ON DISCRETIZATIONS OF BIFURCATION PROBLEMS

Wolf-Jürgen Beyn

It is well known that bifurcation points are usually quite sensitive to perturbations. For example, introducing an imperfection in a bifurcation problem may turn two intersecting branches into two non-intersecting ones. In this paper it is shown that discretizing a nontrivial bifurcation problem may have the same effect. In particular, a sufficient criterion is given which relates the effect to the discretization error of the bifurcation point. The theory is developed in an abstract framework in order to show the general applicability of the results. In the applications the emphasis is on finite difference methods from which also the illustrative and numerical examples are drawn.

1. Introduction and an elementary example

We consider operator equations

(1) $T(\lambda,u) = 0 \quad , \quad \lambda \in \mathbb{R}\ , \ u \in U$

where $T \in C^2(\mathbb{R}\times U,Y)$ and U,Y are Banach spaces. Let us further assume that certain discrete problems

(2) $T_h(\lambda,u_h) = 0 \quad , \quad \lambda \in \mathbb{R}\ , \ u_h \in U_h$

are given where $T_h \in C^2(\mathbb{R}\times U_h,Y_h)$ and $h \in H$ is a discretization parameter tending to zero. The relation between the Banach spaces U_h,Y_h (usually finite dimensional) and the spaces U,Y will be expressed in terms of 'restrictions' or 'projections'
$p_h : U \to U_h,\ q_h : Y \to Y_h (h \in H)$.
The general problem may then be described as follows: given a bifurcation point (λ_o,u_o) of (1), what is the structure of the solution set of the discrete equations (2) in a neighborhood of (λ_o,p_hu_o)?

In the simple case of a linear eigenvalue problem the situation can be visualized as in figure 1.

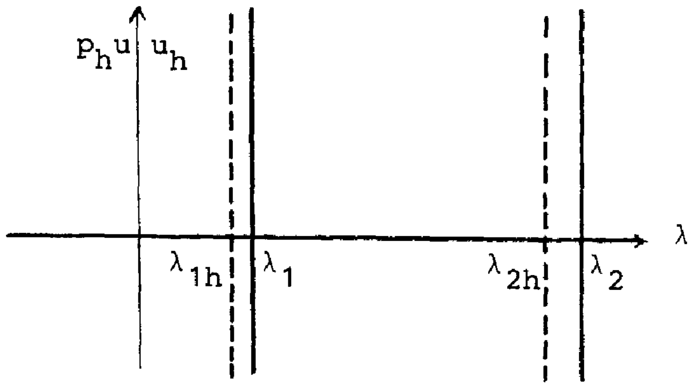


fig. 1 broken lines indicate discrete solutions and solid lines restrictions of continuous solutions, in this figure the λ-axis is also a broken line

The above picture represents the solution curves of the two-point boundary value problem

(3) $u'' + \lambda u = 0$ in $[0,1]$, $u(0) = u(1) = 0$

and the finite difference equations

(4) $D_h^2 u_h + \lambda u_h = 0$, $u_h(0) = u_h(1) = 0$

where u_h is a grid function on $J_h = \{0,h,\dots,1-h,1\}$, $h = (M+1)^{-1}$ $M \in \mathbb{N}$, and

(5) $D_h^2 u_h(x) = h^{-2}(u_h(x-h) - 2u_h(x) + u_h(x+h))$, $x = h,\dots,1-h$.

The solutions to (3) and (4) are known to be

$(\lambda,0)$, $\lambda \in \mathbb{R}$; $(\lambda_n, c\,\varphi_n)$, $c \in \mathbb{R}$, where $\lambda_n = n^2\pi^2$, $\varphi_n(x) = \sin(n\pi x)$ $(n \in \mathbb{N})$ and

$(\lambda,0)$, $\lambda \in \mathbb{R}$; $(\lambda_{nh}, c\,\varphi_{nh})$, $c \in \mathbb{R}$, where $\lambda_{nh} = 2h^{-2}(1-\cos(n\pi h))$, $\varphi_{nh} = [\varphi_n]_h$ $(n = 1,\dots,M)$.

Here $p_h = [\ ]_h$ denotes the restriction to the mesh J_h .

Finally, we have for each fixed $n \in \mathbb{N}$

(6) $|\lambda_n - \lambda_{nh}| = O(h^2)$.

The qualitative behaviour of fig. 1 remains valid for eigenvalue problems of a very general type, e.g. if T is linear in u and depends analytically on λ (see [12,20] for finite difference methods and [14,15,25,26] for abstract results and further references). Exceptional cases occur for multiple eigenvalues which can be splitted by the discretization into several distinct eigenvalues each of which converges of lower order.

It has also been shown for non-linear problems that a discretization method causes a shift of the solution diagram (at least locally) if bifurcation from the trivial solution at simple eigenvalues is considered, cf. [29] (finite difference methods), [1] (collectively compact approximations), [30] (Galerkin methods), [18] (discrete approximations).

A typical picture in this case is

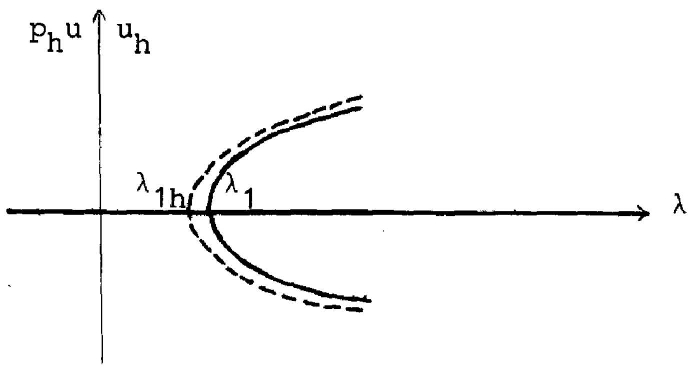


fig. 2

In all problems considered so far, the discrete equations inherited a smooth solution branch from the continuous problem - the trivial one. Although this may also happen for equations with a certain symmetry (see [4,32]), a nontrivial bifurcation point will in general be separated by a discretization method. This has already been observed in [32] for finite element equations associated with problems of nonlinear elasticity.

We can demonstrate the same effect for the simple equation (3) by making the transformation $v = u + \varphi$ where φ is a fixed C^2-function satisfying $\varphi(0) = \varphi(1) = 0$. We obtain the boundary value problem

(7) $v'' + \lambda(v-\varphi) - \varphi'' = 0$ in $[0,1]$, $v(0) = v(1) = 0$

and the finite difference equations

(8) $D_h^2 v_h + \lambda v_h - \lambda[\varphi]_h - [\varphi'']_h = 0, \; v_h(0) = v_h(1) = 0 .$

Let us write the restrictions $[\varphi]_h, [\varphi'']_h$ in terms of the eigenfunctions φ_{nh}, i.e.

$$[\varphi]_h = \sum_{n=1}^{M} \xi_n \varphi_{nh} , \quad [\varphi'']_h = \sum_{n=1}^{M} \eta_n \varphi_{nh} ,$$

then the solutions of (8) are readily computed from the ansatz

$$v_h = \sum_{n=1}^{M} \mu_n \varphi_{nh} \quad \text{as}$$

$$\mu_n(\lambda) \begin{cases} = (\lambda\xi_n + \eta_n)(\lambda - \lambda_{nh})^{-1} , & \lambda \neq \lambda_{nh} , \\ \in \mathbb{R} \text{ arbitrary} , & \lambda = \lambda_{nh} \text{ if } \lambda_{nh}\xi_n + \eta_n = 0 . \end{cases} \qquad (n = 1,\ldots,M)$$

In case $\lambda_{jh}\xi_j + \eta_j \neq 0$ for some $j = 1,\ldots,M$, there is no solution to (8) for $\lambda = \lambda_{jh}$ and the j-th bifurcation point (λ_j, φ) of (7) is separated. If $\varphi = \varphi_1$, for example, the bifurcation point (λ_1, φ_1) is separated by the difference equations while the bifurcation points (λ_n, φ_1), $n \geq 2$, are shifted to $(\lambda_{nh}, (\lambda_{nh} - \lambda_1)(\lambda_{nh} - \lambda_{1h})^{-1} \varphi_{1h})$.

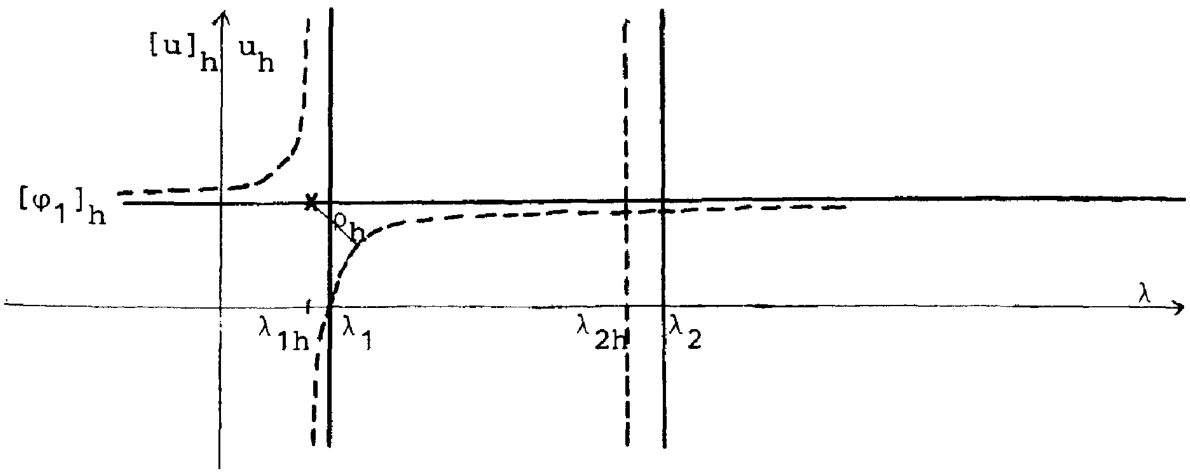


fig. 3

We note that the solution branch

$$(9)\quad \left(\lambda, \frac{\lambda-\lambda_1}{\lambda-\lambda_{1h}}\varphi_{1h}\right) = (\lambda_{1h},\varphi_{1h})+\left(\lambda-\lambda_{1h}, \frac{\lambda_{1h}-\lambda_1}{\lambda-\lambda_{1h}}\varphi_{1h}\right), \lambda \neq \lambda_{1h}$$

can be regarded as a hyperbola with centre $(\lambda_{1h},\varphi_{1h})$ and semi-axis $\rho_h = (2|\lambda_{1h}-\lambda_1|)^{1/2}$ in the space spanned by $(0,\varphi_{1h})$ and $(1,0)$. Moreover, the distance (measured in some norm) of the discrete branch (9) from the restriction (λ_1,φ_{1h}) of the bifurcation point (λ_1,φ_1) is at most $O(h)$. This is half the order of convergence obtained for the shifted bifurcation points (cf. (6)).

Let us return to the case of a general $\varphi \in C^2$ where the n-th bifurcation point (λ_n,φ) was separated if and only if

$$(10)\quad \lambda_{nh}\,\xi_n + \eta_n \neq 0 .$$

This is a generic condition; it may be further elucidated by noting that the discretization error of (λ_n,φ) in the equations (8) is given by

$$-\sum_{i=1}^{M}(\lambda_{ih}\xi_i+\eta_i)\varphi_{ih} .$$

Hence (10) means that this discretization error has a nonzero component with respect to the n-th discrete eigenfunction.

In section 3 we will prove an abstract theorem which generalizes our observations for the simple example (7),(8) to the abstract setting of equations (1) and (2). The result will be developed within the framework of discrete approximations. The basic tool is a quantitave theorem on perturbed bifurcation which will be given in section 2. Finally, the general results are applied to finite difference equations in section 4 and numerical examples are treated in section 5 .

2. Perturbed bifurcation

From fig. 3 it is obvious that there is a close relation between discrete and perturbed bifurcation diagrams. There are quite a few approaches to perturbed bifurcation in the literature (see e.g. [7,11,16,22]). However, these results are not sufficient in our situation for various reasons: usually the knowledge of a primary branch is assumed, sometimes only parts of the perturbed branches are constructed [16] or the branches are parametrized by λ which leads to unnecessary nondegeneracy conditions [7]. Moreover, we need a quantitative theorem which applies to a family of perturbed bifurcation problems (depending on the mesh parameter h) and which ensures neighbourhoods of parameter independent size.

We will follow Crandall, Rabinowitz [9] who suggest to treat (1) in the joint variable $z = (\lambda, u)$. Therefore, we consider an equation

(11) $T(z) = 0, \; z \in Z$

where $T \in C^2(Z,Y)$ and Z,Y are Banach spaces.

Definition 1

$z_o \in Z$ is called a hyperbolic point of T if

(i) $T(z_o) = 0$,

(ii) $\dim N(T'(z_o)) = 2$, $\operatorname{codim} R(T'(z_o)) = 1$
(N = null space, R = range),

(iii) there exist linearly independent $r_1, r_2 \in N(T'(z_o))$ such that
$r \in N(T'(z_o)), T''(z_o)r^2 \in R(T'(z_o)) \iff$
$r = c\,r_1$ or $r = c\,r_2$ for some $c \in \mathbb{R}$.

As is shown in [28] under the assumption $T \in C^3$ (see [9,17,27,32] for related results), the solutions of (11) in a neighbourhood of a hyperbolic point z_o of T consist of two intersecting branches

$z_j(t) = z_o + tr_j + tw_j(t)$, $|t| < \rho$, $j = 1,2$
where $w_j \in C^1((-\rho,\rho),Z)$
and $w_j(0) = 0$ $(j = 1,2)$.
The vectors $r_j = z_j'(0)$ are the bifurcation directions. It can easily be seen that this result also holds under the weaker assumption $T \in C^2$.

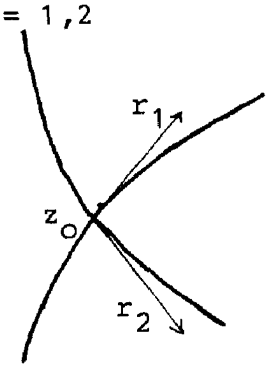


fig. 4

Usually, condition (iii) above is expressed as the indefiniteness of the quadratic form

(12) $g(x,y) = \alpha x^2 + 2\beta xy + \gamma y^2$

where $\alpha = \langle\psi,T''(z_o)p^2\rangle$, $\beta = \langle\psi,T''(z_o)pq\rangle$, $\gamma = \langle\psi,T''(z_o)q^2\rangle$, $N(T'(z_o)) = \text{span}\{p,q\}$ and $\psi \in Y^*$ is a continuous linear functional such that $R(T'(z_o)) = N(\psi)$.

Let us now consider a perturbed problem

(13) $T(z,\tau) = 0$, $T \in C^2(Z \times \mathbb{R}, Y)$

where $z_o \in Z$ is a hyperbolic point of $T(\cdot,0)$.
Then the following assumption is basic in perturbed bifurcation theory (cf. [16])

(14) $T_\tau^o \notin R(T_z^o)$.

Here lower indices denote partial derivatives and the upper index 'o' indicates the argument $(z_o,0)$, e.g. $T_\tau^o = T_\tau(z_o,0)$, $T_z^o = T_z(z_o,0)$ etc. . It will be convenient to normalize the bifurcation directions r_1,r_2 and the functional $\psi \in Y^*$ with $N(\psi) = R(T_z^o)$ in such a way that

(15) $\langle\psi,T_{zz}^o r_1 r_2\rangle = 1$, $\langle\psi,T_\tau^o\rangle = -1$.

Finally, we will only consider the case $\tau > 0$ since the results for $\tau < 0$ can be obtained from the equation $T(z,-\tau) = 0$ after redefining ψ,r_1 and r_2 .

A first approximation to the solutions of (13) is given by the hyperbolas $z_o + tl_s$ where

$tl_s = t(sr_1+s^{-1}r_2)$, $s > o$, $t^2 = \tau$ $(t \in \mathbb{R})$.

Clearly, $<\psi,\frac{1}{2}T^o_{zz}\, l_s^2> = <\psi,T^o_{zz}r_1\, r_2> = -<\psi,T^o_\tau>$ and hence

(16) $\frac{1}{2}\, T^o_{zz}l_s^2 + T^o_\tau \in R(T^o_z)$, $s > o$.

Let us write $Z = N(T^o_z) \oplus V$. Then $T : V \to R(T^o_z)$ is a homeomorphism and by (16) there is a uniquely determined $v_s \in V$ such that

$n_s^2\, T^o_z\, v_s + \frac{1}{2}\, T^o_{zz}\, l_s^2 + T^o_\tau = 0$, $s > 0$,

where $n_s = \frac{1}{2}(s+s^{-1})$ provides two-sided bounds for $\|l_s\|$ by

$$2n_s \inf_{0\le\sigma\le 1} \|\sigma r_1+(1-\sigma)r_2\| \le \|l_s\| \le 2\mathrm{Max}(\|r_1\|, \|r_2\|)n_s , \quad s > o .$$

Now $z = z_o+tl_s + (tn_s)^2 v_s$ is a 'better' approximation to the solutions of (13) since $T(z_o+tl_s+(tn_s)^2v_s,t^2) = O(|tn_s|^3)$ by a Taylor expansion. In general, the solutions of (13) can be obtained from a nonlinear correction to $z_o + tl_s$ as described in the following

Theorem 1:

Let $T \in C^2(Z\times\mathbb{R},Y)$ and let $z_o \in Z$ be a hyperbolic point of $T(\cdot,0)$ such that $T^o_\tau \notin R(T^o_z)$. Then there exists a $\delta > 0$ and solutions of (13) of the form

(17) $z(t,s) = z_o + tl_s + (tn_s)\phi(t,s)$, $\tau = t^2$ $((t,s) \in M_\delta, t \neq 0)$

where $M_\delta = \{(t,s) \in \mathbb{R}^2 : s > 0, |t|n_s < \delta\}$ and $\phi : M_\delta \to Z$ is a continously differentiable function satisfying $\phi(0,s) = 0$, $s>0$. Moreover, there is a $\delta' > 0$ such that every solution of (13) in $\|z-z_o\| \le \delta'$, $0 < \tau \le \delta'$ is given by (17).

A complete proof of theorem 1 is given in [4]. For our purposes it is important to note from the proof that the quantities δ and δ' depend on the following 'data':

(18) a positive lower bound for $\inf_{0 \le \sigma \le 1} \|\sigma r_1 + (1-\sigma) r_2\|$,

(19) upper bounds for $\|r_1\|$, $\|r_2\|$, $\|T_\tau^o\|$, $\|T_z^o\|$, $\|T_{zz}^o\|$ and for $\|T_{z\tau}\|$, $\|T_{\tau\tau}\|$ in a neighbourhood of $(z_o, 0)$,

(20) the modulus of continuity of T_{zz} at $(z_o, 0)$, i.e. $\omega(\varepsilon) = \sup\{\|T_{zz}(z,\tau) - T_{zz}^o\| : \|z - z_o\| \le \varepsilon \ , \ |\tau| \le \varepsilon\}$,

(21) the constant κ in the stability inequality $\|v\| \le \kappa \|T_z^o v\|$ for all $v \in V$.

If these data can be estimated uniformly for a family of operators T and of spaces U,Y then there are common values of δ and δ' for the whole family. Moreover, for each $\bar\varepsilon > 0$ there is a $\bar\delta = \bar\delta(\bar\varepsilon)$, independent of the family, such that

(22) $\|\phi(t,s)\| \le \bar\varepsilon$ if $|t| n_s \le \bar\delta$.

3. Discrete approximations of bifurcation problems

In this section we consider the relation between (1) and (2) in the framework of discrete approximations (cf. [13,24,25,26]). We will use the notations of [26]. To simplify matters we make the following assumptions:

U is separable, $\dim U_h = \dim Y_h < \infty$ $(h \in H)$,

$p_h \in L(U,U_h), q_h \in L(Y,Y_h)$ are bounded linear operators such that

$\|p_h u\| \to \|u\|$ $(h \in H)$ for every $u \in U$,

$\|q_h y\| \to \|y\|$ $(h \in H)$ for every $y \in Y$.

Let us briefly review some standard definitions ([26,§1,2]).

P-convergence : $u_h \overset{P}{\to} u\,(h \in H) \;:\Longleftrightarrow\; \|u_h - p_h u\| \to 0 \;(h \in H)$,

Q-convergence : $y_h \overset{Q}{\to} y\,(h \in H) \;:\Longleftrightarrow\; \|y_h - q_h y\| \to 0 \;(h \in H)$,

P-compactness : a sequence $u_h \in U_h\,(h \in H)$ is called P-compact, iff every subsequence has a P-convergent subsequence.

For bounded linear operators $A \in L(U,Y)$, $A_h \in L(U_h,Y_h)$ we have

PQ-convergence : $A_h \overset{PQ}{\to} A\,(h \in H) \;:\Longleftrightarrow\; \|A_h\| \le C \;(h \in H)$ and $A_h p_h u \overset{Q}{\to} Au\,(h \in H)$ for all $u \in U$

(in the special case $Y_h = \mathbb{R}$, $Y = \mathbb{R}$ the operators A_h, A are linear functionals and the PQ-convergence is written as $A_h \rightharpoonup A\,(h \in H)$),

regular convergence: $A_h \to A$ regular $:\Longleftrightarrow$ $A_h \overset{PQ}{\to} A$ and $(\|u_h\| \le C\,(h \in H),\, A_h u_h$ is Q-compact $\Rightarrow u_h$ is P-compact$)$,

stable convergence: $A_h \to A$ stable $:\Longleftrightarrow$ $A_h \overset{PQ}{\to} A$ and A_h^{-1} exists for almost every $h \in H$ and $\|A_h^{-1}\| \le C$.

First we derive a result on the convergence of simple eigenvalues which may be of interest in itself since we do not require analyticity with respect to the eigenparameter. The proof has some similarities to the methods used in [31] for asymptotic expansions.

Definition 2: Let $\Lambda \subset \mathbb{R}$ be an open set and let $A : \Lambda \to L(U,Y)$ be continuously differentiable. $\lambda_o \in \Lambda$ is called a simple eigenvalue of A iff $N(A(\lambda_o)) = \mathrm{span}\{\varphi\}$ for some $\varphi \in U, \varphi \neq 0$, and $A'(\lambda_o)\varphi \notin R(A(\lambda_o))$.

We do not adopt here the notation of an $A'(\lambda_o)$-simple eigenvalue which is common in bifurcation theory (e.g. [10]), since the above definition is a straightforward generalization of simple eigenvalues for analytic operators A (e.g. [26,§4]).

Lemma 1:

Let $\lambda_o \in \Lambda \subset \mathbb{R}$ be a simple eigenvalue of $A \in C^1(\Lambda, L(U,Y))$ and let operators $A_h \in C^1(\Lambda, L(U_h, Y_h))$ be given such that

(23) $A_h(\lambda_o) \to A(\lambda_o)$ regular, $A_h'(\lambda_o) \overset{PQ}{\to} A'(\lambda_o)$.

Let A_h' be equicontinuous at λ_o, i.e. for each $\bar{\varepsilon} > 0$ there is a $\bar{\delta} > 0$ with $\|A_h'(\lambda) - A_h'(\lambda_o)\| \leq \bar{\varepsilon}$ if $|\lambda - \lambda_o| \leq \bar{\delta}$, $h \in H$. Then there exists a $\delta_o > 0$ such that A_h has exactly one simple eigenvalue λ_h in $[\lambda_o - \delta_o, \lambda_o + \delta_o]$ for a.e. $h \in H$. Moreover there is a corresponding eigenfunction φ_h satisfying

(24) $|\lambda_o - \lambda_h| + \|p_h\varphi - \varphi_h\| \leq C \|A_h(\lambda_o)p_h\varphi - q_hA(\lambda_o)\varphi\|$.

Finally, we can define functionals $\psi \in Y^*, \psi_h \in Y_h^*$ by

$\psi(R(A(\lambda_o))) = \{0\}, \langle\psi,\zeta\rangle = 1$, $\zeta = A'(\lambda_o)\varphi$,

$\psi_h(R(A_h(\lambda_h))) = \{0\}, \langle\psi_h,\zeta_h\rangle = 1, \zeta_h = A_h'(\lambda_h)\varphi_h$

and for these $\psi_h \rightharpoonup \psi \, (h \in H)$ holds .

Proof: Let $U = \mathrm{span}\{\varphi\} \oplus W$ and define $g \in U^*$ by $g(W) = \{0\}$ and $\langle g,\varphi\rangle = 1$. As U is separable there exists a sequence $f_h \in U_h^*$ such that $f_h \rightharpoonup g$ ([26,§1(37)]) . Hence $\langle f_h, p_h\varphi\rangle \to 1$ and the functionals $g_h = \langle f_h, p_h\varphi\rangle^{-1} f_h$ satisfy $g_h \rightharpoonup g, \langle g_h, p_h\varphi\rangle = 1$
Now we define the auxiliary operators
$B : \Lambda \times U \to \mathbb{R} \times Y$, $B(\lambda,u) = (\langle g,u\rangle - 1, A(\lambda)u)$

$B_h : \Lambda \times U_h \to \mathbb{R} \times Y_h, B_h(\lambda,u_h) = (\langle g_h,u_h\rangle - 1, A_h(\lambda)u_h)$

and apply the local convergence theorem [26,§3(14)] to the nonlinear equations $B(\lambda,u) = 0$ and

(25) $B_h(\lambda,u_h) = 0$.

For that purpose we use the projections $\bar{p}_h : \mathbb{R} \times U \to \mathbb{R} \times U_h$, $\bar{p}_h(\lambda,u) = (\lambda, p_hu)$ and $\bar{q}_h : \mathbb{R} \times Y \to \mathbb{R} \times Y_h, \bar{q}_h(\lambda,u) = (\lambda, q_hu)$. The main steps in the proof of the conditions of [26,§3(14)] are:

(i) $B(\lambda_o,\varphi) = 0$,

(ii) $B'(\lambda_o,\varphi) = \begin{pmatrix} 0 & g \\ A'(\lambda_o)\varphi & A(\lambda_o) \end{pmatrix}$ and $N(B'(\lambda_o,\varphi)) = \{0\}$ from the simplicity of the eigenvalue λ_o ,

(iii) $B_h'(\lambda_o,p_h\varphi) = \begin{pmatrix} 0 & g_h \\ A_h'(\lambda_o)p_h\varphi & A_h(\lambda_o) \end{pmatrix} \to B'(\lambda_o,\varphi)$ regular

which is a consequence of (23) .

Therefore, (25) has a unique solution (λ_h,φ_h) for a.e. $h \in H$ in some neighbourhood $|\lambda - \lambda_o| + \|u_h - p_h\varphi\| \le \delta_1$. Moreover, from [26,§3(15)] we have $|\lambda_h - \lambda_o| + \|p_h\varphi - \varphi_h\| \le C\| B_h(\lambda_o,p_h\varphi)\| =$
$= C\|A_h(\lambda_o)p_h\varphi - q_hA(\lambda_o)\varphi\|$ which yields (24) and $\lambda_h \to \lambda_o, \varphi_h \overset{P}{\to} \varphi$.
Also, λ_h is the only eigenvalue of $A_h(\lambda)$ in some neighbourhood $|\lambda - \lambda_o| \le \delta_o$. Let us assume to the contrary that there exists a subsequence $H' \subset H$ and sequences $\mu_h \in \Lambda, u_h \in U_h (h \in H')$ with

(26) $\mu_h \to \lambda_o,\ \|u_h\| = 1\ ,\ A_h(\mu_h)u_h = 0$

such that $\mu_h \neq \lambda_h$ or $u_h \notin \mathrm{span}\{\varphi_h\}$.
Then we have

(27) $\|A_h(\lambda_o)u_h\| \le \|A_h(\lambda_o) - A_h(\mu_h)\|\ \|u_h\| \to 0 \quad (h \in H')$

and $u_h \overset{P}{\to} u\ (h \in H'' \subset H')$ for some $u \in U$, $\|u\| = 1$.
Hence $A(\lambda_o)u = 0$ by (23),(27) and $u = c\varphi$ for some $c \neq 0$.
Moreover, $v_h := \langle g_h,u_h\rangle^{-1} u_h \overset{P}{\to} \varphi\ (h \in H'')$ and (μ_h,v_h) is a solution of (25) as well as $|\mu_h - \lambda_o| + \|p_h\varphi - v_h\| \le \delta_1$ for a.e. $h \in H''$. Therefore $\mu_h = \lambda_h$, $v_h = \varphi_h$ which is a contradiction.

Finally, we obtain from conditions (ii) and (iii) that
$B_h'(\lambda_o,p_h\varphi) \to B'(\lambda_o,\varphi)$ stable, $R(B'(\lambda_o,\varphi)) = \mathbb{R} \times Y$ (cf. [26,§2(60)])
and also $B_h'(\lambda_h,\varphi_h) \to B'(\lambda_o,\varphi)$ stable by the Banach lemma.
Hence we have a stability inequality

(28) $|\lambda| + \|u_h\| \leq C\ \{ |g_h(u_h)| + \|\lambda\zeta_h + A_h(\lambda_h)u_h\| \}$, $(\lambda, u_h) \in \mathbb{R}\times U_h$

for a.e. $h \in H$. (28) shows that $\zeta_h \notin R(A_h(\lambda_h))$ so that λ_h is a simple eigenvalue of A_h . Moreover,
$Y_h = R(A_h(\lambda_h)) \oplus \mathrm{span}\{\zeta_h\}$, $Y = R(A(\lambda_o)) \oplus \mathrm{span}\{\zeta\}$
and $\psi_h \rightharpoonup \psi$ follows upon noting that
$<\psi_h, y_h>$ is the first component of $(B_h'(\lambda_h,\varphi_h))^{-1}(0,y_h)$, $y_h \in Y_h$
and $<\psi, y>$ is the first component of $(B'(\lambda_o,\varphi))^{-1}(0,y)$, $y \in Y$.

q.e.d

Let us return to equation (1) and assume that

(V_1) (λ_o, u_o) is a simple bifurcation point of $T \in C^2(\mathbb{R}\times U, Y)$.

By definition this means that $z_o = (\lambda_o, u_o)$ is a hyperbolic point of the operator T and λ_o is a simple eigenvalue of $T_u(\cdot, u_o)$. Recalling definition 1 and 2 the explicit meaning of (V_1) is:

$T(\lambda_o, u_o) = 0$,

$N(T_u^o) = \mathrm{span}\{\varphi\}$ for some $\varphi \neq 0$ where $T_u^o = T_u(\lambda_o, u_o)$,

codim $R(T_u^o) = 1$, say $R(T_u^o) = N(\psi)$, $\psi \in Y^*$,

$T_{\lambda u}^o \varphi \notin R(T_u^o)$,

$T_\lambda^o \in R(T_u^o)$, e.g.

(29) $T_u^o w = -T_\lambda^o$ where $w \in W$ and $U = \mathrm{span}\{\varphi\} \oplus W$,

$\alpha\gamma - \beta^2 < 0$ where $\alpha = <\psi, T_{uu}^o \varphi^2>$, $\beta = <\psi, T_{\lambda u}^o \varphi + T_{uu}^o\, w\varphi>$,

$$\gamma = <\psi, T_{\lambda\lambda}^o + 2T_{\lambda u}^o\, w + T_{uu}^o\, w^2>.$$

The last condition refers to the quadratic form g of (12) which is determined by ψ and the basis $p = (0,\varphi)$, $q = (1,w)$ of $N(T'^o)$.

Our next two assumptions relate the operators T_h and T to each other.

$$(V_2)\quad \begin{array}{l} T_h^o \xrightarrow{Q} T^o\ ,\ T_{h,\lambda}^o \xrightarrow{Q} T_\lambda^o\ ,\ T_{h,u}^o \to T_u^o \ \text{ regular}, \\ T_{h,\lambda\lambda}^o \xrightarrow{Q} T_{\lambda\lambda}^o\ ,\ T_{h,\lambda u}^o \xrightarrow{PQ} T_{\lambda u}^o\ ,\ T_{h,uu}^o \xrightarrow{P\times P,Q} T_{uu}^o\ , \end{array}$$

(V_3) T_h'' is equicontinuous at $(\lambda_o, p_h u_o)$, i.e. for each $\bar{\varepsilon} > o$ there is a $\bar{\delta} > o$ such that $\|T_h''(\lambda, u_h) - T_h''^{o}\| \leq \bar{\varepsilon}$ if $|\lambda - \lambda_o| + \|u_h - p_h u_o\| \leq \bar{\delta}$ $(h \in H)$.

Here we have used the abbreviations $T_h^o = T_h(\lambda_o, p_h u_o)$, $T_{h,u}^o = T_{h,u}(\lambda_o, p_h u_o)$ etc. and a $P \times P, Q$ - convergence defined by

$\|T_{h,uu}^o\| \leq C$ $(h \in H)$ and $T_{h,uu}^o p_h u \, p_h v \overset{Q}{\to} T_{uu}^o uv$ for all $u, v \in U$.

By (V_1)-(V_3) we can apply lemma 1 to $A = T_u(\cdot, u_o)$, $A_h = T_{h,u}(\cdot, p_h u_o)$, and we will use the notations of lemma 1 throughout.

Our main idea in the treatment of $T_h(\lambda, u_h) = 0$ is to find an auxiliary operator $S_h : \mathbb{R} \times U_h \to Y_h$ with the properties

(i) $(\lambda_o, p_h u_o)$ is a simple bifurcation point of S_h,

(ii) S_h is a small perturbation of T_h.

As we will show this can be achieved by setting

(30) $S_h(\lambda, u_h) = T_h(\lambda, u_h) - T_h^o - (\lambda - \lambda_o)\rho_h \zeta_h - (T_{h,u}^o - T_{h,u}(\lambda_h, p_h u_o))(u_h - p_h u_o)$

where λ_h, φ_h are given by lemma 1 and

(31) $\rho_h = \langle \psi_h, T_{h,\lambda}^o \rangle$, $\zeta_h = T_{h,u}(\lambda_h, p_h u_o)\varphi_h$.

The stability inequality (28) now reads

(32) $|\lambda| + \|u_h\| \leq C\{|g_h(u_h)| + \|\lambda\zeta_h + T_{h,u}(\lambda_h, p_h u_o)u_h\|\}, (\lambda, u_h) \in \mathbb{R} \times U_h$

and there is a unique $w_h \in N(g_h)$ such that $T_{h,u}(\lambda_h, p_h u_o)w_h = \rho_h \zeta_h - T_{h,\lambda}^o$. If we put $\lambda = \rho_h, u_h = p_h w - w_h - \langle g_h, p_h w \rangle \varphi_h$ in (32), with w from (29), then

$|\rho_h| + \|p_h w - w_h - \langle g_h, p_h w \rangle \varphi_h\| \leq C\| T_{h,u}(\lambda_h, p_h u_o) p_h w + T_{h,\lambda}^o \|$
$\leq C(\|T_{h,u}^o p_h w + T_{h,\lambda}^o\| + |\lambda_o - \lambda_h|)$.

Combining this with (24) yields

(33) $|\rho_h| \le C(\|T^o_{h,u}p_h w - q_h T^o_u w\| + \|T^o_{h,\lambda} - q_h T^o_\lambda\| + \|T^o_{h,u}p_h\varphi - q_h T^o_u \varphi\|)$

and $\rho_h \to 0$ by (V_2). Moreover $\langle g_h, p_h w\rangle \to \langle g,w\rangle = 0$ and hence $w_h \overset{p}{\to} w$.

Lemma 2: Let (V_1)-(V_3) be satisfied. Then $(\lambda_o, p_h u_o)$ is a simple bifurcation point of S_h for a.e. $h \in H$.
In particular, we have a decomposition $\mathbb{R} \times U_h = N(S_h'^o) \oplus V_h$ and linearly independent vectors $r_{1h}, r_{2h} \in N(S_h'^o)$ such that for a.e. $h \in H$

(34) $|\lambda| + \|u_h\| \le C \| S_h'^o(\lambda, u_h)\|$ for all $(\lambda, u_h) \in V_h$,

(35) $\langle \psi_h, S_h''^o r_{jh}^2\rangle = 0 \, (j=1,2)$, $\langle \psi_h, S_h''^o r_{1h} r_{2h}\rangle = 1$,

(36) $\|r_{jh}\| \le C \; (j=1,2)$, $0 < C \le \inf_{0 \le \sigma \le 1} \|\sigma r_{1h} + (1-\sigma) r_{2h}\|$.

Proof: By lemma 1 and the definition of S_h we have
$S_h^o = 0$, $N(S^o_{h,u}) = N(T_{h,u}(\lambda_h, p_h u_o)) = \mathrm{span}\{\varphi_h\}$,
$R(S^o_{h,u}) = N(\psi_h)$, $S^o_{h,\lambda} = T^o_{h,\lambda} - \rho_h \zeta_h \in R(S^o_{h,u})$ and
$S^o_{h,\lambda u} \varphi_h \notin R(S^o_{h,u})$ since
$\langle \psi_h, S^o_{h,\lambda u} \varphi_h\rangle = \langle \psi_h, T^o_{h,\lambda u} \varphi_h\rangle \to \langle \psi, T^o_{\lambda u} \varphi\rangle \neq 0$.
Furthermore $N(S_h'^o) = \mathrm{span}\{(0,\varphi_h),(1,w_h)\}$ and we can choose $V_h = \{0\} \times N(g_h)$. (34) then follows directly from (32).
It remains to show (35) and (36), since (35) also yields the condition (iii) of a hyperbolic point.
Now the bifurcation directions r_1, r_2 (see (15)) can be represented as $r_j = \theta_j(0,\varphi) + \sigma_j(1,w) \, (j=1,2)$. If we define

(37) $\bar{r}_{jh} = \theta_j(0,\varphi_h) + \sigma_j(1,w_h) \quad (j=1,2)$,

then $\langle \psi_h, S_h''^o \bar{r}_{jh}^2\rangle = \langle \psi_h, T_h''^o \bar{r}_{jh}^2\rangle \to \langle \psi, T''^o r_j^2\rangle = 0 \quad (j=1,2)$
and $\langle \psi_h, S_h''^o \bar{r}_{1h} \bar{r}_{2h}\rangle \to \langle \psi, T''^o r_1 r_2\rangle = 1$.

Because of these relations it is easy to find a transformation

$r_{1h} = (1+d_h)\bar{r}_{1h} + d_h\bar{r}_{2h},\ r_{2h} = e_h\bar{r}_{2h} + f_h r_{1h}$

such that (35) is satisfied and $d_h \to 0,\ e_h \to 1,\ f_h \to 0$.
But then also (36) is obvious from (37) and the linear independence of r_1 and r_2 . q.e.d.

We now proceed to the basic assumption on the discretization error $T_h(\lambda_o,p_h u_o)$ of the bifurcation point. The following condition is in some sense the generalization of condition (10) from the introduction

(V_4) $T_h^o = h^m q_h e + o(h^m)$ for some $m \in \mathbb{R}$, $m > 0$ and some $e \notin R(T_u^o)$.

(V_4) requires that the principal error term has a coefficient which is not in the range of T_u^o . If bifurcation from the trivial solution is considered then normally $T_h^o = 0$ and (V_4) is violated.
Let $\tau_h = -\langle\psi_h, T_h^o\rangle$,

then from $\psi_h \rightharpoonup \psi$ and (V_4) we have

(38) $\tau_h = -h^m\langle\psi,e\rangle + o(h^m) \neq 0$ for a.e. $h \in H$.

We will apply theorem 1 to the operator

$$\Gamma_h : \begin{array}{l} \mathbb{R} \times U_h \times \mathbb{R} \to Y \\ (\lambda,u_h,\tau) \to S_h(\lambda,u_h) + \frac{\tau}{\tau_h}(T_h(\lambda,u_h) - S_h(\lambda,u_h)) . \end{array}$$

By lemma 2, $(\lambda_o,p_h u_o)$ is a hyperbolic point of $\Gamma_h(\cdot,\cdot,0)=S_h(\cdot,\cdot)$ and $\langle\psi_h,\Gamma_{h,\tau}^o\rangle = \tau_h^{-1}\langle\psi_h,T_h^o - S_h^o\rangle = -1$.
Moreover, with the exception of $\|\Gamma_{h,\tau}^o\|$ and $\|\Gamma_{h,z\tau}^o\|$ the data from (18)-(20) can be estimated uniformly in h . This follows from lemma 2, (V_2),(V_3) and the formulas

$\Gamma_{h,zz}(\cdot,\tau) = T_h''(\cdot)$, $\Gamma_{h,z}^o = S_h'^o = (T_{h,\lambda}^o - \varphi_h\zeta_h, T_{h,u}(\lambda_h,p_h u_o))$, $\Gamma_{h,\tau\tau} = 0$.
Finally, $\|\Gamma_{h,\tau}^o\| = |\tau_h|^{-1}\|T_h^o\| \leq C$ by (V_4),(38), and
$\|\Gamma_{h,z\tau}^o\| = |\tau_h|^{-1}\|T_h'^o - S_h'^o\| =$

$|\tau_h|^{-1}(\|T^o_{h,u} - T_{h,u}(\lambda_h,p_hu_o)\| + \|\rho_h\zeta_h\|) \le C|\tau_h|^{-1}(|\lambda_o-\lambda_h| + |\rho_h|)$

is uniformly bounded if we assume

(V_5) $\|T^o_{h,u}p_hw - q_hT^o_u w\| + \|T^o_{h,\lambda} - q_hT^o_\lambda\| + \|T^o_{h,u}p_h\varphi - q_hT^o_u\varphi\| = O(h^m)$

(compare (24),(33),(38)).

(V_5) requires that certain discretization errors caused by the linearizations $T_{h,u}, T_{h,\lambda}$ are at least of the same order as the discretization error $T^o_h = T^o_h - q_hT^o$. This is quite natural in applications (see section 4).

On the whole we have shown that the values of δ and δ' in theorem 1 can be chosen independently of h. If $\langle\psi,e\rangle < 0$ then $0 < \tau_h \le \delta,\delta'$ for a $h \in H$, and the operator $\Gamma_h(\cdot,\tau_h) = T_h(\cdot)$ is covered by theorem 1.

Therefore, the discrete equations (2) have solutions

(39) $(\lambda,u_h)(s) = (\lambda_o,p_hu_o) \pm \sqrt{|\tau_h|}\, l_{hs} + \sqrt{|\tau_h|}\, n_s\, \phi_{h\pm}(s)\ (s \in M_{h\delta})$

where $M_{h\delta} = \{s > 0 : \sqrt{|\tau_h|}\, n_s < \delta\}$, $\phi_{h\pm} \in C(M_{h\delta}, \mathbb{R} \times U_h)$,

$\phi_{h\pm}(1) \to 0$ and $l_{hs} = sr_{1h} + s^{-1}r_{2h}$, $s > 0$.

If $\langle\psi,e\rangle > 0$ then r_{2h} has to be replaced by $-r_{2h}$.

Moreover, for each $\bar{\varepsilon} > 0$ there is a $\bar{\delta} > 0$, independent of h, such that

(40) $\|\phi_{h\pm}(s)\| \le \bar{\varepsilon}$ if $\sqrt{|\tau_h|}\, n_s \le \bar{\delta}$

(see (22)).

Theorem 2:

Let $(V_1)-(V_5)$ be satisfied. Then the discrete equations $T_h(\lambda,u_h) = 0$ have two solution branches (39) for a.e. $h \in H$ and there is a $\delta' > 0$ such that all solutions in $|\lambda - \lambda_o| + \|u_h - p_hu_o\| \le \delta'$ belong to these branches.

In the vicinity of (λ_o,p_hu_o) the two branches can be written as

(41) $(\lambda_o, p_h u_o) \pm h^{m/2} |\langle\psi, e\rangle|^{1/2} (s\, \bar{p}_h r_1 + s^{-1} \bar{p}_h r_2) + o(h^{m/2})$

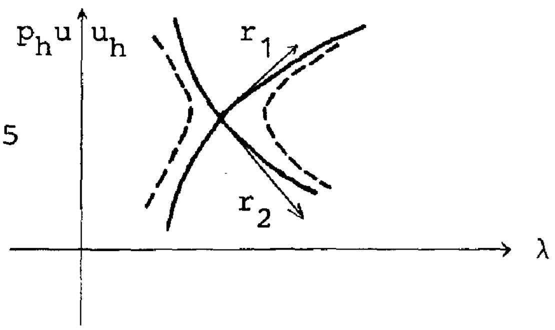
where $0 < s_o \leq s \leq s_1$. In (41) the vectors $r_j = (\lambda_j, u_j)$ (j=1,2) are the bifurcation directions, $\bar{p}_h r_j = (\lambda_j, p_h u_j)$ are their projections and the normalization

(42) $\langle\psi, T''^o r_1 r_2\rangle = 1$, $\langle\psi, e\rangle < 0$

is assumed.

The proof of the representation (41) follows immediately from (39) and (40) since $0 < C_o \leq n_s = \frac{1}{2}(s+s^{-1}) \leq C_1$ for some constants C_o, C_1 independent of h . Note, however, that (41) only represents a portion of the two branches which is close to $(\lambda_o, p_h u_o)$. The situation for the solution set of (2) can be visualized as in fig. 5

fig. 5



A simple estimate using (39) and (40) also shows that the distance of the discrete branches from $(\lambda_o, p_h u_o)$ behaves like $h^{m/2}$, i.e.

(43) $0 < \underline{C}\, h^{m/2} \leq \inf\{ \|(\lambda_o, p_h u_o) - (\lambda, u_h)(s)\| \;:\; s \in M_{h\delta}\} \leq \bar{C}\, h^{m/2}$

for some $0 < \underline{C} \leq \bar{C}$.

4. An application to finite difference equations

Due to its general nature, theorem 2 has far reaching applications to many approximation methods of numerical analysis. For example, we could use theorem 2 for approximations of bifurcation problems which involve Galerkin or finite element methods [30,32] or - more generally - collectively compact operators [1] .

In what follows we will demonstrate the application of theorem 2 to second order systems

(44) $u'' + f(x,u,\lambda) = 0$ in $[a,b]$, $u(a) = u_a$, $u(b) = u_b$

where $u \in C^2([a,b],\mathbb{R}^N)$, $f \in C^2([a,b] \times \mathbb{R}^{N+1},\mathbb{R}^N)$, $u_a, u_b \in \mathbb{R}^N$. The finite difference equations under consideration are

(45) $D_h^2 u_h + K_h F_h(u_h,\lambda) = 0$ in $\mathring{J}_h$, $u_h(a) = u_a, u_h(b) = u_b$

where $J_h = \{a,a+h,\dots,b-h,b\}, h = (b-a)(M+1)^{-1}$, $M \in \mathbb{N}$, and $\mathring{J}_h \doteq \{a+h,\dots,b-h\}$. $u_h \in U_h := (\mathbb{R}^N)^{J_h}$ is the unknown grid function and D_h^2 is the second difference quotient as in (5). Moreover F_h is the nonlinear operator

$$F_h(u_h,\lambda)(x) = f(x,u_h(x),\lambda) \ , \ x \in J_h \ , \ u_h \in U_h$$

and $K_h \in L(U_h,\mathring{U}_h)$, $\mathring{U}_h = (\mathbb{R}^N)^{\mathring{J}_h}$, is a matrix such that either

$K_h u_h(x) = u_h(x)$, $x \in \mathring{J}_h$, case I ($O(h^2)$-method)

or

$K_h u_h(x) = \frac{1}{12}(u_h(x-h) + 10 u_h(x) + u_h(x+h))$, case II
(Hermitian $O(h^4)$-method [8]) .

It can be shown that our theory applies to more general boundary value problems and further finite difference methods [4], for example to the higher order schemes of [5] and to the linear schemes in the sense of [3] if some standard assumptions from the theory of convergence are satisfied ([2,12,20]).
Our equations (44), (45) can easily be subjected to the abstract setting of section 3 if we define:

$U = C^2([a,b],\mathbb{R}^N)$ with the norm $\|u\| = \|u\|_\infty + \|u'\|_\infty + \|u''\|_\infty$,

$\|\ \|_\infty$ always denotes the maximum norm for continuous functions, grid functions, vectors etc. . Further, let

$Y = C([a,b],\mathbb{R}^N) \times \mathbb{R}^{2N}$ with the norm $\|(v,\gamma)\| = \|v\|_\infty + \|\gamma\|_\infty$,

$p_h u = [u]_h$ = restriction of a function $u : [a,b] \to \mathbb{R}^N$ to the grid J_h, $Y_h = \overset{o}{U}_h \times \mathbb{R}^{2N}$ with the norm $\|(v_h,\gamma)\| = \|v_h\|_\infty + \|\gamma\|_\infty$,
$q_h(v,\gamma) = (K_h[v]_h,\gamma)$ for $(v,\gamma) \in Y$.
In U_h we use the norm $\|u_h\| = \|u_h\|_\infty + \|D_h^1 u_h\|_\infty + \|D_h^2 u_h\|_\infty$ where $D_h^1 u_h(x) = h^{-1}(u_h(x+h) - u_h(x))$, $x = a,\dots,b-h$.
The operators T and T_h are given by

$$T(\lambda,u) = (u'' + f(\cdot,u,\lambda), u(a) - u_a, u(b) - u_b)$$

$$T_h(\lambda,u_h) = (D_h^2 u_h + K_h F_h(u_h,\lambda), u_h(a) - u_a, u_h(b) - u_b).$$

With these definitions the assumptions (V_2), (V_3) are satisfied for an arbitrary $(\lambda_o,u_o) \in \mathbb{R} \times U$. The regular convergence $T^o_{h,u} \to T^o_u$ can be proved as in [26,§6] by using the theorem of Arzela and Ascoli.

Let us now assume that (λ_o,u_o) is a simple bifurcation point of the boundary value problem (44). Then (λ_o,u_o) is a solution of (44) and there exists an eigenfunction φ of the linearized equation

$$\varphi'' + A_o\varphi = 0 \text{ in } [a,b],\ \varphi(a) = \varphi(b) = 0, \text{ where } A_o = f_u(\cdot,u_o,\lambda_o)$$

and also an eigenfunction $\varphi^* \in U$ of the adjoint equation

$$\varphi^{*\prime\prime} + A_o^T\varphi^* = 0 \text{ in } [a,b],\ \varphi^*(a) = \varphi^*(b) = 0.$$

The range $R(T_u^o)$ can then be written as the nullspace of a functional

$$\text{(46)}\quad \begin{aligned} &\psi : C([a,b],\mathbb{R}^N) \times \mathbb{R}^{2N} \to \mathbb{R}, \\ &\langle\psi,(v,\gamma)\rangle = \langle\varphi^*,v\rangle + \sum_{i=1}^{2N} c_i^*\gamma_i,\quad \langle\varphi^*,v\rangle = \sum_{i=1}^{N}\int_a^b \varphi_i^*(x)v_i(x)\,dx \end{aligned}$$

for some properly chosen $c_i^* \in \mathbb{R}$.

Finally, there exists a $w \in C^2([a,b], \mathbb{R}^N)$ such that

(47) $w'' + A_o w = - f_\lambda(\cdot, u_o, \lambda_o), \; w(a) = w(b) = 0.$

The principal error term in (V_4) is obtained from the remainder of the above difference formulas (see [8]) as

case I : $m = 2$, $e = \frac{1}{12}(u_o^{(4)}, 0, 0)$ if $u_o \in C^4[a,b]$,

case II : $m = 4$, $e = -\frac{1}{240}(u_o^{(6)}, 0, 0)$ if $u_o \in C^6[a,b]$.

With these values of m one easily verifies condition (V_5) if $w, \varphi \in C^4[a,b]$ in case I ($\in C^6[a,b]$ in case II). Our smoothness assumptions on u_o, w and φ follow from $f \in C^3$ in case I ($f \in C^5$ in case II). Thus we have proved

Theorem 3:

Let (λ_o, u_o) be a simple bifurcation point of the boundary value problem (44) and let $f \in C^3$ in case I ($f \in C^5$ in case II). Assume that

$$\eta = \begin{cases} \frac{1}{12}\langle \varphi^*, u_o^{(4)} \rangle & \text{in case I,} \\ -\frac{1}{240}\langle \varphi^*, u_o^{(6)} \rangle & \text{in case II} \end{cases}$$

is different from zero.
Then the finite difference equations (45) have two solution branches for sufficiently small h

$$(\lambda_o, [u_o]_h) \pm h^{\bar{m}} \sqrt{|\eta|}\,(s(\lambda_1, [u_1]_h) + s^{-1}(\lambda_2, [u_2]_h)) + o(h^{\bar{m}}),$$
$$0 < s_o \leq s \leq s_1,$$

where $\bar{m} = 1$ in case I, $\bar{m} = 2$ in case II and (λ_j, u_j) (j=1,2) are the bifurcation directions of the boundary value problem at (λ_o, u_o).

By (42) the normalization of φ^* and (λ_j, u_j) (j=1,2) is

$$(48)\qquad 0 \begin{cases} > \langle \varphi^*, u_o^{(4)} \rangle & \text{in case I} \\ < \langle \varphi^*, u_o^{(6)} \rangle & \text{in case II, and} \end{cases}$$

$$1 = \langle \varphi^*, f^o_{uu}\, u_1 u_2 + \lambda_1 f^o_{u\lambda}\, u_2 + \lambda_2\, f^o_{u\lambda}\, u_1 + \lambda_1 \lambda_2\, f^o_{\lambda\lambda} \rangle \quad .$$

Theorem 3 clearly shows that the assumptions $(V_2)(V_3)(V_5)$ in theorem 2 are more of technical type while (V_1) and (V_4) play the crucial role in the separation of bifurcation points.

5. Numerical results

Theorem 3 gives a general description of finite difference solutions near nontrivial or secondary bifurcation points of the boundary value problem (44). The crucial condition $\eta \neq 0$ usually cannot be checked explicitly since the bifurcation point (λ_o, u_o) is unknown a-priori.

A simple example, however, where this can be done is the following scalar boundary value problem

$$(49)\quad u'' + \lambda \sin(u - \varphi_1) - \varphi_1'' = 0 \;, \quad u(0) = u(1) = 0$$

where $\varphi_1(x) = \sin \pi x \,,\; 0 \le x \le 1$.

(49) is obtained by the transformation $u = v + \varphi_1$ from the equation of the mathematical pendulum

$$(50)\quad v'' + \lambda \sin v = 0 \;\text{ in } [0,1],\; v(0) = v(1) = 0$$

(compare the illustrative example in the introduction).
Note that equation (50) includes the solutions of

$$v'' + \lambda \sin v = 0 \text{ in } [\tfrac{1}{2},1] \;,\; v'(\tfrac{1}{2}) = v(1) = 0$$

which describes the buckling of a rod (see [6] for numerical results and further references). The exact solutions of (50) and hence those of (49) can be expressed in terms of elliptic integrals [6].

It is easily verified that (π^2,φ_1) is a simple bifurcation point of (49). Moreover, $(-1,0),(0,\varphi_1)$ are the bifurcation directions and with $\varphi^* = -2\varphi_1$ the normalization (48) is satisfied.

The quantity η from theorem 3 is then computed as

$$\eta = -\frac{\pi^4}{12} \quad \text{in case I} \quad \text{and} \quad \eta = -\frac{\pi^6}{240} \quad \text{in case II}$$

Theorem 3 shows that the finite difference equations (45) associated with (49) have solutions in a neighbourhood of $(\pi^2,[\varphi_1]_h)$

$$(51) \qquad (\pi^2,[\varphi_1]_h) \pm \begin{cases} h\,\dfrac{\pi^2}{2\sqrt{3}}\,(-s,s^{-1}[\varphi_1]_h) + o(h) & \text{in case I,} \\[2ex] h^2\,\dfrac{\pi^3}{4\sqrt{15}}\,(-s,s^{-1}[\varphi_1]_h) + o(h^2) & \text{in case II} \end{cases}$$

where $0 \le s_0 \le s \le s_1$.

We have solved these finite difference equations numerically by employing a continuation procedure (for details of the method see [4,17]). Fig. 6 shows the exact solution branches of (49) and the discrete solution branches for the $O(h^2)$-method $(h = \frac{1}{10}, \frac{1}{20}, \frac{1}{40})$.

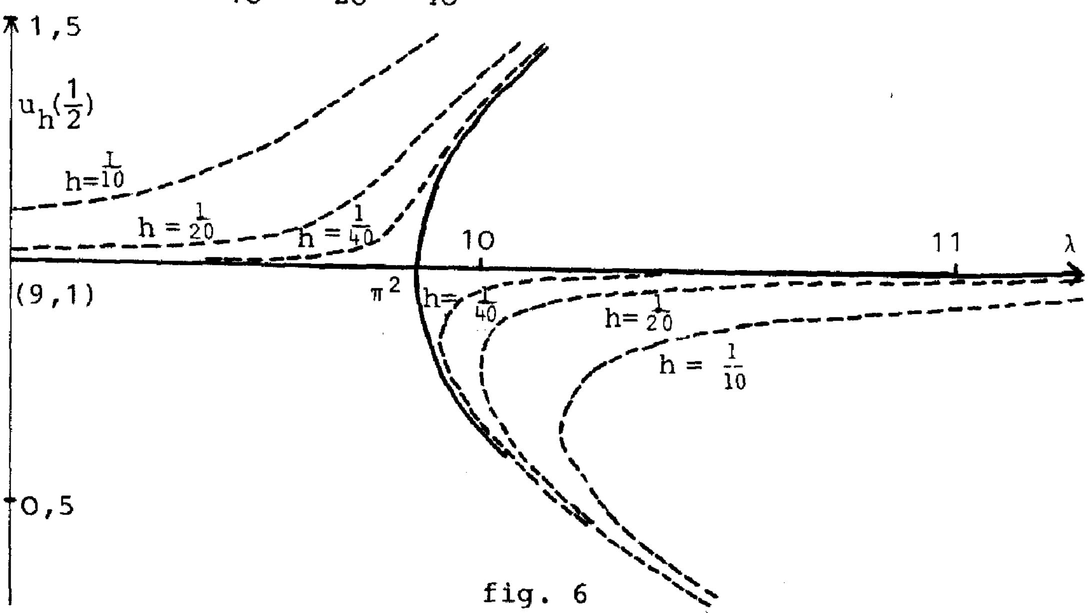


fig. 6

We have also computed the discrete solution branches in case of the $O(h^4)$-method, but these branches are already very close to the continuous ones and would not show up in fig.6. A close-up of the situation near the bifurcation point is given in fig. 7 in the case of the $O(h^4)$-method, $h = \frac{1}{10}$.

fig. 7

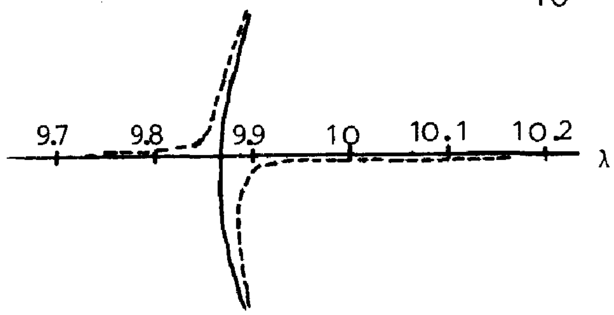


Fig. 7 also shows that a continuation procedure starting on the left upper branch would normally pass to the right lower branch if the stepsize with respect to the continuation parameter is not small enough. Following [17], the change of sign of a certain determinant in this critical step would then indicate a bifurcation point although the solutions consist of two separated branches.

It is worth noting at this point that the finite difference equations of the example from [17] indeed have a nontrivial bifurcation point. This follows from the fact that the example in [17] has a primary branch $(\lambda, q(\lambda)u_o)$ where u_o is a quadratic polynomial. This branch is reproduced exactly by the finite difference equations. In particular, the quantity η of our theorem 3 is zero.

The next table contains the distances of the upper and the lower branches from the restriction $(\pi^2, [\varphi_1]_h)$ of the bifurcation point. The distances were measured in the maximum norm

$$\|(\lambda, u_h)\| = \mathrm{Max}\{|\lambda|, \|u_h\|_o\} .$$

h	$O(h^2)$ - method		$O(h^4)$ - method	
	upper branch	lower branch	upper branch	lower branch
$\frac{1}{10}$	0.278	0.302	0.0200	0.0201
$\frac{1}{20}$	0.140	0.146	0.0050	0.0050
$\frac{1}{40}$	0.071	0.072	0.0012	0.0012

This clearly shows the halving of the order of convergence near the bifurcation point. Moreover, from formula (51) we can estimate these distances as

$$\frac{\pi^2}{2\sqrt{3}}\, h \cong 2.849\, h \quad \text{in case I}, \quad \frac{\pi^3}{4\sqrt{15}}\, h^2 \cong 2.001\, h^2 \quad \text{in case II}$$

which is in good agreement with the numerical values.

A more realistic example with a secondary bifurcation point is given by the Ginzburg-Landau equations for a superconducting slab of thickness d (cf. [19,21,23])

$$\begin{aligned} u_1'' &= K^2 u_1 (u_1^2 - 1 + \lambda u_2^2) \\ u_2'' &= u_1^2 u_2 \end{aligned} \quad \text{in } [-\tfrac{d}{2}, \tfrac{d}{2}], \quad \begin{aligned} u_1'(\pm \tfrac{d}{2}) &= 0 \\ u_2'(\pm \tfrac{d}{2}) &= 1 . \end{aligned}$$

Here K^2 is the Ginzburg-Landau parameter, u_1 is called the order parameter and $\sqrt{\lambda}\, u_2$ is one component of the potential of the magnetic field. The bifurcation parameter λ is the square of the external field. This example not only exhibits a bifurcation from the trivial solution $u_1 = 0, u_2(x) = x$, but also a secondary bifurcation. The separation of this secondary bifurcation point by the finite difference method depends on the formulas used for the boundary conditions. For a detailed numerical study of this example we refer to [4].

References

[1] Atkinson, K.E.: The numerical solution of a bifurcation problem. SIAM J. Numer. Anal. 14, 584-599 (1977).

[2] Beyn, W.-J.: Zur Stabilität von Differenzenverfahren für Systeme linearer gewöhnlicher Randwertaufgaben. Numer. Math. 29, 209-226, (1978).

[3] Beyn, W.-J.: On the convergence of the finite difference method for nonlinear ordinary boundary value problems, pp. 9-19. Constructive Methods for Nonlinear Oscillations J. Albrecht, L. Collatz, K. Kirchgässner (Eds.), ISNM 48 Birkhäuser Verlag Basel, (1979).

[4] Beyn, W.-J.: Zur Approximation von Lösungszweigen nichtlinearer Randwertaufgaben mit dem Differenzenverfahren, Manuscript 1980.

[5] Bohl, E., Lorenz, J.: Inverse monotonicity and difference schemes of higher order. A summary for two-point boundary value problems. Aequ. Math. 19, 1-36, (1979).

[6] Bohl, E.: On the numerical treatment of a class of discrete bifurcation problems, to appear in IAC, Istituto par le Applicazioni del Calcolo "Mauro Picone", Publicazioni Serie III (1980).

[7] Chow, S.-N., Hale, J.K., Mallet Paret J.: Applications of generic bifurcation I. Arch. Rat. Mech. Anal. 59,159-188, (1975).

[8] Collatz, L.: The numerical treatment of differential equations, 3rd ed. Berlin-Göttingen-Heidelberg, Springer 1966.

[9] Crandall, M.G., Rabinowitz, P.H.: Bifurcation from simple eigenvalues. J. of Funct. Anal. 8, 321-340 (1971).

[10] Crandall, M.G., Rabinowitz, P.H.: Bifurcation, perturbation of simple eigenvalues, and linearized stability. Arch. Rat. Mech. Anal. 52, 161-180 (1973).

[11] Golubitsky, M., Schaeffer, D.: A theory for imperfect bifurcation via singularity theory. Comm. Pure Appl. Math. 32, 21-98 (1979).

[12] Grigorieff, R. D.: Die Konvergenz des Rand- und Eigenwertproblems linearer gewöhnlicher Differenzengleichungen. Numer. Math. 15, 15-48, (1970).

[13] Grigorieff, R.D.: Zur Theorie linearer approximationsregulärer Operatoren I, Math. Nachr. 55, 233-249, (1972), II, Math. Nachr. 55, 251-263, (1972).

[14] Grigorieff, R.D., Jeggle, H.: Approximation von Eigenwertproblemen bei nichtlinearer Parameterabhängigkeit. Manuscripta math. 10, 245-271, (1973).

[15] Grigorieff, R.D.: Diskrete Approximation von Eigenwertproblemen I, Qualitative Konvergenz, Numer. Math. 24,355-374, (1975), II Konvergenzordnung, Numer.Math. 24,415-433, (1975).

[16] Keener, J.P., Keller, H.B.: Perturbed bifurcation theory, Arch. Rat. Mech. Anal. 50, 159-175, (1974).

[17] Keller, H.B.: Numerical solution of bifurcation and nonlinear eigenvalue problems in: Rabinowitz, P.H. (Ed.), Applications of bifurcation theory, Proceedings of an advanced seminar conducted by the Mathematics Research Center, The University of Wisconsin at Madison, New York, Academic Press (1977).

[18] Kornhuber, R.: Approximation von Verzweigungsproblemen mit Anwendungen auf Differential- und Integralgleichungen. Diplomarbeit, Technische Universität Berlin, 1979.

[19] Koslowski, R.: Verzweigungsverhalten von Lösungen eines eindimensionalen Supraleitungsproblems . Universität zu Köln, Math. Inst.,Diplomarbeit 1975.

[20] Kreiss, H.-O.: Difference approximations for boundary and eigenvalue problems for ordinary differential equations. Math. of Comp. 26, 605-624 (1972).

[21] Odeh, F.: Existence and bifurcation theorems for the Ginzburg-Landau Equations. J. Math. Phys. 8, 2351-2356 (1967).

[22] Potier-Ferry, M.: Perturbed bifurcation theory, Journal of Diff. Equ. 33, 112-146 (1979).

[23] Seydel, R.: Numerical computation of branch points in ordinary differential equations. Numer. Math. 32,51-68 (1979).

[24] Stummel, F.: Diskrete Konvergenz linearer Operatoren I, Math. Ann. 190, 45-92 (1970).

[25] Vainikko, G.: Über die Konvergenz und Divergenz von Näherungsmethoden bei Eigenwertproblemen. Math. Nachr. 78, 145-164 (1977).

[26] Vainikko, G.: Funktionalanalysis der Diskretisierungsmethoden. Teubner Texte zur Mathematik, Leipzig, Teubner Verlag 1976.

[27] Weber, H.: Numerische Behandlung von Verzweigungsproblemen bei gewöhnlichen Randwertaufgaben, pp. 176-190, in: Albrecht, J., Collatz, L., Kirchgässner, K. (Eds.) Constructive methods for nonlinear boundary value problems ISNM 48, Birkhäuser Verlag, Stuttgart 1979.

[28] Weber, H.: Numerische Behandlung von Verzweigungsproblemen bei Randwertaufgaben gewöhnlicher Differentialgleichungen, Ph. D. Thesis, Mainz (1978).

[29] Weiss, R.: Bifurcation in difference approximations to two point boundary value problems. Math. Comp. 29, 746-760 (1975).

[30] Westreich, D., Vaarol, Y.L.: Applications of Galerkin's method to bifurcation and two point boundary value problems. J. Math. Anal. Appl. 70, 399-422 (1979).

[31] Wolf, R.: Asymptotische Entwicklungen für Eigenwerte und Eigenvektoren bei der Approximation parameternichtlinearer Eigenwertaufgaben. Numer. Math. 30, 207-226 (1978).

[32] Yamaguti, M., Fujii, H.: On numerical deformation of singularities in nonlinear elasticity, 267-278, in R. Glowinski, J.L. Lions (Eds.): Computing Methods in Applied Sciences and Engineering 1977, 1. Springer Lecture Notes in Mathematics 704.

Wolf-Jürgen Beyn
Universität Konstanz
Fachbereich Mathematik
Postfach 7733
D-7750 Konstanz