that\*\*\**

$$\|v - w\| \leq K_\rho \|F(v) - F(w)\|$$

for all  $v, w \in S_\rho(u)$ .

(b) *A solution  $x = u$  of (2.1) is stable iff  $F(\cdot)$  is stable on  $S_\rho(u)$  for some  $\rho > 0$ .*

Trivially, we note that a stable solution is also unique in  $S_\rho(u)$ . If  $F(\cdot)$  is linear and stable, then our definition implies (for any  $\rho > 0$ ) Lipschitz continuous dependence

\*\*To simplify matters, we assume that the domain of  $F$  is  $\mathcal{B}_1$ ; the restriction to a proper subset offers no difficulty.

\*\*\*We do not distinguish notationally between norms on  $\mathcal{B}_1$  and  $\mathcal{B}_2$ . Rather, we adopt the convention that  $\|x\| \equiv \|x\|_{\mathcal{B}_\nu}$  if  $x \in \mathcal{B}_\nu$ .

of the solution of  $F(x) = g$  on the inhomogeneous data  $g$  (which is the usual definition for linear problems).

The Fréchet derivative of  $F$  at  $x$  will be denoted by  $L(x)$ . This is a linear operator  $L(x): \mathcal{B}_1 \rightarrow \mathcal{B}_2$  which is such that

$$(2.3) \quad r(x, y) \equiv \frac{\|F(x + y) - [F(x) + L(x)y]\|}{\|y\|} \rightarrow 0 \quad \text{as } \|y\| \rightarrow 0.$$

In terms of the Fréchet derivative, we introduce the

(2.4) *Definition.* A solution  $u$  of (2.1) is isolated iff  $L(u)$  exists and is nonsingular; that is: if  $L(u)y = 0$ , then  $y = 0$ .

Now we show that stability implies isolation as in

(2.5) **THEOREM.** Let  $u$  be a stable solution of (2.1). Then, if  $L(u)$  exists,  $u$  is an isolated solution.

*Proof.* Suppose  $L(u)y = 0$  and  $\|y\| \neq 0$ . Then, for all positive scalars  $\alpha < \rho/\|y\|$ , it follows that  $v(\alpha) \equiv u + \alpha y \in S_\rho(u)$ . By the stability of  $F(\cdot)$  on  $S_\rho(u)$ , it follows that

$$\begin{aligned} \|u - v(\alpha)\| &\leq K_\rho \|F(u) - F(v)\| \\ &\leq K_\rho \{ \|L(u)\alpha y\| + \|\alpha y\| r(u, \alpha y) \}. \end{aligned}$$

Thus,  $\alpha\|y\| \leq K_\rho r(u, \alpha y)\alpha\|y\|$  and, if  $\alpha > 0$  is chosen so small that  $K_\rho r(u, \alpha y) < 1$ , we must have  $\|y\| = 0$ .

Stability is such a strong condition that it implies  $L(x)$  nonsingular wherever it exists in the interior of  $S_\rho(u)$ . The proof is identical to that of Theorem (2.5) since the latter does not employ  $F(u) = 0$ . Thus, it is not surprising that a form of converse to Theorem (2.5) requires the existence of  $L(x)$  in some sphere about  $u$ . Indeed, we require even more in stating

(2.6) **THEOREM.** Let  $L(u)$  be nonsingular for some  $u \in \mathcal{B}_1$ . Let  $L(x)$  exist and be Lipschitz continuous on  $S_{\rho_0}(u)$  for some  $\rho_0 > 0$ ; that is: for some constant  $K_L > 0$ ,  $\|L(x) - L(y)\| \leq K_L \|x - y\|$  for all  $x, y \in S_{\rho_0}(u)$ . Then,  $F(\cdot)$  is stable on  $S_\rho(u)$  for any  $\rho < (K_L \|L^{-1}(u)\|)^{-1}$ , and the stability constant is

$$K_\rho = \|L^{-1}(u)\| (1 - \rho K_L \|L^{-1}(u)\|)^{-1}.$$

*Proof.* For any  $x, y \in S_\rho(u)$  with  $\rho \leq \rho_0$ , we have  $F(x) - F(y) = \hat{L}(x, y)(x - y)$ , where

$$\hat{L}(x, y) \equiv \int_0^1 L(tx + [1 - t]y) dt.$$

Write  $\hat{L}(x, y) = L(u) + [\hat{L}(x, y) - L(u)]$  and note that

$$\begin{aligned} \|\hat{L}(x, y) - L(u)\| &\leq \int_0^1 \|L(tx + [1 - t]y) - L(tu + [1 - t]u)\| \\ &\leq K_L \int_0^1 \|t(x - u) + [1 - t](y - u)\| dt \leq \rho K_L. \end{aligned}$$

Thus, if  $\rho$  is so small that  $\rho K_L \|L^{-1}(u)\| < 1$ , the Banach lemma implies that  $\hat{L}(x, y)$  is nonsingular and

$$\|\hat{L}^{-1}(x, y)\| \leq \|L^{-1}(u)\| / (1 - \rho K_L \|L^{-1}(u)\|).$$

Now stability follows from  $(x - y) = \hat{L}^{-1}(x, y)[F(x) - F(y)]$  with the indicated  $K_\rho$ .  $\square$

The hypothesis in Theorem (2.6) can be weakened to require only local Hölder continuity in the restricted form, with some  $0 < \alpha \leq 1$ :

$$\|L(x) - L(u)\| \leq K_L \|x - u\|^\alpha \quad \text{for all } x \in S_{\rho_0}(u).$$

In this case,  $\rho$  must be restricted by  $\rho^\alpha < (K_L \|L^{-1}(u)\|)^{-1}$ , and the stability constant must be suitably altered.

Finally, we note that if  $u$  is an isolated solution of (2.1), then  $L(u)$  is nonsingular. So if, in addition,  $L(x)$  exists and is Lipschitz continuous in some  $S_{\rho_0}(u)$ , then  $u$  is a stable solution, by Theorem (2.6). This is the essential converse of Theorem (2.5).

**3. Approximation Problems.** On a family of Banach spaces,  $\{\mathcal{B}_1^h, \mathcal{B}_2^h\}$ , we consider the family of approximating problems, for  $0 < h \leq h_0$ :

$$(3.1) \quad F_h(x_h) = 0,$$

where  $F_h: \mathcal{B}_1^h \rightarrow \mathcal{B}_2^h$ . To relate problems (2.1) and (3.1), we require that there exist a family of linear mappings  $\{P_1^h, P_2^h\}$  where<sup>†</sup>

$$(3.2) \quad (a) P_\nu^h: \mathcal{B}_\nu \rightarrow \mathcal{B}_\nu^h, \quad (b) \lim_{h \rightarrow 0} \|P_\nu^h x\| = \|x\| \quad \forall x \in \mathcal{B}_\nu.$$

We find it convenient to use the notation

$$P_\nu^h x \equiv [x]_h, \quad \nu = 1, 2,$$

where, of course,  $[x]_h \in \mathcal{B}_\nu^h$  if  $x \in \mathcal{B}_\nu$ . The Fréchet derivative of  $F_h$  at  $x_h$  is denoted by  $L_h(x_h)$  and  $S_\rho(x_h)$  is the sphere in  $\mathcal{B}_1^h$  of radius  $\rho$  about  $x_h$ . We introduce several concepts.

(3.2) *Definition.* The family  $\{F_h(\cdot)\}$  is stable for  $u \in \mathcal{B}_1$  iff for some  $h_0 > 0$ ,  $\rho > 0$  and some constant  $M_\rho$ , independent of  $h$ .

$$\|x_h - y_h\| \leq M_\rho \|F_h(x_h) - F_h(y_h)\|$$

for all  $x_h, y_h \in S_\rho([u]_h)$  and all  $h \in (0, h_0]$ .

(3.3) *Definition.* The family  $\{F_h(\cdot)\}$  is consistent of order  $p$  with  $F(\cdot)$  on  $S_\rho(u)$  if and only if

$$\|F_h([x]_h) - [F(x)]_h\| \equiv \|\tau_h(x)\| \leq M(x)h^p,$$

for all  $x \in S_\rho(u)$  and some bounded functional  $M(x) \geq 0$  independent of  $h$ .

---

<sup>†</sup> Again norms are those of the relevant spaces;  $\|x_h\| \equiv \|x_h\|_{\mathcal{B}_\nu^h}$  if  $x_h \in \mathcal{B}_\nu^h$ .

The significance of these definitions is well known and, indeed, best summarized in the

(3.4) THEOREM. *Let  $F(u) = 0$  and  $F_h(v_h) = 0$  for some  $v_h \in S_\rho([u]_h)$ ,  $\rho > 0$  and all  $h \in (0, h_0]$ . Let  $\{F_h(\cdot)\}$  be stable for  $u$  and consistent of order  $p$  on  $S_0(u)$ . Then*

$$\|[u]_h - v_h\| \leq M_\rho M(u) h^p.$$

*Proof.* By (3.2) with  $x_h = [u]_h$  and  $y_h = v_h$ ,

$$\|[u]_h - v_h\| \leq M_\rho \|F_h([u]_h) - F_h(v_h)\|.$$

Using  $F_h(v_h) = 0$  and  $[F(u)]_h = 0$  in the above and recalling (3.3) with  $\rho = 0$  yields the result.  $\square$

We are thus faced with the basic problems: (i) to be insured that the approximating problems actually have solutions in some sphere about  $[u]_h$ ; (ii) to be able to verify stability; and (iii) to determine the order of consistency. For many explicit difference schemes, it is trivial to verify (i), but, for implicit schemes and projection methods, this is frequently quite difficult. Again, for most difference approximation schemes, the order of consistency is determined by simple Taylor expansions. However, for projection or expansion methods, this is by no means a trivial task. The stability verification for nonlinear problems of great generality is also not a standard procedure. It is usually reduced to a study of the linearized problems. We present such a result as

(3.5) LEMMA. *Let the family of mappings  $\{F_h(\cdot)\}$  have Fréchet derivatives (i.e., linearizations)  $\{L_h(x_h)\}$  on some family of spheres  $S_{\rho_0}(z_h)$  and satisfy for all  $h \in (0, h_0]$ :*

(a)  *$\{L_h(z_h)\}$  have uniformly bounded inverses at the centers of the spheres; that is, for some constant  $K_0 > 0$ ,  $\|L_h^{-1}(z_h)\| \leq K_0$ .*

(b)  *$\{L_h(x_h)\}$  are uniformly Lipschitz continuous on  $S_{\rho_0}(z_h)$ ; that is, for some constant  $K_L > 0$ ,*

$$\|L_h(x_h) - L_h(y_h)\| \leq K_L \|x_h - y_h\|$$

for all  $x_h, y_h \in S_{\rho_0}(z_h)$ .

*If  $z_h = [u]_h$  for some  $u \in \mathcal{B}_1$ , then the family  $\{F_h(\cdot)\}$  is stable for  $u$ .*

*Proof.* The proof is essentially identical to that of Theorem 2.6.  $\square$

To insure the existence of a family of solutions  $\{v_h\}$  approximating a solution  $u$  of (2.1), we need only adjoin consistency to the above. More precisely, we have

(3.6) THEOREM. *Let  $x = u$  be a solution of  $F(x) = 0$ . Let the family  $\{F_h(\cdot)\}$  be consistent of order  $p$  with  $F(\cdot)$  on  $S_0(u)$ . Let the hypothesis (a) and (b) of Lemma (3.5) hold with  $z_h = [u]_h$ . Then, for  $\rho_0$  and  $h_0$  sufficiently small and for each  $h \in (0, h_0]$ , the problem  $F_h(x_h) = 0$  has a unique solution  $x_h = v_h \in S_{\rho_0}([u]_h)$ . These solutions satisfy*

$$\|[u]_h - v_h\| \leq M_{\rho_0} M(u) h^p.$$

*Proof.* We define the family of mappings  $\{G_h(x_h)\}$  by

$$G_h(x_h) \equiv x_h - L_h^{-1}([u]_h)F_h(x_h)$$

and shall show that they are uniformly contracting on  $S_{\rho_0}([u]_h)$ , provided  $\rho_0$  and  $h_0$  are sufficiently small. Since the sphere is convex, we have, for any  $x_h, y_h \in S_{\rho_0}([u]_h)$ :

$$\begin{aligned} G_h(x_h) - G_h(y_h) &= L_h^{-1}([u]_h) \{L_h([u]_h)(x_h - y_h) - (F(x_h) - F(y_h))\} \\ &= L_h^{-1}([u]_h) \{L_h([u]_h) - \hat{L}_h(x_h, y_h)\}(x_h - y_h). \end{aligned}$$

Here, as in Theorem (2.6), we used the generalized mean value theorem and

$$\hat{L}_h(x_h, y_h) \equiv \int_0^1 L_h(tx_h + [1 - t]y_h) dt.$$

From (3.5b), it follows that

$$\|L_h([u]_h) - \hat{L}_h(x_h, y_h)\| \leq K_L \rho_0,$$

and thus, by (3.5a),

$$\|G_h(x_h) - G_h(y_h)\| \leq \alpha \|x_h - y_h\|, \quad \alpha = K_0 K_L \rho_0.$$

At the center of the sphere,  $x_h = [u]_h$ , we have, by consistency (3.3) and since  $F(u) = 0$ :

$$\|[u]_h - G_h([u]_h)\| \leq K_0 \|F_h([u]_h) - [F(u)]_h\| \leq K_0 M(u) h^p.$$

Now if  $\alpha < 1$  and  $K_0 M(u) h^p \leq (1 - \alpha) \rho_0$ , the Contraction Mapping Theorem applied to  $x_h = G(x_h)$  implies the existence of a unique solution in  $S_{\rho_0}([u]_h)$ .

The error estimate follows from Theorem (3.4), which is now applicable.  $\square$

Obviously, the iteration scheme implied in the proof of Theorem (3.6) cannot be used to compute the approximate solutions since  $[u]_h$  is not known. However, Newton's method is frequently applicable for this purpose as we show in

(3.7) **THEOREM.** *Let the hypothesis of Theorem (3.6) hold. Then, for any  $h \in (0, h_0]$ , if  $\rho_0, h_0$  and  $\rho_1 \leq \rho_0$  are sufficiently small, the Newton iterates  $\{v_h^{(v)}\}$  defined by:*

(a)  $v_h^{(0)} \in S_{\rho_1}([u]_h),$

(b)  $L_h(v_h^{(v)}) [v_h^{(v+1)} - v_h^{(v)}] = -F_h(v_h^{(v)}), v = 0, 1, 2, \dots,$

*converge quadratically to the unique solution of  $F_h(x_h) = 0$  in  $S_{\rho_0}([u]_h)$ .*

*Proof.* By writing

$$\begin{aligned} L_h(x_h) &= L_h([u]_h) + [L_h(x_h) - L_h([u]_h)] \\ &= L_h([u]_h) \{I + L_h^{-1}([u]_h) [L_h(x_h) - L_h([u]_h)]\} \end{aligned}$$

and using (3.5a, b) with  $K_0 K_L \rho_0 < 1$ , the Banach lemma yields that  $L_h(x_h)$  is non-singular and, in fact,

(a)  $\|L_h^{-1}(x_h)\| \leq K_0/(1 - K_0K_{L\rho_0}) = K_{\rho_0}$  for all  $x_h \in S_{\rho_0}([u]_h)$ .

For any  $v_h^{(0)} \in S_{\rho_1}([u]_h)$ , we have from (3.7b) with  $\nu = 0$ , as in the proof of Theorem (3.6):

$$v_h^{(1)} - v_h^{(0)} = -L_h^{-1}(v_h^{(0)})F_h([u]_h) + L_h^{-1}(v_h^{(0)})\hat{L}_h(v_h^{(0)}, [u]_h)([u]_h - v_h^{(0)}).$$

However, the identity

$$L_h^{-1}(v_h^{(0)})\hat{L}_h(v_h^{(0)}, [u]_h) = I + L_h^{-1}(v_h^{(0)})(\hat{L}_h(v_h^{(0)}, [u]_h) - L_h(v_h^{(0)}))$$

implies by (a) above and (3.5b) that for some constant  $C > 0$ :

$$\|L_h^{-1}(v_h^{(0)})\hat{L}_h(v_h^{(0)}, [u]_h)\| \leq C.$$

Now, consistency in (3.3), recalling  $[F(u)]_h = 0$ , and the above yield

(b)  $\|v_h^{(1)} - v_h^{(0)}\| \leq K_{\rho_0}M(u)h^p + C_{\rho_1}.$

Together with (3.5b), conditions (a) and (b) above are sufficient for the quadratic convergence of Newton’s method, provided  $h_0$  and  $\rho_1$  are sufficiently small (see for instance [2] or [6]).□

The basic difficulty in applying the above theory is to verify (3.5a); that is, the establishment of the stability of the family of linear operators  $\{L_h([u]_h)\}$ . Frequently, this can be done by showing that there are some close approximations to  $L_h([u]_h)$ , say  $L_h([u]_h)$ , which have uniformly bounded inverses. That is, if

$$\|L_h([u]_h) - L_h([u]_h)\| = O(h)$$

and  $\|L_h^{-1}([u]_h)\| \leq K_0$  for all  $h \leq h_0$ , then, by the Banach lemma, (3.5a) holds with some modified constant,  $K_0'$ . This technique is illustrated for two-point boundary value problems in Section 4; but, of course, the problem is just modified to show the stability of  $L_h([u]_h)$ . If the problem  $F(x) = 0$  has an isolated solution  $x = u$ , then we know that  $L(u)$  has a bounded inverse. In Section 4, this fact and some additional assumptions on related (initial-value) problems are used to show that the consistent approximation  $L_h([u]_h)$  is stable. In another important class of problems,  $L(u)$  is selfadjoint and, say positive definite. Then if the same is true, uniformly in  $h$ , of  $L_h([u]_h)$ , the stability may easily follow. This technique is very close to that used by R. B. Simpson in [8]. Finally, the technique devised by H.-O. Kreiss is perhaps most powerful; see, for example, [5]. Here, since  $L(u)$  does not have the eigenvalue zero, Kreiss shows that  $\{L_h([u]_h)\}$  must have eigenvalues bounded away from zero. This assumes the consistency of  $L_h([u]_h)$  with  $L(u)$  and employs a contradiction obtained by using an appropriate map from  $B_h$  to  $B$  of the normalized solutions of  $L_h([u]_h)\phi_h = 0$  which can be shown to converge to  $\phi$ , a nontrivial solution of  $L(u)\phi = 0$ .

**4. Nonlinear Two-Point Boundary Value Problems.** We assume the nonlinear two-point boundary value problem

$$(4.1) \quad \begin{aligned} (a) \quad & Ny(t) \equiv y'(t) - f(t, y(t)) = 0, \quad a < t < b, \\ (b) \quad & g(y(a), y(b)) = 0, \end{aligned}$$

has an isolated solution,  $y(t)$ . That is, the linear problem

$$(4.2) \quad \begin{aligned} (a) \quad & L[y]\phi(t) \equiv \phi'(t) - A(t)\phi(t) = 0, \quad a < t < b, \\ (b) \quad & B[y]\phi \equiv B_a\phi(a) + B_b\phi(b) = 0, \end{aligned}$$

where

$$(4.2) \quad (c) \quad A(t) \equiv \frac{\partial f(t, y(t))}{\partial y}; \quad B_z \equiv \frac{\partial g(y(a), y(b))}{\partial y(z)}, \quad z = a, b;$$

has only the trivial solution,  $\phi(t) \equiv 0$ . We shall apply the previous theory to justify some fairly general difference schemes for approximating this solution of (4.1).

A family of nets is considered of which the general one is

$$(4.3) \quad \begin{aligned} (a) \quad & t_0 = a: \quad t_j = t_{j-1} + h_j, \quad 1 \leq j \leq J; \quad t_J = b, \\ (b) \quad & h \equiv \max_j h_j \leq \lambda \min_j h_j, \end{aligned}$$

where  $\lambda$  is a fixed constant and on which  $h \rightarrow 0$  in some manner. For each such net, a difference scheme, determined by the coefficients  $\{\alpha_{jk}(h), \beta_{jk}(h)\}$ , is defined by:

$$(4.4) \quad \begin{aligned} (a) \quad & N_h v_j \equiv \sum_{k=0}^J \{\alpha_{jk} v_k - \beta_{jk} f(t_k, v_k)\} = 0, \quad 1 \leq j \leq J; \\ (b) \quad & g(v_0, v_J) = 0. \end{aligned}$$

(4.5) Our main assumption on the numerical method is that: *The family of schemes (4.4a), with  $v_0 = u_0$ , is consistent of order  $p$  and stable for all sufficiently smooth initial-value problems of the form:*

$$(4.6) \quad u' = F(t, u), \quad a < t < b; \quad u(a) = u_0.$$

For example, to satisfy (4.5), the scheme (4.4a) could be a one-step scheme such as Euler's method, centered Euler or the trapezoidal rule. It could equally well be some standard multistep scheme on a uniform net including a prescribed starting scheme on a refined net (to maintain uniform accuracy). Our present formulation does not include Runge-Kutta or implicit Runge-Kutta schemes, but this is essentially a notational simplification as we shall show later.

To apply the theory of Section 3, we introduce the family of Banach spaces  $\mathcal{B}_1^h \equiv \mathcal{B}_2^h \equiv E^{n(J+1)}$  and, say,  $\mathcal{B}_1 \equiv C_{p+1}[a, b]$ ,  $\mathcal{B}_2 \equiv C_p[a, b]$ . Then we require that  $f(t, z) \in C_p[[a, b] \times E^n]$  and  $F(x) = 0$  represents the boundary value problem (4.1a, b) with  $x \equiv y(t)$ . For  $x_h \equiv \{v_j\}_0^J$ , the family of problems  $F_h(x_h) = 0$  represents the family of difference equations (4.4a, b). The mapping  $[ ]_h$  on  $\mathcal{B}_v$  into  $\mathcal{B}_v^h$  is defined by  $[y]_h = \{y(t_j)\}_0^J$ .



We now easily get the basic

(4.7) THEOREM. *Let (4.1) have an isolated solution  $y(t) \in C_{p+1}[a, b]$ . Let  $f(t, z) \in C_p([a, b] \times E^n)$  and  $g(z, w) \in C_{1, \text{Lip}}(E^n \times E^n)$ . If the scheme (4.3)–(4.4) satisfies (4.5), then for some  $\rho_0 > 0, h_0 > 0$  sufficiently small and all  $h \leq h_0$ :*

(i) *The difference equations (4.4) have a unique solution  $\{v_j\}_0^J$  in  $\|v_j - y(t_j)\| \leq \rho_0$ .*

(ii)  $\|v_j - y(t_j)\| \leq Mh^p, 0 \leq j \leq J$ .

(iii) *The difference solution can be computed by Newton's method which converges quadratically for any initial iterate  $\{v_j^0\}_0^J$  in  $\|v_j^0 - y(t_j)\| \leq \rho_1$ , provided  $\rho_1 \leq \rho_0$  is sufficiently small.*

*Proof.* For (i) and (ii), we need only verify the hypothesis of Theorem (3.6). Clearly, by (4.5), the scheme (4.4) is consistent of order  $p$  with (4.1). It remains to verify (a) and (b) of Lemma (3.5) with  $z_h = [y]_h$ . The linearized difference operators obtained from the nonlinear difference operators in (4.4) are, say, applied to  $\phi_h \equiv \{\phi_j\}_0^J$ :

$$(4.8) \quad (a) \quad L_h(v_h)\phi_j \equiv \sum_{k=0}^J \left[ \alpha_{jk}I - \beta_{jk} \frac{\partial f}{\partial y}(t_k, v_k) \right] \phi_k, \quad 1 \leq j \leq J;$$

$$(b) \quad B_h(v_h)\phi_h \equiv \frac{\partial g}{\partial v_0}(v_0, v_J)\phi_0 + \frac{\partial g}{\partial v_J}(v_0, v_J)\phi_J.$$

Now let us apply the difference scheme of (4.4) to the linear problem (4.2). This yields

$$(4.9) \quad (a) \quad \sum_{k=0}^J [\alpha_{jk}I - \beta_{jk}A(t_k)] \phi_k = 0, \quad 1 \leq j \leq J;$$

$$(b) \quad B_a\phi_0 + B_b\phi_J = 0.$$

Recalling (4.2), it follows from (4.8) that (4.9) is just

$$(4.10) \quad (a) \quad L_h([y]_h)\phi_j = 0, \quad 1 \leq j \leq J; \quad (b) \quad B_h([y]_h)\phi_h = 0.$$

However, Corollary (3.13) of [4] states essentially that: *if (4.2) has a unique solution, then the (linear) difference scheme (4.10) is stable and consistent for (4.2) if and only if (4.10a) with  $\phi_0 = c$  is stable and consistent with (4.2a) and  $\phi(a) = c$ .* But the latter part follows from the assumption (4.5). Thus, since  $y(t)$  is isolated, and hence (4.2) has a unique solution, the scheme in (4.10) is stable. Hence, (a) of Lemma (3.5) is established. (We point out that the operator corresponding to  $L_h([y]_h)$  of (3.5) is just that represented by the coefficient matrix of the difference equations in (4.10a, b). Stability of this scheme is shown in [4] to be equivalent to the uniform boundedness of the inverses of these coefficient matrices for all  $h \leq h_0$ .)

From the Lipschitz continuity of the first derivatives of  $f(t, y)$  and  $g(y(a), y(b))$ , it follows, using (4.6b), that (b) of Lemma (3.5) with  $z_h = [y]_h$  holds for the scheme

in (4.10). Thus, (i) and (ii) of our theorem are established by an application of Theorem (3.6).

To establish (iii), we need only apply Theorem (3.7).□

We now show how to extend the above result to more general difference schemes, say including all Runge-Kutta and implicit Runge-Kutta types. Thus, in place of (4.4), we consider:

$$(4.11) \quad \begin{aligned} (a) \quad N_h \mathbf{v}_j &\equiv \sum_{k=0}^J \left\{ \alpha_{jk} \mathbf{v}_k - \beta_{jk} \mathbf{f} \left( t_{jk}, \sum_{l=0}^J \gamma_{jkl} \mathbf{v}_l \right) \right\} = 0, \quad 1 \leq j \leq J; \\ (b) \quad &\mathbf{g}(\mathbf{v}_0, \mathbf{v}_J) = 0. \end{aligned}$$

We again impose (4.5), now with (4.4a) replaced by (4.11a). The consistency requirement in (4.5) implies that for all sufficiently smooth functions  $\mathbf{u}(t)$ :

$$(4.12) \quad \left\| \mathbf{u}(t_{jk}) - \sum_{l=0}^J \gamma_{jkl} \mathbf{u}(t_l) \right\| \leq C(u)h, \quad 1 \leq j \leq J, 0 \leq k \leq J.$$

Note that this is also a restriction on the choice of the points  $t_{jk}$ .

The linearized difference equations obtained from (4.11) are:

$$(4.13) \quad \begin{aligned} (a) \quad L_h(v_h)\phi_j &\equiv \sum_{k=0}^J \left\{ \alpha_{jk} I - \sum_{s=0}^J \beta_{js} \frac{\partial \mathbf{f}}{\partial \mathbf{y}} \left( t_{js}, \sum_{l=0}^J \gamma_{jls} \mathbf{v}_l \right) \gamma_{jks} \right\} \phi_k = 0, \\ (b) \quad &B_h(v_h)\phi_h = 0. \end{aligned}$$

However, when applied to the linear problem (4.2), our more general difference scheme (4.11a) yields

$$(4.14) \quad \begin{aligned} (a) \quad L_h([y]_h)\phi_j &\equiv \sum_{k=0}^J \left\{ \alpha_{jk} I - \sum_{s=0}^J \beta_{js} A(t_{js}) \gamma_{jks} \right\} \phi_k = 0, \\ (b) \quad &B_h([y]_h)\phi_h = 0. \end{aligned}$$

Since  $\mathbf{y}(t)$  is isolated and (4.11a) is stable and consistent for initial-value problems, it follows, by the above cited Corollary (3.13) of [4], that the linear difference scheme in (4.14) is stable. Now use (4.12) to observe that  $\|L_h([y]_h) - L_h([y]_h)\| = O(h)$ , provided  $\partial \mathbf{f} / \partial \mathbf{y}$  and  $\mathbf{y}(t)$  are sufficiently smooth. From the Banach lemma, it easily follows that the linear difference scheme in (4.13) is also stable. Thus, with no difficulty, we see that Theorem (4.7) goes over for difference schemes of the form (4.11).

The analog of Theorem (4.7) for implicit Runge-Kutta schemes has previously been demonstrated in [11] by R. K. Weiss. In place of  $L_h([y]_h)$  given by (4.14a), Weiss employs the centered Euler (Box scheme) whose stability was demonstrated in [1]. It is a bit more involved to show the ‘‘consistency’’ of the Box scheme with  $L_h([y]_h)$ .

The analog of Theorem (4.7) for “gap-schemes” has been illustrated in [12] by A. B. White. These are high-order accurate two-point difference schemes in which the local truncation error has leading term  $\mathcal{O}(h^{2m})$ , and the higher-order terms proceed in higher powers of  $h^2$ . The lowest-order part of these gap-difference schemes is just that obtained from the trapezoidal rule, and the corrections are bounded perturbations of order  $h^2$ . Thus, the stability proof could easily be obtained as above by showing consistency with the (stable) trapezoidal scheme.

Applied Mathematics  
California Institute of Technology  
Pasadena, California 91109

1. H. B. KELLER, “Accurate difference methods for linear ordinary differential systems subject to linear constraints,” *SIAM J. Numer. Anal.*, v. 6, 1969, pp. 8–30. MR 40 #6776.
2. H. B. KELLER, “Newton’s method under mild differentiability conditions,” *J. Comput. System Sci.*, v. 4, 1970, pp. 15–28. MR 40 #3710.
3. H. B. KELLER, “Accurate difference methods for nonlinear two point boundary value problems,” *SIAM J. Numer. Anal.*, v. 11, 1974, pp. 305–320.
4. H. B. KELLER & A. B. WHITE, “Difference methods for boundary value problems in ordinary differential equations,” *SIAM J. Numer. Anal.*, v. 12, 1975. (To appear.)
5. H.-O. KREISS, “Difference approximations for boundary and eigenvalue problems for ordinary differential equations,” *Math. Comp.*, v. 26, 1972, pp. 605–624.
6. J. M. ORTEGA & W. C. RHEINBOLDT, *Iterative Solution of Nonlinear Equations in Several Variables*, Academic Press, New York, 1970. MR 42 #8686.
7. R. D. RICHTMYER, *Difference Methods for Initial-Value Problems*, Interscience Tracts in Pure and Appl. Math., no. 4, Interscience, New York, 1957. MR 20 #438.
8. R. B. SIMPSON, “Existence and error estimates for solutions of a discrete analog of nonlinear eigenvalue problems,” *Math. Comp.*, v. 26, 1972, pp. 359–375. MR 47 #4466.
9. H. J. STETTER, “Asymptotic expansions for the error of discretization algorithms for non-linear functional equations,” *Numer. Math.*, v. 7, 1965, pp. 18–31. MR 30 #5505.
10. H. J. STETTER, “Stability of nonlinear discretization algorithms,” *Numerical Solution of Partial Differential Equations* (Proc. Sympos. Univ. Maryland, 1965), Academic Press, New York, 1966, pp. 111–123. MR 34 #5322.
11. R. K. WEISS, “The application of implicit Runge-Kutta and collocation methods to boundary value problems,” *Math. Comp.*, v. 28, 1974, pp. 449–464.
12. A. B. WHITE, *Numerical Solution of Two Point Boundary Value Problems*, Ph. D. Thesis, Calif. Inst. of Technology, Pasadena, 1974.
13. V. PEREYRA, “Iterated deferred corrections for nonlinear operator equations,” *Numer. Math.*, v. 10, 1967, pp. 316–323.