MATHEMATICS OF COMPUTATION, VOLUME 29, NUMBER 131
JULY 1975, PAGES 746–760

Bifurcation in Difference Approximations to Two-Point Boundary Value Problems

By Richard Weiss*

Abstract. Numerical methods for bifurcation problems of the form

$$(*) \qquad Ly = \lambda f(y), \quad By = 0,$$

where $f(0) = 0$ and $f'(0) \neq 0$, are considered. Here y is a scalar function, λ is a real scalar, L is a linear differential operator and $By = 0$ represents some linear homogeneous two-point boundary conditions. Under certain assumptions, it is shown that if (*) is replaced by an appropriate difference scheme, then there exists a unique branch of nontrivial solutions of the discrete problem in a neighborhood of a branch of nontrivial solutions of (*) bifurcating from the trivial solution and that the discrete branch converges to the continuous one. Error estimates are derived and an illustrative numerical example is included.

1. Introduction. One of the key assumptions in the analysis of numerical methods for nonlinear problems is that the desired solution be isolated, i.e., the linearized problem be nonsingular. This implies that the nonlinear problem is locally (at the desired solution) well posed. For instance, the assumption of isolation is fundamental to the theory of difference approximations for nonlinear boundary value problems in ordinary differential equations given in Keller [4].

In this paper, we investigate the application of difference methods in a situation where the condition of isolation is not satisfied, namely, that of bifurcation from the trivial solution in certain nonlinear two-point boundary value problems.

In particular, we consider problems of the form

$$(1.1a) \qquad Ly \equiv \sum_{j=0}^{m} a_j(t) d^j y/dt^j = \lambda f(y), \quad 0 \leqslant t \leqslant 1,$$

$$(1.1b) \qquad By = 0,$$

where $y(t)$, $a_j(t)$ and $f(y)$ are real-valued scalar functions, λ is a real scalar and (1.1b) are m real linear homogeneous boundary conditions which contain derivatives of y up to order $m-1$ at $t = 0$ and $t = 1$. We shall assume that $a_m(t) \equiv 1$, $t \in [0, 1]$, $a_j(t)$, $0 \leqslant j \leqslant m-1$, are continuous on $[0, 1]$ and that there exists a complex value $\tilde{\lambda}$ such that the homogeneous problem $[L - \tilde{\lambda}]y = 0$, $By = 0$ has only the trivial solution. Furthermore, we require that f satisfy a certain smoothness condition, $f(0) = 0$ and $f'(0) \neq 0$.

Clearly, $y = 0$ is a solution of (1.1) for all λ. Let λ_0 be a value for which the

Received April 1, 1974.
AMS (MOS) subject classifications (1970). Primary 65L10.
Key words and phrases. Ordinary differential equations, boundary value problems, bifurcation, difference methods.

*Present address: Mathematics Research Center, University of Wisconsin-Madison, 610 Walnut Street, Madison, Wisconsin 53706.

linearized problem

$$[L - \lambda_0 f'(0)]\varphi = 0, \qquad B\varphi = 0$$

has a nontrivial solution. If, as will be assumed throughout the paper, the nullspace associated with λ_0 is one dimensional and the index of λ_0 is one, then a branch of nontrivial solutions of (1.1) bifurcates from the trivial solution at $\lambda = \lambda_0$.

For computational purposes, (1.1) is replaced by a family of difference equations

$$L_h y_h = \lambda F_h(y_h), \qquad B_h y_h = 0, \qquad h > 0. \tag{1.2}$$

The aim of the paper is to investigate the behavior of y_h for λ in a neighborhood of λ_0. Under natural conditions on the discretization (1.2) we shall show that there is a branch of nontrivial solutions of (1.2) bifurcating at a value λ_{0h} "close" to λ_0 and that, as $h \to 0$, $\lambda_{0h} \to \lambda_0$ and the branch of (1.2) "converges" to that of (1.1).

The organization of the paper is as follows. In Section 2 we consider the continuous problem (1.1) in more detail, while the existence and uniqueness of solutions of (1.2) is discussed in Section 3. Error estimates are derived in Section 4. In Section 5, we shall indicate how the results can be extended to equations in which the parameter λ appears nonlinearly, i.e.,

$$Ly = f(\lambda, y), \qquad By = 0.$$

Finally, numerical results illustrating the theory are given in Section 6.

Recently, Atkinson [1] examined bifurcation from the trivial solution in collectively compact approximations to nonlinear compact operators. The connections between his theory and the results derived here will be discussed at the end of Section 4.

2. Bifurcation in the Differential Equation. In this section we will give a proof of the existence of a branch of nontrivial solutions of (1.1) bifurcating from the trivial solution. The reason for including this proof is that it will aid in the understanding of the continuous problem as well as the discrete problem (1.2) and that it will allow a less detailed treatment of the discrete case. The proof will be based on the constructive theory developed in Keller and Langford [5].

We shall first collect some results on linear boundary value problems of the form

$$[L - \lambda a]u = g, \qquad Bu = 0, \tag{2.1}$$

where L, B are as in Section 1, $a \neq 0$, real, λ complex and $u, g \in C_c$, where C_c is the Banach space of elements $x = u + iv$, $u, v \in C[0, 1] \equiv C$, over the complex field with the norm $\|x\|_c = \|u\| + \|v\|$. ($\|\cdot\|$ is the maximum norm in C.) Let $\bar{C}_c$ be the subspace of C_c whose elements satisfy the boundary conditions, C_c^m the subspace of m times continuously differentiable functions on [0, 1], and $\bar{C}_c^m = \bar{C}_c \cap C_c^m$. Then we can write (2.1) in operator form

$$(L - \lambda aI)u = g, \qquad u \in \bar{C}_c^m. \tag{2.2}$$

In the sequel, we shall use the notation $N(A)$ and $R(A)$ for the nullspace and the range of an operator A.

LEMMA 2.1. *Let λ_0 be an eigenvalue of* (2.2) *with index one. Then*

(i) $C_c = N(L - \lambda_0 aI) \oplus R(L - \lambda_0 aI)$.

(ii) *The projection from C_c to $N(L - \lambda_0 aI)$ corresponding to* (i) *is given by*

$$P_c = -\frac{1}{2\pi i} a \int_{\Gamma_0} (L - \lambda aI)^{-1} \, d\lambda,$$

where $\Gamma_0 = \{\lambda \mid |\lambda - \lambda_0| = \delta_0 > 0\}$ with δ_0 so small that there is no other eigenvalue λ with $|\lambda - \lambda_0| \leqslant \delta_0$.

(iii) *The mapping*

$$(L - \lambda_0 aI): \; R(L - \lambda_0 aI) \cap \bar{C}_c^m \longrightarrow R(L - \lambda_0 aI)$$

has a bounded inverse which is given by the restriction to $R(L - \lambda_0 aI)$ of the operator

$$G_c = \frac{1}{2\pi i} \int_{\Gamma_0} \frac{1}{\lambda - \lambda_0} (L - \lambda aI)^{-1} \, d\lambda.$$

Proof. See Dunford and Schwartz [3, Chapter VII].

So far we employed the complex space C_c. However, since (1.1) is a real problem, we shall have to work in C, $\bar{C} = \bar{C}_c \cap C$ and $\bar{C}^m = C_c^m \cap C$, respectively. If λ_0 is real, this provides no difficulty since we can then assume that $N(L - \lambda_0 aI)$ is spanned by an element of C. As an operator on $\bar{C}^m$, $(L - \lambda_0 aI)$ has the nullspace $N = N(L - \lambda_0 aI) \cap C$, the range $R = R(L - \lambda_0 aI) \cap C$ and $C = N \oplus R$. The corresponding projection from C to N, P, is the restriction of P_c to C and the mapping $(L - \lambda_0 aI)$: $R \cap \bar{C}^m \longrightarrow R$ has a bounded inverse G given by the restriction of G_c to R.

In the sequel, we shall not distinguish explicitly between complex and real spaces, but assume that the reader uses the appropriate interpretation.

We now return to the nonlinear problem

$$Ly = \lambda f(y), \qquad By = 0 \tag{2.3}$$

and make the following assumptions.

A2.1. $f \in C^{2+p}(U)$, $p \geqslant 0$ where $U = \{s \mid s \text{ real}, |s| \leqslant M = \text{const} > 0\}$, $f(0) = 0$, $f'(0) \neq 0$ and $a_j \in C^p[0, 1]$, $j = 0, \ldots, m - 1$.

A2.2. For λ_0 real and $\varphi \in C$ with $\|\varphi\| = 1$ we have

$$[L - \lambda_0 a]\varphi = 0, \qquad B\varphi = 0,$$

where $a = f'(0)$, λ_0 has index one and $\dim N = 1$ (i.e. $N = \text{span}\{\varphi\}$).

We then rewrite (2.3) as

$$[L - \lambda_0 a]y = \lambda f(y) - \lambda_0 ay, \qquad By = 0. \tag{2.4}$$

This problem has a solution if and only if $\lambda f(y) - \lambda_0 ay \in R$. By Lemma 2.1, this is the case if and only if

$$\lambda P f(y) = \lambda_0 P ay. \tag{2.5}$$

Hence, if $Pf(y) \neq 0$, then $\lambda = \Lambda(y)$ is uniquely determined and instead of (2.4) we may consider

$$[L - \lambda_0 a]y = \Lambda(y)f(y) - \lambda_0 ay, \qquad By = 0. \tag{2.6}$$

We now proceed by considering for some positive constants ϵ_0 and ρ and all real ϵ with $0 < |\epsilon| \leqslant \epsilon_0$ elements of the form

$$w = \epsilon(\varphi + \epsilon v), \qquad v \in V_\rho = \{u \mid u \in \overline{C} \cap R, \ \|u\| \leqslant \rho\}. \tag{2.7}$$

Then we obtain the following theorem which is the main result of this section.

THEOREM 2.1. *Let L, f and λ_0 satisfy the conditions* A2.1, A2.2. *Then there are positive constants ϵ_0 and ρ such that for each ϵ with $0 < |\epsilon| \leqslant \epsilon_0$ there exists a unique pair $[\lambda(\epsilon), y(\epsilon)]$ where*

$$\lambda(\epsilon) = \lambda_0 + \epsilon\tilde{\lambda}(\epsilon), \qquad |\tilde{\lambda}(\epsilon)| \leqslant K_0, \qquad K_0 = \text{const} > 0,$$

and $y(\epsilon)$ has the form (2.7) *and is a nontrivial solution of* (2.3) *with $\lambda = \lambda(\epsilon)$.*

Before we can prove this theorem, we need certain estimates for $\Lambda(w)$ and $S(w) = \Lambda(w)f(w) - \lambda_0 aw$, which are collected in the following lemma.

LEMMA 2.2. *Let $w = \epsilon(\varphi + \epsilon v)$, $w' = \epsilon(\varphi + \epsilon v')$ where $v, v' \in V_\rho$. Then, for ϵ_0 sufficiently small, $Pf(w) \neq 0$ and*

(i) $|\Lambda(w) - \lambda_0| \leqslant K_1 |\epsilon|$,

(ii) $|\Lambda(w) - \Lambda(w')| \leqslant |\epsilon|^2 K_2 \|v - v'\|$,

(iii) $\|S(w)\| \leqslant |\epsilon|^2 K_3$,

(iv) $\|S(w) - S(w')\| \leqslant |\epsilon|^3 K_4 \|v - v'\|$,

where K_1, K_2, K_3, K_4 are positive constant.

Proof. (i) By Taylor's theorem,

$$f(x) = ax + r(x), \qquad \|x\| \leqslant M$$

with

$$\|r(u) - r(u')\| \leqslant K_5(\|u\| + \|u'\|)(\|u - u'\|), \qquad \|u\|, \|u'\| \leqslant M, \qquad K_5 = \text{const}. \tag{2.8}$$

According to (2.5),

$$\lambda P(a\epsilon(\varphi + \epsilon v) + r(w)) = \lambda_0 Pa\epsilon(\varphi + \epsilon v). \tag{2.9}$$

Let

$$\gamma(\epsilon, v)\varphi = Pr(w). \tag{2.10}$$

Then from (2.8),

$$|\gamma(\epsilon, v)| \leqslant |\epsilon|^2 K_5(1 + \epsilon_0\rho)^2 \|P\|$$

and if $|\epsilon_0| K_5(1 + \epsilon_0\rho)^2 \|P\|/|a| \leqslant \frac{1}{2}$, then (2.9) yields $Pf(w) \neq 0$ and (i) with $K_1 = 4K_5(1 + \epsilon_0\rho)^2 \|P\|/|a|$.

(ii) From (2.10),

$$(\gamma(\epsilon, w) - \gamma(\epsilon, w'))\varphi = P[r(w) - r(w')],$$

and (2.8) yields

$$|\gamma(\epsilon, w) - \gamma(\epsilon, w')| \leqslant K_6 \epsilon^3 \|v - v'\|, \qquad K_6 = \text{const},$$

which, by (2.9), implies (ii).

(iii) Using (2.7), (2.8) and (i) to estimate $S(w) = [\Lambda(w) - \lambda_0]aw + \Lambda(w)r(w)$ yields the result.

(iv) The estimate follows from

$$S(w) - S(w') = [\Lambda(w) - \Lambda(w')]aw + (\Lambda(w') - \lambda_0)a[w - w'] + [\Lambda(w) - \Lambda(w')]r(w) + \Lambda(w')[r(w) - r(w')]$$

and (2.7), (2.8), (i) and (ii).

Proof of Theorem 2.1. If $y = \epsilon(\varphi + \epsilon v)$, $v \in \bar{C}^m \cap V_\rho$, then

$$[L - \lambda_0 a]\epsilon^2 v = S(y), \qquad Bv = 0, \tag{2.11}$$

or, equivalently,

$$v = GS(y)/\epsilon^2 \equiv H(v), \qquad v \in V_\rho. \tag{2.12}$$

We shall now show that H is contracting on V_ρ for $0 < |\epsilon| \leqslant \epsilon_0$ and appropriate ϵ_0, ρ. From Lemma 2.2(iii), $\|H(v)\| \leqslant \|G\|K_3$. Looking at the explicit form of K_3 in terms of ϵ_0 and ρ (as was illustrated for K_1), we see that by making ρ sufficiently large and ϵ_0 sufficiently small we can obtain $\|G\|K_3 < \rho$. From Lemma 2.2(iv),

$$\|H(v) - H(v')\| \leqslant \|G\|\epsilon K_4\|v - v'\|,$$

and hence the theorem holds if $|\epsilon_0|\|G\|K_4 < 1$.

Certain additional information about the branch constructed above is available. In particular, the following two statements follow from Crandall and Rabinowitz [2, Theorems 1.7, 1.18]: (i) If we define $\lambda(\epsilon) = 0$ and $y(\epsilon) = 0$ for $\epsilon = 0$, then $\lambda(\epsilon)$, $y(\epsilon)$ are k times continuously differentiable with respect to ϵ for $|\epsilon| \leqslant \epsilon_0$ if $f \in C^{k+1}[U]$. (ii) There is a $\hat{\lambda} > 0$ and a sphere $B = \{x | x \in C, \|x\| \leqslant \hat{\delta} > 0\}$ such that for $\lambda_0 - \hat{\lambda} \leqslant \lambda \leqslant \lambda_0 + \hat{\lambda}$ the set of all solutions of (2.3) contained in B consists of the trivial solution and the branch constructed in Theorem 2.1.

For the analysis of the following sections, a knowledge of the smoothness of the solution of (2.1) and of $y(\epsilon)$ as functions of t is of importance. Clearly, assumption A2.1 implies that if λ is not an eigenvalue of (2.1) and $g \in C^p[0, 1]$, then $u \in C^{p+m}[0, 1]$. In addition, it follows easily from the analysis given that $\varphi \in C^{m+p}[0, 1]$, $v(\epsilon) \in C^{m+p}[0, 1]$ and $\|d^l v(\epsilon)/dt^l\| \leqslant E_l$, $0 < |\epsilon| \leqslant \epsilon_0$, $l = 0, \ldots, m + p$, $E_l =$ const.

3. Bifurcation in the Difference Equation. We shall first briefly consider the algebraic eigenvalue problem

$$[A - \lambda B]x = 0, \tag{3.1}$$

where A, B are $q \times q$ matrices and $x \in X^q$, the usual q-dimensional vector space. We assume that A is nonsingular. (This is no restriction, since, if $\tilde{\lambda}$ is such that $A - \tilde{\lambda}B$ is nonsingular, then we may rewrite (3.1) as $[\tilde{A} - \mu B]x = 0$, where $\tilde{A} = A - \tilde{\lambda}B$, $\mu = \lambda - \tilde{\lambda}$ and $\tilde{A}$ is nonsingular.)

Instead of (3.1), we can then consider the ordinary eigenvalue problem

$$[A^{-1}B - \mu I]x = 0, \qquad \mu = 1/\lambda. \tag{3.2}$$

The following lemma contains some results on (3.1), (3.2) which will be required further on.

LEMMA 3.1. *Let μ_0 be an eigenvalue of* (3.2) *with index one. Then*

(i) $X^q = N(A^{-1}B - \mu_0 I) \oplus R(A^{-1}B - \mu_0 I)$.

(ii) *The corresponding projection from X^q to $N(A^{-1}B - \mu_0 I)$ is given by*

$$P = -\frac{1}{2\pi i}\int_{\Gamma_0}[A - \lambda B]^{-1}B\, d\lambda$$

with Γ_0 defined analogous to Lemma 2.1(ii).

(iii) $x \in R(A - \lambda_0 B)$, $\lambda_0 = 1/\mu_0$, *if and only if* $BQx = 0$, *where*

$$Q = -\frac{1}{2\pi i}\int_{\Gamma_0}[A - \lambda B]^{-1}\, d\lambda.$$

Also, $\dim R(BQ) = \dim N(A^{-1}B - \mu_0 I)$.

(iv) *The mapping $A - \lambda_0 B$: $R(A^{-1}B - \mu_0 I) \longrightarrow R(A - \lambda_0 B)$ is one-to-one and onto; its inverse is given by the restriction of*

$$G = \frac{1}{2\pi i}\int_{\Gamma_0}\frac{1}{\lambda - \lambda_0}[A - \lambda B]^{-1}\, d\lambda$$

to $R(A - \lambda_0 B)$.

Proof. (i) This follows immediately from the fact that μ_0 has index one.

(ii) It is well known (see, for instance, Dunford and Schwartz [3, Chapter VII]) that

$$P = \frac{1}{2\pi i}\int_{\Gamma_{\mu_0}}[\mu I - A^{-1}B]^{-1}\, d\mu \tag{3.3}$$

for an appropriate curve Γ_{μ_0}. The result follows from the identity

$$\mu^2(\mu I - A^{-1}B)^{-1} = \mu I + (A - \lambda B)^{-1}B$$

which can be derived proceeding as in Dunford and Schwartz [3, pp. 600–601] and a change of variables $\mu = 1/\lambda$ in (3.3).

(iii) Clearly $x \in R(A - \lambda_0 B)$ if and only if $A^{-1}x \in R(A^{-1}B - \mu_0 I)$, i.e., $PA^{-1}x = 0$. But

$$(A - \lambda B)^{-1}BA^{-1} = ((I - \lambda BA^{-1})A)^{-1}BA^{-1} = A^{-1}(I - \lambda BA^{-1})^{-1}BA^{-1}$$

$$= A^{-1}BA^{-1}(I - \lambda BA^{-1})^{-1} = A^{-1}B(A - \lambda B)^{-1},$$

which yields the first result. The second statement is obvious.

(iv) The mapping

$$A^{-1}B - \mu_0 I\colon\ R(A^{-1}B - \mu_0 I) \longrightarrow R(A^{-1}B - \mu_0 I)$$

is one-to-one and onto and its inverse is the restriction of

$$H = \frac{1}{2\pi i}\int_{\Gamma_{\mu_0}} \frac{1}{\mu - \mu_0}[A^{-1}B - \mu I]^{-1}\, d\mu$$

$$= \frac{1}{2\pi i}\int_{\Gamma_{\mu_0}} \frac{1}{\mu - \mu_0}[B - \mu A]^{-1}A\, d\mu$$

to $R(A^{-1}B - \mu_0 I)$. Since

$$[A - \lambda_0 B]x = y \quad \text{if and only if } [I - \lambda_0 A^{-1}B]x = A^{-1}y,$$

or equivalently $[\mu_0 - A^{-1}B]x = \mu_0 A^{-1}y$, it follows that $x = Gy$ where

$$G = -\frac{\mu_0}{2\pi i}\int_{\Gamma_{\mu_0}} \frac{1}{\mu - \mu_0}[B - \mu A]^{-1}\, d\mu.$$

The change of variables $\mu = 1/\lambda$ now yields the result.

To derive a difference method for (2.3), we introduce a grid $\pi_I = \{t_0, t_1, \ldots, t_I | t_j = jh, h = 1/I\}$ on $[0, 1]$.

We shall denote net functions $(z_0, \ldots, z_I)^T$ by $z_h \in X^{I+1}$. For later purposes, we define two linear operators mapping C to X^{I+1} and vice versa. Firstly, let Δ_h be the usual discretization operator, i.e. for $x \in C$, $\Delta_h x = x_h = (x(t_0), \ldots, x(t_I))^T \in X^{I+1}$. Secondly, assign to each z_h a function $z = z(t, h) \in C^m[0, 1]$ such that

$$z(t_j, h) = z_j, \qquad j = 0, \ldots, I,$$

$$\|d^\nu z/dt^\nu\| \leqslant d_\nu(\|z_h\|_h + \|D_+^\nu z_h\|_h), \qquad \nu = 0, \ldots, m,$$

where d_ν are constants, D_+ is the forward divided-difference operator and $\|\cdot\|_h$ is the maximum norm on X^{I+1}. Such a function can be constructed by Hermite interpolation as in Kreiss [6, Lemma 2.1]. We shall denote it by $z = \operatorname{Int} z_h$.

The differential equation is replaced by the scheme

$$\sum_{\nu=-r}^{s} c_\nu(t_j, h)y_{j+\nu} = h^m \lambda \tilde{f}(h; y_{j-r}, \ldots, y_j, \ldots, y_{j+s}), \qquad j = r, \ldots, I - s, \tag{3.4}$$

where y_j denotes an approximation to $y(t_j)$, r and s are natural numbers with $r + s \geqslant m$, c_ν are continuous functions of t and h and $\tilde{f}$ satisfies the following condition.

A3.1. $\tilde{f}(h; 0, \ldots, 0) = 0$ for $h \leqslant h_0 = \text{const} > 0$ and $\tilde{f}$ is twice continuously differentiable with respect to $(s_1, \ldots, s_{r+s+1})$ on $\tilde{U} = \{(s_1, \ldots, s_{r+s+1}), |s_l| \leqslant M, l = 1, \ldots, r + s + 1\}$ for $h \leqslant h_0$. All derivatives are uniformly bounded in h.

In addition to (3.4), $r + s$ linear homogeneous boundary conditions are prescribed. We write (3.4) after division by h^m plus the boundary conditions as

$$L_h y_h = \lambda F_h(y_h), \qquad B_h y_h = 0, \tag{3.5}$$

with obvious definitions of L_h, F_h and B_h. Together with (3.5), we have to consider the problem obtained by linearizing (3.5) at the trivial solution,

$$[L_h - \lambda a E_h]x_h = 0, \qquad B_h x_h = 0, \tag{3.6}$$

where $E_h = F_h'(0)/a$, and the related inhomogeneous scheme

$$[L_h - \lambda a E_h]u_h = E_h \Delta_h g, \qquad B_h z_h = 0, \qquad g \in C. \tag{3.7}$$

Denoting by $\overline{X}^{I+1}$ the subspace of X^{I+1} whose elements satisfy the boundary conditions, we can write (3.7) in operator form

$$[L_h - \lambda a E_h]u_h = E_h \Delta_h g, \qquad z_h \in \overline{X}^{I+1}.$$

The following hypotheses will be required.

A3.2. Let Ω denote a compact set in the complex λ plane which does not contain any eigenvalue of $a^{-1}L$. Then for $\lambda \in \Omega$ and $h \leqslant h_0$ the problem

$$[L_h - \lambda a E_h]x_h = b, \qquad B_h x_h = 0$$

has a unique solution for all $b \in X^{I-r-s+1}$ and

$$\|x_h\|_h \leqslant K_\Omega^1 \|b\|_1, \qquad K_\Omega^1 = \text{const},$$

where $\|\cdot\|_1$ denotes the maximum norm on $X^{I-r+s+1}$.

This condition implies that

$$\sup_{\lambda \in \Omega, h \leqslant h_0} \|(L_h - \lambda a E_h)^{-1}\|_h \leqslant K_\Omega^1.$$

A3.3. For every $x \in C$, $\lambda \in \Omega$,

$$\lim_{h \to 0} \|\mathrm{Int}(L_h - \lambda a E_h)^{-1} E_h \Delta_h x - (L - \lambda a I)^{-1} x\| = 0.$$

A3.4. Let $h_\mu \longrightarrow 0$, $g^{(\mu)} \in C$ with $\sup_\mu \|g^{(\mu)}\| < \infty$. Then the sequence

$$w^{(\mu)} = \mathrm{Int}(L_h - \lambda a E_h)^{-1} E_h \Delta_h g^{(\mu)}, \qquad \lambda \in \Omega, h = h_\mu,$$

has a convergent subsequence.

A3.2, A3.3 and A3.4 are the stability, convergence and compactness assumptions as used in Kreiss [6]. A3.2 and A3.3 imply the convergence of the "eigenvalues" of (3.6) to the eigenvalues of $a^{-1}L$. A3.4 guarantees that the invariant subspaces also converge. Kreiss [6] has shown that if $r + s = m$ and (3.7) is consistent with (2.1) then A3.2–A3.4 are satisfied. For $r + s > m$, he provides conditions for A3.2 and A3.4 in terms of the roots of a polynomial associated with (3.7). A3.3 then follows from consistency and A3.2.

The following two conditions will also be required.

A3.5. Let $g \in C^p[0, 1]$, $p \geqslant 1$, and

$$x = (L - \lambda a I)^{-1} g, \quad x_h = (L_h - \lambda a E_h)^{-1} E_h \Delta_h g, \qquad \lambda \in \Omega, \ h \leqslant h_0.$$

Then

$$\|\Delta_h x - x_h\|_h \leqslant K_\Omega^2 \max_{0 \leqslant l \leqslant p} \|d^l g/dt^l\| h^p, \qquad K_\Omega^2 = \text{const}.$$

A3.6. For $z_h \in X^{I+1}$, $u \in C$, define $r_h(z_h) = F_h(z_h) - a E_h z_h$, $r(u) = f(u) - au$ and let

$$x = (L - \lambda a I)^{-1} r(u), \quad x_h = (L_h - \lambda a E_h)^{-1} r_h(\Delta_h u), \qquad \lambda \in \Omega, \ h \leqslant h_0.$$

If $u \in C^p[0, 1]$, $p \geqslant 1$, and $f \in C^p(U)$, then

$$\|\Delta_h x - x_h\|_h \leqslant K_\Omega^3 \max_{0 \leqslant l \leqslant p} \|d^l r(u)/dt^l\| h^p, \qquad K_\Omega^3 = \text{const.}$$

We now return to the problem of constructing a family of nontrivial solutions of (3.5). From the assumptions A2.1, A2.2, A3.1–A3.4 and Kreiss [6], there is a unique "eigenvalue" λ_{0h} of $(L_h - \lambda a E_h)$ in a neighborhood of λ_0 independent of h and by A3.5,

$$|\lambda_0 - \lambda_{0h}| \leqslant C_1 h^p, \qquad h \leqslant h_0, \quad C_1 = \text{const.} \tag{3.8}$$

The invariant subspace N_h associated with λ_{0h} has dimension one and is given by $N_h = P_h X^{I+1}$, where

$$P_h = -\frac{1}{2\pi i} a \int_{\Gamma_0} [L_h - \lambda a E_h]^{-1} E_h \, d\lambda.$$

The space N_h is spanned by $\varphi_h = P_h \Delta_h \varphi$ and

$$\|\varphi_h - \Delta_h \varphi\|_h \leqslant C_2 h^p, \qquad h \leqslant h_0, \quad C_2 = \text{const.} \tag{3.9}$$

Proceeding as in Section 3, we rewrite (3.5) in the form

$$[L_h - \lambda_{0h} a E_h] y_h = \lambda F_h(y_h) - \lambda_{0h} a E_h y_h, \qquad B_h y_h = 0. \tag{3.10}$$

From Lemma 3.1(iii) this problem has a solution if and only if

$$\lambda E_h Q_h F_h(y_h) = \lambda_{0h} E_h a Q_h E_h y_h \tag{3.11}$$

where

$$Q_h = -\frac{1}{2\pi i} \int_{\Gamma_0} [L_h - \lambda a E_h]^{-1} \, d\lambda.$$

We now look for solutions of the form

$$w_h = \epsilon(\varphi_h + \epsilon v_h) \tag{3.12}$$

where $v_h \in \widetilde{V}_{\tilde{\rho}} = \{u_h | u_h \in R(I - \lambda_{0h} a L_h^{-1} E_h) \subset \overline{X}^{I+1}, \ \|u_h\|_h \leqslant \tilde{\rho} > 0\}$. Clearly, from Lemma 3.1(i) every element of $\overline{X}^{I+1}$ can be represented in the form (3.12).

We then obtain the following theorem which is the discrete analogue of Theorem 2.1.

THEOREM 3.1. *Let the conditions* A2.1, A2.2 *and* A3.1–A3.4 *be satisfied. Then there exist positive constants* $\tilde{\epsilon}_0, \tilde{\rho}$ *and* h_0 *such that for all* ϵ *with* $0 < |\epsilon| \leqslant \tilde{\epsilon}_0$ *and all* $h \leqslant h_0$ *there exists a unique pair* $[\lambda_h(\epsilon), y_h(\epsilon)]$ *where*

$$\lambda_h(\epsilon) = \lambda_{0h} + \epsilon \tilde{\lambda}_h(\epsilon), \qquad |\tilde{\lambda}_h(\epsilon)| \leqslant C_3, \quad C_3 = \text{const},$$

and $y_h(\epsilon)$ *is of the form* (3.12) *and is a nontrivial solution of* (3.5) *with* $\lambda = \lambda_h(\epsilon)$.

Proof. The proof proceeds as for Theorem 2.1 and is therefore only sketched. For $\tilde{\epsilon}_0$ sufficiently small, it follows as in Lemma 2.2(i) that $E_h Q_h F_h(w_h) \neq 0$; and hence, $\lambda = \lambda_h = \Lambda_h(w_h)$ is uniquely determined. Defining

$$S_h(w_h) = \Lambda_h(w_h) F_h(w_h) - \lambda_{0h} a E_h w_h$$

and using (3.12), we write (3.10) as

$$[L_h - \lambda_{0h} a E_h] \epsilon^2 v_h = S_h(y_h), \qquad B_h v_h = 0.$$

Then, by Lemma 3.1(iv),

$$v_h = G_h S_h(y_h)/\epsilon^2 \equiv H_h(v_h) \tag{3.13}$$

where

$$G_h = \frac{1}{2\pi i}\int_{\Gamma_0} \frac{1}{\lambda - \lambda_{0h}} [L_h - \lambda E_h]^{-1}\, d\lambda.$$

It is straightforward to establish a discrete equivalent of Lemma 2.2 and to prove that $H_h(v_h)$ is contracting on $\tilde{V}_{\tilde{\rho}}$ for appropriate $\tilde{\rho}$ and $\tilde{\epsilon}_0$.

From Crandall and Rabinowitz [2], we can obtain results about the smoothness of λ_h and v_h as functions of ϵ analogous to those quoted at the end of Section 2. Also, a uniqueness result corresponding to the one stated there holds.

4. Error Estimates. The main result of this section is contained in the following theorem.

THEOREM 4.1. *Let the conditions* A3.1, A3.2, A4.1–A4.6 *be satisfied. Then there are positive constants* h_1 *and* ϵ_1 *such that for* $|\epsilon| \leqslant \epsilon_1$ *and* $h \leqslant h_1$,

$$\|v_h(\epsilon) - \Delta_h v(\epsilon)\|_h \leqslant D_1 h^p, \qquad |\lambda_h(\epsilon) - \lambda(\epsilon)| \leqslant D_2 h^p, \qquad D_1, D_2 = \text{const.}$$

Proof. Firstly, recall the results on the smoothness of φ and v as functions of t stated at the end of Section 2.

We shall now derive an estimate for $(\lambda(\epsilon) - \lambda_0) - (\lambda_h(\epsilon) - \lambda_0 h)$. From (2.5),

$$(\lambda(\epsilon) - \lambda_0) P a y = -\lambda(\epsilon) P r(y) \tag{4.1}$$

and, from (3.11),

$$(\lambda_h(\epsilon) - \lambda_{0h}) E_h P_h a y_h = -\lambda_h(\epsilon) E_h a Q_h r_h(y_h). \tag{4.2}$$

Premultiplying (4.1) by $P_h \Delta_h$ and (4.2) by aQ_h yields

$$(\lambda(\epsilon) - \lambda_0) a\epsilon\varphi_h = -\lambda(\epsilon) P_h \Delta_h P r(y),$$

$$(\lambda_h(\epsilon) - \lambda_{0h}) a\epsilon\varphi_h = -\lambda_h(\epsilon) P_h a Q_h r_h(y_h),$$

and hence,

$$\begin{aligned}[(\lambda(\epsilon) - \lambda_h(\epsilon)) - (\lambda_0 - \lambda_{0h})] a\epsilon\varphi_h \\ = -\{(\lambda(\epsilon) - \lambda_h(\epsilon)) P_h \Delta_h P r(y) + \lambda_h(\epsilon) P_h [\Delta_h P r(y) - a Q_h r_h(y_h)]\}.\end{aligned} \tag{4.3}$$

Clearly, $P_h \Delta_h P r(y) = \gamma_h(\epsilon, v_h)\varphi_h$, where

$$|\gamma_h(\epsilon, v_h)| \leqslant D_3 |\epsilon|^2, \qquad D_3 = \text{const.} \tag{4.4}$$

Also,

$$\Delta_h P r(y) - a Q_h r_h(y_h) = \delta_h^1 + \delta_h^2 = \delta_h$$

with

$$\delta_h^1 = \Delta_h P r(y) - a Q_h r_h(\Delta_h y), \qquad \delta_h^2 = a Q_h r_h(\Delta_h y) - a Q_h r_h(y_h).$$

By A2.1 and A3.6,

$$\|\delta_h^1\|_h \leqslant D_4 \epsilon^2 h^p, \qquad D_4 = \text{const} \tag{4.5}$$

and using A3.1,

$$\|\delta_h^2\|_h \leqslant D_5 \epsilon^2 h^p + D_6 \epsilon^3 \|v_h - \Delta_h v\|_h, \qquad D_5, D_6 = \text{const}. \tag{4.6}$$

From (4.3), (4.4), (4.5) and (4.6),

$$(\lambda(\epsilon) - \lambda_h(\epsilon))a(1 + \gamma_h(\epsilon, v_h)/\epsilon a) = (\lambda_0 - \lambda_{0h})a + \eta_h, \tag{4.7}$$

where

$$|\eta_h| \leqslant \epsilon D_7 (h^p + \epsilon \|v_h - \Delta_h v\|_h), \qquad D_7 = \text{const}.$$

We thus obtain the desired estimate,

$$\lambda_h(\epsilon) - \lambda_{0h} = \lambda(\epsilon) - \lambda_0 + O(\epsilon(h^p + \epsilon \|v_h - \Delta_h v\|_h)). \tag{4.8}$$

From (2.12) and (3.13),

$$v_h - \Delta_h v = (G_h S_h(y_h) - \Delta_h G S(y))/\epsilon^2, \tag{4.9}$$

where

$$S_h(y_h) = a(\lambda_h(\epsilon) - \lambda_{0h}) E_h y_h + \lambda_h(\epsilon) r_h(y_h)$$

and

$$S(y) = a(\lambda(\epsilon) - \lambda_0) y + \lambda(\epsilon) r(y).$$

From (3.9) and (4.8),

$$S_h(y_h) = a(\lambda(\epsilon) - \lambda_0) E_h \Delta_h y + \lambda(\epsilon) r_h(\Delta_h y) + O(\epsilon^2(h^p + \epsilon \|v_h - \Delta_h v\|_h)). \tag{4.10}$$

Also, from (3.8),

$$G_h S_h(y_h) = \frac{1}{2\pi i} \int_{\Gamma_0} \frac{1}{\lambda - \lambda_0} [L_h - \lambda a E_h]^{-1} S_h(y_h) d\lambda + O(h^p \epsilon^2).$$

Hence, using (4.9) and (4.10),

$$v_h - \Delta_h v = \frac{1}{2\pi i \epsilon^2} \int_{\Gamma_0} \frac{1}{\mu - \lambda_0} [(L_h - \mu a E_h)^{-1} (a(\lambda(\epsilon) - \lambda_0) E_h \Delta_h y + \lambda(\epsilon) r_h(\Delta_h y))$$

$$- \Delta_h (L - \mu a I)^{-1} (a(\lambda(\epsilon) - \lambda_0) y + \lambda(\epsilon) r(y))] \, d\mu$$

$$+ O(h^p + \epsilon \|v_h - \Delta_h v\|_h),$$

which, by A3.5 and A3.6, yields

$$\|v_h - \Delta_h v\|_h \leqslant D_8 (h^p + \epsilon \|v_h - \Delta_h v\|_h), \qquad D_8 = \text{const}.$$

Thus, if $|\epsilon| \leqslant D_8/2$, then $\|v_h - \Delta_h v\|_h \leqslant D_1 h^p$.

The second statement of the theorem follows from (4.7).

Combining the results of Section 3 and Theorem 4.1 we find that for $h \leqslant h_1$ and $0 < |\epsilon| \leqslant \epsilon_1$,

$$\|y_h(\epsilon) - \Delta_h y(\epsilon)\|_h/\epsilon \leqslant D_9 h^p, \qquad D_9 = C_2 + \epsilon_1 D_1, \tag{4.11}$$

$$|\lambda_h(\epsilon) - \lambda(\epsilon)| \leqslant D_2 h^p. \tag{4.12}$$

Thus the parameterization of the solutions of (2.3) and (3.5) by ϵ has led to a satisfactory convergence theory.

In computations, one usually determines the solution y_h of (3.5) for a given value of λ and not $y_h(\epsilon)$, $\lambda_h(\epsilon)$ for a given ϵ. But we can still apply our theory once we observe that there is a unique ϵ such that $y_h = y_h(\epsilon)$, $\lambda = \lambda_h(\epsilon)$.

Under appropriate conditions, it is possible to show that $\lambda_h(\epsilon)$ and $y_h(\epsilon)$ possess asymptotic expansion in powers of h (or h^2) for a fixed ϵ with coefficients which are continuous in ϵ. However, one cannot, in general, make such a statement for the case when λ is kept fixed.

Under the assumption that $f \in C^{p+2+k}[U]$ and $\tilde{f} \in C^{p+2+k}[\tilde{U}]$, $k \geqslant 1$, one can extend Theorem 4.1 to obtain $O(h^p)$ convergence of the first k derivatives with respect to ϵ of $\lambda_h(\epsilon)$ and $y_h(\epsilon)$ to the corresponding derivatives of $\lambda(\epsilon)$ and $y(\epsilon)$. This is accomplished by differentiating (2.6) and (3.10) with respect to ϵ and basically repeating the proof of Theorem 4.1.

Denoting $\mathcal{D} = (L - \tilde{\lambda}I)^{-1}$, we may write (2.3) as an integral equation bifurcation problem

$$\mu y = \mathcal{D}(y + ((1 + \mu\tilde{\lambda})/a)r(y)) \equiv \mathcal{D}\psi(\mu, y),$$

where $\mu = 1/(\lambda a - \tilde{\lambda})$. Similarly, with $\mathcal{D}_h = (L_h - \tilde{\lambda}E_h)^{-1}$, problem (3.5) can be written as

$$\mu y_h = \mathcal{D}_h(E_h y_h + ((1 + \mu\tilde{\lambda})/a)r_h(y_h)) \equiv \mathcal{D}_h\psi_h(\mu, y_h).$$

The assumptions of Section 3 (with condition A3.6 modified to include the case where $u(t)$ is only continuous) imply that the family of operators

$$\text{Int } \mathcal{D}_h\psi_h(\mu, \Delta_h x): \; C \longrightarrow C, \tag{4.13}$$

forms a collectively compact sequence which converges pointwise to $\mathcal{D}\psi(\mu, x)$. The theory of Atkinson [1] can be applied to (4.13) and yields results corresponding to Theorems 3.1 and 4.1. However, an accurate interpolation procedure Int (consistent with the order h^p) combined with a refinement of Atkinson's techniques, is needed to obtain exactly our results. (Note that the approximations on the grid are, of course, independent of the interpolation procedure.) The alternative approach chosen in this paper has the advantage of being completely within the framework of differential equations and difference methods.

5. Extensions. A generalization of (1.1) is given by the problem

$$Ly = f(\lambda, y), \qquad 0 \leqslant t \leqslant 1, \qquad By = 0, \tag{5.1}$$

where the nonlinear operator $f(\lambda, y)$ has the decomposition $f(\lambda, y) = g(\lambda)y + r(\lambda, y)$.

For $\lambda \in D$, an open interval, and all $s \in B = \{s | s \text{ real}, |s| \leqslant M > 0\}$ the following conditions are assumed to hold.

(i) $g(\lambda) \in C^2(D)$,

(ii) $r(\lambda, s) \in C^2(D \times B)$,

(iii) $r(\lambda, 0) = 0, r_s(\lambda, 0) = 0, r_\lambda(\lambda, 0) = 0, r_{s\lambda}(\lambda, 0) = 0, r_{\lambda\lambda}(\lambda, 0) = 0$.

For some $\lambda_0 \in D$ with $g_\lambda(\lambda_0) \neq 0$, let the problem

$$L\varphi = g(\lambda_0)\varphi, \quad B\varphi = 0, \quad \|\varphi\| = 1,$$

have a solution and let the related invariant subspace have dimension one. Then it is straightforward to extend Theorem 2.1 to (5.1). The only difference in the proof is that the equation corresponding to (2.5) is now nonlinear and must be treated in the same way as equation (2.21b) in Keller and Langford [5].

The difference scheme for (5.1) is assumed to have the form

$$L_h y_h = g(\lambda)E_h y_h + r_h(\lambda, y_h), \quad B_h y_h = 0, \tag{5.2}$$

where $r_h(\lambda, y_h)$ satisfies conditions analogous to (ii) and (iii) above. If the assumptions corresponding to A3.2–A3.6 are satisfied, then the arguments of Sections 3 and 4 immediately generalize to include (5.1), (5.2). The details may be safely left to the reader.

The results of this paper are extendable to systems of the form

$$Ly = f(\lambda, t, y, y^{(1)}, \ldots, y^{(m-1)}), \quad 0 \leqslant t \leqslant 1, \quad By = 0,$$

where L is a linear differential operator of order m, f and y are vector valued functions and $By = 0$ are linear homogeneous boundary conditions involving derivatives up to order $m - 1$. Details will be given elsewhere.

6. Numerical Results. We report some calculations with the problem

$$y'' = \lambda[y + y^2], \quad y(0) = y(1) = 0. \tag{6.1}$$

The difference method used is

$$\frac{y_{i+1} - 2y_i + y_{i-1}}{h^2} = \lambda[y_i + y_i^2], \quad i = 1, \ldots, I - 1; \; h = 1/I, \tag{6.2}$$

$$y_0 = 0, \quad y_I = 0.$$

This scheme satisfies the assumptions of Section 3 ($p = 2$).

The eigenvalues and eigenfunctions of the linearized problems corresponding to (6.1) and (6.2) are of course well known.

In Table 1, we give the value of y_h at $t = 0.5$ on the branch corresponding to the eigenvalue $\lambda_0 = -\pi^2$ for $\lambda = -\pi^2 + \Delta\lambda$ using various h and $\Delta\lambda$. Similarly, Table 2 contains the value of y_h at $t = 0.25$ on the branch corresponding to the eigenvalue $\lambda_0 = -4\pi^2$. In the last row of Tables 1 and 2, we give $\lambda_{0h} - \lambda_0$, i.e., the difference between the discrete and the continuous eigenvalue.

In the case of $\lambda_0 = -4\pi^2$, there are two nontrivial solutions y_h for the negative

$\Delta\lambda$ and all h. Nontrivial solutions for the positive $\Delta\lambda$ exist only for large enough h. There are two nontrivial solutions for $\Delta\lambda = 0.1$, $\Delta\lambda = 0.2$ and $h = 1/20$, but none for the smaller h. (Where no solutions exist, we have left a blank field.) The reason for this is that $\lambda(\epsilon) - \lambda_0 = O(\epsilon^2)$, i.e., the linear term in ϵ is missing. This difference between the two cases is best made apparent by plotting the entries of Tables 1 and 2 as is done in Figs. 1 and 2 respectively for certain values of h.

TABLE 1

$\Delta\lambda$ \ h	1/20	1/40	1/80	1/160
-0.2	-2.5763 E-2	-2.3985 E-2	-2.3540 E-2	-2.3429 E-2
-0.1	-1.4210 E-2	-1.2414 E-2	-1.1965 E-2	-1.1852 E-2
0.1	9.6146 E-3	1.1448 E-2	1.1907 E-2	1.2022 E-2
0.2	2.1902 E-2	2.3755 E-2	2.4219 E-2	2.4335 E-2
$\lambda_{0h}-\lambda_0$	2.0277 E-2	5.0723 E-3	1.2683 E-3	3.1708 E-4

TABLE 2

$\Delta\lambda$ \ h	1/20	1/40	1/80	1/160
-0.2	1.2013 E-1	8.9102 E-2	7.9178 E-2	7.6473 E-2
	-1.3064 E-1	-9.4773 E-2	-8.3627 E-2	-8.0616 E-2
-0.1	1.0881 E-1	7.2156 E-2	5.9091 E-2	5.5304 E-2
	-1.1733 E-1	-7.5815 E-2	-6.1524 E-2	-5.7430 E-2
0.1	8.0329 E-2			
	-8.4838 E-2			
0.2	6.0315 E-2			
	-6.2811 E-2			
$\lambda_{0h}-\lambda_0$	3.3236 E-1	8.1108 E-2	2.0289 E-2	5.0731 E-3

In Figs. 1 and 2, we do not know the curves $\epsilon = \text{const}$ exactly; but it is clear from the theory that they are nearly parallel to the $\Delta\lambda$ axis. Particularly Fig. 2 demonstrates the usefulness of expressing convergence via (4.11), (4.12).

The nonlinear system 6.2 was solved by Newton's method. Partial pivoting was used for the resulting linear equations. Accurate starting iterates can be obtained by the following consideration: For a value λ close to λ_{0h} (either $\lambda < \lambda_{0h}$ or $\lambda > \lambda_{0h}$ or both), there is a value $\epsilon_{\lambda h}$ such that $y_h(\lambda) \approx \epsilon_{\lambda h}\varphi_h$, i.e., the net function $\epsilon_{\lambda h}\varphi_h$ is a very good starting iterate for this particular λ. If Newton iteration is performed with different starting iterates $\epsilon\varphi_h$, $\epsilon = \pm\Delta\epsilon, \pm 2\Delta\epsilon, \ldots$, where $\Delta\epsilon$ is small, some ϵ will be close to $\epsilon_{\lambda h}$ and the iteration will converge. Once the solution is known for a certain λ, one can use continuation with respect to λ to obtain starting iterates for other λ values. Similarly, one can use continuation with respect to h for obtaining starting iterates for different h values. This strategy was used successfully in all our calculations. For all $\Delta\lambda$ and h, 5 to 7 iterations were needed to give the solution to 10 digits. It should be noted however that a theoretical analysis of iterative schemes for the solution of (3.5) for λ "close" to λ_{0h} has yet to be given.

To employ the above method for obtaining starting iterates in the general situation, it is first necessary to solve the eigenvalue problem for the linearized equation.

FIGURE 1

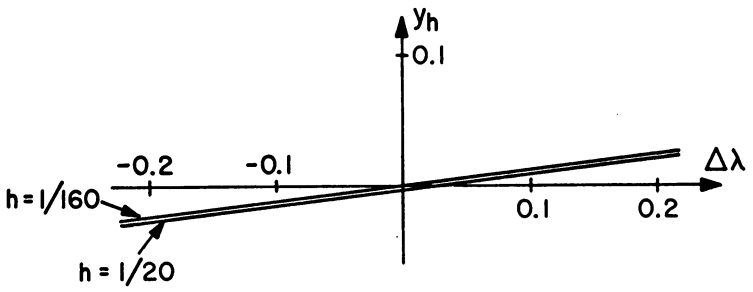
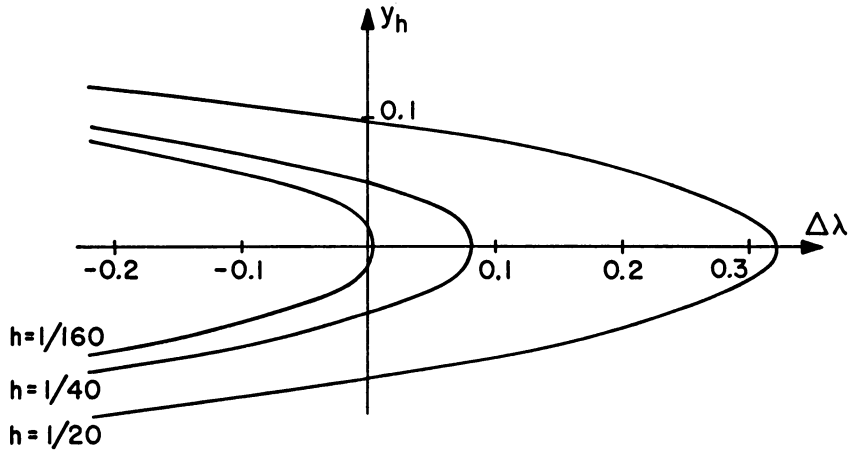


FIGURE 2



Remark. The computations were done in double-precision arithmetic on the IBM 360/158 computer at the California Institute of Technology.

Acknowledgment. I am very grateful to Professor H. B. Keller for many stimulating discussions during the course of this work.

Department of Applied Mathematics
California Institute of Technology
Pasadena, California 91109

1. K. E. ATKINSON, "The numerical solution of some bifurcation problems" (submitted).

2. M. G. CRANDALL & P. H. RABINOWITZ, "Bifurcation from simple eigenvalues," *J. Functional Analysis*, v. 8, 1971, pp. 321–340. MR 44 #5836.

3. N. DUNFORD & J. T. SCHWARTZ, *Linear Operators*. Vol. 1: *General Theory*, Pure and Appl. Math., vol. 7, Interscience, New York, 1958. MR 22 #8302.

4. H. B. KELLER, "Approximation methods for nonlinear problems" (submitted).

5. H. B. KELLER & W. F. LANGFORD, "Iterations, perturbations and multiplicities for nonlinear bifurcation problems," *Arch. Rational Mech. Anal.*, v. 48, 1972, pp. 83–108.

6. H.-O. KREISS, "Difference approximations for ordinary differential equations," *Math. Comp.*, v. 26, 1972, pp. 605–631.