Numer. Math. 38, 1–30 (1981)

Numerische Mathematik


# Finite Dimensional Approximation of Nonlinear Problems

## Part III: Simple Bifurcation Points

F. Brezzi[1]*, J. Rappaz[2]**, and P.-A. Raviart[3]

[1] Istituto di Matematica Applicata and Istituto di Analisi Numerica del C.N.R., Università di Pavia, I-27100 Pavia, Italy
[2] Centre de Mathématiques Appliquées (ERA/CNRS 747), Ecole Polytechnique, F-91128 Palaiseau Cedex, France
[3] Analyse Numérique, Université P. et M. Curie, 4, Place Jussieu, F-75230 Paris Cedex 05, France

**Summary.** In the first two papers of this series [4, 5], we have studied a general method of approximation of nonsingular solutions and simple limit points of nonlinear equations in a Banach space. We derive here general approximation results of the branches of solutions in the neighborhood of a simple bifurcation point. The abstract theory is applied to the Galerkin approximation of nonlinear variational problems and to a mixed finite element approximation of the von Kármán equations.

*Subject Classifications:* AMS(MOS): 65N30; CR: 5.17.

## 1. Introduction

Consider a nonlinear problem of the form:

$$F(\lambda, u) = 0, \tag{1.1}$$

where $F$ is a smooth function from $\mathbb{R} \times V$ into $V$ for some Banach space $V$. In the first two papers of this series [4, 5], we have studied the numerical approximation of branches of solutions of problem (1.1) in a neighborhood of a nonsingular point and of a limit point. We now turn to the approximation of solutions of (1.1) in a neighborhood of a simple bifurcation point $(\lambda_0, u_0)$ of $F$.

More precisely, we assume that the point $(\lambda_0, u_0) \in \mathbb{R} \times V$ is a simple critical point of $F$ in the sense that:

$$F(\lambda_0, u_0) = 0; \tag{1.2}$$

* The work of F. Brezzi has been completed during his stay at the Université P. et M. Curie and at the Ecole Polytechnique

** The work of J. Rappaz has been supported by the Fonds National Suisse de la Recherche Scientifique

0029-599X/81/0038/0001/$06.00

$$D_u F(\lambda_0, u_0) \quad \text{is singular and} \\ \dim \operatorname{Ker}(D_u F(\lambda_0, u_0)) = \operatorname{codim} \operatorname{Range}(D_u(F(\lambda_0, u_0)) = 1; \tag{1.3}$$

$$D_\lambda F(\lambda_0, u_0) \in \operatorname{Range}(D_u F(\lambda_0, u_0)). \tag{1.4}$$

Moreover, we suppose an additional condition on the 2nd derivatives of $F$ (cf. (2.14)) which ensures that, in a neighborhood of the simple critical point $(\lambda_0, \mu_0)$, the solutions of (1.1) consist exactly of two branches which intersect transversally at $(\lambda_0, u_0)$.

As in [4, 5], we shall present a fairly general analysis which includes a variety of classical approximation schemes such as conforming finite element methods, mixed finite element methods, spectral methods...

An outline of the paper is a follows. In Sect. 2, we recall some standard local results on simple bifurcation points in a form which is well adapted to the study of the numerical methods. Section 3 is devoted to the general theory of approximation of simple bifurcation points: we prove that, in a fixed neighborhood of $(\lambda_0, u_0)$, there exist two branches of solutions of the numerical problem which may not intersect, and we give sharp error estimates of the distance between the set of solutions of problem (1.1) and the set of solutions of the approximate problem. In Sect. 4, we consider two important particular situations: bifurcation from the trivial branch and symmetry-breaking bifurcation; in these two cases, we prove that the approximate problem has indeed a simple bifurcation point $(\lambda_h^0, u_h^0)$ and we obtain optimal results for $|\lambda_h^0 - \lambda_0|$ and $\|u_h^0 - u_0\|_V$.

In Sect. 5, we apply the above results to the Galerkin approximations of nonlinear variational problems; we then improve and generalize results of [8, 14, 16]. Finally, Sect. 6 is devoted to the study of the Hellan-Herrmann-Johnson mixed finite element scheme for the von Kármán equations yet considered in [4]. Note that the corresponding analysis can be easily extended to a variety of mixed finite element approximations of the von Kármán equations or other nonlinear problems. For partial results previously obtained in that direction, see [15].

A more detailed analysis of the behaviour of the branches of approximate solutions will appear in [19]. The methods of this paper have been extended so as to cover the case of Hopf bifurcation [3]. Some of our results have also been generalized to the case of multiple bifurcation [18].

Throughout the paper, we shall constantly use the following notations. Given a smooth function $\Phi$ from $X \times Y$ into $Z$ for some normed spaces $X$, $Y$, $Z$, we denote by $D^m \Phi(x, y) \in \mathscr{L}_m(X \times Y; Z)$ the $m$-th total derivative of $\Phi$ at the point $(x, y) \in X \times Y$, where $\mathscr{L}_m(X \times Y; Z)$ is the space of all continuous $m$-linear mappings from $X \times Y$ into $Z$. We denote by $D_x \Phi(x, y)$, $D_y \Phi(x, y)$, $D^2_{xx} \Phi(x, y)$, $D^2_{xy} \Phi(x, y)$, $D^2_{yy} \Phi(x, y), \ldots$ the corresponding partial derivatives and we set for brevity

$$D^2_{xx} \Phi(x, y) \cdot (\xi)^2 = D^2_{xx} \Phi(x, y) \cdot (\xi, \xi),$$

$$D^3_{xxx} \Phi(x, y) \cdot (\xi)^3 = D^3_{xxx} \Phi(x, y) \cdot (\xi, \xi, \xi), \qquad \xi \in X.$$

We shall also use the classical Sobolev spaces $W^{m,p}(\Omega)$ and $H^m(\Omega) = W^{m,2}(\Omega)$ with the norms $\| \ \|_m = \| \ \|_{H^m(\Omega)}$.

## 2. Local Analysis of Simple Bifurcation Points

Let $V$ and $W$ be two (real) Banach spaces with norms $\|.\|_V$ and $\|.\|_W$ respectively. We introduce a $C^p$ mapping ($p \geqq 2$) $G$: $\mathbb{R} \times V \to W$ and a linear *compact* operator $T \in \mathscr{L}(W; V)$. We set:

$$F(\lambda, u) = u + TG(\lambda, u). \tag{2.1}$$

We assume that $(\lambda_0, u_0) \in \mathbb{R} \times V$ is *a simple critical point* of $F$ in the sense that

(i) $F^0 \equiv F(\lambda_0, u_0) = 0$;
(ii) $D_u F^0 \equiv D_u F(\lambda_0, u_0) = I + TD_u G(\lambda_0, u_0) \in \mathscr{L}(V; V)$ is singular and $-1$ is an eigenvalue of the compact operator $TD_u G(\lambda_0, u_0)$ with the algebraic multiplicity 1; (2.2)
(iii) $D_\lambda F^0 \equiv D_\lambda F(\lambda_0, u_0) \in \text{Range}(D_u F^0)$.

We want to solve the equation:

$$F(\lambda, u) = 0 \tag{2.3}$$

in a neighborhood of the simple critical point $(\lambda_0, u_0)$.

As a consequence of (2.2)(ii) and the classical Riesz-Schauder theory on compact operators, there exists $\varphi_0 \in V$ such that

$$\begin{aligned} &D_u F^0 \cdot \varphi_0 = 0, \qquad \|\varphi_0\|_V = 1, \\ &V_1 \equiv \text{Ker}(D_u F^0) = \mathbb{R}\, \varphi_0. \end{aligned} \tag{2.4}$$

Moreover, denoting by $V'$ the dual space of $V$ and by $\langle .,. \rangle$ the duality pairing between the spaces $V$ and $V'$, there exists $\varphi_0^* \in V'$ such that

$$\begin{aligned} &(D_u F^0)^* \cdot \varphi_0^* = 0, \qquad \langle \varphi_0, \varphi_0^* \rangle = 1, \\ &V_2 \equiv \text{Range}(D_u F^0) = \{v \in V; \langle v, \varphi_0^* \rangle = 0\}. \end{aligned} \tag{2.5}$$

Finally, we have:

$$V = V_1 \oplus V_2,$$

and $D_u F^0$ is an isomorphism of $V_2$. We denote by $L \in \mathscr{L}(V_2; V_2)$ the inverse isomorphism of $D_u F^0|_{V_2}$.

Using (2.5), we notice that the condition (2.2)(iii) becomes:

$$\langle D_\lambda F^0, \varphi_0^* \rangle = 0. \tag{2.6}$$

Let us next define the projection operator $Q: V \to V_2$ by

$$Qv = v - \langle v, \varphi_0^* \rangle \varphi_0. \tag{2.7}$$

Then, the Eq. (2.3) is equivalent to the system

$$\begin{aligned} &QF(\lambda, u) = 0, \\ &(I - Q) F(\lambda, u) = \langle F(\lambda, u), \varphi_0^* \rangle \varphi_0 = 0. \end{aligned} \tag{2.8}$$

Now, by the implicit function theorem, there exist (cf. [5, Lemma 2]) two positive constants $\xi_0$, $\alpha_0$ and a unique $C^p$ mapping $v$: $[-\xi_0, \xi_0] \times [-\alpha_0, \alpha_0] \to V_2$ such that

$$\begin{aligned} &QF(\lambda_0 + \xi, u_0 + \alpha\varphi_0 + v(\xi, \alpha)) = 0, \\ &v(0,0) = 0. \end{aligned} \tag{2.9}$$

Hence, solving the Eq. (2.3) in a neighborhood of the critical point $(\lambda_0, u_0)$ amounts to solve the *bifurcation equation*

$$f(\xi, \alpha) \equiv \langle F(\lambda_0 + \xi, u_0 + \alpha\varphi_0 + v(\xi, \alpha)), \varphi_0^* \rangle = 0 \tag{2.10}$$

in a neighborhood of the origin.

By elementary calculations, we obtain

$$f(0,0) = \frac{\partial f}{\partial \xi}(0,0) = \frac{\partial f}{\partial \alpha}(0,0) = 0, \tag{2.11}$$

so that the origin is a critical point of the function $f$. Furthermore, we have

$$\frac{\partial^2 f}{\partial \xi^2}(0,0) = C_0, \qquad \frac{\partial^2 f}{\partial \xi\, \partial \alpha}(0,0) = B_0, \qquad \frac{\partial^2 f}{\partial \alpha^2}(0,0) = A_0 \tag{2.12}$$

where

$$\begin{aligned} &A_0 = \langle D^2_{uu} F^0 \cdot (\varphi_0)^2, \varphi_0^* \rangle, \\ &B_0 = \langle D^2_{\lambda u} F^0 \cdot \varphi_0 + D^2_{uu} F^0 \cdot (\varphi_0, -L D_\lambda F^0), \varphi_0^* \rangle, \\ &C_0 = \langle D^2_{\lambda\lambda} F^0 + 2 D^2_{\lambda u} F^0 \cdot (-L D_\lambda F^0) + D^2_{uu} F^0 \cdot (-L D_\lambda F^0)^2, \varphi_0^* \rangle. \end{aligned} \tag{2.13}$$

From now on, we shall assume that $(\lambda_0, u_0)$ is *a simple bifurcation point* of $F$, i.e. a simple critical point of $F$ which satisfies in addition

$$B_0^2 - A_0 C_0 > 0. \tag{2.14}$$

Then, by Morse lemma (cf. [17] for instance), we find that, in a neighborhood of the origin, the solutions of the Eq. (2.10) consist in two curves of class $C^{p-2}$ which are transverse at the origin. Hence, we have

**Theorem 1.** *Assume that* $(\lambda_0, u_0)$ *is a simple bifurcation point of* $F$. *Then, in a neighborhood of* $(\lambda_0, u_0)$, *the solution of* (2.3) *consist in two* $C^{p-2}$ *branches which intersect transversally at the point* $(\lambda_0, u_0)$.

In fact, we need to derive some appropriate parametrization of these two branches of solutions. We are looking for a general representation of the form

$$\begin{aligned}\lambda &= \lambda_0 + \xi(t),\\ u &= u_0 + \alpha(t)\,\varphi_0 + v(\xi(t), \alpha(t)),\end{aligned} \tag{2.15}$$

with

$$\xi(t) = t\,\sigma(t), \qquad \alpha(t) = t\,a(t). \tag{2.16}$$

Let us then consider the function $\mathscr{F}$: $(t, \sigma, a) \in [-1, 1] \times [-\xi_0, \xi_0] \times [-\alpha_0, \alpha_0] \to \mathscr{F}(t, \sigma, a) \in \mathbb{R}^2$ defined by

$$\mathscr{F}(t, \sigma, a) = (t^{-2} f(t\sigma, ta), \sigma^2 + a^2 - 1). \tag{2.17}$$

We thus have to solve the equation

$$\mathscr{F}(t, \sigma, a) = 0. \tag{2.18}$$

Note that the second scalar equation $\sigma^2 + a^2 - 1 = 0$ appearing in (2.18) is simply a normalization condition. Let us check

**Lemma 1.** *There exist a positive constant $t_0$ and two pairs of $C^{p-2}$ functions $t \in [-t_0, t_0] \to (\sigma_i(t), a_i(t)) \in \mathbb{R}^2$, $i = 1, 2$, such that*

$$\mathscr{F}(t, \sigma_i(t), a_i(t)) = 0, \qquad i = 1, 2. \tag{2.19}$$

*Proof.* Since $f$ is a $C^p$ function ($p \geqq 2$) and 0 is a critical point of $f$, we have

$$f(t\sigma, ta) = \frac{t^2}{2}(A_0 a^2 + 2B_0 a\sigma + C_0 \sigma^2) + o(t^2), \qquad t \to 0.$$

Hence $\mathscr{F}$ is a $C^{p-2}$ function and we get

$$\mathscr{F}(0, \sigma, a) = (\tfrac{1}{2}(A_0 a^2 + 2B_0 a\sigma + C_0 \sigma^2), \sigma^2 + a^2 - 1).$$

It follows from the condition (2.14) that

$$A_0 a^2 + 2B_0 a\sigma + C_0 \sigma^2 = 0 \tag{2.20}$$

is the equation in the $(\sigma, a)$-plane of two straight lines which intersect at the origin. Therefore, there exist two distinct pairs $(\sigma_i^0, a_i^0)$, $i = 1, 2$, of real numbers such that

$$\mathscr{F}(0, \sigma_i^0, a_i^0) = 0, \qquad \sigma_i^0 > -1, \qquad a_i^0 \geqq 0, \qquad i = 1, 2.$$

On the other hand, $D_{(\sigma, a)}\mathscr{F}$ is a $C^{p-2}$ function and we have

$$D_{(\sigma, a)}\mathscr{F}(0, \sigma, a) = \begin{pmatrix} B_0 a + C_0 \sigma & 2\sigma \\ A_0 a + B_0 \sigma & 2a \end{pmatrix}$$

so that

$$\det(D_{(\sigma, a)}\mathscr{F}(0, \sigma, a)) = 2[B_0(a^2 - \sigma^2) + (C_0 - A_0) a\sigma].$$

By noticing that $B_0(a^2-\sigma^2)+(C_0-A_0)a\sigma=0$ is the equation of the axes of the degenerate hyperbola (2.20), we have

$$\det(D_{(\sigma,a)}\mathscr{F}(0,\sigma_i^0,a_i^0))\neq 0, \qquad i=1,2.$$

Hence, we may apply the implicit function theorem to the function $\mathscr{F}$ at each point $(0,\sigma_i^0,a_i^0)$: for $i=1,2$, there exists a unique pair of $C^{p-2}$ real functions $t\to(\sigma_i(t),a_i(t))$ defined for $|t|\leqq t_0$ such that

$$\mathscr{F}(t,\sigma_i(t),a_i(t))=0,$$
$$\sigma_i(0)=\sigma_i^0, \qquad a_i(0)=a_i^0. \quad \blacksquare$$

We shall denote by $\{(\xi_i(t)=t\sigma_i(t),\ \alpha_i(t)=t a_i(t));\ |t|\leqq t_0\}$, $i=1,2$, the branches of solutions of (2.10) which intersect at the origin and by $\{(\lambda_i(t),\ u_i(t));\ |t|\leqq t_0\}$, $i=1,2$, the corresponding branches of solutions of (2.3).

*Remark* 1. These branches of solutions of (2.3) are usually parametrized in another way. Let us consider separately the two cases $A_0\neq 0$ and $A_0=0$.

*1st case.* $A_0\neq 0$ (*asymmetric bifurcation point* or *fold bifurcation*). Here, we observe that:

$$\frac{d\xi_i}{dt}(0)=\sigma_i(0)\neq 0, \qquad i=1,2.$$

Hence, by the inverse function theorem and for $i=1,2$, we can express $t$ as a $C^{p-2}$ function of $\xi$. The two branches of solutions of (2.3) are therefore of the form

$$\begin{aligned}&\lambda=\lambda_0+\xi,\\ &u=u_i(\xi)=u_0+\alpha_i(\xi)\varphi_0+v(\xi,\alpha_i(\xi)), \qquad i=1,2,\end{aligned} \tag{2.21}$$

where each function $\xi\to a_i(\xi)$ is a $C^{p-2}$ function defined in $[-\xi_0,\xi_0]$, $\xi_0$ small enough, such that

$$\alpha_i(0)=0, \qquad \frac{d\alpha_i}{d\xi}(0)=\frac{a_i(0)}{\sigma_i(0)}.$$

One can easily check that the second equation (2.21) may be written

$$u=u_0+\xi\left(\frac{-B_0\pm\sqrt{B_0^2-A_0C_0}}{A_0}\varphi_0-LD_\lambda F^0\right)+o(\xi), \qquad \xi\to 0. \tag{2.22}$$

*2nd case.* $A_0=0$ (*symmetric bifurcation point* or *cusp bifurcation*). Here we find that

$$\frac{d\xi_1}{dt}(0)=\sigma_1(0)\neq 0, \qquad \frac{d\xi_2}{dt}(0)=\sigma_2(0)=0.$$

Hence, there still exists a branch of solutions of (2.3) of the form (2.21) with $i=1$ which can also be written

$$\begin{aligned}&\lambda=\lambda_0+\xi\\ &u=u_1(\xi)=u_0-\xi\left(\frac{C_0}{2B_0}\varphi_0+LD_\lambda F^0\right)+o(\xi), \qquad \xi\to 0.\end{aligned} \tag{2.23}$$

On the other hand, we have $\frac{d\alpha_2}{d\xi}(0)=a_2(0)=1$. Thus, in that case, we can express $t$ as a $C^{p-2}$ function of $\alpha$ and we find a second branch of solutions of (2.3) given by

$$\begin{aligned}\lambda&=\lambda_2(\xi)=\lambda_0+\xi_2(\alpha),\\ u&=u_2(\alpha)=u_0+\alpha\varphi_0+v(\xi_2(\alpha),\alpha),\end{aligned}\tag{2.24}$$

where $\alpha\to\xi_2(\alpha)$ is a $C^{p-2}$ function defined in $[-\alpha_0,\alpha_0]$, $\alpha_0$ small enough, such that

$$\xi_2(0)=\frac{d\xi_2}{d\alpha}(0)=0.$$

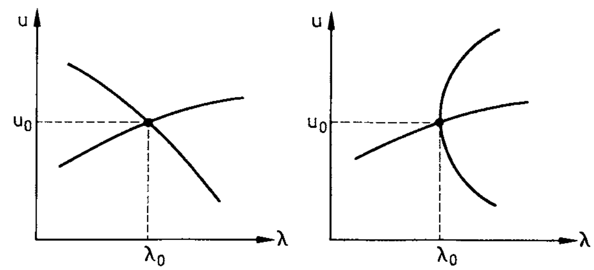
Assume in addition $p\geqq 3$ and set

$$D_0=\frac{\partial^3 f}{\partial\xi^3}(0,0)=\langle D^3_{uuu}F^0\cdot(\varphi_0)^3-3D^2_{uu}F^0\cdot(\varphi_0,-LD^2_{uu}F^0\cdot(\varphi_0)^2),\varphi_0^*\rangle.\tag{2.25}$$

Then, by straightforward calculations, we may write (2.24) in the form

$$\begin{aligned}\lambda&=\lambda_0-\frac{1}{6}\frac{D_0}{B_0}\alpha^2+o(\alpha^2),\\ u&=u_0+\alpha\varphi_0+o(\alpha^2).\end{aligned}\qquad \alpha\to 0\tag{2.26}$$

We summarize the previous results in the following diagram




fold bifurcation: $A_0\neq 0$

cusp bifurcation: $A_0=0$

$D_0\neq 0$, $\frac{D_0}{B_0}<0$ ■

**Fig. 1**

## 3. Approximation of Simple Bifurcation Points

We now turn to the finite-dimensional approximation of Eq. (2.3) in the neighborhood of the simple bifurcation point $(\lambda_0,u_0)$. As in the first two papers of this series, for each value of a parameter $h>0$ which will tend to zero, we introduce a finite-dimensional subspace $V_h$ of the space $V$ and an operator $T_h\in\mathscr{L}(W;V_h)$. We set:

$$F_h(\lambda,u)=u+T_hG(\lambda,u),\qquad \lambda\in\mathbb{R},\ u\in V.\tag{3.1}$$

We then consider the approximate problem: Find the pairs $(\lambda, u_h) \in \mathbb{R} \times V_h$ solutions of the equation

$$F_h(\lambda, u_h) = 0. \tag{3.2}$$

If a pair $(\lambda, u) \in \mathbb{R} \times V$ satisfies $F_h(\lambda, u) = 0$, we observe that $u = -T_h G(\lambda, u) \in V_h$ so that we can equivalently solve the approximate problem (3.2) in $\mathbb{R} \times V$.

As in the previous section, the Eq. (3.2) is equivalent to the following system

$$\begin{aligned} &Q F_h(\lambda, u_h) = 0, \\ &\langle F_h(u, u_h), \varphi_0^* \rangle = 0. \end{aligned} \tag{3.3}$$

Let us briefly recall some of the results obtained in [5] concerning the properties of the system (3.3) which will be constantly used in the sequel. We make the further assumptions:

the mapping $D^p G$ is bounded on all bounded subsets of $\mathbb{R} \times V$ with $p \geqq 3$, (3.4)

and

$$\lim_{h \to 0} \|T - T_h\|_{\mathscr{L}(W; V)} = 0. \tag{3.5}$$

Then, under the hypotheses (2.2)(i), (ii), (3.4) and (3.5), there exist by [5, Theorem 2] three positive constants $\xi_0$, $\alpha_0$, $a$ and, for $h \leqq h_0$[1] small enough, a unique $C^p$ mapping $v_h$: $(\xi, \alpha) \in R(\xi_0, \alpha_0) \to v_h(\xi, \alpha) \in V_2$ such that

$$\begin{aligned} &Q F_h(\lambda_0 + \xi, u_0 + \alpha \varphi_0 + v_h(\xi, \alpha)) = 0, \\ &\|v_h(\xi, \alpha)\|_V \leqq a, \qquad \forall (\xi, \alpha) \in R(\xi_0, \alpha_0), \end{aligned} \tag{3.6}$$

where $R(\xi_0, \alpha_0) = [-\xi_0, \xi_0] \times [-\alpha_0, \alpha_0]$.

Hence, solving problem (3.2) in the neighborhood of the point $(\lambda_0, u_0)$ amounts to solve the approximate bifurcation equation

$$f_h(\xi, \alpha) \equiv \langle F_h(\lambda_0 + \xi, u_0 + \alpha \varphi_0 + v_h(\xi, \alpha)), \varphi_0^* \rangle = 0 \tag{3.7}$$

in a neighborhood of the origin.

We have been able in [5, Lemma 4] to estimate the distance between the functions $f_h$ and $f$. Setting

$$J(\xi, \alpha) = G(\lambda_0 + \xi,\ u_0 + \alpha \varphi_0 + v(\xi, \alpha)) \tag{3.8}$$

and denoting by $K, K_1, \ldots, K_i, \ldots$ (or $C_i$) various positive constants independent of $h$, we have

**Lemma 2.** *Assume the hypotheses* (2.2)(i), (ii), (3.4) *and* (3.5). *Then, for all integer* $m$ *with* $0 \leqq m \leqq p-1$, *there exists a constant* $K_m$ *such that, for all* $(\xi, \alpha)$, $(\xi^*, \alpha^*) \in R(\xi_0, \alpha_0)$, *we have:*

---

[1] The positive constants $\xi_0$, $\alpha_0$, $h_0$ will be chosen smaller and smaller in the sequel of this paper

(i) $$\|D^m f_h(\xi^*, \alpha^*) - D^m f(\xi, \alpha)\|_{\mathscr{L}_m(\mathbb{R}^2;\mathbb{R})} \leq K_m \left\{ |\xi^* - \xi| + |\alpha^* - \alpha| + \sum_{l=0}^{m} \|(T - T_h) D^l J(\xi, \alpha)\|_{\mathscr{L}_l(\mathbb{R}^2; V)} \right\}, \tag{3.9}$$

(ii) $$\|D^p f_h(\xi, \alpha)\|_{\mathscr{L}_p(\mathbb{R}^2;\mathbb{R})} \leq K. \tag{3.9}$$

In fact, we shall need a more specific version of the previous result also proved in [5, Lemma 5]. We introduce a pair of $C^r$ $(0 \leq r \leq p)$ real functions $t \in [-t_0, t_0] \to (\xi(t), \alpha(t)) \in R(\xi_0, \alpha_0)$ and, for $h \leq h_0$, we are given a pair of $C^r$ functions $t \to (\xi_h^*(t), \alpha_h^*(t))$ defined for $|t| \leq t_0$ with values in $R(\xi_0, \alpha_0)$. We assume that

$$\begin{aligned} &\text{(i)} \quad \lim_{h \to 0} \sup_{|t| \leq t_0} \left\{ \left| \frac{d^m}{dt^m} (\xi_h^*(t) - \xi(t)) \right| + \left| \frac{d^m}{dt^m} (\alpha_h^*(t) - \alpha(t)) \right| \right\} = 0, \quad 0 \leq m \leq r-1, \\ &\text{(ii)} \quad \sup_{|t| \leq t_0} \left\{ \left| \frac{d^r}{dt^r} \xi_h^*(t) \right| + \left| \frac{d^r}{dt^r} \alpha_h^*(t) \right| \right\} \leq c. \end{aligned} \tag{3.10}$$

**Lemma 3.** *Assume the hypotheses of Lemma 2 together with* (3.10). *Then, for all integer m with* $0 \leq m \leq r-1$, *there exists a constant* $K_m$ *such that, for all* $|t| \leq t_0$, *we have:*

$$\begin{aligned} &\left| \frac{d^m}{dt^m} (f_h(\xi_h^*(t), \alpha_h^*(t)) - f(\xi(t), \alpha(t))) \right| \leq K_m \sum_{l=0}^{m} \left\{ \left| \frac{d^l}{dt^l} (\xi(t) - \xi_h^*(t)) \right| + \right. \\ &\left. \left| \frac{d^l}{dt^l} (\alpha(t) - \alpha_h^*(t)) \right| + \left\| (T - T_h) \frac{d^l}{dt^l} J(\xi(t), \alpha(t)) \right\|_V \right\}. \end{aligned} \tag{3.11}$$

We shall analyze the behaviour of the branches of solutions of (3.7) in a neighborhood of the origin in terms of the following quantities

$$\mathscr{E}_i^{(m)}(h, t) = \sum_{l=0}^{m} \left\| (T - T_h) \frac{d^l}{dt^l} J(\xi_i(t), \alpha_i(t)) \right\|_V, \quad i = 1, 2, \tag{3.12}$$

where $\{(\xi_i(t), \alpha_i(t));\ |t| \leq t_0\}$, $i = 1, 2$, are the branches of solutions of (2.10) which intersect at the origin.

Now, in order to perform on the approximate problem a similar analysis to that of Sect. 2, we would need the existence of a critical point of the function $f_h$ in a neighborhood of the origin. However such a point does not necessarily exist at least in general but we can state the following result.

**Lemma 4.** *Assume the hypotheses* (2.2), (3.4) *and* (3.5). *Then, there exists for* $h \leq h_0$ *small enough a unique point* $(\xi_h^0, \alpha_h^0) \in R(\xi_0, \alpha_0)$ *such that*

$$Df_h(\xi_h^0, \alpha_h^0) = 0. \tag{3.13}$$

*Moreover, we have the estimates*

$$|\xi_h^0| + |\alpha_h^0| \leq c_1 \|Df_h(0,0)\|_{\mathscr{L}(\mathbb{R}^2;\mathbb{R})} \leq c_2 (\mathscr{E}_1^{(1)}(h, 0) + \mathscr{E}_2^{(1)}(h, 0)). \tag{3.14}$$

*Proof.* It follows from Lemma 2 that the fucntion $Df_h$ converges uniformly in $R(\xi_0, \alpha_0)$ together with its first derivative to the function $Df$ and that its second derivative $D^3 f_h$ is bounded. On the other hand, we have $Df(0,0)=0$ and, due to the condition (2.14), $D^2 f(0,0)$ is invertible. Therefore, we can apply Theorem 1 of [5] in the following particular situation:

$$\Phi(x,y)=Df(y), \qquad \Phi_h(x,y)=Df_h(y), \qquad y=(\xi,\alpha), \qquad g(x)=0.$$

There exists a unique point $(\xi_h^0, \alpha_h^0) \in R(\xi_0, \alpha_0)$ such that (3.13) holds and we have the first inequality (3.14). Applying again the estimate (3.9)(i) with $m=1$ gives

$$\|Df_h(0,0)\|_{\mathscr{L}(\mathbb{R}^2;\mathbb{R})} \leqq c \sum_{l=0}^{1} \|(T-T_h) D^l J(0,0)\|_{\mathscr{L}_l(\mathbb{R}^2;V)}.$$

But, since the two branches $\{(\xi_i(t), \alpha_i(t));\ |t| \leqq t_0\}$, $i=1,2$, intersect transversally at the origin, we have

$$\|(T-T_h) DJ(0,0)\|_{\mathscr{L}(\mathbb{R}^2;V)} \leqq c \sum_{i=1}^{2} \left\| (T-T_h) \frac{d}{dt} J(\xi_i(t), \alpha_i(t))|_{t=0} \right\|_V$$

which implies the second inequality (3.14). ■

We now set

$$D^m f_h^0 = D^m f_h(\xi_h^0, \alpha_h^0), \qquad m=0,1,2,\ldots \tag{3.15}$$

*In general*, we have $f_h^0 \neq 0$ so that $(\xi_h^0, \alpha_h^0)$ is not a critical point of the function $f_h$. Therefore, we set

$$\mathscr{K}(h) = f_h^0 = f_h(\xi_h^0, \alpha_h^0) \tag{3.16}$$

and we introduce the function $\tilde{f}_h$: $R(\xi_0, \alpha_0) \to \mathbb{R}$ defined by

$$\tilde{f}_h(\xi,\alpha) = f_h(\xi,\alpha) - \mathscr{K}(h). \tag{3.17}$$

**Lemma 5.** *Assume the hypotheses of Lemma* 4. *Then, we have the estimates:*

$$\begin{aligned} |\mathscr{K}(h)| &\leqq |f_h(0,0)| + c_1 \sum_{i=1}^{2} \mathscr{E}_i^{(1)}(h,0)^2 \\ &\leqq c_2 \left( \mathscr{E}^{(0)}(h,0) + \sum_{i=1}^{2} \mathscr{E}_i^{(1)}(h,0)^2 \right). \end{aligned} \tag{3.18}$$

*Proof.* Using Lemma 2 with $m=2$, we find that the function $D^2 f_h$ is uniformly bounded so that

$$|f_h^0| \leqq |f_h(0,0)| + \|Df_h(0,0)\|_{\mathscr{L}(\mathbb{R}^2;\mathbb{R})} (|\xi_h^0| + |\alpha_h^0|) + c(|\xi_h^0|^2 + |\alpha_h^0|^2).$$

Next, applying the first inequality (3.14) gives

$$|f_h^0| \leqq |f_h(0,0)| + c\, \|Df_h(0,0)\|^2_{\mathscr{L}(\mathbb{R}^2;\mathbb{R})}.$$

The inequalities (3.18) then follow from Lemma 2 again. ■

*Remark 2.* We have shown in the proof of [5, Theorem 5] that

$$|f_h(0,0)| \leqq |\langle (T-T_h)G^0, \varphi_0^* \rangle| + \|(T-T_h)G^0\|_V (\|(T-T_h)G^0\|_V + \|[(T-T_h)D_u G^0]^* \varphi_0^*\|_{V'}), \tag{3.19}$$

where $G^0 = G(\lambda_0, u_0)$. In many applications and, in particular, in the examples of Sects. 5 and 6, we can observe that the bound (3.19) leads to an order to convergence for $|f_h(0,0)|$ which is roughly the same as that of $\mathscr{E}_i^{(1)}(h,0)^2$. ■

Since $\tilde{f}_h(\xi_h^0, \alpha_h^0) = 0$ and $D^m \tilde{f}_h(\xi_h^0, \alpha_h^0) = D^m f_h^0$ for $m \geqq 1$, $(\xi_h^0, \alpha_h^0)$ is a critical point of the function $\tilde{f}_h$. We set:

$$A_h = \frac{\partial^2 f_h^0}{\partial \alpha^2}, \quad B_h = \frac{\partial^2 f_h^0}{\partial \xi \, \partial \alpha}, \quad C_h = \frac{\partial^2 f_h^0}{\partial \xi^2}. \tag{3.20}$$

Assuming again $p \geqq 3$, it follows from Lemmata 2 and 4 that

$$D^2 f_h(\xi_h^0, \alpha_h^0) \to D^2 f(0,0) \quad \text{as } h \to 0.$$

Hence, for $h \leqq h_0$ small enough, we have $B_h^2 - A_h C_h > 0$. Using Morse lemma, we obtain that in a neighborhood of the point $(\xi_h^0, \alpha_h^0)$, the solutions of the equation

$$\tilde{f}_h(\xi, \alpha) = 0 \tag{3.21}$$

consist in two $C^{p-2}$ branches which intersect transversally at $(\xi_h^0, \alpha_h^0)$.

More precisely, we have

**Lemma 6.** *Assume that $(\lambda_0, u_0)$ is a simple bifurcation point of $F$ and that the hypotheses* (3.4) *with $p \geqq 4$ and* (3.5) *hold. Then, there exist four positive constants $\xi_0$, $\alpha_0$, $t_0$, $h_0$ such that, for $h \leqq h_0$, the branches of solutions of* (3.21) *contained in $R(\xi_0, \alpha_0)$ may be parametrized in the form $\{(\tilde{\xi}_h^i(t), \tilde{\alpha}_h^i(t));\ |t| \leqq t_0\}$, $i = 1, 2$, where the $C^{p-2}$ functions $\tilde{\xi}_h^i$, $\tilde{\alpha}_h^i$ satisfy*

$$\tilde{\xi}_h^i(0) = \xi_h^0, \quad \tilde{\alpha}_h^i(0) = \alpha_h^0, \quad i = 1, 2. \tag{3.22}$$

*Moreover, for all integer $m$ with $0 \leqq m \leqq p-3$, there exists a constant $K_m$ such that:*

$$\begin{aligned} &\sup_{|t| \leqq t_0} \left\{ \left| \frac{d^m}{dt^m} (\tilde{\xi}_h^i(t) - \xi_i(t)) \right| + \left| \frac{d^m}{dt^m} (\tilde{\alpha}_h^i(t) - \alpha_i(t)) \right| \right\} \\ &\quad \leqq K_m \{ \mathscr{E}_1^{(1)}(h,0) + \mathscr{E}_2^{(1)}(h,0) + \sup_{|t| \leqq t_0} \mathscr{E}_i^{(m+1)}(h,t) \}. \end{aligned} \tag{3.23}$$

*Proof.* By choosing the constants $\xi_0, \alpha_0, h_0$ small enough, we may consider for $h \leqq h_0$ the function $\mathscr{F}_h$: $(t, \sigma, a) \in [-1, 1] \times [-\xi_0, \xi_0] \times [-\alpha_0, \alpha_0] \to \mathscr{F}_h(t, \sigma, a) \in \mathbb{R}^2$ defined by

$$\mathscr{F}_h(t, \sigma, a) = (t^{-2} \tilde{f}_h(\xi_h^0 + t\sigma, \alpha_h^0 + ta), \sigma^2 + a^2 - 1).$$

Since $(\xi_h^0, \alpha_h^0)$ is a critical point of the function $\tilde{f}_h$, we have

$$\mathscr{F}_h(t, \sigma, a) = \left( \int_0^1 (1-s) D^2 f_h(\xi_h^0 + st\sigma, \alpha_h^0 + sta) \cdot (\sigma, a)^2 \, ds,\ \sigma^2 + a^2 - 1 \right)$$

so that $\mathscr{F}_h$ is a $C^{p-2}$ function.

On the other hand, the function $\mathscr{F}$ being defined by (2.17), we have also

$$\mathscr{F}(t,\sigma,a)=\left(\int_0^1(1-s)D^2f(st\sigma,sta)\cdot(\sigma,a)^2\,ds,\ \sigma^2+a^2-1\right).$$

Hence, using Lemmata 2 and 4 together with (3.5), we obtain that the function $\mathscr{F}_h$ converges uniformly to $\mathscr{F}$ together with all its derivatives of order $\leqq p-3$ and $D^{p-2}\mathscr{F}_h$ is bounded as $h$ tends to zero. Moreover, we have observed in the proof of Lemma 1 that

$$\det(D_{\sigma,a)}\mathscr{F}(0,\sigma_i(0),a_i(0)))\neq 0,\qquad i=1,2.$$

Therefore, by choosing $t_0$ small enough, we have for all $|t|\leqq t_0$

$$\|D_{(\sigma,a)}\mathscr{F}(t,\sigma_i(t),a_i(t))^{-1}\|_{\mathscr{L}(\mathbb{R}^2;\mathbb{R}^2)}\leqq c.$$

Then, using again Lemma 1, we may apply [5, Theorem 1]: there exist two pairs of $C^{p-2}$ real functions $t\to(\sigma_h^i(t),\ a_h^i(t))$, $i=1,2$, defined for $|t|\leqq t_0$ such that

$$\mathscr{F}_h(\sigma_h^i(t),a_h^i(t))=0,\qquad i=1,2.$$

Furthermore, we have for all integer $m$ with $0\leqq m\leqq p-3$ and all $|t|\leqq t_0$

$$\begin{aligned}&\left|\frac{d^m}{dt^m}(\sigma_h^i(t)-\sigma_i(t))\right|+\left|\frac{d^m}{dt^m}(a_h^i(t)-a_i(t))\right|\\&\qquad\leqq c\sum_{l=0}^{m}\left\|\frac{d^l}{dt^l}(\mathscr{F}(t,\sigma_i(t),a_i(t))-\mathscr{F}_h(t,\sigma_i(t),a_i(t)))\right\|_{\mathbb{R}^2}.\end{aligned}\tag{3.24}$$

If we set:

$$\tilde{\xi}_h^i(t)=\xi_h^0+t\sigma_h^i(t),\qquad \tilde{\alpha}_h^i(t)=\alpha_h^0+ta_h^i(t),$$

we find that $\{(\tilde{\xi}_h^i(t),\ \tilde{\alpha}_h^i(t));\ |t|\leqq t_0\}$, $i=1,2$, are the branches of solutions of (3.21) in $R(\xi_0,\alpha_0)$ and we have (3.22). We now prove the estimate (3.23). Setting

$$e_h^{(m)}(t)=\left|\frac{d^m}{dt^m}(\tilde{\xi}_h^i(t)-\xi_i(t))\right|+\left|\frac{d^m}{dt^m}(\tilde{\alpha}_h^i(t)-\alpha_i(t))\right|,\qquad 0\leqq m\leqq p-3,$$

we get

$$\begin{aligned}e_h^{(m)}(t)\leqq{}&|\xi_h^0|+|\alpha_h^0|+\left|t\frac{d^m}{dt^m}(\sigma_h^i(t)-\sigma_i(t))\right|+\left|t\frac{d^m}{dt^m}(a_h^i(t)-a_i(t))\right|\\&+m\left|\frac{d^{m-1}}{dt^{m-1}}(\sigma_h^i(t)-\sigma_i(t))\right|+m\left|\frac{d^{m-1}}{dt^{m-1}}(a_h^i(t)-a_i(t))\right|.\end{aligned}$$

Then, using (3.24), we easily obtain

$$\begin{aligned}e_h^{(m)}(t)\leqq c\Bigg\{&|\xi_h^0|+|\alpha_h^0|+\sum_{l=0}^{m-1}\left\|\frac{d^l}{dt^l}(\mathscr{F}(t,\sigma_i(t),a_i(t))-\mathscr{F}_h(t,\sigma_i(t),a_i(t)))\right\|_{\mathbb{R}^2}\\&+\sum_{l=0}^{m}\left\|\frac{d^l}{dt^l}(t(\mathscr{F}(t,\sigma_i(t),a_i(t))-\mathscr{F}_h(t,\sigma_i(t),a_i(t))))\right\|_{\mathbb{R}^2}\Bigg\}.\end{aligned}$$

Let us next set

$$\varphi_h^i(t) = \tilde{f}_h(\xi_h^0 + t\sigma_i(t), \alpha_h^0 + t a_i(t)) - f(t\sigma_i(t), t a_i(t)), \qquad i = 1, 2.$$

We have for $0 \leqq l \leqq m-1$

$$\frac{d^l}{dt^l}(t^{-2}\varphi_h^i(t)) = \frac{d^l}{dt^l}\int_0^1 (1-s) D^2 \varphi_h^i(st)\, ds = \int_0^1 s^l (1-s) D^{l+2} \varphi_h^i(st)\, ds$$

and for $0 \leqq l \leqq m$

$$\frac{d^l}{dt^l}(t^{-1}\varphi_h^i(t)) = \frac{d^l}{dt^l}\int_0^1 D\varphi_h^i(st)\, ds = \int_0^1 s^l D^{l+1} \varphi_h^i(st)\, ds,$$

so that

$$\sup_{|t| \leqq t_0} e_h^m(t) \leqq c\left\{|\xi_h^0| + |\alpha_h^0| + \sum_{l=1}^{m+1} \sup_{|t| \leqq t_0} \left|\frac{d^l}{dt^l}\varphi_h^i(t)\right|\right\}. \tag{3.25}$$

By using Lemma 3 with $r = p-2$ and

$$\xi(t) = \xi_i(t) = t\sigma_i(t), \qquad \xi_h^*(t) = \xi_h^0 + \xi_i(t),$$
$$\alpha(t) = \alpha_i(t) = t a_i(t), \qquad \alpha_h^*(t) = a_h^0 + \alpha_i(t),$$

we get, for $0 \leqq l \leqq p-3$

$$\sup_{|t| \leqq t_0}\left|\frac{d^l}{dt^l}\varphi_h^i(t)\right| \leqq c\left\{|\xi_h^0| + |\alpha_h^0| + \sup_{|t| \leqq t_0} \sum_{k=0}^{l} \left\|(T - T_h)\frac{d^k}{dt^k} J(\xi_i(t), \alpha_i(t))\right\|_V\right\}.$$

Carrying this estimate in (3.25) and using Lemma 4, we obtain the desired estimate (3.23). ■

After having studied the branches of solutions of the perturbed Eq. (3.21), it remains to come back to the original equation (3.7). There is no extra work if $f_h = \tilde{f}_h$, i.e. if $f_h^0 = 0$; we shall give in Sect. 4 sufficient practical conditions for this situation to occur. Let us now consider the general case where

$$\mathscr{K}(h) = f_h^0 \neq 0. \tag{3.26}$$

Denote by $S_h$ the set of solutions of (3.7) contained in $R(\xi_0, \alpha_0)$ and by $\tilde{S}_h$ the set of solutions of (3.21) also contained in $R(\xi_0, \alpha_0)$. We want to estimate the distance $d(S_h, \tilde{S}_h)$ of the two sets $S_h$ and $\tilde{S}_h$, where in general the distance $d(A, B)$ of two closed sets $A$ and $B$ in a normed space is defined by

$$d(A, B) = \max\left(\sup_{x \in A} \inf_{y \in B} \|x - y\|, \sup_{y \in B} \inf_{x \in A} \|x - y\|\right).$$

**Lemma 7.** *Assume the hypotheses of Lemma 6 and also the condition* (3.26). *Then, the set $S_h$ is $C^{p-2}$-diffeomorphic to (a part of) a nondegenerate hyperbola*[2]. *Moreover, the distance between the sets $S_h$ and $\tilde{S}_h$ may be estimated by*

$$d(S_h, \tilde{S}_h) \leqq c\sqrt{|\mathscr{K}(h)|}. \tag{3.27}$$

[2] By this, we mean that $S_h$ is $C^{p-2}$-diffeomorphic to the set $\{(\xi, \alpha) \in \mathbb{R}^2;\ \xi\alpha = 1,\ \xi^2 + \alpha^2 = 2\}$

*Proof.* Again, since $(\xi_h^0, \alpha_h^0)$ is a critical point of the function $\tilde{f}_h$, we may apply Morse Lemma. In fact, using Lemma 2 and slightly generalizing the proof of [2, Theorem 6.5.4, p. 355], one can easily show that, for $\xi_0$, $\alpha_0$, $h_0$ small enough, there exists for all $h \leqq h_0$ a $C^{p-2}$ mapping $(\xi, \alpha) \in R(\xi_0, \alpha_0) \to (\bar{\xi}, \bar{\alpha}) = \varphi_h(\xi, \alpha) \in \mathbb{R}^2$ such that

$$\varphi_h(\xi_h^0, \alpha_h^0) = (0, 0), \tag{3.28}$$

$$D\varphi_h(\xi_h^0, \alpha_h^0) = I \quad \text{(identity)}, \tag{3.29}$$

$$\tilde{f}_h(\xi, \alpha) = \tfrac{1}{2} D^2 f_h^0 \cdot (\bar{\xi}, \bar{\alpha})^2. \tag{3.30}$$

In addition, we have for all $(\xi, \alpha) \in R(\xi_0, \alpha_0)$ and all integer $m$ with $0 \leqq m \leqq p-2$

$$\|D^m \varphi_h(\xi, \alpha)\|_{\mathscr{L}_m(\mathbb{R}^2; \mathbb{R}^2)} \leqq c. \tag{3.31}$$

Since $B_h^2 - A_h C_h$ is $> 0$, the function $\tilde{f}_h$ vanishes in the $(\bar{\xi}, \bar{\alpha})$-plane along two straight lines which intersect at the origin. Using (3.17) and (3.30), we have

$$f_h(\xi, \alpha) = \mathscr{K}(h) + \tfrac{1}{2} D^2 f_h^0 \cdot (\bar{\xi}, \bar{\alpha})^2,$$

so that the function $f_h$ vanishes in the $(\bar{\xi}, \bar{\alpha})$-plane along a nondegenerate hyperbola ($\mathscr{K}(h) \neq 0$). Now, the distance of the images $\bar{S}_h$ and $\bar{\tilde{S}}_h$ of $S_h$ and $\tilde{S}_h$ respectively in the $(\bar{\xi}, \bar{\alpha})$-plane coincide with the distance of the origin to this hyperbola and it is an easy matter to check that

$$d(\bar{S}_h, \bar{\tilde{S}}_h) \leqq \{\mathscr{K}(h) \, \|(D^2 f_h^0)^{-1}\|_{\mathscr{L}(\mathbb{R}^2; \mathbb{R}^2)}\}^{\frac{1}{2}} \leqq c\sqrt{|\mathscr{K}(h)|}. \tag{3.32}$$

Next, using (3.29) and (3.31) with $m = 2$ and assuming again that the constants $\xi_0$, $\alpha_0$, $h_0$ are small enough, we get for all $(\eta, \beta) \in R(\xi_0, \alpha_0)$ and all $h \leqq h_0$

$$\|D\varphi_h(\eta, \beta)^{-1}\|_{\mathscr{L}(\mathbb{R}^2; \mathbb{R}^2)} \leqq 2 \quad \text{(say!)},$$

so that we have for all $(\xi, \alpha)$, $(\xi^*, \alpha^*) \in R(\xi_0, \alpha_0)$

$$|\xi - \xi^*| + |\alpha - \alpha^*| \leqq 2 \, \|\varphi_h(\xi, \alpha) - \varphi_h(\xi^*, \alpha^*)\|_{\mathbb{R}^2}.$$

Hence we obtain

$$d(S_h, \tilde{S}_h) \leqq 2 d(\bar{S}_h, \bar{\tilde{S}}_h),$$

Together with (3.32), this proves the estimate (3.27). ∎

We can now state the main result of this section.

**Theorem 2.** *Assume that $(\lambda_0, u_0)$ is a simple bifurcation point of $F$ and in addition that the hypotheses* (3.4) *with $p \geqq 4$ and* (3.5) *hold. Then, there exist a neighborhood $\mathcal{O}$ of the point, $(\lambda_0, u_0)$ in $\mathbb{R} \times V$ and a positive constant $h_0$ such that, for $h \leqq h_0$, the set $\mathscr{S}_h$ of the solutions of* (3.2) *contained in $\mathcal{O}$ consists of two $C^{p-2}$ branches.*

*If these two branches intersect at a point $(\lambda_h^0, u_h^0) \in \mathcal{O}$, they can be parametrized in the form $\{(\lambda_h^i(t), u_h^i(t));\ |t| \leqq t_0\}$, $i=1,2$, where on the one hand*

$$\lambda_h^i(0)=\lambda_h^0, \qquad u_h^i(0)=u_h^0, \qquad i=1,2, \tag{3.33}$$

*and on the other hand we have for all integer $m$ with $0 \leqq m \leqq p-3$:*

$$\begin{aligned} &\sup_{|t| \leqq t_0} \left\{ \left| \frac{d^m}{dt^m}(\lambda_h^i(t)-\lambda_i(t)) \right| + \left\| \frac{d^m}{dt^m}(u_h^i(t)-u_i(t)) \right\|_V \right\} \\ &\leqq K_m \{\mathscr{E}_1^{(1)}(h,0)+\mathscr{E}_2^{(1)}(h,0)+\sup_{|t| \leqq t_0} \mathscr{E}_i^{(m+1)}(h,t)\} \qquad i=1,2. \end{aligned} \tag{3.34}$$

*Otherwise, the set $\mathscr{S}_h$ is $C^{p-2}$ diffeomorphic to (a part of) a nondegenerate hyperbola and the distance between $\mathscr{S}_h$ and the set $\mathscr{S}$ of the solutions of* (2.3) *contained in $\mathcal{O}$ can be estimated by*

$$d(\mathscr{S}_h, \mathscr{S}) \leqq c \left\{ \sqrt{|\mathscr{K}(h)|} + \sup_{|t| \leqq t_0} \left( \sum_{i=1}^{2} \mathscr{E}_i^{(1)}(h,t) \right) \right\}. \tag{3.35}$$

*Proof.* Let us define the two curves $\{(\tilde{\lambda}_h^i(t), \tilde{u}_h^i(t));\ |t| \leqq t_0\}$, $i=1,2$ by

$$\begin{aligned} &\tilde{\lambda}_h^i(t)=\lambda_0+\tilde{\xi}_h^i(t), \\ &\tilde{u}_h^i(t)=u_0+\tilde{\alpha}_h^i(t)\,\varphi_0+v_h(\tilde{\xi}_h^i(t), \tilde{\alpha}_h^i(t)), \end{aligned} \tag{3.36}$$

where the functions $\tilde{\xi}_h^i$, $\tilde{\alpha}_h^i$ are given by Lemma 6. Using (3.21), we have

$$\tilde{\lambda}_h^i(0)=\lambda_h^0=\lambda_0+\xi_h^0, \qquad \tilde{u}_h^i(0)=u_h^0=u_0+\alpha_h^0\,\varphi_0+v_h(\xi_h^0, \alpha_h^0), \qquad i=1,2.$$

Moreover, it follows from (3.23) and [5, Lemma 3] that we get for $0 \leqq m \leqq p-3$

$$\begin{aligned} &\left| \frac{d^m}{dt^m}(\tilde{\lambda}_h^i(t)-\lambda_i(t)) \right| + \left\| \frac{d^m}{dt^m}(\tilde{u}_h^i(t)-u_i(t)) \right\|_V \\ &\leqq K_m \{\mathscr{E}_1^{(1)}(h,0)+\mathscr{E}_2^{(1)}(h,0)+\sup_{|t| \leqq t_0} \mathscr{E}_i^{(m+1)}(h,t)\}. \end{aligned} \tag{3.37}$$

If $\mathscr{K}(h)=f_h^0=0$, the curves $\{(\tilde{\lambda}_h^i(t), \tilde{u}_h^i(t));\ |t| \leqq t_0\}$ are the branches of solutions of problem (3.2) in a neighborhood of $(\lambda_0, u_0)$; they intersect at the point $(\lambda_h^0, u_h^0)$ and (3.34) holds.

If $\mathscr{K}(h) \neq 0$, it follows from Lemma 7 that $\mathscr{S}_h$ is $C^{p-2}$-diffeomorphic to an hyperbola and thus consists of two $C^{p-2}$ branches which do not intersect in a neighborhood of $(\lambda_0, u_0)$. Furthermore, we have by (3.37)

$$d(\mathscr{S}, \tilde{\mathscr{S}}_h) \leqq c \sup_{|t| \leqq t_0} \left\{ \sum_{i=1}^{2} \mathscr{E}_i^{(1)}(h,t) \right\},$$

where $\tilde{\mathscr{S}}_h$ is the union of the two curves (3.36). Hence, (3.27) implies the estimate (3.35). ∎

Hence, in the general case $\mathscr{K}(h) \neq 0$, the approximate problem (3.2) does not bifurcate; we obtain a numerical imperfect bifurcation. However, when $\mathscr{K}(h) = 0$, the approximate problem still bifurcates. We summarize the conclusions of Theorem 2 in the following diagram

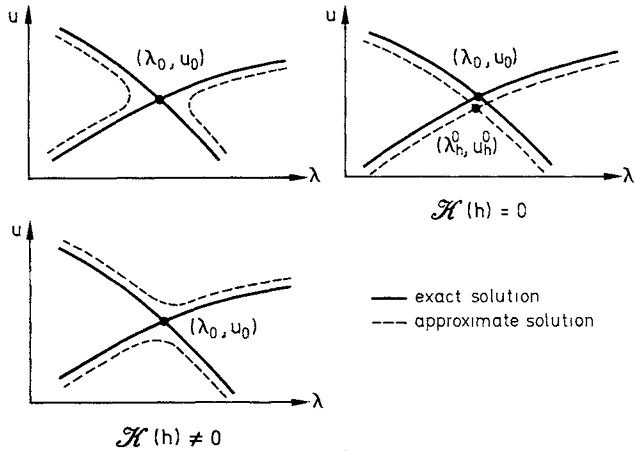



**Fig. 2**

*Remark 3.* So far, we have restricted ourselves to the case where the condition (2.14) holds. Let us notice that the results of this section can be easily extended to the case $B_0^2 - A_0 C_0 < 0$. In fact, when this condition holds, the set $\mathscr{S}$ of solutions of (2.3) consists of an *isolated point* $(\lambda_0, u_0)$. Now, we remark that, in the proof of Lemma 4, we only used $B_0^2 - A_0 C_0 \neq 0$ so that the conclusions of this Lemma are still valid. Hence, applying Morse lemma to the function $\tilde{f}_h$, we get the following result: If $\mathscr{K}(h) = 0$, the set $\mathscr{S}_h$ of solutions of (3.2) consists of the isolated point $(\lambda_h^0, u_h^0)$ where

$$\lambda_h^0 = \lambda_0 + \xi_h^0, \qquad u_h^0 = u_0 + \alpha_h^0 \varphi_0 + v_h(\xi_h^0, \alpha_h^0).$$

Otherwise, the set $\mathscr{S}_h$ is either empty or $C^{p-2}$-diffeomorphic to an ellipse according to the sign of $\mathscr{K}(h)$. In the later case, we obtain the formation of an "isola" (cf. [13]) whose diameter is bounded by $\sqrt{\mathscr{K}(h)}$. ■

## 4. Bifurcating Approximate Problems

We want to show in this section that the approximate problem (3.2) does indeed bifurcate in a number of interesting practical cases: bifurcation from the trivial branch, symmetry-breaking bifurcation.

We begin with the first case. Thus, we assume that

$$G(\lambda, 0) = 0 \qquad \text{for all } \lambda \in \mathbb{R}, \tag{4.1}$$

so that $(\lambda, 0)$ is a solution of (2.3) and also a solution of (3.2); the trivial branch is then a branch of solutions of both problems (2.3) and (3.2).

**Theorem 3.** *Assume that the condition* (4.1) *holds and that* $(\lambda_0, 0)$ *is a simple bifurcation point of* $F$. *Then, under the hypotheses* (3.4) *and* (3.5) *with* $p \geqq 3$, *there exists a neighborhood* $\Lambda$ *of* $\lambda_0$ *in* $\mathbb{R}$ *and, for* $h \leqq h_0$ *small enough, a unique point* $\lambda_h^0 \in \Lambda$ *such that* $(\lambda_h^0, 0)$ *is a bifurcation point of* $F_h$.

*Proof.* It follows from (4.1) that we have for all $\xi \in \mathbb{R}$

$$\frac{\partial^m f}{\partial \xi^m}(\xi, 0) = 0, \qquad m = 0, 1, 2, \dots.$$

Since, by (2.14), $\dfrac{\partial^2 f}{\partial \xi \, \partial \alpha}(0,0) \neq 0$, the function $\xi \to \dfrac{\partial f}{\partial \alpha}(\xi, 0)$ changes its sign at the point $\xi = 0$. Hence, using Lemma 2 with $m = 1, 2$ and (3.5), we find that there exist a neighborhood $\Xi$ of the origin in $\mathbb{R}$ and, for $h \leqq h_0$ small enough, a unique point $\xi_h^0 \in \Xi$ such that

$$\frac{\partial f_h}{\partial \alpha}(\xi_h^0, 0) = 0.$$

Since for all $\xi \in \mathbb{R}$

$$\frac{\partial^m f}{\partial \xi^m}(\xi, 0) = 0, \qquad m = 0, 1, 2, \dots,$$

we have that $(\xi_h^0, 0)$ is a bifurcation point of $f_h$. Hence, the conclusion of the theorem follows with $\lambda_h^0 = \lambda_0 + \xi_h^0$ and $\Lambda = \lambda_0 + \Xi$. ∎

Let us notice that, in this case, the first inequality (3.14) becomes

$$|\lambda_h^0 - \lambda_0| \leqq c \left| \frac{\partial f_h}{\partial \alpha}(0, 0) \right|. \tag{4.2}$$

We now consider the case of a symmetry-breaking bifurcation. Let us briefly describe such a situation (for more details and the relationship with group theory, we refer to [8, 20]). We assume that there exists an isometry $H \in \mathscr{L}(V; V)$ such that for all $\lambda \in \mathbb{R}$ and all $u \in V$

$$H F(\lambda, u) = F(\lambda, H u), \tag{4.3}$$

and moreover that $(\lambda_0, u_0)$ is a simple bifurcation point of $F$ such that

$$H u_0 = u_0. \tag{4.4}$$

Differentiating (4.3) with respect to $u$ at the point $(\lambda_0, u_0)$ and using (4.4) gives for all $v \in V$

$$H D_u F^0 \cdot v = D_u F^0 \cdot H v, \tag{4.5}$$

so that

$$H \varphi_0 = \varepsilon \varphi_0, \qquad \varepsilon = \varepsilon(H) = \pm 1. \tag{4.6}$$

Let us now state a classical result whose proof is given for reader's convenience.

**Lemma 8.** *Assume the hypotheses* (4.3), (4.4) *and* (4.6). *Then, we get:*

$$f(\xi, \varepsilon\alpha) = \varepsilon f(\xi, \alpha). \tag{4.7}$$

*Proof.* We first observe that

$$H^* \varphi_0^* = \varepsilon \varphi_0^*. \tag{4.8}$$

In fact, we get from (4.5)

$$H^*(D_u F^0)^* \cdot g = (D_u F^0)^* \cdot H^* g, \qquad \forall g \in V',$$

so that $H^* \varphi_0^* = \kappa \varphi_0^*$. But, on the other hand, using (4.6), we have

$$\kappa = \langle \varphi_0, \kappa \varphi_0^* \rangle = \langle \varphi_0, H^* \varphi_0^* \rangle = \langle H \varphi_0, \varphi_0^* \rangle = \varepsilon \langle \varphi_0, \varphi_0^* \rangle = \varepsilon$$

and (4.8) holds.

Next, using (4.6) and (4.8), we may write for all $v \in V$

$$HQv = Hv - \langle v, \varphi_0^* \rangle H \varphi_0 = Hv - \langle v, H^* \varphi_0^* \rangle \varphi_0,$$

and therefore

$$HQv = QHv. \tag{4.9}$$

Then it follows from (4.3), (4.4), (4.6) and (4.8) that

$$\begin{aligned} HQF(\lambda_0 + \xi, u_0 + \alpha\varphi_0 + v(\xi, \alpha)) &= QHF(\lambda_0 + \xi, u_0 + \alpha\varphi_0 + v(\xi, \alpha)) \\ &= QF(\lambda_0 + \xi, H(u_0 + \alpha\varphi_0 + v(\xi, \alpha))) = QF(\lambda_0 + \xi, u_0 + \varepsilon\alpha\varphi_0 + Hv(\xi, \alpha)), \end{aligned}$$

so that

$$QF(\lambda_0 + \xi, u_0 + \varepsilon\alpha\varphi_0 + Hv(\xi, \alpha)) = 0.$$

Moreover, by (4.9) we have $Hv(\xi, \alpha) \in V_2$. Hence, by the uniqueness of the function $(\xi, \alpha) \to v(\xi, \alpha)$ defined by (2.9), we obtain

$$Hv(\xi, \alpha) = v(\xi, \varepsilon\alpha). \tag{4.10}$$

Finally, using (4.3), (4.4), (4.6), (4.8) and (4.10), we get

$$\begin{aligned} f(\xi, \alpha) &= \langle F(\lambda_0 + \xi, u_0 + \alpha\varphi_0 + v(\xi, \alpha), \varphi_0^* \rangle \\ &= \langle HF(\lambda_0 + \xi, u_0 + \alpha\varphi_0 + v(\xi, \alpha), H^* \varphi_0^* \rangle \\ &= \varepsilon \langle F(\lambda_0 + \xi, u_0 + \varepsilon\alpha\varphi_0 + v(\xi, \varepsilon\alpha), \varphi_0^* \rangle = \varepsilon f(\xi, \varepsilon\alpha)), \end{aligned}$$

which proves the desired property. ■

Note that Lemma 8 provides a nontrivial property for the bifurcation equation only when in (4.6) $\varepsilon = -1$: this is precisely a symmetry-breaking bifurcation.

Let us consider next the approximate problem; we assume again that for all $\lambda \in \mathbb{R}$ and all $u \in V$

$$HF_h(\lambda, u) = F_h(\lambda, Hu). \tag{4.11}$$

We have

**Theorem 4.** *Assume that for some isometry $H \in \mathscr{L}(V; V)$ the hypothesis* (4.3) *holds and that $(\lambda_0, u_0)$ is a simple bifurcation point of $F$ which satisfies the properties* (4.4) *and* (4.6) *with $\varepsilon = -1$ (symmetry-breaking). Then, under the assumptions* (3.4) *with $p \geqq 3$,* (3.5) *and* (4.11), *there exists a neighborhood $\mathcal{O}$ of $(\lambda_0, u_0)$ in $\mathbb{R} \times V$ and, for $h \leqq h_0$ small enough, a unique point $(\lambda_h^0, u_h^0) \in \mathcal{O}$ which is a bifurcation point of $F_h$.*

*Proof.* Using (4.4) and (4.6) with $\varepsilon = -1$ and (4.11), we show exactly as in Lemma 8 that

$$f_h(\xi, -\alpha) = -f_h(\xi, \alpha).$$

Applying Lemma 8, we thus obtain

$$f(\xi, 0) = f_h(\xi, 0) = 0, \qquad |\xi| \leqq \xi_0, \tag{4.12}$$

so that we may parallel the proof of Theorem 3: we get a unique bifurcation point $(\xi_h^0, 0)$ of $f_h$ in the neighborhood of the origin.

Setting

$$\lambda_h^0 = \lambda_0 + \xi_h^0, \qquad u_h^0 = u_0 + v_h(\xi_h^0, 0),$$

$(\lambda_h^0, u_h^0)$ is the desired bifurcation point of $F_h$. ■

*Remark 4.* Assume the hypotheses of Theorem 4. Then, by (4.12) we may parametrize with $\xi = \lambda - \lambda_0$ a branch of solutions of (2.3) and the corresponding approximating branch of solutions of (3.2); we obtain:

$$u_1(\xi) = u_0 + v(\xi, 0), \qquad u_h^1(\xi) = u_0 + v_h(\xi, 0).$$

Let $\tilde{H}$ be any other isometry $\in \mathscr{L}(V, V)$ such that the analogue of properties (4.3) and (4.4) hold and $\tilde{\varepsilon} = \varepsilon(\tilde{H}) = 1$. Then, using (4.4) and (4.10) for both isometries $H$ and $\tilde{H}$, we have

$$H u_1(\xi) = \tilde{H} u_1(\xi) = u_1(\xi), \qquad H u_h^1(\xi) = \tilde{H} u_h^1(\xi) = u_h^1(\xi),$$

so that these branches retain the symmetry properties of the problem.

On the other hand, concerning the other branches $\{(\lambda_2(t), u_2(t)); |t| \leqq t_0\}$ and $\{(\lambda_h^2(t), u_h^2(t); |t| \leqq t_0\}$, we get by (4.4), (4.6) and (4.10)

$$\tilde{H} u_2(t) = u_2(t), \qquad \tilde{H} u_h^2(t) = u_h^2(t),$$

while for $t \neq 0$

$$H u_2(t) \neq u_2(t), \qquad H u_h^2(t) \neq u_h^2(t).$$

Hence these branches loose the symmetry properties corresponding to the isometry $H$. ■

*Remark 5.* Assume again the hypotheses of Theorem 4. Then, we may write

$$\begin{aligned} A_0 &= \langle D_{uu}^2 F^0 \cdot (\varphi_0)^2, \varphi_0^* \rangle = -\langle D_{uu}^2 F^0 \cdot (\varphi_0)^2, H^* \varphi_0^* \rangle \\ &= -\langle H D_{uu}^2 F^0 \cdot (\varphi_0)^2, \varphi_0^* \rangle = -\langle D_{uu}^2 F^0 \cdot (H \varphi_0)^2, \varphi_0^* \rangle = -A_0, \end{aligned}$$

so that we have $A_0=0$. If in addition $D_0$ is $\neq 0$, we obtain a cusp bifurcation (cf. Remark 1). Now, assuming that (4.11) holds and using Remark 4, one can easily prove that $(\lambda_h^0, u_h^0)$ is indeed a symmetry-breaking bifurcation point of $F_h$. Therefore, when $D_0$ is $\neq 0$, the approximate problem has for $h \leqq h_0$ small enough a cusp-bifurcation. ■

## 5. Application I: Galerkin Approximation of Nonlinear Problems

As in [5], let us apply the above results to a class of conforming approximations of variationally posed nonlinear problems. Let $V$ and $H$ be two (real) Hilbert spaces with scalar products $((.,.))$, $(.,.)$ and norms $\|.\|$, $|.|$ respectively. We suppose that $V \subset H$ with densely continuous imbedding. Identifying $H$ with its dual space $H'$, we have $V \subset H \subset V'$ and the scalar product $(.,.)$ can also represent the duality pairing between the spaces $V$ and $V'$. Let moreover $W$ be a reflexive Banach space such that $H \subset W \subset V'$ with continuous imbeddings. We assume that the canonical injection of $W$ into $V'$ is *compact*.

Next, we are given a bilinear continuous form $a(.,.)$: $V \times V \to \mathbb{R}$ which is $V$-elliptic in the sense that

$$a(v,v) \geqq \gamma \|v\|^2, \quad \forall v \in V, \quad \gamma > 0, \tag{5.1}$$

and a $C^p$ mapping $(p \geqq 4)$ $G: \mathbb{R} \times V \to W$. Then, we consider the nonlinear problem: Find $(\lambda, u) \in \mathbb{R} \times V$ such that

$$a(u,v) + (G(\lambda, u), v) = 0, \quad \forall v \in V. \tag{5.2}$$

We now introduce a sequence $\{V_h\}$ of finite-dimensional subspaces of $V$ and we consider the approximate problem: Find $(\lambda, u_h) \in \mathbb{R} \times V_h$ such that

$$a(u_h, v) + (G(\lambda, u_h), v) = 0, \quad \forall v \in V_h. \tag{5.3}$$

In order to put problems (5.2) and (5.3) into the framework of Sects. 2 and 3, we define the operators $T \in \mathscr{L}(V'; V)$ and $T_h \in \mathscr{L}(V'; V_h)$ by

$$a(Tf, v) = (f, v), \quad \forall v \in V, \ \forall f \in V'$$

and

$$a(T_h f, v) = (f, v), \quad \forall v \in V_h, \ \forall f \in V'.$$

If we introduce the elliptic projection operator $\Pi_h \in \mathscr{L}(V; V_h)$ defined by

$$a(\Pi_h u - u, v) = 0, \quad \forall v \in V_h, \ \forall u \in V,$$

we have $T_h = \Pi_h T$. Then, assuming that for all $v \in V$

$$\lim_{h \to 0} \inf_{v_h \in V_h} \|v - v_h\| = 0, \tag{5.4}$$

we get

$$\lim_{h \to 0} \|v - \Pi_h v\| = 0, \quad \forall v \in V,$$

and, by the compactness of the operator $T \in \mathscr{L}(W; V)$, we obtain

$$\lim_{h \to 0} \|T - T_h\|_{\mathscr{L}(W; V)} = \lim_{h \to 0} \|(I - \Pi_h) T\|_{\mathscr{L}(W; V)} = 0.$$

Clearly, problems (5.2) and (5.3) are now equivalent to

$$F(\lambda, u) \equiv u + TG(\lambda, u) = 0 \tag{5.5}$$

and

$$F_h(\lambda, u) = u + T_h G(\lambda, u) = 0 \tag{5.6}$$

respectively.

Now, it is an easy matter to check that the vectors $\varphi_0$ and $\varphi_0^*$ which appear in (2.4) and (2.5) are given by

$$a(\varphi_0, v) + (D_u G^0 \cdot \varphi_0, v) = 0, \qquad \forall v \in V \tag{5.7}$$

and

$$\begin{aligned} & a(v, \psi_0^*) + (D_u G^0 \cdot v, \psi_0^*) = 0, \qquad \forall v \in V, \ \psi_0^* \in V, \\ & \varphi_0^* = (T^*)^{-1} \psi_0^*, \end{aligned} \tag{5.8}$$

where $T^* \in \mathscr{L}(V'; V)$ is the adjoint operator of $T$, i.e.

$$a(v, T^* f) = (v, f), \qquad \forall v \in V, \ \forall f \in V'.$$

Hence, we have

$$V_2 = \{v \in V; a(v, \psi_0^*) = 0\}$$

and

$$\begin{aligned} A_0 &= (D_{uu}^2 G^0 \cdot (\varphi_0)^2, \psi_0^*), \\ B_0 &= (D_{\lambda u}^2 G^0 \cdot \varphi_0 + D_{uu}^2 G^0 \cdot (\varphi_0, -L D_\lambda G^0), \psi_0^*), \\ C_0 &= (D_{\lambda\lambda}^2 G^0 + 2 D_{\lambda u}^2 G^0 \cdot (-L D_\lambda G^0) + D_{uu}^2 G^0 \cdot (-L D_\lambda G^0)^2, \psi_0^*), \end{aligned}$$

where

$$L = [(I + T D_u G^0)|_{V_2}]^{-1}.$$

Next, we assume that $(\lambda_0, u_0) \in \mathbb{R} \times V$ is a simple bifurcation point of $F$. Note that the condition (2.2)(iii) (or (2.6)) becomes

$$(D_\lambda G^0, \psi_0^*) = 0.$$

We may apply the results of Sect. 2: we obtain that, in a neighborhood of $(\lambda_0, u_0)$, the solutions of problem (5.2) consist of two $C^{p-2}$ branches $\{(\lambda_i(t), u_i(t)); |t| \leqq t_0\}$, $i = 1, 2$, which intersect at $(\lambda_0, u_0)$. Moreover, Theorem 2 gives here:

**Theorem 5.** *Assume that $G$ is a $C^p$ mapping ($p \geqq 4$) and the mapping $D^p G$ is bounded on all bounded subsets of $\mathbb{R} \times V$. Assume in addition that $(\lambda_0, u_0)$ is a simple bifurcation point of* (5.5) *and that the approximation property* (5.4) *holds. Then, there exist a neighborhood $\mathcal{O}$ of the point $(\lambda_0, u_0)$ in $\mathbb{R} \times V$ and a positive constant $h_0$ such that, for $h \leqq h_0$, the set $\mathscr{S}_h$ of solutions of* (5.3) *contained in $\mathcal{O}$ consist of two $C^{p-2}$ branches.*

*If these two branches intersect at a point $(\lambda_h^0, u_h^0) \in \mathcal{O}$, they can be parametrized in the form $\{(\lambda_h^i(t), u_h^i(t)), |t| \leqq t_0\}$, $i=1,2$, so that we have for all integer $m$ with $0 \leqq m \leqq p-3$:*

$$\begin{aligned}
&\sup_{|t|\leqq t_0} \left\{ \left| \frac{d^m}{dt^m} (\lambda_h^i(t) - \lambda_i(t)) \right| + \left\| \frac{d^m}{dt^m} (u_h^i(t) - u_i(t)) \right\| \right\} \\
&\quad \leqq K_m \left\{ \sum_{j=1}^{2} \left( \inf_{v_h \in V_h} \left\| \frac{du_j}{dt}(0) - v_h \right\| \right) \right. \\
&\qquad \left. + \sup_{|t|\leqq t_0} \sum_{l=0}^{m+1} \left( \inf_{v_h \in V_h} \left\| \frac{d^l}{dt^l} u_i(t) - v_h \right\| \right) \right\}, \quad i=1,2.
\end{aligned} \tag{5.9}$$

*Otherwise, the set $\mathcal{S}_h$ is $C^{p-2}$ diffeomorphic to a nondegenerate hyperbola and the distance between $\mathcal{S}_h$ and the set $\mathcal{S}$ of solutions of (5.2) contained in $\mathcal{O}$ is bounded by*

$$\begin{aligned}
d(\mathcal{S}_h, \mathcal{S}) \leqq c &\left\{ \sup_{|t|\leqq t_0} \left( \sum_{i=1}^{2} \sum_{l=0}^{1} \inf_{v_h \in V_h} \left\| \frac{d^l}{dt^l} u_i(t) - v_h \right\| \right) \right. \\
&\left. + \left( \inf_{\psi_h \in V_h} \|\psi_0^* - \psi_h\| \right)^{\frac{1}{2}} \left( \inf_{v_h \in V_h} \|u_0 - v_h\| \right)^{\frac{1}{2}} \right\}.
\end{aligned} \tag{5.10}$$

*Proof.* Let us check that the estimates (5.9) and (5.10) are consequences of (3.34) and (3.35) respectively. On the one hand, we observe as in the proof of [5, Theorem 6] that

$$\mathcal{E}_i^m(h,t) \leqq c_1 \sum_{l=0}^{m} \left( \inf_{v_h \in V_h} \left\| \frac{d^l}{dt^l} u_i(t) - v_h \right\| \right). \tag{5.11}$$

Hence (5.9) follows from (3.34) and (5.11).

On the other hand, for evaluating $\mathcal{K}(h)$, we use the first inequality (3.18) and the inequality (3.19). In fact, we have shown in the proof of [5, Theorem 7] that

$$|\langle (T - T_h) G^0, \varphi_0^* \rangle| \leqq c_2 \left( \inf_{v_h \in V_h} \|u_0 - v_h\| \right) \left( \inf_{\psi_h \in V_h} \|\psi_0^* - \psi_h\| \right)$$

and

$$\|[(T - T_h) D_u G^0]^* \varphi_0^*\|_{V'} \leqq c_3 \inf_{\psi_h \in V_h} \|\psi_0^* - \psi_h\|.$$

Since

$$\|(T - T_h) G^0\| = \mathcal{E}^0(h, 0) \leqq c_4 \inf_{v_h \in V_h} \|u_0 - v_h\|,$$

we obtain

$$\mathcal{K}(h) \leqq c_5 \left\{ \left( \inf_{v_h \in V_h} \|u_0 - v_h\| \right) \left( \inf_{\psi_h \in V_h} \|\psi_0^* - \psi_h\| \right) + \sum_{i=1}^{2} \mathcal{E}_i^{(1)}(h, 0)^2 \right\}, \tag{5.12}$$

and (5.10) follows from (3.35), (5.11) and (5.12). ■

We now turn to the case of a bifurcation from the trivial branch. Then, Theorem 3 applies and the approximate problem (5.3) (or (5.6)) has a unique bifurcation point $(\lambda_h^0, 0)$ in a neighborhood $\mathcal{O}$ of $(\lambda_0, 0)$. In this case, we can derive an improved estimate of $|\lambda_h^0 - \lambda_0|$.

**Theorem 6.** *Assume the hypotheses of Theorem 5. Assume in addition that the hypotheses of Theorem 3 hold. Then, we have:*

$$|\lambda_h^0-\lambda_0|\leqq c\left\{\inf_{\varphi_h\in V_h}\|\varphi_0-\varphi_h\|+\sum_{i=1}^{2}\sum_{l=0}^{1}\inf_{v_h\in V_h}\left\|\frac{d^l u_i}{dt^l}(0)-v_h\right\|\right\}$$
$$\times(\inf_{\psi_h\in V_h}\|\psi_0^*-\psi_h\|). \tag{5.13}$$

*Proof.* In this case, we have $f_h(\xi,0)=0$ so that the estimate (3.14) gives

$$|\lambda_h^0-\lambda_0|=|\xi_h^0|\leqq c_1\left|\frac{\partial f_h}{\partial\alpha}(0,0)\right|.$$

Since $v_h(0,0)=0$, we may write for all $\psi_h\in V_h$

$$\frac{\partial f_h}{\partial\alpha}(0,0)=\left((T_h-T)D_uG^0\cdot\left(\varphi_0+\frac{\partial v_h}{\partial\alpha}(0,0)\right),\varphi_0^*\right)$$
$$=a\left((T_h-T)D_uG^0\cdot\left(\varphi_0+\frac{\partial v_h}{\partial\alpha}(0,0)\right),\psi_0^*-\psi_h\right).$$

Hence, we obtain

$$\left|\frac{\partial f_h}{\partial\alpha}(0,0)\right|\leqq c_2\left(\|T-T_h)D_uG^0\cdot\varphi_0\|\right.$$
$$\left.+\left\|(T-T_h)D_uG^0\cdot\frac{\partial v_h}{\partial\alpha}(0,0)\right\|\right)\cdot\inf_{\psi_h\in V_h}\|\psi_0^*-\psi_h\|.$$

But

$$\|(T-T_h)D_uG^0\cdot\varphi_0\|=\|(I-\Pi_h)\varphi_0\|\leqq c_3\inf_{\varphi_h\in V_h}\|\varphi_0-\varphi_h\|.$$

Moreover, we have by [5, Theorem 2]

$$\left\|\frac{\partial v_h}{\partial\alpha}(0,0)\right\|=\left\|\frac{\partial v_h}{\partial\alpha}(0,0)-\frac{\partial v}{\partial\alpha}(0,0)\right\|\leqq c_4\sum_{i=1}^{2}\mathscr{E}_i^1(h,0).$$

Thus, we get

$$\left|\frac{\partial f_h}{\partial\alpha}(0,0)\right|\leqq c_5\left\{\inf_{\varphi_h\in V_h}\|\varphi_0-\varphi_h\|+\sum_{i=1}^{2}\mathscr{E}_i^1(h,0)\right\}(\inf_{\psi_h\in V_h}\|\psi_0^*-\psi_h\|),$$

and by (5.11) the desired inequality (5.13) follows at once. ■

## 6. Application II: A Mixed Finite Element Method for the Kármán Equations

Consider now the von Kármán equations for a clamped plate: Find $\boldsymbol{\psi}=(\psi^1,\psi^2)$ such that

$$\begin{aligned}
&\Delta^2\psi^1=-\tfrac{1}{2}[\psi^2,\psi^2] && \text{in } \Omega,\\
&\Delta^2\psi^2=[\psi^1+\lambda\bar{\psi},\psi^2]+f && \text{in } \Omega,\\
&\psi^1=\frac{\partial\psi^1}{\partial n}=\psi^2=\frac{\partial\psi^2}{\partial n}=0 && \text{on } \Gamma,
\end{aligned} \tag{6.1}$$

where $\Omega$ is a bounded *plane* domain with boundary $\Gamma$, the functions $f$ and $\bar{\psi}$ are given in $H^{-1}(\Omega)$ and $H^3(\Omega)$ respectively and

$$[\varphi,\psi]=\frac{\partial^2\varphi}{\partial x_1^2}\frac{\partial^2\psi}{\partial x_2^2}+\frac{\partial^2\varphi}{\partial x_2^2}\frac{\partial^2\psi}{\partial x_1^2}-2\frac{\partial^2\varphi}{\partial x_1\partial x_2}\frac{\partial^2\psi}{\partial x_1\partial x_2}. \tag{6.2}$$

Then, as it is well known, for any $\lambda\in\mathbb{R}$ problem (6.1) has at least one solution $\boldsymbol{\psi}\in H_0^2(\Omega)^2$.

Now we assume that $(\lambda_0,\boldsymbol{\psi}_0)$ is a simple bifurcation point of (6.1) and we want to study the behaviour of the solutions of the "Hellan-Herrmann-Johnson" mixed finite element approximation scheme (cf. for instance [10, 11, 12]) in a neighborhood of $(\lambda_0,\boldsymbol{\psi}_0)$. Let us briefly recall how this scheme may be put into the framework of Sects. 2 and 3. We set:

$$\begin{aligned}
&\sigma^k=(\sigma_{ij}^k)=\left(\frac{\partial^2\psi^k}{\partial x_i\partial x_j}\right)_{1\le i,j\le 2}, \quad k=1,2,\\
&\bar{\sigma}=(\bar{\sigma}_{ij})=\left(\frac{\partial^2\bar{\psi}}{\partial x_i\partial x_j}\right)_{1\le i,j\le 2}.
\end{aligned} \tag{6.3}$$

Our new unknown will be the pair $u=(\boldsymbol{\psi},\boldsymbol{\sigma})$, $\boldsymbol{\sigma}=(\sigma^1,\sigma^2)$.

Next, let $q$ be a real number with $q>2$; we define the spaces

$$V=W_0^{1,q}(\Omega)^2\times(L^2(\Omega))_s^4,\quad W=((W_0^{1,q}(\Omega))')^2=W^{-1,q'}(\Omega)^2,\quad \left(\frac{1}{q}+\frac{1}{q'}=1\right) \tag{6.4}$$

where

$$L^2(\Omega)_s^4=\{\tau=(\tau_{ij})_{1\le i,j\le 2};\ \tau_{ij}\in L^2(\Omega),\ \tau_{12}=\tau_{21}\}.$$

Let $T\colon \mathbf{g}=(g^1,g^2)\in H^{-2}(\Omega)^2\to v=(\boldsymbol{\varphi},\boldsymbol{\tau})=T\mathbf{g}\in H_0^2(\Omega)^2\times(L^2(\Omega)_s^4)^2$ be defined by

$$\Delta^2\varphi^k=g^k \text{ in } \Omega,\quad \varphi^k=\frac{\partial\varphi^k}{\partial n}=0 \text{ on } \Gamma,\quad \tau_{ij}^k=\frac{\partial^2\varphi^k}{\partial x_i\partial x_j},\quad 1\le i,j,k\le 2. \tag{6.5}$$

Assuming very weak regularity hypotheses on the domain $\Omega$, it is easy to see that $T$ is a linear compact operator from $W$ into $V$.

Now, we define the $C^\infty$ mapping $G\colon(\lambda,v=(\boldsymbol{\varphi},\boldsymbol{\tau}))\in\mathbb{R}\times V\to G(\lambda,v)\in W$ by

$$G(\lambda,v)=(\tfrac{1}{2}[\tau^2,\tau^2],\ -[\tau^1+\lambda\bar{\sigma},\tau^2]-f) \tag{6.6}$$

where

$$[\sigma,\tau]=\sigma_{11}\tau_{22}+\sigma_{22}\tau_{11}-2\sigma_{12}\tau_{12},\qquad \sigma,\tau\in L^2(\Omega)_s^4. \tag{6.7}$$

Then, problem (6.1) can be equivalently written in the form

$$F(\lambda,u)\equiv u+TG(\lambda,u)=0. \tag{6.8}$$

Let us consider the "Hellan-Herrmann-Johnson" finite element scheme. We assume for simplicity that $\Omega$ is a *convex* polygonal domain so that the operator $T$ is continuous from $H^{-1}(\Omega)^2$ into $H^3(\Omega)^2 \times (H^1(\Omega)^4)^2$. Let $(\mathscr{T}_h)$ be a family of triangulations of $\bar{\Omega}$ made with triangles $K$ whose diameters are $\leqq h$. We assume the usual angle condition. For each integer $l \geqq 1$, we introduce the finite-dimensional spaces

$$\Phi_h = \{\varphi \in C^0(\bar{\Omega});\ \varphi|_K \in P_l\ \forall K \in \mathscr{T}_h,\ \varphi = 0 \text{ on } \Gamma\} \tag{6.9}$$

and

$$\Sigma_h = \{\tau \in (L^2(\Omega))_s^4;\ \tau_{ij}|_K \in P_{l-1}\ \forall K \in \mathscr{T}_h \text{ and } M_n(\tau) \text{ is continuous across the interelement boundaries}\}, \tag{6.10}$$

where $M_n(\tau) = \tau_{ij} n_j n_i$ [3] and $n = (n_1, n_2)$ is the unit outward normal along $\partial K$. Then, we set:

$$V_h = \Phi_h^2 \times \Sigma_h^2. \tag{6.11}$$

Next, we consider the bilinear forms

$$a(\sigma, \tau) = \int_\Omega \sigma_{ij}^k \tau_{ij}^k \, dx, \tag{6.12}$$

$$b(\tau, \varphi) = \sum_{K \in \mathscr{T}_h} \left\{ \int_K \frac{\partial \tau_{ij}^k}{\partial x_j} \frac{\partial \varphi^k}{\partial x_i} \, dx - \int_{\partial K} \frac{\partial \varphi^k}{\partial t} M_{nt}(\tau^k) \, ds \right\}, \tag{6.13}$$

where $t = (-n_2, n_1)$ is the unit tangent along $\partial K$ and $M_{nt}(\tau) = \tau_{ij} n_j t_i$.

We define the linear operator $T_h$: $\mathbf{g} \in W \to (\psi_h, \sigma_h) = T_h \mathbf{g} \in V_h$ by

$$\begin{aligned} a(\sigma_h, \tau) + b(\tau, \psi_h) &= 0 \qquad \forall \tau \in \Sigma_h^2, \\ b(\sigma_h, \varphi) &= -\langle \mathbf{g}, \varphi \rangle \qquad \forall \varphi \in \Phi_h^2, \end{aligned} \tag{6.14}$$

where $\langle \cdot, \cdot \rangle$ denotes the duality pairing between $W^{-1,q'}(\Omega)^2$ and $W_0^{1,q}(\Omega)^2$.

Now, the approximate problem consists in finding the solutions $(\lambda, u_h) \in \mathbb{R} \times V_h$ of the equation

$$F_h(\lambda, u_h) \equiv u_h + T_h G(\lambda, u_h) = 0. \tag{6.15}$$

Let us recall some approximation properties of the operator $T_h$. We have [9] (cf. also [1, 6, 7] for slightly weaker results).

**Lemma 9.** *Assume that the polygonal domain is convex. Then, we have for all* $q > 2$

$$\lim_{h \to 0} \|T - T_h\|_{\mathscr{L}(W;V)} = 0. \tag{6.16}$$

*Moreover, if* $u = (\psi, \sigma) = T\mathbf{g}$ *satisfies* $\psi \in H^{k+2}(\Omega)^2$ *for some* $k$ *with* $1 \leqq k \leqq l$, *we get*

$$\|(T - T_h)\mathbf{g}\|_V \leqq C h^k \|\psi\|_{H^{k+2}(\Omega)}. \tag{6.17}$$

On the other hand, it is easy to check that $(\lambda_0, \psi_0)$ is a simple bifurcation point of (6.1) if and only if $(\lambda_0, u_0)$ is a simple bifurcation point of (6.8) where

---

[3] Here and in the sequel, we use the classical summation convention

$u_0=(\psi_0,\sigma_0)$, $(\sigma_0)_{ij}=\dfrac{\partial^2\psi_0}{\partial x_i\partial x_j}$, $1\leqq i,j\leqq 2$. We shall also need to characterize the eigenvectors $\varphi_0$ and $\varphi_0^*$ of the operators $D_uF^0$ and $(D_uF^0)^*$ respectively. We denote by $\chi_0\in H_0^2(\Omega)^2$ an eigenfunction of the linearized von Kármán operator

$$\chi=(\chi^1,\chi^2)\to(\Delta^2\chi^1+[\psi_0^2,\chi^2],\Delta^2\chi^2-[\psi_0^2,\chi^1]-[\psi_0^1+\lambda\bar\psi,\chi^2]),$$

corresponding to the zero eigenvalue. Similarly, we denote by $\chi_0^*\in H_0^2(\Omega)^2$ an eigenfunction of the formal adjoint of the above operator corresponding to the zero eigenvalue; this means that

$$\Delta^2\chi_0^*+D_2^*\mathscr{H}^0(\chi_0^*)=0, \tag{6.18}$$

where

$$\mathscr{H}^0(\chi)=(\chi^2A\sigma_0^2,\chi^1A\sigma_0^2-\chi^2A(\sigma_0^1+\lambda_0\bar\sigma)),$$

$$A\tau=\begin{pmatrix}\tau_{22} & -\tau_{12}\\ -\tau_{12} & \tau_{11}\end{pmatrix},$$

and

$$(D_2^*\tau)=\left(\frac{\partial^2\tau_{ij}^1}{\partial x_i\partial x_j},\frac{\partial^2\tau_{ij}^2}{\partial x_i\partial x_j}\right).$$

Then, we have the following result whose elementary but lengthy proof is left to the reader.

**Lemma 10.** *We have (up to a normalizing factor):*

$$\varphi_0=(\chi_0,\omega_0),\qquad (\omega_0)_{ij}=\frac{\partial^2\chi_0}{\partial x_i\partial x_j},\qquad 1\leqq i,j\leqq 2, \tag{6.19}$$

*and*

$$\varphi_0^*=(0,\eta_0^*),\qquad \eta_0^*=-\mathscr{H}^0(\chi_0^*).\quad\blacksquare \tag{6.20}$$

As a consequence of (6.18) and (6.20), we find that

$$D_2^*\eta_0^*=\Delta^2\chi_0^*. \tag{6.21}$$

Now, for applying the results of the previous sections, we need to evaluate $\mathscr{K}(h)$. We denote by $\{(\lambda_i(t),\ \psi_i(t));\ |t|\leqq t_0\}$, $i=1,2$, the two branches of solutions of (6.1) which intersect at the point $(\lambda_0,\psi_0)$ (corresponding to $t=0$).

**Lemma 11.** *Assume that, for some $k$ with $1\leqq k\leqq l$, we have $\bar\psi\in H^{k+2}(\Omega)$ and $\psi_0$, $\dfrac{d\psi_i}{dt}(0)$, $\chi_0^*\in H^{k+2}(\Omega)^2$. Then, there exists a constant $C$ such that*

$$|\mathscr{K}(h)|\leqq Ch^{2k-1}. \tag{6.22}$$

*On the other hand, when $l=1$ and $\bar\psi\in H^3(\Omega)$, $\psi_0\in H^4(\Omega)^2$, $\dfrac{d\psi_i}{dt}(0)$, $\chi_0^*\in H^3(\Omega)^2$, we get*

$$|\mathscr{K}(h)|\leqq Ch^2. \tag{6.23}$$

*Proof.* We make use of the inequalities (3.18) and (3.19). By (3.19), we have

$$|f_h(0,0)| \leqq |\langle (T-T_h)G^0, \varphi_0^* \rangle| + \mathscr{E}^{(0)}(h,0)(\mathscr{E}^{(0)}(h,0) + \|[(T-T_h)D_u G^0]^* \varphi_0^*\|_{V'}). \tag{6.24}$$

We begin by estimating the first term in the right-hand side of (6.24). We set

$$TG^0 = (\boldsymbol{\psi}, \boldsymbol{\sigma}) \equiv -(\boldsymbol{\psi}_0, \boldsymbol{\sigma}_0), \qquad T_h G^0 = (\tilde{\boldsymbol{\psi}}_h, \tilde{\boldsymbol{\sigma}}_h).$$

It follows from (6.14) that

$$a(\boldsymbol{\sigma} - \tilde{\boldsymbol{\sigma}}_h, \boldsymbol{\tau}) + b(\boldsymbol{\tau}, \boldsymbol{\psi} - \tilde{\boldsymbol{\psi}}_h) = 0 \qquad \forall \boldsymbol{\tau} \in \Sigma_h^2 \tag{6.25}$$

$$b(\boldsymbol{\sigma} - \tilde{\boldsymbol{\sigma}}_h, \boldsymbol{\varphi}) = 0 \qquad \forall \boldsymbol{\varphi} \in \Phi_h^2. \tag{6.26}$$

Since $\boldsymbol{\psi}_0 \in H^{k+2}(\Omega)^2$, we have

$$\|\boldsymbol{\psi} - \tilde{\boldsymbol{\psi}}_h\|_{W^{1,q}(\Omega)} + \|\boldsymbol{\sigma} - \tilde{\boldsymbol{\sigma}}_h\|_0 \leqq c_1 h^k \|\boldsymbol{\psi}_0\|_{k+2}. \tag{6.27}$$

Next, using the results of [6] and [1, Sect. 4c], there exist two linear operators $\Pi_h$: $(H^1(\Omega)_s^4)^2 \to \Sigma_h^2$ and $r_h \in W_0^{1,q}(\Omega)^2 \to \Phi_h^2$ such that

$$b((I-\Pi_h)\boldsymbol{\tau}, \boldsymbol{\varphi}) = 0 \qquad \forall \boldsymbol{\varphi} \in \Phi_h^2, \ \forall \boldsymbol{\tau} \in (H^1(\Omega)_s^4)^2, \tag{6.28}$$

$$b(\boldsymbol{\tau}, (I-r_h)\boldsymbol{\varphi}) = 0 \qquad \forall \boldsymbol{\tau} \in \Sigma_h^2, \ \forall \boldsymbol{\varphi} \in W_0^{1,q}(\Omega)^2. \tag{6.29}$$

Moreover, we have for any $k$ with $1 \leqq k \leqq l$

$$\|(I-\Pi_h)\boldsymbol{\tau}\|_0 \leqq c_2 h^k \|\boldsymbol{\tau}\|_k, \qquad \forall \boldsymbol{\tau} \in (H^k(\Omega)_s^4)^2, \tag{6.30}$$

and

$$|b((I-\Pi_h)\boldsymbol{\tau}, (I-r_h)\boldsymbol{\varphi})| \leqq \begin{cases} c_3 h^{2k-1} \|\boldsymbol{\tau}\|_k \|\boldsymbol{\varphi}\|_{k+1} & \forall \boldsymbol{\varphi} \in (H^{k+1}(\Omega) \cap H_0^1(\Omega))^2, \\ & \forall \boldsymbol{\tau} \in (H^k(\Omega)_s^4)^2, \quad l \geqq 1, \\ c_4 h^2 \|D_2^* \boldsymbol{\tau}\|_0 \|\boldsymbol{\varphi}\|_2, & \forall \boldsymbol{\varphi} \in (H^2(\Omega) \cap H_0^1(\Omega))^2, \\ & \forall \boldsymbol{\tau} \in (H^2(\Omega)_s^4)^2, \quad l = 1. \end{cases} \tag{6.31}$$

Now it follows from (6.20) that

$$\langle (T-T_h)G^0, \varphi_0^* \rangle = a(\boldsymbol{\sigma} - \tilde{\boldsymbol{\sigma}}_h, \boldsymbol{\eta}_0^*) = a(\boldsymbol{\sigma} - \tilde{\boldsymbol{\sigma}}_h, (I-\Pi_h)\boldsymbol{\eta}_0^*) + a(\boldsymbol{\sigma} - \tilde{\boldsymbol{\sigma}}_h, \Pi_h \boldsymbol{\eta}_0^*).$$

Then, using (6.25) and (6.28), we may write

$$a(\boldsymbol{\sigma} - \boldsymbol{\sigma}_h, \Pi_h \boldsymbol{\eta}_0^*) = b((I-\Pi_h)\boldsymbol{\eta}_0^*, (I-r_h)\boldsymbol{\psi}) + b(\boldsymbol{\eta}_0^*, \tilde{\boldsymbol{\psi}}_h - \boldsymbol{\psi}).$$

But, as a consequence of (6.21), we get for all $\boldsymbol{\varphi} \in W_0^{1,q}(\Omega)^2$

$$b(\boldsymbol{\eta}_0^*, \boldsymbol{\varphi}) = -\langle D_2^* \boldsymbol{\eta}_0^*, \boldsymbol{\varphi} \rangle = -\langle \Delta^2 \boldsymbol{\chi}_0^*, \boldsymbol{\varphi} \rangle = b(D_2 \boldsymbol{\chi}_0^*, \boldsymbol{\varphi})$$

where

$$D_2 \boldsymbol{\chi} = \left( \frac{\partial^2 \chi^1}{\partial x_i \partial x_j}, \frac{\partial^2 \chi^2}{\partial x_i \partial x_j} \right)_{1 \leqq i,j \leqq 2}$$

Hence, we obtain

$$b(\boldsymbol{\eta}_0^*, \tilde{\boldsymbol{\psi}}_h - \tilde{\boldsymbol{\psi}}) = b(D_2 \boldsymbol{\chi}_0^*, \tilde{\boldsymbol{\psi}}_h - \boldsymbol{\psi})$$

so that by (6.28)

$$b(\eta_0^*, \tilde{\psi}_h - \psi) = b((I - \Pi_h) D_2 \chi_0^*, (r_h - I)\psi) + b(\Pi_h D_2 \chi_0^*, \tilde{\psi}_h - \psi).$$

Next, applying again (6.25) gives

$$b(\Pi_h D_2 \chi_0^*, \tilde{\psi}_h - \psi) = a(\sigma - \tilde{\sigma}_h, (\Pi_h - I) D_2 \chi_0^*) + a(\sigma - \tilde{\sigma}_h, D_2 \chi_0^*).$$

Note that

$$a(\sigma - \tilde{\sigma}_h, D_2 \chi_0^*) = b(\tilde{\sigma}_h - \sigma, \chi_0^*)$$

and therefore by (6.26) and (6.29)

$$a(\sigma - \tilde{\sigma}_h, D_2 \chi_0^*) = b((\Pi_h - I)\sigma, (I - r_h)\chi_0^*).$$

Thus, we get

$$\langle (T - T_h) G^0, \varphi_0^* \rangle = A_h + B_h$$

where

$$A_h = a(\sigma - \tilde{\sigma}_h, (I - \Pi_h)(\eta_0^* - D_2 \chi_0^*)),$$
$$B_h = b((I - \Pi_h)(\eta_0^* - D_2 \chi_0^*), (I - r_h)\psi) + b((\Pi_h - I)\sigma, (I - r_h)\chi_0^*).$$

If we assume that $\psi_0$, $\chi_0^* \in H^{k+2}(\Omega)^2$ and $\eta_0^* \in H^k(\Omega)^2$ for some $k$ with $1 \leqq k \leqq l$, we have by using (6.27) and (6.30)

$$|A_h| \leqq c_5 h^{2k} \|\psi_0\|_{k+2} (\|\eta_0^*\|_k + \|\chi_0^*\|_{k+2}).$$

On the other hand, it follows from (6.31) that

$$|B_h| \leqq c_6 h^{2k-1} \|\psi_0\|_{k+2} (\|\eta_0^*\|_k + \|\chi_0^*\|_{k+2}).$$

In the case $l=1$, assuming that $D_2^* \sigma_0 = \Delta^2 \psi_0 \in L^2(\Omega)^2$, we obtain as a consequence of (5.21) and (6.30)

$$|B_h| \leqq c_7 h^2 \|\Delta^2 \psi_0\|_0 \|\chi_0^*\|_2.$$

If in addition $\bar{\psi} \in H^{k+2}(\Omega)$, we deduce from (6.20) and the Sobolev imbedding theorem that

$$\|\eta_0^*\|_k \leqq c_8 \|\chi_0^*\|_{k+2} (\|\psi_0\|_{k+2} + \|\bar{\psi}\|_{k+2}).$$

Hence, we find if $\bar{\psi} \in H^{k+2}(\Omega)$, $\psi_0, \chi_0^* \in H^{k+2}(\Omega)^2$ for some $k$ with $1 \leqq k \leqq l$

$$|\langle (T - T_h) G^0, \varphi_0^* \rangle| \leqq c_9 h^{2k-1}, \qquad c_9 = c_9(\bar{\psi}, \psi_0, \chi_0^*). \tag{6.32}$$

In the case $l=1$, we get if $\bar{\psi} \in H^3(\Omega)$, $\psi_0 \in H^4(\Omega)^2$, $\chi_0^* \in H^3(\Omega)^2$

$$|\langle (T - T_h) G^0, \varphi_0^* \rangle| \leqq c_{10} h^2, \qquad c_{10} = c_{10}(\bar{\psi}, \psi_0, \chi_0^*). \tag{6.33}$$

We now estimate the second term in the right-hand side of (6.24). First, (6.17) gives

$$\mathscr{E}^{(0)}(h, 0) = \|(T - T_h) G^0\|_V \leqq c_{11} h^k \|\psi_0\|_{k+2}, \tag{6.34}$$

and using the same method of proof as above, we obtain estimates similar to (6.32) and (6.33) for the term $\mathscr{E}^{(0)}(h,0)\,\|[(T-T_h)D_uG^0]^*\,\varphi_0^*\|_{V'}$.

Analogously, assuming that $\dfrac{d\psi_i}{dt}(0)\in H^{k+2}(\Omega^2)$, $i=1,2$, we have

$$\begin{aligned}\mathscr{E}^{(1)}(h,0)&=\mathscr{E}^{(0)}(h,0)+\sum_{i=1}^{2}\left\|\frac{d}{dt}(T-T_h)\,G(\lambda_i(t),u_i(t))|_{t=0}\right\|_V\\&\leqq c_{12}h^k\left(\|\psi_0\|_{k+2}+\sum_{i=1}^{2}\left\|\frac{d\psi_i}{dt}(0)\right\|_{k+2}\right).\end{aligned}\tag{6.35}$$

The desired results then follow from (3.18), (6.24), (6.32), (6.33), (6.34) and (6.35). ■

We are now able to apply to our problem the results of Theorem 2.

**Theorem 7.** *Assume that $(\lambda_0,\psi_0)$ is a simple bifurcation point of (6.1). Then there exist a neighborhood $\mathcal{O}$ of the point $(\lambda_0,u_0)$ in $\mathbb{R}\times V$ and a positive constant $h_0$ such that for $h\leqq h_0$ the set $\mathscr{S}_h$ of solutions of (6.15) consists of two $C^\infty$ branches.*

*If these two branches intersect at a point $(\lambda_h^0,u_h^0)\in\mathcal{O}$, they can be parametrized in the form $\{(\lambda_h^i(t),u_h^i(t))=(\psi_h^i(t),\sigma_h^i(t));\ |t|\leqq t_0\}$ with $\lambda_h^i(0)=\lambda_h^0$, $u_h^i(0)=u_h^0$, $i=1,2$. Moreover, if $t\to\psi_i(t)$ is a $C^{m+1}$ function from $|t|\leqq t_0$ into $H^{k+2}(\Omega)^2$ for some integer $m\geqq 0$ and some $k$ with $1\leqq k\leqq l$, we get the estimate*

$$\begin{aligned}\sup_{|t|\leqq t_0}\Bigg\{&\left|\frac{d^m}{dt^m}(\lambda_h^i(t)-\lambda_i(t))\right|+\left\|\frac{d^m}{dt^m}(\psi_h^i(t)-\psi_i(t))\right\|_{W_0^{1,q}(\Omega)}\\&+\left\|\frac{d^m}{dt^m}(\sigma_h^i(t)-\sigma_i(t))\right\|_0\Bigg\}\leqq K_m h^k.\end{aligned}\tag{6.36}$$

*Otherwise, $\mathscr{S}_h$ is $C^\infty$ diffeomorphic to (a part of) a nondegenerate hyperbola. If $\bar\psi\in H^{k+2}(\Omega)$, $\chi_0^*\in H^{k+2}(\Omega)^2$ and $t\to\psi(t)$ is a $C^2$ function from $|t|\leqq t_0$ into $H^{k+2}(\Omega)^2$ for some $k$ with $1\leqq k\leqq l$, the distance between $\mathscr{S}_h$ and the set $\mathscr{S}$ of solutions of (6.8) contained in $\mathcal{O}$ is bounded by*

$$d(\mathscr{S}_h,\mathscr{S})\leqq ch^{k-\frac{1}{2}}.\tag{6.37}$$

*When $l=1$, we have if $\bar\psi\in H^3(\Omega)$, $\chi_0^*\in H^3(\Omega)^2$, $\psi_0\in H^4(\Omega)^2$ and $t\to\psi(t)$ is a $C^2$ function from $|t|\leqq t_0$ in $H^3(\Omega)^2$*

$$d(\mathscr{S}_h,\mathscr{S})\leqq ch.\tag{6.38}$$

*Remark 6.* Consider the case of a bifurcation from the trivial branch at the point $(\lambda_0,0)$. Then, as in Sect. 5, it is possible to prove that the approximate problem bifurcates at a point $(\lambda_h^0,0)$ which satisfies the bound

$$|\lambda_h^0-\lambda_0|\leqq\begin{cases}Ch^{2l-1}, & l\geqq 2,\\ Ch^2, & l=1,\end{cases}$$

under reasonable smoothness properties. ■

## References


1. Babuska, I., Osborn, J., Pitkäranta, J.: Analysis of mixed methods using mesh dependent norms. MRC Rech. Summary Report # 2003 (1979), University of Wisconsin, Madison
2. Berger, M.S.: Nonlinearity and functional analysis. New York: Academic Press 1977
3. Bernardi, C.: Approximation of Hopf bifurcation. Numer. Math. (in press, 1981)
4. Brezzi, F., Rappaz, J., Raviart, P.-A.: Finite-dimensional approximation of nonlinear problems. Part I: Branches of nonsingular solutions. Numer. Math. **36**, 1–25 (1980)
5. Brezzi, F., Rappaz, J., Raviart, P.-A.: Finite-dimensional approximation of nonlinear problems. Part II: Limit points. Numer. Math. **37**, 1–28 (1981)
6. Brezzi, F., Raviart, P.-A.: Mixed finite element methods for 4th order elliptic equations. Topics in numerical analysis III (J.J.H. Miller ed.), pp. 33–56. London: Academic Press 1976
7. Falk, R.S., Osborn, J.: Error estimates for fixed methods. RAIRO Numer. Anal. **14**, 249–277 (1980)
8. Fujii, H., Yamaguti, M.: Structure of singularities and its numerical realization in nonlinear elasticity. J. Math. Kyoto Univ. **20**, 489–590 (1980)
9. Girault, V., Raviart, P.-A.: An analysis of some upwind schemes for the Navier-Stokes equations SIAM. J. Numer. Anal. (in press)
10. Herrmann, L.R.: Finite element bending analysis for plates. J. Eng. Mech. Div. ASCE EM 5, **93**, 49–83 (1967)
11. Herrmann, L.R.: A bending analysis for plates. Proc. Conf. on Matrix Methods in Structural Mechanics, AFFDL-TR-66-88, 577–604
12. Johnson, C.: On the convergence of a mixed finite element method for plate bending problems. Numer. Math. **21**, 43–62 (1973)
13. Keller, H.B.: (To appear)
14. Kesavan, S.: La méthode de Kikuchi appliquée aux équations de van Kármán. Numer. Math. **32**, 209–232 (1979)
15. Kesavan, S.: Une méthode d'éléments finis mixte pour les équations de von Kármán. RAIRO Numer. Anal. **14**, 149–173 (1980)
16. Kikuchi, F.: An iterative finite element scheme for bifurcation analysis of semilinear elliptic equations. Report Inst. Space. Aero. Sc. # 542 (1976), Tokyo University
17. Nirenberg, L.: Topics in nonlinear functional Analysis. Courant Institute of Mathematical Science, New York University (1974)
18. Rappaz, J., Raugel, G.: Finite-dimensional approximation of bifurcation problems at a double eigenvalue. (To appear)
19. Raugel, G.: (To appear)
20. Sattinger, D.H.: Group theoretic methods in bifurcation theory. Lecture Notes in Mathematics Vol. 762. Berlin Heidelberg New York: Springer 1979


Received September 16, 1980