RAIRO. ANALYSE NUMÉRIQUE

JEAN DESCLOUX
JACQUES RAPPAZ

Approximation of solution branches of nonlinear equations

RAIRO. Analyse numérique, tome 16, nº 4 (1982), p. 319-349
<http://www.numdam.org/item?id=M2AN_1982__16_4_319_0>

© AFCET, 1982, tous droits réservés.

L'accès aux archives de la revue « RAIRO. Analyse numérique » implique l'accord avec les conditions générales d'utilisation (http://www.numdam.org/conditions). Toute utilisation commerciale ou impression systématique est constitutive d'une infraction pénale. Toute copie ou impression de ce fichier doit contenir la présente mention de copyright.

NUMDAM
Article numérisé dans le cadre du programme Numérisation de documents anciens mathématiques
http://www.numdam.org/

R A I R O Analyse numérique/Numerical Analysis
(vol 16, n° 4, 1982, p 319 à 349)

APPROXIMATION OF SOLUTION BRANCHES OF NONLINEAR EQUATIONS (*), (**)

by Jean DESCLOUX (1) and Jacques RAPPAZ (1)

Communicated by P G CIARLET

Abstract — *We present a general theory for the approximation of regular and bifurcating branches of solutions of nonlinear equations It can be applied to numerous problems, including differential equations on unbounded domains, in connection with various numerical algorithms, for example Galerkin methods with numerical integration*

Résumé — *On présente une théorie générale de l'approximation de branches, régulières ou avec bifurcation, de solutions d'équations non linéaires Cette théorie s'applique à de nombreux problèmes, y compris les équations différentielles sur des domaines non bornes, résolus par des méthodes numériques variées, par exemple des méthodes de Galerkin avec intégration numerique*

1. INTRODUCTION

In their three papers [1], [2], [3], Brezzi, Rappaz and Raviart consider the approximation of nonlinear equations of the type

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

by a family of equations of the form

$$u + T_h G(\lambda, u) = 0 \; ; \tag{1.2}$$

here $G : \mathbb{R} \times V \to W$ is a regular nonlinear mapping, $T : W \to V$ and $T_h : W \to V_h$ are bounded linear operators ; V and W are real Banach spaces,

(*) Received in december 1981
(**) This work was supported by the Fonds National Suisse de la Recherche Scientifique
(1) Département de Mathématiques, École Polytechnique Fédérale, Lausanne, Suisse.

© Bordas-Dunod

$\{ V_h \}_h$ is a family of finite dimensional subspaces of V As a main hypothesis connecting (1 1) and (1 2), they suppose that

$$\lim_{h \to 0} \| T - T_h \|_{\mathscr{L}(W V)} = 0 , \tag{1 3}$$

which implies in particular that T is compact

Brezzi, Rappaz and Raviart have limited their investigations to regular branches of solutions [1], limit points [2] and simple bifurcation points [3], whereas in [12], [13] Rappaz and Raugel have considered in the same context bifurcation at multiple eigenvalues

The purpose of this paper is to generalize in an unified treatment some of the main results contained in the references mentioned above In particular, our theory includes the possibility to analyse two new situations

a) in (1 1), T is non compact, *b*) the approximation is of Galerkin type with numerical integration

Also most concrete problems can be written naturally in the form (1 1), we have found suitable to adopt the following framework Let X and Y be real Banach spaces, $F \; X \to Y$ be a sufficiently regular nonlinear map, $x_0 \in X$ be such that $F(x_0) = 0$ In a neighborhood of x_0, we consider the equation

$$F(x) = 0 , \tag{1 4}$$

we shall suppose that $F \,(x_0) \; X \to Y$ is a Fredholm operator of index 1, however we shall assume *no compactness* hypothesis on $F \,(x_0)$

Several authors have considered bifurcation problems in the general form (1 4) (see for example Magnus [10]), since with it the parameter $\lambda \in \mathbb{R}$ does not appear explicitly, a simple limit point cannot be distinguished from a regular point, in the same way, the " double limit point " introduced by Decker and Keller in [5], can be treated as an usual simple bifurcation point (see Descloux, Rappaz [8])

For approximating the equation $F(x) = 0$ we consider two families of finite dimensional subspaces $\{ X_h \}_h$, $\{ Y_h \}_h$ of X and Y respectively, nonlinear mappings $F_h \; X_h \to Y_h$ and the equations

$$F_h(x) = 0 , \tag{1 5}$$

instead of (1 3), we impose on F and F_h a *consistency condition* and a *stability condition* which are given by relations (3 5) and (3 7) in Section 3, remark that they do not suppose that $F \,(x_0)$ possesses any property of compactness

As for the analysis of problems (1 1), (1 2) by Brezzi, Rappaz, Raviart and Raugel, our investigation will be based on the implicit function theorem In

Section 1, we recall a version of this theorem and prove a basic error estimate (Theorem 2.2).

Section 3 deals with " regular points ", i.e. we require that $F'(x_0)$ is surjective. Theorem 3.1 contains the general results. In Theorem 3.2, we suppose that X is of the form $\mathbb{R} \times V$, and that $x_0 = (\lambda_0, u_0) \in X$ is a turning point; for Galerkin methods, as in [2], we obtain an improved bound for the parameter λ at the approximate turning point.

The main results of this paper are contained in Section 4 which is devoted to bifurcation points. Although not impossible, we have found it very complicated to work with the approximate problem (1.5) when F_h operates on finite dimensional subspaces; for this reason we require that F_h admits a suitable extension $\tilde{F}_h : X \to Y$ such that, in particular, the equations $F_h(x) = 0$ and $\tilde{F}_h(x) = 0$ possess the same solutions. As in Section 3, we impose to F and F_h (where now F_h denotes the extended operator $X \to Y$) a consistency condition (4.4) and a stability condition (4.6). Supposing that the dimension of the kernel of $F'(x_0)$ is $n + 1$ with $n \geqslant 1$, we apply the Lyapunov-Schmidt procedure to F and F_h (Theorem 4.1) and reduce problems (1.4) and (1.5) to equations of the form $f(\sigma) = 0$ and $f_h(\sigma) = 0$, where f and f_h operate on the same finite dimensional subspaces. Theorems 4.2 and 4.3 are based on the following hypothesis :

$$f(0) = f'(0) = \cdots = f^{(q-1)}(0) = 0\,, \; f_h(0) = f_h'(0) = \cdots = f_h^{(q-1)}(0) = 0$$

for some $q \geqslant 2$ and there exists a non degenerate characteristic ray (Hypotheses (4.16), (4.17)); they show the existence of a branch Γ of solutions of the exact problem (1.4) passing through x_0 and tangent to the characteristic ray at x_0 and, on the other side, the existence of a branch Γ_h of solutions of the approximate problem (1.5) converging to Γ. The particular case of a simple bifurcation point is treated in Theorems 4.4 and 4.5 which give error estimates similar to those obtained in [3].

The aim of Section 5 is to show how the results of Section 4 can be applied to the following classical problem

$$(\lambda, u) \in \mathbb{R} \times H_0^1(\Omega), \quad -\Delta u - \lambda u + u^3 = 0\,;$$

here Ω is the unit square $0 < x, y < 1$ and we are interested by solutions in the neighborhood of $(\lambda_0, 0)$ where $\lambda_0 = 5\,\Pi^2$ is a double eigenvalue of the eigenvalue equation $\Delta u + \lambda u = 0$. The approximate problem is obtained by the Galerkin method *with numerical integration.* The exact problem can be written in the form (1.1) with T compact; however, due to the presence of numerical integration and to the fact that $H_0^1(\Omega)$ is not imbedded in $C^0(\bar{\Omega})$, the approximate problem cannot be put in the form (1 2) Note that this difficulty can be overcome by replacing $H_0^1(\Omega)$ by $W_0^{1,p}(\Omega)$ for $p > 2$, but, then, the estimate $0(h)$ in (5 25) should be replaced by $0(h^{1-\varepsilon})$, $\varepsilon > 0$ (see [11])

Except for a part of Section 5, all the results of this paper are contained in our Report [8] in which however some further questions are discussed, for example bifurcation in presence of symmetry, the situation of imperfect numerical bifurcation (i e using the above notations, f satisfies the relations $f(0) = f'(0) = \dots = f^{(q-1)}(0) = 0$, but f_h does not satisfy the relations $f_h(0) = f_h'(0) = \dots = f^{(q-1)}(0) = 0$) Let us also mention the analysis of a nonlinear Sturm-Liouville eigenvalue problem on the infinite interval $(0, \infty)$, here, the exact problem can be written in the form (1 1), but with T non compact (see also [9])

2. NOTATIONS. PRELIMINARIES

We first introduce some notations Let X, Y, Z be real Banach spaces For the sake of simplicity, we shall denote by $\| \cdot \|$ the various norms in X, Y, Z, $\mathscr{L}_m(X, Y)$, $\mathscr{L}_m(X \times Y, Z)$, where $\mathscr{L}_m(X, Y)$ is the space of continuous m-linear mappings of X^m into Y In the same way, for any space, $B(a, \rho)$ denotes the open ball of center a and radius ρ The norm in $X \times Y$ is defined by the relation $\| (x, y) \| = \| x \| + \| y \|$ For a map $G : \Omega \subset X \to Y$, $D^m G$ or $G^{(m)} : \Omega \to \mathscr{L}_m(X, Y)$ represents the m-th Frechet derivative of G, for $x \in \Omega$, $\xi = (\xi_1, \dots, \xi_m) \in X^m$, we use the notations $G^{(m)}(x) \in \mathscr{L}_m(X, Y)$,

$$G^{(m)}(x)\,\xi = G^{(m)}(x)\,(\xi_1, \dots, \xi_m) \in Y\,,$$

if $\xi_1 = \xi_2 = \dots = \xi_{m-1}$, or if $\xi_1 = \xi_2 = \dots = \xi_m$, we may also write $G^{(m)}(x)\,\xi = G^{(m)}(x)\,\xi_1^{m-1}\,\xi_m$ or $G^{(m)}(x)\,\xi = G^{(m)}(x)\,\xi_1^m$, respectively For a map $G : \Omega \subset X \times Y \to Z$, $D_x G$, $D_y G$, $D_{xx}^2 G$, $D_{xy}^2 G$ will denote the partial derivatives

As in [1], [2], [3], the essential tool of this work will be the implicit function theorem We quote here a particular version of it, for the proof, see, for exemple [8]

THEOREM 2 1 *Let X, Y and Z be Banach spaces, $x_0 \in X$, $y_0 \in Y$, δ be a positive number, $\Omega = B(x_0, \delta) \times B(y_0, \delta) \subset X \times Y$, $G : \Omega \to Z$ be a C^p mapping with $p \geqslant 2$ We suppose that $D_y G(x_0, y_0)$ is an isomorphism from Y onto Z and that there exist the numbers $c_0, c_1, \dots, c_p$ such that*

$$\| D_y G(x_0, y_0)^{-1} \| \leqslant c_0\,,$$

$$\| G^{(k)}(x, y) \| \leqslant c_k \quad \forall (x, y) \in \Omega\,,\ k = 1, 2, \dots, p$$

Then, there exist positive numbers a, b, d depending only on δ, c_0, c_1, c_2 and for $k = 1, 2, \ldots, p$, the numbers M_k depending only on $c_0, c_1, \ldots, c_k$ such that :

i) *for any $(x, y) \in B(x_0, a) \times B(y_0, b)$, $D_y G(x, y)$ is an isomorphism from Y onto Z such that $\| D_y G(x, y)^{-1} \| \leqslant 2 c_0$,*

ii) *if $\| G(x_0, y_0) \| < d$, there exists a C^p mapping $g : B(x_0, a) \to B(y_0, b)$ such that, for any $x \in B(x_0, a)$, $y = g(x)$ is the unique solution of the equation $G(x, y) = 0$, $y \in B(y_0, b)$, i.e.*

$$G(x, g(x)) = 0 \,, \quad g(x) \in B(y_0, b) \quad \forall x \in B(x_0, a) \,;$$

furthermore

$$\| g^{(k)}(x) \| \leqslant M_k \quad \forall x \in B(x_0, a), 1 \leqslant k \leqslant p \,.$$

The next theorem provides a key result for error estimates.

THEOREM 2.2 : *We consider the situation given by Theorem 2.1 with $\| G(x_0, y_0) \| < d$. Let W be a real Banach space, $\Delta \subset W$ be open, $s : B(x_0, a) \to B(y_0, b)$ and $\alpha : \Delta \to B(x_0, a)$ be C^{p-1} mappings, $\tilde{g} = g \circ \alpha$ and $\tilde{s} = s \circ \alpha : \Delta \to B(y_0, b)$. We suppose that there exist constants*

$$e_1, e_2, \ldots, e_{p-1} \,, \quad \gamma_1, \gamma_2, \ldots, \gamma_{p-1}$$

such that

$$\| s^{(k)}(x) \| \leqslant e_k \quad \forall x \in B(x_0, a) \,, \quad \| \alpha^{(k)}(t) \| \leqslant \gamma_k \quad \forall t \in \Delta \,,$$
$$k = 1, 2, \ldots, p - 1 \,.$$

Then, for $k = 0, 1, \ldots, p - 1$, there exist constants K_k depending only on $c_0, c_1, \ldots, c_{k+1}, e_1, e_2, \ldots, e_k, \gamma_1, \gamma_2, \ldots, \gamma_k$ such that

$$\| \tilde{g}^{(k)}(t) - \tilde{s}^{(k)}(t) \| \leqslant K_k \sum_{l=0}^{k} \| H^{(l)}(t) \| \quad \forall t \in \Delta \,, \quad 0 \leqslant k \leqslant p - 1 \,, \quad (2.1)$$

where $H : \Delta \to Z$ is the C^{p-1} mapping defined by $H(t) = G(\alpha(t), \tilde{s}(t))$.

Proof : Since $G(\alpha(t), \tilde{g}(t)) = 0$, we obtain by the fundamental theorem of calculus :

$$H(t) = G(\alpha(t), \tilde{s}(t)) - G(\alpha(t), \tilde{g}(t)) = E(t) \, (\tilde{s}(t) - \tilde{g}(t)) \,, \quad t \in \Delta \,, \quad (2.2)$$

where

$$E(t) = \int_0^1 D_y G(\alpha(t), \tilde{g}(t) + \tau(\tilde{s}(t) - \tilde{g}(t)) \, \mathrm{d}\tau \,; \quad (2.3)$$

by the same theorem again, we have that $\| E(t) - D_y G(x_0, y_0) \| \leqslant c_2(a + b)$. Without restriction of generality we can assume in Theorem 2.1 that $2 c_0 c_2(a + b) \leqslant 1$; writting

$$E(t) = D_y G(x_0, y_0) (I - D_y G(x_0, y_0)^{-1} (D_y G(x_0, y_0) - E(t))),$$

we see that for any $t \in \Delta$, $E(t)$ is an isomorphism from Y onto Z with inverse bounded by $2 c_0$. By differentiating (2.3) j times, $1 \leqslant j \leqslant p - 1$, and by using the bound $\| g^{(k)}(x) \| \leqslant M_k$ of Theorem 2.1, we see that $\| E^{(j)}(t) \|$ is bounded, uniformly with respect to t, by a constant depending only on $c_0, c_1, ..., c_{j+1}, e_1, e_2, ..., e_j, \gamma_1, ..., \gamma_j$. We now prove Theorem 2.2 by induction ; for $k = 0$, we have by (2.2) :

$$\| \tilde{s}(t) - \tilde{g}(t) \| \leqslant \| E(t)^{-1} \| \, \| H(t) \| \leqslant 2 c_0 \| H(t) \|$$

which proves (2.1) for $k = 0$ with $K_0 = 2 c_0$. Now suppose (2.1) true for $1 \leqslant k \leqslant q - 1$, $1 \leqslant q \leqslant p - 1$; by differentiating (2.2) q times we obtain for any $\xi = (\xi_1, ..., \xi_q) \in W^q$ and any $t \in \Delta$:

$$E(t) (\tilde{s}^{(q)}(t) - \tilde{g}^{(q)}(t)) \xi = H^{(q)}(t) \xi - \sum_{j=1}^{q} \sum_{\Pi_j} D^j E(t) (\eta(\Pi_j)) \cdot D^{q-j}(\tilde{s}(t) - \tilde{g}(t)) (\zeta(\Pi_j)),$$

where Π_j is any partition of the set $\{ \xi_1, ..., \xi_q \}$ in two subsets $\eta(\Pi_j)$ and $\zeta(\Pi_j)$ containing respectively j and $(q - j)$ elements; by the hypothesis of induction, this proves (2.1) for $k = q$. ■

We conclude this section by recalling a classical result we shall use frequently in the following.

THEOREM 2.3 : *Let X and Y be Banach spaces, D be a relatively compact subset of X, f and f_n, $n \in \mathbb{N}$, be maps from D into Y. We suppose :*

a) $\lim_{n \to \infty} f_n(x) = f(x)$ $\forall x \in D$, b) *there exists a constant L such that*

$$\| f_n(x) - f_n(\xi) \| \leqslant L \| x - \xi \| \quad \forall x, \quad \xi \in D, \quad \forall n \in \mathbb{N}.$$

Then $\lim_{n \to \infty} f_n = f$ *uniformly.*

3. REGULAR POINTS

Let X and Y be two real Banach spaces, $F : X \to Y$ be a C^p mapping with $p \geqslant 2$ and $x_0 \in X$ be such that $F(x_0) = 0$. We suppose that x_0 is a regular point in the following sense

$$F'(x_0) : X \to Y \text{ is a Fredholm operator of index } 1\,, \tag{3.1}$$

$$\text{Range } F'(x_0) = Y\,. \tag{3.2}$$

Hypotheses (3.1) and (3.2) imply that the kernel of $F'(x_0)$ is one-dimensional and consequently is spanned by some vector $\varphi_0 \in X$, $\varphi_0 \neq 0$. Let $\psi_0 \in X^*$ be such that $\langle \varphi_0, \psi_0 \rangle \neq 0$, where X^* denotes the dual space of X and $\langle \cdot, \cdot \rangle$ the duality pairing between X and X^*. As we shall see in Theorem 3.1, there is an unique branch of solutions of the equation $F(x) = 0$ passing through x_0 which can be parametrized by a function $x(t)$ satisfying the relations $F(x(t)) = 0$, $\langle x(t) - x_0, \psi_0 \rangle - t = 0$.

In order to approximate this branch of solutions let $\{ X_h \}_h$ and $\{ Y_h \}_h$ be two families of finite dimensional subspaces of X and Y respectively, $\{ F_h \}_h$ be a family of C^p functions mapping X_h into Y_h and $\{ \Pi_h \}_h$ be a family of projectors mapping X onto X_h, here h is a parameter which tends to zero. We suppose:

a) $\lim_{h \to 0} \Pi_h x = x \quad \forall x \in X\,,$ (3.3)

b) $\text{dimension } X_h = \text{dimension } Y_h + 1\,,$ (3.4)

c) for any $0 \leqslant k \leqslant p - 1$ and for any fixed $x, \xi_1, \dots, \xi_k \in X$, we have

$$\lim_{h \to 0} \| F^{(k)}(x)(\xi_1, \dots, \xi_k) - F_h^{(k)}(\Pi_h x)(\Pi_h \xi_1, \dots, \Pi_h \xi_k) \| = 0\,, \tag{3.5}$$

d) there exist the positive constants δ and c such that

$$\| F_h^{(k)}(x) \| \leqslant c, \quad \forall x \in X_h \quad \text{with} \quad \| x - \Pi_h x_0 \| < \delta\,, \quad \forall h,\ 1 \leqslant k \leqslant p\,, \tag{3.6}$$

e) there is a positive constant μ such that

$$\| F_h'(\Pi_h x_0)\, \xi \| \geqslant \mu \, \| \xi \| \ \forall \xi \in W_h\,, \quad \forall h\,, \tag{3.7}$$

where $W_h = \{ x \in X_h \mid \langle x, \psi_0 \rangle = 0 \}$.

(3.3) implies that the projectors Π_h are uniformly bounded, (3.4) is the discrete analogous of (3.1), (3.5) is a relation of pointwise convergence which can be interpreted as a condition of consistency whereas (3.7) will appear as a condition of stability allowing the use of the implicit function theorem.

Remark 3.1. Let $W = \{ x \in X \mid \langle x, \psi_0 \rangle = 0 \}$, by Hypotheses (3.1), (3.2) and by Banach's theorem, $F'(x_0)$ defines an isomorphism from W onto Y

and consequently there exists a positive constant c such that

$$\| F(x_0)\xi \| \geqslant c \| \xi \| \quad \forall \xi \in W$$

It follows that a *sufficient* condition which insures the stability hypothesis (3 7) is the following one

$$\lim_{h \to 0} \sup_{\xi \in X_h \; \|\xi\| = 1} \| (F_h'(\Pi_h x_0)\xi - F(x_0)\xi \| = 0, \qquad (3\ 8)$$

such a condition has been used for example in [7] in connection with eigenvalue problems

Let $\mathscr{G} \; \mathbb{R} \times X \to \mathbb{R} \times Y$ and $\mathscr{G}_h \; \mathbb{R} \times X_h \to \mathbb{R} \times Y_h$ be defined by the relations

$$\mathscr{G}(t, x) = (\langle x - x_0, \psi_0 \rangle - t, F(x)), \quad \mathscr{G}_h(t, x) = (\langle x - x_0, \psi_0 \rangle - t, F_h(x)) \qquad (3\ 9)$$

LEMMA 3 1 *Assume Hypotheses* (3 1) *to* (3 7) *Then* a) $D_x \mathscr{G}(0, x_0)$ *is an isomorphism from X onto $\mathbb{R} \times Y$* b) *For h small enough, $D_x \mathscr{G}_h(0, \Pi_h x_0)$ is an isomorphism from X_h onto $\mathbb{R} \times Y_h$ with uniformly (with respect to h) bounded inverse*

Proof By Hypotheses (3 1), (3 2) and the fact that $\langle \omega_0, \psi_0 \rangle \neq 0$, part a) of lemma 3 1 follows immediately from Banach's theorem In the following, c will denote a positive generic constant independent of h, since X_h and $\mathbb{R} \times Y_h$ have the same finite dimension, it suffices, for proving part b), to show, for h small enough, that $\| D_x \mathscr{G}_h(0, \Pi_h x_0)\xi \| \geqslant c \| \xi \| \quad \forall \xi \in X_h$ Let $\omega_{0h} = \Pi_h \omega_0$, by Hypotheses (3 3) and (3 5) we have that $\lim_{h \to 0} \omega_{0h} = \omega_0$ and $\lim_{h \to 0} F_h(\Pi_h x_0)\omega_{0h} = 0$ Any $\xi \in X_h$ can be decomposed as $\xi = \alpha\omega_{0h} + w$, $\alpha \in \mathbb{R}$, $w \in W_h$ and we obtain by Hypothesis (3 7)

$$\begin{aligned} \| D_x \mathscr{G}_h(0, \Pi_h x_0)\xi \| &= | \alpha \langle \omega_{0h}, \psi_0 \rangle | + \| F_h'(\Pi_h x_0)(\alpha\omega_{0h} + w) \| \geqslant \\ &\geqslant c | \alpha | + c \| w \| - | \alpha | \; \| F_h(\Pi_h x_0)\omega_{0h} \| \\ &\geqslant c \{ | \alpha | + \| w \| \} \geqslant c \| \xi \| \qquad \blacksquare \end{aligned}$$

THEOREM 3 1 *Assume Hypotheses* (3 1) *to* (3 7) *Then there exist positive constants h_0, t_0, α, K and two unique maps $x(t) \in X$ and $x_h(t) \in X_h$, $| t | < t_0$ satisfying respectively the conditions*

$$\mathscr{G}(t, x(t)) = 0, \quad \| x(t) - x_0 \| < \alpha, \quad \textit{for } | t | < t_0, \qquad (3\ 10)$$

$$\mathscr{G}_h(t, x_h(t)) = 0, \quad \| x_h(t) - \Pi_h x_0 \| < \alpha, \quad \textit{for } | t | < t_0 \textit{ and } h < h_0, \qquad (3\ 11)$$

moreover $x(0) = x_0$, $x'(0) \neq 0$, $x(.)$ *and* $x_h(.)$ *are of class* C^p *with bounded derivatives of order* 0, 1, , p *where the bounds are uniform with respect to* t *and* $h < h_0$, *and we have*

$$\lim_{h \to 0} \sup_{|t| < t_0} \| x^{(k)}(t) - x_h^{(k)}(t) \| = 0 , \quad k = 0, 1, \quad , p - 1 , \tag{3 12}$$

$$\| x^{(k)}(t) - x_h^{(k)}(t) \| \leqslant K \sum_{l=0}^{k} \left\{ \left\| \frac{d^l}{dt^l} F_h(\Pi_h x(t)) \right\| + \| (I - \Pi_h) x^{(l)}(t) \| \right\},$$
$$| t | < t_0 , \quad h < h_0 , \quad 0 \leqslant k \leqslant p - 1 \tag{3 13}$$

Proof By Hypotheses (3 1)-(3 7) and by Lemma 3 1, relations (3 10), (3 11) and the boundedness of the derivatives of $x(t)$ and of $x_h(t)$ follow easily from Theorem 2 1 applied to $\mathcal{G}$ and $\mathcal{G}_h$ By applying Theorem 2 2 to $\mathcal{G}_h$, with $W = \mathbb{R}$, $\alpha(t) = t$, $\tilde{s}(t) = \Pi_h x(t)$, we obtain for $0 \leqslant k \leqslant p - 1$, $| t | < t_0$ and some constant c

$$\| x_h^{(k)}(t) - \Pi_h x^{(k)}(t) \| \leqslant c \sum_{l=0}^{k} \left\| \frac{d^l}{dt^l} \mathcal{G}_h(t, \Pi_h x(t)) \right\| ,$$

from which, by using (3 10), (3 13) follows immediately Hypotheses (3 3), (3 5) together with the fact that $F(x(t)) = 0$ imply that the right member of (3 13) converges, for each t, to zero as h tends to zero , in fact, by Theorem 2 3, the convergence is uniform with respect to t , this proves (3 12) ∎

Besides Hypotheses (3 1)-(3 7), we shall assume from now on that we have the following particular situation $X = \mathbb{R} \times V$, where V is a real Banach space, $X_h = \mathbb{R} \times V_h$, where V_h is a subspace of V , an element of $\mathbb{R} \times V$ will be denoted by (λ, u), $\lambda \in \mathbb{R}$, $u \in V$ and we shall write $F(\lambda, u)$ for $F(x)$ and $F_h(\lambda, u)$ for $F_h(x)$, we set $x_0 = (\lambda_0, u_0)$, $x(t) = (\lambda(t), u(t))$, $x_h(t) = (\lambda_h(t), u_h(t))$, $| t | < t_0$, where $x(t)$ and $x_h(t)$ are defined by Theorem 3 1 , we suppose that

$$\text{Range } D_u F(\lambda_0, u_0) \text{ is closed and of codimension 1 in } Y , \tag{3 14}$$

$$D_\lambda F(\lambda_0, u_0) \notin \text{Range } D_u F(\lambda_0, u_0) , \tag{3 15}$$

note that (3 14) and (3 15) are consistent with (3 2) , in fact it is easy to prove (see Appendix I of [8]) that (3 1) implies that $D_u F(\lambda_0, u_0)$ is a Fredholm operator of index 0 so that (3 14) is a consequence of (3 1), (3 2) and (3 15)

Let a $Y \times Y \to \mathbb{R}$ be a *continuous* and *coercive* bilinear form , we assume that F_h is the Galerkin approximation of F with respect to a i e

$$a(F_h(\lambda, u), y) = a(F(\lambda, u), y) \quad \forall v \in Y_h , \quad \forall (\lambda, u) \in X_h , \quad \forall h , \tag{3 16}$$

$$\lim_{h \to 0} \inf_{z \in Y_h} \| y - z \| = 0 , \quad \forall v \in Y \tag{3 17}$$

By differentiating the relation $F(\lambda(t), u(t)) = 0$ at $t = 0$, and by taking in account (3 15), we obtain that $\lambda'(0) = 0$, i e (λ_0, u_0) is a limit point, we shall assume furthermore that it is a turning point i e

$$\lambda''(0) \neq 0 \tag{3 18}$$

Our purpose is to show that the approximate branch parametrized by $(\lambda_h(t), u_h(t))$ has also a turning point for some $t = t_h$ near $t = 0$ and to give an « improved » estimate for the quantity $\lambda_0 - \lambda_h(t_h)$

Hypothesis (3 14) implies the existence of an element $y_0 \in Y$ such that

$$y_0 \neq 0\,, \quad a(D_u\, F(\lambda_0, u_0)\, v, y_0) = 0 \quad \forall v \in V \tag{3 19}$$

THEOREM 3 2 *Assume Hypotheses* (3 1)-(3 7), (3 14)-(3 18) *and suppose that* $p \geqslant 3$ *Then there exist positive constants* t_1 *and* h_1 *such that for* $h < h_1$ *there exists an unique* $t_h \in (-t_1, t_1)$ *with* $\lambda_h'(t_h) = 0$, $\lambda_h''(t_h) \neq 0$, *furthermore there exists a constant c such that, for h small enough, we have*

$$|\lambda_h(t_h) - \lambda_0| \leqslant c\,\{\, \| x_h'(0) - x'(0) \|^2 + \| x_h(0) - x_0 \| \times \\ \times (\| x_h(0) - x_0 \| + \inf_{y \in Y_h} \| y_0 - y \|)\} \tag{3 20}$$

Proof In the following, c will denote a generic positive constant independent of h We use estimate (3 12) of Theorem 3 1 for $k = 1, 2$ Since $\lambda'(0) = 0$, we obtain that $\lim_{h \to 0} \lambda_h'(0) = 0$, by (3 18), there exists $t_1 > 0$ such that, for h small enough, $|\lambda_h''(t)| \geqslant c$, $|t| < t_1$, consequently there exists $h_1 > 0$ and for $h < h_1$ an unique $t_h \in (-t_1, t_1)$ such that $\lambda_h'(t_h) = 0$, furthermore, we have the estimates

$$|t_h| \leqslant c\,|\lambda_h'(0)| = c\,|\lambda_h'(0) - \lambda'(0)| \leqslant c\,\| x_h'(0) - x'(0) \|\,, \tag{3 21}$$

$$\lambda_h(0) = \lambda_h(t_h) - (\lambda_h'(t_h))\, t_h + 0(t_h^2) = \lambda_h(t_h) + 0(t_h^2) \tag{3 22}$$

In order to prove (3 20), let $z_h \in Y_h$ such that $\| y_0 - z_h \| = \inf_{z \in Y_h} \| y_0 - z \|$

By (3 14), (3 15), (3 19), $a(D_\lambda\, F(\lambda_0, u_0), y_0) \neq 0$ and consequently, by (3 17), we shall have that $|a(D_\lambda\, F(\lambda_0, u_0), z_h)| \geqslant c$ for h small enough By (3 16), $a(F(\lambda_h(0), u_h(0)), z_h) = a(F_h(\lambda_h(0), u_h(0)), z_h) = 0$ and by Taylor's expansion we have

$$\begin{aligned} 0 &= a(F(\lambda_h(0), u_h(0)), z_h) - a(F(\lambda_0, u_0), z_h) \\ &= (\lambda_h(0) - \lambda_0)\; a(D_\lambda\, F(\lambda_0, u_0), z_h) + a(D_u\, F(\lambda_0, u_0)\,(u_h(0) - u_0), z_h) + \\ &\qquad + 0(\| x_h(0) - x_0 \|^2)\,, \end{aligned}$$

by (3 19),

$$a(D_u F(\lambda_0, u_0)(u_h(0) - u_0), z_h) = a(D_u F(\lambda_0, u_0)(u_h(0) - u_0), (z_h - y_0))$$

and, for h small enough, we deduce the estimate

$$|\lambda_h(0) - \lambda_0| \leqslant c \| x_h(0) - x_0 \| (\| x_h(0) - x_0 \| + \| z_h - y_0 \|), \qquad (3\ 23)$$

combining (3 21), (3 22) and (3 23), we obtain (3 20) ∎

Remark 3 2 (3 21) and (3 22) are independent of the fact that F_h is a Galerkin approximation of F, whereas (3 23) is independent of the condition $\lambda''(0) \neq 0$

4. BIFURCATION POINTS

Let X and Y be two real Banach spaces, $F \quad X \to Y$ be a C^p mapping with $p \geqslant 2$ and $x_0 \in X$ be such that $F(x_0) = 0$ We suppose that x_0 is a *critical point* of order $n \geqslant 1$ in the following sense

$a)$ $F'(x_0) \quad X \to Y$ is a Fredholm operator of index 1, (4 1)

$b)$ codimension Range $F'(x_0) = n$ (4 2)

Hypotheses (4 1), (4 2) imply that $X_1 = \text{Ker}\, F\,(x_0)$ has dimension $n + 1$ and, if we set $Y_2 = \text{Range}\, F'(x_0)$, there exist two closed subspaces $X_2 \subset X$ and $Y_1 \subset Y$ such that

$$X = X_1 \oplus X_2, \quad Y = Y_1 \oplus Y_2, \qquad (4\ 3)$$

clearly dimension $Y_1 = n$ and the restriction of $F'(x_0)$ to X_2 defines an isomorphism from X_2 onto Y_2

Let $Q \quad Y \to Y_2$ and $I\text{-}Q \quad Y \to Y_1$ be the projectors associated with the decomposition $Y = Y_1 \oplus Y_2$

In order to approximate the solutions of the equation $F(x) = 0$ in a neighborhood of x_0, we consider a family $\{ F_h \}_h$ of C^p operators $F_h \quad X \to Y$, where h is a positive parameter tending to zero, in applications, F_h will appear as a suitable extension of a function defined on a finite dimensional subspace X_h of X with values in a finite dimensional subspace Y_h of Y We suppose

$c)$ for any $0 \leqslant k \leqslant p - 1$ and for any fixed $x, \xi_1, \quad, \xi_k \in X$, we have

$$\lim_{h \to 0} \| F^{(k)}(x)(\xi_1, \xi_2, \quad, \xi_k) - F_h^{(k)}(x)(\xi_1, \xi_2, \quad, \xi_k) \| = 0, \qquad (4\ 4)$$

$d)$ there exist the positive constants δ and c such that

$$\| F_h^{(k)}(x) \| \leqslant c \quad \forall x \in X \quad \text{with} \quad \| x - x_0 \| < \delta, \quad \forall h, \quad 1 \leqslant k \leqslant p, \qquad (4\ 5)$$

e) $QF_h'(x_0)$ is an isomorphism from X_2 onto Y_2 with uniformly bounded inverse with respect to h. (4.6)

(4.4) is a relation of pointwise convergence which can be interpreted as a condition of consistency whereas (4.6) will appear as a condition of stability which will allow, in Theorem 4.1, the Lyapunov-Schmidt procedure for F_h; clearly (4.4) and (4.6) are analogous to Hypotheses (3.5) and (3.7) introduced in the preceding section.

THEOREM 4.1 : *We suppose that Hypotheses a) to e) are satisfied. Then there exist positive constants h_0, ζ, α, K and two unique maps $v : B(0,\zeta) \subset X_1 \to X_2$, $v_h : B(0,\zeta) \subset X_1 \to X_2$ such that :*

$$QF(x_0 + \sigma + v(\sigma)) = 0\,, \quad \| v(\sigma) \| < \alpha \quad \forall \sigma \in B(0,\zeta)\,, \tag{4.7}$$

$$QF_h(x_0 + \sigma + v_h(\sigma)) = 0\,, \quad \| v_h(\sigma) \| < \alpha \quad \forall \sigma \in B(0,\zeta)\,, \quad \forall h < h_0\,; \tag{4.8}$$

v and v_h are C^p mappings with bounded derivatives of order 0, 1, ..., *p where the bounds are uniform with respect to $\sigma \in B(0,\zeta)$ and $h < h_0$; furthermore, we have :*

$$\lim_{h\to 0} \sup_{\sigma \in B(0,\zeta)} \| v^{(k)}(\sigma) - v_h^{(k)}(\sigma) \| = 0\,, \quad k = 0, 1, \dots, p-1\,, \tag{4.9}$$

$$\| v^{(k)}(\sigma) - v_h^{(k)}(\sigma) \| \leqslant K \sum_{j=0}^{k} \left\| \frac{d^j}{d\sigma^j} QF_h(x_0 + \sigma + v(\sigma)) \right\|, \tag{4.10}$$
$$0 \leqslant k \leqslant p-1\,, \quad \sigma \in B(0,\zeta)\,, \quad h < h_0\,.$$

Proof : We apply Theorems 2.1 and 2.2 to $G : X_1 \times X_2 \to Y_2$ and $G_h : X_1 \times X_2 \to Y_2$, where $G(\sigma, v) = QF(x_0 + \sigma + v)$, $G_h(\sigma, v) = QF_h(x_0 + \sigma + v)$ from which (4.7), (4.8) and (4.10) follow immediately; then (4.9) is a consequence of Theorem 2.3, of (4.10), of Hypotheses (4.4), (4.5) and from the fact that X_1 is finite dimensional. ■

By Theorem 4.1, the equations $F(x) = 0$ and $F_h(x) = 0$ are reduced, in a neighborhood of x_0, to the equation $f(\sigma) = 0$ and $f_h(\sigma) = 0$ in a neighborhood of 0, where f and f_h are the *bifurcation functions* defined by :

$$f : B(0,\zeta) \to Y_1\,, \quad f(\sigma) = (I\text{-}Q)\, F(x_0 + \sigma + v(\sigma))\,, \tag{4.11}$$

$$f_h : B(0,\zeta) \to Y_1\,, \quad f_h(\sigma) = (I\text{-}Q)\, F_h(x_0 + \sigma + v_h(\sigma))\,. \tag{4.12}$$

The following relations are either obvious or easy to verify :

$$v(0) = 0\,, \quad v'(0) = 0\,, \quad f(0) = 0\,, \quad f'(0) = 0\,, \tag{4.13}$$

$$F'(x_0)\,\xi = 0 \quad \forall \xi \in X_1\,, \quad (I\text{-}Q)\, F'(x_0) = 0\,. \tag{4.14}$$

We now introduce the following new hypotheses $p \geqslant 4$, there exist $\sigma_0 \in X_1$ and the integer q with $2 \leqslant q \leqslant p/2$ such that

f) $f^{(k)}(0) = 0 \quad 2 \leqslant k \leqslant q - 1$, (4 15)

g) $f^{(q)}(0)\,\sigma_0^q = 0$, (4 16)

h) the relations $\sigma \in X_1$, $f^{(q)}(0)\,\sigma_0^{q-1}\,\sigma = 0$ imply the existence of $\tau \in \mathbb{R}$ with $\sigma = \tau\sigma_0$, (4 17)

i) $f_h^{(k)}(0) = 0 \quad 0 \leqslant k \leqslant q - 1$ (4 18)

Remark 4 1 Consider the conditions α) $(I\text{-}Q)\,F^{(k)}(x_0) = 0, 2 \leqslant k \leqslant q - 1$, β) the restriction of $F^{(k)}(x_0)$ to X_1^k vanishes for $2 \leqslant k \leqslant q - 1$, then it is easy to verify that α) *or* β) is a sufficient condition for obtaining (4 15), furthermore, if α) *or* β) is satisfied, then $f^{(q)}(0)$ is equal to the restriction of $(I\text{-}Q)\,F^{(q)}(x_0)$ to X_1^q, which allows to express (4 16) and (4 17) in terms of F directly

Remark 4 2 We could replace (4 18) by the more general hypothesis γ) there exists $\eta_h \in X_1$ such that $\lim\limits_{h \to 0} \eta_h = 0$ and $f_h^{(k)}(\eta_h) = 0$ for $0 \leqslant k \leqslant q - 1$, in fact, with minor modifications, all the following results of this section would remain valid However, it is possible to reduce γ) to (4 18) in the following way, let $z_h = \eta_h + v_h(\eta_h)$ and $\tilde{F}_h(x) = F_h(x + z_h)$, then by applying Theorem 4 1 to $\tilde{F}_h$ we obtain a map $\tilde{v}_h \quad B(0, \zeta) \subset X_1 \to X_2$ such that

$$Q\tilde{F}_h(x_0 + \sigma + \tilde{v}_h(\sigma)) \equiv 0$$

and a new bifurcation function $\tilde{f}_h(\sigma) = (I\text{-}Q)\,\tilde{F}_h(x_0 + \sigma + \tilde{v}_h(\sigma))$, it is possible to verify that $\tilde{f}_h(\sigma) = f_h(\sigma + \eta_h)$ and consequently, by γ), $\tilde{f}_h^{(k)}(0) = 0$ for $0 \leqslant k \leqslant q - 1$, note also that $\tilde{F}_h(x_0) = 0$

Under the above hypotheses, we shall show the existence of a C^{p-q} branch of solutions of the equation $F(x) = 0$ passing through x_0 and of a corresponding approximate branch for the equation $F_h(x) = 0$, the « exact » branch will be parametrized by a function $x(t) \in X$ such that $x(0) = x_0$, $x'(0) = \sigma_0$ To this end, let $\psi_0 \in X^*$ be such that $\langle \sigma_0, \psi_0 \rangle \neq 0$, where X^* denotes the dual of X and $\langle \cdot, \cdot \rangle$ the duality pairing between X and X^*, we introduce the following mappings

$$\mathscr{G}(t, \sigma) = \left(\langle \sigma - \sigma_0, \psi_0 \rangle, \frac{1}{t^q} f(t\sigma)\right), \tag{4 19}$$

$$\mathscr{G}_h(t, \sigma) = \left(\langle \sigma - \sigma_0, \psi_0 \rangle, \frac{1}{t^q} f_h(t\sigma)\right), \tag{4 20}$$

$\mathcal{G}$ and $\mathcal{G}_h$ (h small enough) are defined on some neighborhood $\Omega \subset \mathbb{R} \times X_1$ of $(0, \sigma_0)$ with values in $\mathbb{R} \times Y_1$; Ω is independent of h ; by (4.13) and Hypotheses (4.15) and (4.18), $\mathcal{G}$ and $\mathcal{G}_h$ are C^{p-q} mappings where we recall that $p - q \geqslant 2$.

THEOREM 4.2 : *We assume that Hypotheses a) to i) are satisfied. Then there exist positive constants h_0, t_0, β, M and two unique maps $\sigma(.) : (-t_0, t_0) \to X_1$, $\sigma_h(.) : (-t_0, t_0) \to X_1$ such that*

$$\mathcal{G}(t, \sigma(t)) = 0\,, \quad \| \sigma(t) - \sigma_0 \| \leqslant \beta\,, \quad |t| < t_0\,, \tag{4.21}$$

$$\mathcal{G}_h(t, \sigma_h(t)) = 0\,, \quad \| \sigma_h(t) - \sigma_0 \| \leqslant \beta\,, \quad |t| < t_0\,, \quad h < h_0\,; \tag{4.22}$$

$\sigma(.)$ and $\sigma_h(.)$ are C^{p-q} mappings with bounded derivatives of order $0, 1, \ldots, p - q$ where the bounds are uniform with respect to $|t| < t_0$ and $h < h_0$; furthermore $\mathcal{G}(0, \sigma_0) = 0$ and we have for $0 \leqslant k \leqslant p - 2q + 1$:

$$\lim_{h \to 0} \sup_{|t| < t_0} \left\| \frac{d^k}{dt^k} (t\sigma(t) - t\sigma_h(t)) \right\| = 0\,, \tag{4.23}$$

$$\sup_{|t| < t_0} \left\| \frac{d^k}{dt^k} (t\sigma(t) - t\sigma_h(t)) \right\| \leqslant M \sum_{j=0}^{k+q-1} \sup_{|t| < t_0} \left\| \frac{d^j}{dt^j} f_h(t\sigma(t)) \right\|\,, \quad h < h_0\,. \tag{4.24}$$

Proof : By (4.13) and (4.15) we have for any $\sigma \in X_1$:

$$D_\sigma \mathcal{G}(0, \sigma_0)\, \sigma = \left(\langle \sigma, \psi_0 \rangle, \frac{1}{(q-1)!} f^{(q)}(0)\, \sigma_0^{q-1}\, \sigma \right);$$

by (4.17) and the fact that $\langle \sigma_0, \psi_0 \rangle \neq 0$, we see that $D_\sigma \mathcal{G}(0, \sigma_0)$ is injective ; since X_1 and $\mathbb{R} \times Y_1$ have the same finite dimension $n + 1$, we conclude that $D_\sigma \mathcal{G}(0, \sigma_0)$ defines an isomorphism between these two spaces. Moreover, by (4.13), (4.15) and (4.16), we have $\mathcal{G}(0, \sigma_0) = 0$. By (4.4), (4.5), (4.9) and Theorem 2.3, we obtain

$$\lim_{h \to 0} \sup_{\sigma \in B(0,\zeta)} \| f^{(k)}(\sigma) - f_h^{(k)}(\sigma) \| = 0\,, \quad 0 \leqslant k \leqslant p - 1\,; \tag{4.25}$$

furthermore $f_h^{(p)}$ is bounded on $B(0, \zeta)$ uniformly with respect to h ; by (4.13), (4.15), (4.18) and Taylor's formula, we obtain :

$$\mathcal{G}(t, \sigma) - \mathcal{G}_h(t, \sigma) = \left(0, \frac{1}{(q-1)!} \int_0^1 (1 - s)^{q-1} \left(f^{(q)}(st\sigma) - f_h^{(q)}(st\sigma) \right) \sigma^q \, ds \right); \tag{4.26}$$

together with (4.25), this proves

$$\lim_{h \to 0} \sup_{(t,\sigma) \in \Omega} \| \mathcal{G}^{(k)}(t, \sigma) - \mathcal{G}_h^{(k)}(t, \sigma) \| = 0 , \quad 0 \leqslant k \leqslant p - q - 1 , \quad (4.27)$$

and that $\mathcal{G}_h^{(p-q)}$ is bounded on Ω uniformly with respect to h. In particular we have : $\lim_{h \to 0} \| D_\sigma \mathcal{G}(0, \sigma_0) - D_\sigma \mathcal{G}_h(0, \sigma_0) \| = 0$, which shows that, for h small enough, $D_\sigma \mathcal{G}_h(0, \sigma_0)$ is an isomorphism from X_1 onto $\mathbb{R} \times Y_1$ with inverse uniformly bounded with respect to h.

In the following, c will denote a generic constant independent of h. By using the above preliminary results, in particular (4.25), we can apply the implicit function Theorem 2.1 to $\mathcal{G}$ and $\mathcal{G}_h$, from which we deduce (4.21), (4.22) and the boundedness of the first $(p - q)$ derivatives of $\sigma(.)$ and of $\sigma_h(.)$. Applying furthermore Theorem 2.2 to $\mathcal{G}_h$, we obtain, for h small enough, $0 \leqslant k \leqslant p - q - 1$ and $| t | < t_0$, the estimate

$$\| \sigma^{(k)}(t) - \sigma_h^{(k)}(t) \| \leqslant c \sum_{j=0}^{k} \left\| \frac{d^j}{dt^j} \mathcal{G}_h(t, \sigma(t)) \right\| ; \quad (4.28)$$

since $\mathcal{G}(t, \sigma(t)) = 0$, it follows that $\langle \sigma(t) - \sigma_0, \psi_0 \rangle = 0$ and consequently that $\mathcal{G}_h(t, \sigma(t)) = (0, t^{-q} e_h(t))$, where we set $e_h(t) = f_h(t\sigma(t))$; then, by (4.28), we have for h small enough, $0 \leqslant k \leqslant p - q - 1$ and $| t | < t_0$:

$$\begin{aligned} \left\| \frac{d^k}{dt^k} (t\sigma(t) - t\sigma_h(t)) \right\| &\leqslant c \left\{ | t | \left\| \frac{d^k}{dt^k} (t^{-q} e_h(t)) \right\| + \right. \\ &\left. + \sum_{j=0}^{k-1} \left\| \frac{d^j}{dt^j} (t^{-q} e_h(t)) \right\| \right\} \leqslant c \left\{ \left\| \frac{d^k}{dt^k} (t^{-q+1} e_h(t)) \right\| \right. \\ &\left. + \sum_{j=0}^{k-1} \left\| \frac{d^j}{dt^j} (t^{-q} e_h(t)) \right\| \right\} ; \end{aligned} \quad (4.29)$$

using once more Hypothesis (4.18) and Taylor's formula, we have for $0 \leqslant m \leqslant q$ and, in particular for $m = q - 1, m = q$:

$$\frac{1}{t^m} e_h(t) = \frac{1}{(m-1)!} \int_0^1 (1 - s)^{m-1} e_h^{(m)}(st) \, ds ,$$

by replacing this expression in (4.29), we obtain the desired estimate (4.24). ■

We now can define the parametrizations of the two above announced branches of solutions for the equations $F(x) = 0$ and $F_h(x) = 0$. We set :

$$x(.) : (-t_0, t_0) \to X ; \quad x(t) = x_0 + t\sigma(t) + v(t\sigma(t)) ; \quad (4.30)$$

$$x_h(.) : (-t_0, t_0) \to X ; \quad x_h(t) = x_0 + t\sigma_h(t) + v_h(t\sigma_h(t)) , \quad (4.31)$$

where the mappings $v(.)$, $v_h(.)$, $\sigma(.)$ and $\sigma_h(.)$ are defined by Theorems 4.1 and 4.2.

THEOREM 4.3 : *We assume that Hypotheses a) to i) are satisfied; let $x(.)$ and $x_h(.)$ be given by* (4.30) *and* (4.31). *Then there exist positive constants h_0 and L such that*

i) *$x(.)$ and $x_h(.)$ are C^{p-q} mappings with uniformly bounded derivatives (with respect to $|t| < t_0$ and $h < h_0$) of order $0, 1, \ldots, p-q \geqslant 2$;*

ii) $F(x(t)) = 0,\ F_h(x_h(t)) = 0,\ |t| < t_0,\ h < h_0;\ x(0) = x_0;\ x'(0) = \sigma_0;$ (4.32)

iii) $$\lim_{h\to 0} \sup_{|t|<t_0} \| x^{(k)}(t) - x_h^{(k)}(t) \| = 0, \quad 0 \leqslant k \leqslant p - 2q + 1; \tag{4.33}$$

iv) *for $0 \leqslant k \leqslant p - 2q + 1$ and $h < h_0$, we have the error estimate*

$$\sup_{|t|<t_0} \| x^{(k)}(t) - x_h^{(k)}(t) \| \leqslant L \sum_{j=0}^{k+q-1} \sup_{|t|<t_0} \left\| \frac{d^j}{dt^j} F_h(x(t)) \right\| . \tag{4.34}$$

Proof : Clearly, by Theorems 4.1 and 4.2, we have to verify only (4.34). In the following, c will denote a generic constant independent of h. By using the boundedness of the derivatives of v_h of order $\leqslant p$ (Theorem 4.1), we can write for $0 \leqslant k \leqslant p - 2q + 1$:

$$\begin{aligned} \| x^{(k)}(t) - x_h^{(k)}(t) \| \leqslant & \left\{ \left\| \frac{d^k}{dt^k} (t\sigma(t) - t\sigma_h(t)) \right\| + \left\| \frac{d^k}{dt^k} (v(t\sigma(t)) - v_h(t\sigma(t))) \right\| \right. \\ & \left. + \left\| \frac{d^k}{dt^k} (v_h(t\sigma(t)) - v_h(t\sigma_h(t))) \right\| \right\} \leqslant c \left\{ \left\| \frac{d^k}{dt^k} (t\sigma(t) - t\sigma_h(t)) \right\| \right. \\ & \left. + \left\| \frac{d^k}{dt^k} (v(t\sigma(t)) - v_h(t\sigma(t))) \right\| \right\} . \end{aligned} \tag{4.35}$$

As for the proof of Theorem 4.1, we now apply Theorem 2.2 to $G : X_1 \times X_2 \to Y_2$, $G(\sigma, v) = QF_h(x_0 + \sigma + v)$; with $\tilde{g}(t) = v_h(t\sigma(t))$ and $\tilde{s}(t) = v(t\sigma(t))$, we obtain, for $0 \leqslant j \leqslant p - q$ and $|t| < t_0$, the estimate

$$\begin{aligned} \left\| \frac{d^j}{dt^j} (v(t\sigma(t)) - v_h(t\sigma(t))) \right\| & \leqslant c \sum_{l=0}^{j} \left\| \frac{d^l}{dt^l} QF_h(x(t)) \right\| \leqslant \\ & \leqslant c \sum_{l=0}^{j} \left\| \frac{d^l}{dt^l} F_h(x(t)) \right\| . \end{aligned} \tag{4.36}$$

By (4.24), the definition (4.12) of f_h and (4.5), we have :

$$\left\| \frac{d^k}{dt^k}(t\sigma(t) - t\sigma_h(t)) \right\| \leqslant c \sum_{j=0}^{k+q-1} \sup_{|\tau|<t_0} \left\| \frac{d^j}{d\tau^j} F_h(x_0 + \tau\sigma(\tau) + v_h(\tau\sigma(\tau))) \right\| \leqslant$$

$$\leqslant c \sum_{j=0}^{k+q-1} \left\{ \sup_{|\tau|<t_0} \left\| \frac{d^j}{d\tau^j} F_h(x(\tau)) \right\| + \sup_{|\tau|<t_0} \left\| \frac{d^j}{d\tau^j} (v(\tau\sigma(\tau)) - v_h(\tau\sigma(\tau))) \right\| \right\} : \tag{4.37}$$

(4.34) is then a direct consequence of (4.35), (4.36) and (4.37). ■

Remark 4.3 : If $\sigma_0 \neq 0$ satisfies condition (4.16), we shall say that σ_0 is a *characteristic ray*; if σ_0 satisfies conditions (4.16) and (4.17), we shall say that σ_0 is a *non-degenerate characteristic ray.* Let $\Sigma \subset X_1$ be the set of characteristic rays with norm 1; by using a compactness argument bounded to the fact that X_1 is finite dimensional, it is easy to establish the following result : if all characteristic rays are non-degenerate, then Σ is a finite set.

Remark 4.4 : Let us denote by P_1 and P_2 the projectors associated with the decomposition $X = X_1 \oplus X_2$; let

$$\Gamma = \{ x(t) \mid |t| < t_0 \}, \ \Gamma_h = \{ x_h(t) \mid |t| < t_0 \}$$

be the branches of solutions given by Theorem 4.3; for the constant β introduced in Theorem 4.2, we consider the cone

$$C = \{ \sigma \in X_1 \mid \| \langle \sigma_0, \psi_0 \rangle \sigma - \langle \sigma, \psi_0 \rangle \sigma_0 \| < \beta \, | \langle \sigma, \psi_0 \rangle | \} ;$$

by the uniqueness of the maps $\sigma(.)$ and $\sigma_h(.)$ in Theorem 4.2, it is fairly easy to establish the existence of positive constants γ and h_0 such that

$$\{ x \in X \mid F(x) = 0 , \ \| x - x_0 \| < \gamma , \ P_1(x - x_0) \in C \} \subset \Gamma ,$$

$$\{ x \in X \mid F_h(x) = 0 , \ \| x - x_0 \| < \gamma , \ P_1(x - x_0) \in C \} \subset \Gamma_h , \quad h < h_0 .$$

Furthermore, let Σ be the set of characteristic rays with norm 1 and suppose that all characteristic rays are non-degenerate; by Remark 4.3, Σ is a finite set with elements $\xi, \xi_2, \ldots, \xi_n$, say; by Theorem 4.3, to each ξ_i corresponds a branch of solutions Γ_i of the equation $F(x) = 0$ and a branch Γ_{ih} of the equation $F_h(x) = 0$, for h small enough; note that for each i corresponds j such that $\xi_i = - \xi_j$, so that $\Gamma_i = \Gamma_j$ and $\Gamma_{ih} = \Gamma_{jh}$; it is then possible to show by elementary means the existence of positive constants γ and h_0 such that

$$\{ x \in X \mid F(x) = 0 , \quad \| x - x_0 \| < \gamma \} \subset \bigcup_{i=1}^{n} \Gamma_i ,$$

$$\{ x \in X \mid F_h(x) = 0 , \quad \| x - x_0 \| < \gamma \} \subset \bigcup_{i=1}^{n} \Gamma_{ih} , \quad h < h_0$$

We shall conclude this section by discussing the particular case of *simple bifurcation points*, i e essentially the case $n = 1, q = 2$ Note that our analysis will include implicitly the treatment the « double limit bifurcation point » [5] Specifically we shall suppose, from now on, that

j) $n =$ codimension Range $F'(x_0) = 1, p \geqslant 4$, (4 38)

with $q = 2$, then Hypothesis (4 15) is void, furthermore, by Remark 4 1, $f''(0)(\xi, \eta) = (I - Q) F''(x_0)(\xi, \eta)$ for all $\xi, \eta \in X_1$, so that (4 16) and (4 17) are equivalent, in this case, to the existence of $\sigma_0 \in X_1$ such that

k) $(I - Q) F''(x_0)(\sigma_0, \sigma_0) = 0$, (4 39)

l) the relations $\sigma \in X_1$, $(I - Q) F''(x_0)(\sigma_0, \sigma) = 0$ imply the existence of $\tau \in \mathbb{R}$ with $\sigma = \tau\sigma_0$ (4 40)

$X_1 = \operatorname{Ker} F'(x_0)$ has dimension 2 and Y_1 has dimension 1, let e_1, e_2 be a basis of X_1 and let $g \neq 0$ be an element of Y_1, for any $\sigma = \varepsilon_1 e_1 + \varepsilon_2 e_2 \in X_1$, ε_1 and $\varepsilon_2 \in \mathbb{R}$, we can write

$$f''(0)(\sigma, \sigma) = (I - Q) F''(x_0)(\sigma, \sigma) = R(\varepsilon_1, \varepsilon_2) g , \tag{4 41}$$

where $R \quad \mathbb{R}^2 \to \mathrm{R}$ is a quadratic form, as easily verified, (4 39) together with (4 40) are equivalent to the property that R is indefinite and non-degenerate, i e that the determinant of the matrix associated with R is negative, consequently, (4 39) and (4 40) imply the existence of a non-degenerate characteristic ray σ_1 linearly independent of σ_0 such that any characteristic ray is parallel to σ_0 or to σ_1

In connection with Remark 4 2, we state the following result

THEOREM 4 4 *Let Hypotheses a) to e) and j) to l) be satisfied, we suppose the existence of two C^1 mappings $x(.)$ and $x_h(.)$ $(-t_0, t_0) \subset \mathbb{R} \to X_1$, $t_0 > 0$, such that $F(x(t)) = F_h(x_h(t)) = 0$ for $| t | < t_0$, $x(0) = x_0$, $x'(0) = \sigma_0$,*

$$\lim_{h \to 0} \sup_{|t| < t_0} \| x^{(k)}(t) - x_h^{(k)}(t) \| = 0 \quad \textit{for} \quad k = 0, 1$$

Then there exist positive constants h_0 *and* θ *and, for* $h < h_0$, *an unique point* $\eta_h \in X_1$ *such that*

$$\| \eta_h \| < \theta, \; f_h(\eta_h) = 0 \,, \quad f_h'(\eta_h) = 0 \quad \text{for} \quad h < h_0 \quad \text{and} \quad \lim_{h \to 0} \eta_h = 0 \,, \tag{4.42}$$

furthermore, there exists a constant c such that

$$\| \eta_h \| < c \, \| f_h'(0) \| \tag{4.43}$$

Proof Recalling the definitions (4 11) and (4 12) of f and f_h, we set $\omega = f'$ and $\omega_h = f_h'$ $B(0, \zeta) \to \mathscr{L}(X_1, Y_1)$, the non-degenerancy of the quadratic form R in (4 41) implies that $\omega'(0) \in \mathscr{L}(X_1, \mathscr{L}(X_1, Y_1))$ is an isomorphism, by using (4 25), which is valid under the sole Hypotheses $a)$ to $e)$, (4 13) and Theorems 2 1, 2 2, we obtain for h small enough and in some neighborhood of 0 the existence of an unique $\eta_h \in X_1$ satisfying the relation $\omega(\eta_h) = f'(\eta_h) = 0$ and the estimate (4 43) It remains to show that $f_h(\eta_h) = 0$, to this end, we decompose $x(t)$ in the form $x(t) = x_0 + \theta(t) + w(t)$ where $\theta(t) \in X_1$, $w(t) \in X_2$, since $F'(x_0)\, x'(0) = 0$, we have $w'(0) = 0$ and $\theta'(0) = \sigma_0$, in the same way we write $x_h(t) = x_0 + \theta_h(t) + w_h(t)$, $\theta_h(t) \in X_1$, $w_h(t) \in X_2$ and we set $\varphi_h(t) = f_h'(\theta_h(t))\, \xi$ where $\xi \in X_1$ is a fixed element linearly independent of σ_0, by (4 13), (4 25) and (4 40) we have $\lim_{h \to 0} \varphi_h(0) = 0$,

$$\lim_{h \to 0} \varphi_h'(0) = f''(0)\,(\xi, \sigma_0) \neq 0 \,,$$

since φ_h' and φ_h'' are bounded in a neighborhood of 0, uniformly with respect to h, there exists, for h small enough, t_h such that $\varphi_h(t_h) = 0$ with $\lim_{h \to 0} t_h = 0$, clearly $f_h(\theta_h(t)) = 0$ which implies that $f_h'(\theta_h(t_h))\, \theta_h'(t_h) = 0$ and $f_h(\theta_h(t_h)) = 0$, since $\lim_{h \to 0} \theta_h'(t_h) = \sigma_0$, it follows that $f_h'(\theta_h(t_h)) = 0$ for h small enough, by the uniqueness of η_h we have $\eta_h = \theta_h(t_h)$ for h small enough ■

Remark 4 5 In Theorem 4 4, the existence of the map $x(.)$ is clearly insured by Theorem 4 3, the existence of the map x_h can be obtained, in « practical » situations, in two cases $a)$ when $x(.)$ parametrizes a « trivial branch » and then $x_h(t) = x(t)$, $b)$ in presence of symmetries (for more details see [3], [8])

THEOREM 4 5 *Let Hypotheses a) to e) and j) to l) be satisfied and let* f_h *be defined by* (4 12), *we suppose the existence, for h small enough, of* $\eta_h \in X_1$ *such that* $f_h(\eta_h) = 0$, $f_h'(\eta_h) = 0$ *and* $\lim_{h \to 0} \eta_h = 0$ *Then, for some* $t_0 > 0$ *and* $h_0 > 0$ *we have*

i) *there exist* C^{p-2} *mappings* $x_i(.) : (-t_0, t_0) \to X$, $i = 0, 1$, *such that*

$$F(x_i(t)) = 0 \,, \ |t| < t_0 \,, \ x_i(0) = x_0 \,, \ x_i'(0) \neq 0 \,, \ i = 0, 1 \,, \ x_0'(0) \neq x_1'(0) \,; \tag{4.44}$$

ii) *there exist* C^{p-2} *mappings* $x_{ih}(.) : (-t_0, t_0) \to X, i = 0, 1$, *such that*

$$F_h(x_{ih}(t)) = 0 \,, \quad |t| < t_0 \,, \quad h < h_0 \,, \quad i = 0, 1 \,, \tag{4.45}$$

$$\lim_{h \to 0} \sup_{|t| < t_0} \| x_i^{(k)}(t) - x_{ih}^{(k)}(t) \| = 0 \,, \quad 0 \leqslant k \leqslant p - 3 \,, \quad i = 0, 1 \,; \tag{4.46}$$

iii) *there exists a constant c such that*

$$\| \eta_h \| < c \sum_{i=0}^{1} \sum_{j=0}^{1} \left\| \frac{d^j}{dt^j} F_h(x_i(t)) |_{t=0} \right\| \,, \quad h < h_0 \,; \tag{4.47}$$

iv) *there exists a constant c such that for* $0 \leqslant k \leqslant p - 3$, $i = 0, 1$;

$$\sup_{|t| < t_0} \| x_i^{(k)}(t) - x_{ih}^{(k)}(t) \| \leqslant$$
$$\leqslant c \left\{ \| \eta_h \| + \sum_{j=0}^{k+1} \sup_{|t| < t_0} \left\| \frac{d^j}{dt^j} F_h(x_i(t)) \right\| \right\} \,, \quad h < h_0 \,. \tag{4.48}$$

Proof: In the following, c will denote a generic constant. Following Remark 4.2, we set $z_h = \eta_h + v_h(\eta_h)$, $\tilde{F}_h(x) = F_h(x + z_h)$, where v_h is defined by (4.8); since $\lim_{h<0} \eta_h = 0$, $\tilde{F}_h$ will satisfy the same Hypotheses c), d) and e) as F_h, for h small enough; by applying Theorem 4.1 to $\tilde{F}_h$, we obtain a bifurcation function $\tilde{f}_h$ such that $\tilde{f}_h(0) = 0$, $\tilde{f}_h'(0) = 0$. By (4.39) and (4.40), there exist two lineary independent characteristic non-degenerate rays σ_0 and σ_1. By applying Theorem 4.3 to F and $\tilde{F}_h$, σ_0 and σ_1, we obtain C^{p-2} mappings $x_i(.)$, $\tilde{x}_{ih}(.)$: $(-t_0, t_0) \to X$, $i = 0, 1$ verifying (4.44), $x_i'(0) = \sigma_i$ and the relations

$$\tilde{F}_h(\tilde{x}_{ih}(t)) = 0 \quad \text{for} \quad |t| < t_0 \,, \quad i = 0, 1 \,;$$

furthermore, for h small enough, $0 \leqslant k \leqslant p - 3$ and $i = 0, 1$, we have

$$\lim_{h \to 0} \sup_{|t| < t_0} \| x_i^{(k)}(t) - \tilde{x}_{ih}^{(k)}(t) \| = 0 \,, \tag{4.49}$$

$$\sup_{|t| < t_0} \| x_i^{(k)}(t) - \tilde{x}_{ih}^{(k)}(t) \| \leqslant c \sum_{j=0}^{k+1} \sup_{|t| < t_0} \left\| \frac{d^j}{dt^j} \tilde{F}_h(x(t)) \right\| \,. \tag{4.50}$$

Let us define the C^{p-2} mappings $x_{ih}(.) : (-t_0, t_0) \to X, i = 0, 1$, by

$$x_{ih}(t) = \tilde{x}_{ih}(t) + z_h \,, \quad |t| < t_0 \,, \quad i = 0, 1 \,; \tag{4.51}$$

clearly (4 45) is then satisfied By (4 10),

$$\| v_h(\eta_h) - v(\eta_h) \| \leqslant c \, \| F_h(x_0 + \eta_h + v(\eta_h) \| ,$$

since $v(0) = 0$ (4 13), with Hypothesis (4 5), we obtain the estimate

$$\| v_h(\eta_h) \| \leqslant c \, \{ \| \eta_h \| + \| F_h(x_0) \| \}$$

and consequently $\| z_h \| \leqslant c \, \{ \| \eta_h \| + \| F_h(x_0) \| \}$, in particular, by Hypothesis (4 4), we have $\lim_{h \to 0} z_h = 0$, then (4 46) and (4 48) follow immediately from (4 49) and (4 50) It remains to prove (4 47), clearly the hypotheses of Theorem 4 4 are satisfied and by (4 43), it suffices to estimate $\| f_h'(0) \|$, since σ_0 and σ_1 form a fixed basis of X_1, the proof of Theorem 4 5 will be achieved, if we show, for h small enough, the estimate

$$\| f_h'(0) \, \sigma_i \| \leqslant c \sum_{j=0}^{1} \left\| \frac{d^j}{dt^j} F_h(x_i(t)) \, |_{t=0} \right\| , \quad i = 0, 1 , \tag{4 52}$$

we prove (4 52) for $i = 0$, by (4 30) and (4 32), $x_0(t)$ is of the form

$$x_0(t) \equiv x(t) = x_0 + t\sigma(t) + v(t\sigma(t)) ,$$

with $\sigma(0) = \sigma_0$, by definition (4 12) of f_h we have

$$\| f_h'(0) \, \sigma_0 \| = \left\| \frac{d}{dt} f_h(t\sigma(t)) \, |_{t=0} \right\| \leqslant c \left\| \frac{d}{dt} F_h(x_0 + t\sigma(t) + v_h(t\sigma(t))) \, |_{t=0} \right\| , \tag{4 53}$$

by using the estimate (4 36) (which is valid without Hypothesis i)), we easily deduce (4 52) from (4 53) ■

5. AN EXAMPLE

Let $\Omega = (0, 1) \times (0, 1)$ be the unit square in $\mathbb{R}^2 . H_0^1 = H_0^1(\Omega)$ will denote the set of square integrable functions on Ω, vanishing on $\partial\Omega$ and possessing square integrable first partial derivatives

We consider the classical nonlinear eigenvalue problem of finding $(\lambda, u) \in \mathbb{R} \times H_0^1$ such that

$$- \Delta u - \lambda u + u^3 = 0 \text{ in } \Omega \tag{5 1}$$

Let us introduce the symmetric bilinear forms a and b by

$$a \quad H_0^1 \times H_0^1 \to \mathbb{R}\,, \quad a(u, v) = \int_\Omega (\partial_x u\, \partial_x v + \partial_y u\, \partial_y v)\,, \tag{5 2}$$

$$b \quad L^2 \times L^2 \to \mathbb{R}\,, \quad b(u, v) = \int_\Omega u.v\,, \tag{5 3}$$

clearly (5 1) is equivalent to the problem of finding $(\lambda, u) \in \mathbb{R} \times H_0^1$ such that

$$a(u, v) - \lambda b(u, v) + b(u^3, v) = 0 \qquad \forall v \in H_0^1 \tag{5 4}$$

For n, positive integer, we divide the closure $\overline{\Omega}$ of Ω in n^2 closed equal squares of side $h = 1/n$, let $\mathscr{D}_h$ be the set of these squares and

$$V_h = \{ f \in H_0^1 \quad f|_K \in Q_1 \quad \forall K \in \mathscr{D}_h \}\,, \tag{5 5}$$

where Q_1 is the set of polynomials of the form $axy + bx + cv + d$ We introduce the a-projector Π_h, the interpolatory projector P_h and the symmetric bilinear form b_h defined by

$$\Pi_h \quad H_0^1 \to V_h\,, \quad a(\Pi_h u - u, v) = 0 \quad \forall v \in V_h\,, \quad u \in H_0^1\,, \tag{5 6}$$

$$P_h \quad C_0^0 \to V_h\,, \quad (P_h u)(ih, jh) = u(ih, jh)\,, \quad 1 \leqslant i\,, j \leqslant n - 1\,, \quad u \in C_0^0\,, \tag{5 7}$$

$$b_h \quad C_0^0 \times C_0^0 \to \mathbb{R}\,, \quad b_h(u, v) = \int_\Omega P_h(u.v)\,, \tag{5 8}$$

here, $C_0^0 = C_0^0(\overline{\Omega})$ denotes the set of continuous functions on $\overline{\Omega}$, vanishing on $\partial\Omega$ As an approximation of (5 1) or (5 4), we consider the problem of finding $(\lambda, u) \in \mathbb{R} \times V_h$ such that

$$a(u, v) - \lambda b_h(u, v) + b_h(u^3, v) = 0 \quad \forall v \in V_h \tag{5 9}$$

Note that, if $f \in C_0^0$, we have

$$\int_\Omega P_h f = h^2 \sum_{i\,j=1}^{n-1} f(ih, jh)\,, \tag{5 10}$$

so that (5 9) is equivalent to an explicit system of $(n - 1)^2$ nonlinear equations for the $(n - 1)^2 + 1$ unknowns λ, $u(ih, jh)$, $1 \leqslant i, j \leqslant n - 1$

In the following, we shall be concerned with the approximation of solutions of Problem (5.1) in the neighborhood of $(\lambda_0, u_0) = (5\,\Pi^2, 0) \in \mathbb{R} \times H_0^1$; let $Y_1 \subset H_0^1$ be the two-dimensional subspace spanned by

$$\varphi_1(x, y) = \frac{2}{\sqrt{\lambda_0}} \sin \Pi x \sin 2\,\Pi y\,, \quad \varphi_2(x, y) = \frac{2}{\sqrt{\lambda_0}} \sin 2\,\Pi x \sin \Pi y\,: \quad (5.11)$$

as easily verified, we have $a(\varphi_i, \varphi_j) = \delta_{ij}$, $1 \leqslant i, j \leqslant 2$; furthermore $\lambda_0 = 5\,\Pi^2$ is a double eigenvalue of the problem of finding $(\lambda, u) \in \mathbb{R} \times H_0^1$ such that $-\Delta u = \lambda u$; Y_1 is the corresponding eigenspace. Concerning the approximate problem of finding $(\lambda, u) \in \mathbb{R} \times V_h$ such that

$$a(u, v) = \lambda b_h(u, v) \quad \forall v \in V_h\,, \qquad (5.12)$$

we have the following result :

THEOREM 5.1 : *There exists* $\varepsilon > 0$ *and, for h small enough, there exists an unique eigenvalue* $\lambda_h \in (\lambda_0 - \varepsilon, \lambda_0 + \varepsilon)$ *of Problem* (5.12) ; λ_h *is a double eigenvalue with error estimate* : $|\lambda_0 - \lambda_h| = 0(h^2)$.

For convenience, we shall first present all the results, and delay a sketch of their proofs to the end of the section.

We define by Lax-Milgram Theorem the operators T and T_h as

$$T : L^2 \to H_0^1\,, \quad a(Tu, v) = b(u, v) \quad \forall v \in H_0^1\,, \quad \forall u \in L^2\,, \qquad (5.13)$$

$$T_h : C_0^0 \to V_h\,, \quad a(T_h u, v) = b_h(u, v) \quad \forall v \in V_h\,, \quad \forall u \in C_0^0\,; \qquad (5.14)$$

then (5.4) is equivalent to the equation

$$u + T(-\lambda u + u^3) = 0\,, \quad (\lambda, u) \in \mathbb{R} \times H_0^1\,,$$

whereas (5.9) is equivalent to the equation $u + T_h(-\lambda u + u^3) = 0$, $(\lambda, u) \in \mathbb{R} \times V_h$; since the range of T_h is V_h, (5.9) is also equivalent to the problem of finding $(\lambda, u) \in \mathbb{R} \times H_0^1$ such that

$$u + T_h(-\lambda \Pi_h u + (\Pi_h u)^3) = 0\,;$$

in this last expression, note that it is not possible to suppress the introduction of the projector Π_h defined in (5.6) ; indeed T_h is defined in C_0^0 and H_0^1 is not imbedded in C_0^0.

In order to use the results of Section 4, we set $X = \mathbb{R} \times H_0^1$ and $Y = H_0^1$.

Followıng Sattınger [14], we ıntroduce, for $\zeta = \pm 1$, the functıons

$$F \quad X \to Y, \quad F(x) = u + T(-(\lambda_0 + \zeta s^2) u + u^3), \qquad x = (s, u), \quad (5\ 15)$$

$$F_h \quad X \to Y, \quad F_h(x) = u + T_h(-(\lambda_h + \zeta s^2) \Pi_h u + (\Pi_h u)^3), \qquad x = (s, u), \quad (5\ 16)$$

clearly, by the change of varıable $\lambda = \lambda_0 + \zeta s^2$, solvıng (5 1) ıs equıvalent to solve the equatıon $F(x) = 0$ for $\zeta = 1$ *and* $\zeta = -1$, ın the same way, by the change of varıable $\lambda = \lambda_h + \zeta s^2$, solvıng (5 9) ıs equıvalent to solve the equatıon $F_h(x) = 0$ for $\zeta = 1$ *and* $\zeta = -1$, sınce, by Theorem 5 1,

$$\lambda - \lambda_h = 0(h^2),$$

ıt ıs reasonable to compare, separately for $\zeta = 1$ and $\zeta = -1$, the solutıons of $F(x) = 0$ wıth the solutıons of $F_h(x) = 0$ ın the neıghborhood of $x_0 \equiv 0$

THEOREM 5 2 *F and F_h are C^∞ mappıngs from X ınto Y For any $k \geqslant 0$ and any bounded subset $B \subset X$, we have*

$$\lim_{h \to 0} \sup_{x \in B} \| D^k F(x) - D^k F_h(x) \| = 0 \qquad (5\ 17)$$

We consıder $Y = H_0^1$ as a Hılbert space equıpped wıth the scalar product $a(.,.)$

Let Y_2 be the orthogonal complement of Y_1 ın Y where Y_1 *ıs* defıned ın (5 11)

Let $\Phi_0 = (1, 0), \Phi_1 = (0, \omega_1), \Phi_2 = (0, \omega_2) \in \mathbb{R} \times H_0^1 = X$ and let $X_1 \subset X$ be the three-dımensıonal subspace spanned by Φ_0, Φ_1, Φ_2 Settıng

$$X_2 = \{0\} \times Y_2 \subset X,$$

we see that $X = X_1 \oplus X_2$ Fınally let $Q \quad Y \to Y$ be the orthogonal projector from Y onto Y_2 By (5 15) we have

$$F'(0)(s_1, u_1) = u_1 - \lambda_0 T u_1, \quad \forall (s_1, u_1) \in X, \qquad (5\ 18)$$

sınce T, defined by (5 13), ıs selfadjoınt ın H_0^1 wıth respect to $a(.,.)$, we ımmedıately obtaın

$$X_1 = \ker F'(0), \quad Y_2 = \text{Range } F'(0) \qquad (5\ 19)$$

Clearly, for $x_0 = 0$, F satısfies Hypotheses (4 1) and (4 2) of Sectıon 4 wıth $n = 2$, furthermore Theorem 5 2 ınsures that F and F_h verıfy Hypotheses (4.4), (4.5) and (4.6). Consequently, we can apply Theorem 4.1 and define maps $v(.)$ and $v_h(.)$ satisfying (4.7) and (4.8); we deduce the bifurcation functions f and f_h defined by (4.11) and (4.12).

THEOREM 5.3 : *a)* $f^{(k)}(0) = 0, f_h^{(k)}(0) = 0, k = 0, 1, 2,$ *h small enough*; (5.20)

b) for $\sigma = s\Phi_0 + \eta_1 \Phi_1 + \eta_2 \Phi_2 \in X_1$, *we have*

$$f'''(0)\,\sigma^3 = (A\zeta s^2 \eta_1 + B\eta_1^3 + C\eta_1 \eta_2^2)\,\varphi_1 + \\ + (A\zeta s^2 \eta_2 + B\eta_2^3 + C\eta_1^2 \eta_2)\,\varphi_2\,, \quad (5.21)$$

where

$$A = -\frac{6}{\lambda_0}, \quad \beta = \frac{27}{2\lambda_0^2}, \quad C = \frac{18}{\lambda_0^2}.$$

We remark that (5.20) implies that Hypotheses (4.15) and (4.18) are satisfied with $q = 3$. By using (5.21), it is now easy to solve the equation $f'''(0)\,\sigma^3 = 0$ and to determine a maximal set of linearly independent characteristic rays (see Remark 4.3); using the notation $\sigma = (r_0, r_1, r_2)$ for

$$\sigma = r_0 \Phi_0 + r_1 \Phi_1 + r_2 \Phi_2\,,$$

we obtain :

$$\zeta = -1 : \sigma_{00} = (1, 0, 0)\,; \quad (5.22)$$

$$\begin{aligned} \zeta = 1 : \sigma_{00} &= (1, 0, 0)\,, \quad \sigma_{01} = (1, \alpha, 0)\,, \quad \tilde\sigma_{01} = (1, -\alpha, 0)\,, \\ \sigma_{02} &= (1, 0, \alpha)\,, \quad \tilde\sigma_{02} = (1, 0, -\alpha)\,, \quad \sigma_{03} = (1, \beta, \beta)\,, \\ \tilde\sigma_{03} &= (1, -\beta, -\beta)\,, \quad (5.23) \\ \sigma_{04} &= (1 - \beta, \beta)\,, \quad \tilde\sigma_{04} = (1, \beta, -\beta)\,, \end{aligned}$$

where

$$\alpha = \frac{2}{3}\sqrt{\lambda_0}\,, \quad \beta = \sqrt{\frac{4\lambda_0}{21}}\,;$$

it is also easy to check that all the rays given in (5.22) and (5.23) are not degenerated, i.e. they satisfy Hypothesis (4.17).

THEOREM 5.4 : *Let* $\sigma_0 \in X_1$ *be one of the characteristic rays given in* (5.22) *and* (5.23). *Then there exist positive constants* s_0, h_0 *and the* C^∞ *mappings* $u(.)$, $u_h(.) : (-s_0, s_0) \to H_0^1$ *such that by setting*

$$x(s) = (s, u(s))\,, \quad x_h(s) = (s, u_h(s)) \in X\,,$$

we have

a) $F(x(s)) = 0,\ F_h(x_h(s)) = 0,\ |s| < s_0,\ h < h_0,\ x(0) = x_h(0) = 0,$

$$x'(0) = \sigma_0, \quad (5\ 24)$$

b) for any $k = 0, 1, 2,$, $u^{(k)}(.)$ *and* $u_h^{(k)}(.)$ *are uniformly bounded with respect to s and h*,

c) for any $k = 0, 1, 2,$, *we have*

$$\sup_{|s| < s_0} \| u^{(k)}(s) - u_h^{(k)}(s) \|_{H^1(\Omega)} = 0(h) \quad (5\ 25)$$

Consider $u(s)$ and $u_h(s)$ defined by Theorem 5 4 and set $\lambda(s) = \lambda_0 + \zeta s^2$, $\lambda_h(s) = \lambda_h + \zeta s^2$ where λ_h is given by Theorem 5 1, clearly, by definitions (5 15), (5 16) of F and F_h, $(\lambda(s), u(s))$ and $(\lambda_h(s), u_h(s)) \in \mathbb{R} \times H_0^1$ will be respectively solutions of our original problems (5 1) and (5 9) If $\sigma_0 = \sigma_{00}$ with $\zeta = -1$, we obtain the branch $\{ (\lambda, 0) \mid \lambda \leqslant \lambda_0 \}$ and the approximate branch $\{ (\lambda, 0) \mid \lambda \leqslant \lambda_h \}$, if $\sigma_0 = \sigma_{00}$ with $\zeta = 1$, we obtain respectively the branches $\{ (\lambda, 0) \mid \lambda \geqslant \lambda_0 \}$ and $\{ (\lambda, 0) \mid \lambda \geqslant \lambda_h \}$, together, this gives the trivial branch, solution of both the exact and approximate problems Suppose now that $u(s)$ and $u_h(s)$ correspond to σ_{01} and let $\tilde{u}(s)$ and $\tilde{u}_h(s)$ correspond to $\tilde{\sigma}_{01}$, from the relations

$$f(s\Phi_0 + \eta_1 \Phi_1 + \eta_2 \Phi_2) = f(-s\Phi_0 + \eta_1 \Phi_1 + \eta_2 \Phi_2),$$
$$f_h(s\Phi_0 + \eta_1 \Phi_1 + \eta_2 \Phi_2) = f_h(-s\Phi_0 + \eta_1 \Phi_1 + \eta_2 \Phi_2) \text{ for all } s, \eta_1, \eta_2 \in \mathbb{R},$$

we deduce that, for s small enough, $u(s) = \tilde{u}(-s)$ and $u_h(s) = \tilde{u}_h(-s)$, consequently the branches parametrized by $(\lambda_0 + s^2, u(s))$, $(\lambda_h + s^2, u_h(s))$, $|s| < s_0$, are respectively identical to the branches parametrized by

$$(\lambda_0 + s^2, \tilde{u}(s)), \quad (\lambda_h + s^2, \tilde{u}_h(s)), \quad |s| < s_0$$

The same argument works for the pairs $\sigma_{0i}, \tilde{\sigma}_{0i}$, $i = 2, 3, 4$ so that each of the original exact and approximate problems possess in fact four non-trivial branches in the neighborhood of $(\lambda_0, 0)$, it is easy to verify that these four branches are different and supercritical, that the problems possess no further solution is a consequence of Remark 4 4 By Theorem 5 4, we then deduce our final result

THEOREM 5 5 *In a neighborhood of* $(\lambda_0, 0) \in \mathbb{R} \times H_0^1$, *the exact problem* (5 1) *and the approximate problem* (5 9) *possess each four non-trivial branches which can be parametrized respectively in the form* $(\lambda_0 + s^2, u_i(s))$,

$(\lambda_h + s^2, u_{ih}(s))$, $|s| < s_0$, $i = 1, 2, 3, 4$, *where* u_i *and* i_{ih} : $(-s_0, s_0) \to H_0^1$ *are* C^∞ *mappings such that*

$$\sup_{|s|<s_0} \| u_i^{(k)}(s) - u_{ih}^{(k)}(s) \|_{H^1(\Omega)} = 0(h)\,, \quad i = 1, 2, 3, 4\,, \quad k = 0, 1, 2, \ldots$$

As announced, we shall only sketch the proofs of Theorems 5.1 to 5.5 which essentially rely on the results of Section 4. To this end, we need some preliminary lemmas.

We first introduce some further notations. For $u \in H^1 = H^1(\Omega)$,

$$\| u \| = \left(\int_\Omega u^2 + (\partial_x u)^2 + (\partial_y u)^2 \right)^{1/2}, \quad | u | = \left(\int_\Omega (\partial_x u)^2 + (\partial_y u)^2 \right)^{1/2}.$$

For $u \in L^p = L^p(\Omega)$ and $1 \leqslant p < \infty$,

$$\| u \|_p = \left(\int_\Omega | u |^p \right)^{1/p}.$$

As before Q_1 is the set of polynomials of the form $axy + bx + cy + d$. $P : C^0(\overline{\Omega}) \to Q_1$ is the interpolatory operator at the four vertices of $\overline{\Omega}$. Furthermore, c will denote a generic constant.

LEMMA 5.1 : *Let* $u_i \in Q_1$, $1 \leqslant i \leqslant 4$; *then*

$$a) \quad \left| \int_\Omega P\left(\prod_{i=1}^4 u_i \right) \right| \leqslant c \prod_{i=1}^4 \| u_i \|_{p_i}, \text{ where } \sum_{i=1}^4 \frac{1}{p_i} = 1 \ (c \textit{ depends on } p_i); \tag{5.26}$$

$$b) \quad \left| \int_\Omega \prod_{i=1}^4 u_i - \int_\Omega P\left(\prod_{i=1}^4 u_i \right) \right| \leqslant c \sum_{i=1}^4 | u_i | \prod_{j \neq i} \| u_i \|_6\,; \tag{5.27}$$

$$c) \quad \left| \int_\Omega u_1 u_2 - \int_\Omega P(u_1 u_2) \right| \leqslant c \, | u_1 | \, | u_2 |\,. \tag{5.28}$$

Proof : (5.26) is a consequence of the equivalence of the norms in finite dimensional spaces and of Hölder's inequality. Let μ_i be the mean value of u_i, i.e. $\mu_i = \int_\Omega u_i$; we have $\| u_i - \mu_i \|_2 \leqslant c \, | u_i |$ (see for example Ciarlet [4], page 115); then we deduce (5.27) from (5.26) and from the relations

$$\int \prod_{i=1}^4 u_i - \int P\left(\prod_{i=1}^4 u_i \right) = \int (u_1 - \mu_1) u_2 u_3 u_4 + \int \mu_1 (u_2 - \mu_2) u_3 u_4 + \\ + \cdots + \int \mu_1 \mu_2 \mu_3 (u_4 - \mu_4) - \int P((u_1 - \mu_1) u_2 u_3 u_4)$$

$$- \cdots - \int P(\mu_1 \, \mu_2 \, \mu_3 (u_4 - \mu_4))$$
$$+ \left(\int \mu_1 \, \mu_2 \, \mu_3 \, \mu_4 - \int P(\mu_1 \, \mu_2 \, \mu_3 \, \mu_4) \right),$$

since the last parenthesis vanishes. In the same way, (5.28) follows from the identity

$$\int u_1 \, u_2 - \int P(u_1 \, u_2) = \int (u_1 - \mu_1)(u_2 - \mu_2) - \int P((u_1 - \mu_1)(u_2 - \mu_2)) +$$
$$+ \left(\int \mu_1 (u_2 - \mu_2) - \int P(\mu_1 (u_2 - \mu_2)) \right) + \left(\int \mu_2 \, u_1 - \int P(\mu_2 \, u_1) \right),$$

since the two last parentheses vanish.

By the standard argument of the « reference element » (see Ciarlet [4]) and the continuous injection of H^1 in $L^p \, (1 \leqslant p < \infty)$, we deduce easily from Lemma 5.1 :

LEMMA 5.2 : *Let* $u_1, u_2, u_3, v \in V_h$; *then*

a) $| \, b_h(u_1 \, u_2 \, u_3, v) \, | \leqslant c \, \| u_1 \| \, \| u_2 \| \, \| u_3 \| \, \| v \|$,

$$| \, b_h(u_1, v) \, | \leqslant c \, \| u_1 \| \, \| v \| , \quad (5.29)$$

b) $| \, b(u_1 \, u_2 \, u_3, v) - b_h(u_1 \, u_2 \, u_3, v) \, | \leqslant ch \, \| u_1 \| \, \| u_2 \| \, \| u_3 \| \, \| v \|$, (5.30)

c) $| \, b(u_1, v) - b_h(u_1, v) \, | \leqslant ch^2 \, \| u_1 \| \, \| v \|$. (5.31)

It is well-known that T, defined in (5.13), maps continuously L^2 into $H^2 = H^2(\Omega)$: by using the standard techniques of finite elements in connection with numerical integration, we obtain by (5.30) and (5.31).

LEMMA 5.3 :

$$\| \, T(u_1 \, u_2 \, u_3) - T_h(u_1 \, u_2 \, u_3) \, \| \leqslant ch \, \| u_1 \| \, \| u_2 \| \, \| u_3 \| \quad \forall u_1, u_2, u_3 \in V_h \, ; \quad (5.32)$$

$$\| \, Tu - T_h \, u \, \| \leqslant ch \, \| u \| \qquad \forall u \in V_h \, . \quad (5.33)$$

Proof of Theorem 5.1 : By Lemma 5.3 and classical results on spectral approximation (see for example [6], p. 140), there exist exactly two eigenvalues (repeated following their multiplicity) λ_{1h} and λ_{2h} of Problem (5.12) which converge to λ_0 ; from symmetry arguments, it is easy to show that, in fact,

$\lambda_{1h} = \lambda_{2h}$ for h small enough and we set $\lambda_h = \lambda_{1h}$ By [6], again, there exists an eigenfunction $\omega_{1h} \in V_h$ corresponding to λ_h such that $a(\omega_{1h}, \omega_{1h}) = 1$ and $\| \omega_1 - \omega_{1h} \| = 0(h)$, where ω_1 is given in (5 11) Let $\mu_0 = \lambda_0^{-1}$, $\mu_h = \lambda_h^{-1}$, we have $\mu_0\, a(\omega_1, v) = b(\omega_1, v)$ $\forall v \in H_0^1$ and $\mu_h\, a(\omega_{1h}, v) = b_h(\omega_{1h}, v)$ $\forall v \in V_h$ from which we deduce

$$\begin{aligned} \mu_h - \mu_0 &= b_h(\omega_{1h}, \omega_{1h}) - \mu_0\, a(\omega_{1h}, \omega_{1h}) \\ &= (b_h(\omega_{1h}, \omega_{1h}) - b(\omega_{1h}, \omega_{1h})) + (b(\omega_{1h} - \omega_1, \omega_{1h} - \varphi_1) - \\ &\qquad - \mu_0\, a(\omega_{1h} - \omega_1, \omega_{1h} - \omega_1)), \end{aligned}$$

by (5 31), we obtain $\mu_h - \mu_0 = 0(h^2)$ and consequently $\lambda_h - \lambda_0 = 0(h^2)$ ■

Proof of Theorem 5 2 The fact that F is a C^∞ mapping is well-known, it is based on the continuous, in fact compact, injection of H_0^1 into L^p, $1 \leqslant p < \infty$ Since V_h if finite-dimensional, the restriction of F_h to V_h, equipped with the norm $\| \,.\, \|$, is clearly a C^∞ mapping, since Π_h is a continuous linear operator in H_0^1, F_h is also a C^∞ mapping Let J $H_0^1 \to L^6$ denote the injection and J^* $(L^6)^* \to H_0^1$ be its dual operator, J and consequently J^* are compact, for any $u \in H_0^1$, $\lim\limits_{h \to 0} \Pi_h u = u$, in H_0^1 equipped with the scalar product $a(.,.)$, Π_h is selfadjoint, by a classical result, it follows that $(I - \Pi_h) J^*$ and consequently its dual operator $J(I - \Pi_h)$ converge in norm to zero, i e

$$\lim_{h \to 0} \| I - \Pi_h \|_{\mathscr{L}(H_0^1\, L^6)} = 0 \tag{5 34}$$

Setting $x = (s, u)$ as in (5 15), (5 16), we can write

$$\begin{aligned} \| D^k F(s, u) - D^k F_h(s, u) \| &\leqslant \| D^k F(s, u) - D^k F(s, \Pi_h u) \| + \\ &\quad + \| D^k F(s, \Pi_h u) - D^k F_h(s, u) \|, \end{aligned} \tag{5 35}$$

with (5 35), (5 34) and Lemma 5 3, we can deduce (5 17) by elementary calculations ■

Proof of Theorem 5 3 By (4 13), $f(0) = 0$, $f'(0) = 0$, since $F''(0) = 0$, we also have $f'(0) = 0$, by Remark 4 1, $f'''(0)$ is equal to the restriction of $(I - Q) F'''(0)$ which allows to obtain (5 21) by elementary calculations In the same way, we have $F_h(0) = 0$, $F_h''(0) = 0$ which imply that $f_h(0) = 0$, $f_h''(0) = 0$, by Theorem 5 1, the kernel of $F_h'(0)$ is two-dimensional from which follows that $f_h'(0) = 0$ ■

Proof of Theorem 5 4 F, F_h and σ_0 satisfy Hypotheses *a*) to *i*) with $q = 3$ of Section 4 at $x_0 = 0$, we apply Theorem 4 3 and remark that the parametrization by t is determined by the choice of $\psi_0 \in X^*$ in (4 19), (4 20), by

(5 22) and (5 23), we can choose ψ_0 such that $\langle \Phi_0, \psi_0 \rangle = 1, \langle \Phi_i, \psi_0 \rangle = 0$, $i = 1, 2$, then, for $x(t) = (s(t), u(t))$ and $x_h(t) = (s_h(t), u_h(t))$ given by Theorem 4 3, we obtain $s(t) = s_h(t) = t$ It remains to verify the estimate (5 25), by (4 34), it suffices to prove that for any $j = 0, 1, 2,$, we have

$$\sup_{|s|<s_0} \left\| \frac{d^j}{ds^j} F_h(x(s)) \right\| = \sup_{|s|<s_0} \left\| \frac{d^j}{ds^j} (F_h(x(s)) - F(x(s))) \right\| = 0(h), \quad (5\ 36)$$

since T maps continuously L^2 into H^2, for any $k > 0$, $u^{(k)}(s)$ is uniformly bounded in H^2 with respect to $|s| < s_0$, by standard approximation results (see Ciarlet [4]), $\sup_{|s|<s_0} \| (I - \Pi_h) u^{(k)}(s) \| = 0(h)$, together with Lemma 5 3 and the arguments already used in the proof of Theorem 5 2, the estimate (5 36) follows easily ■

REFERENCES

[1] F BREZZI, J RAPPAZ and P A RAVIART, *Finite Dimensional Approximation of Nonlinear Problems, Part I Branches of Nonsingular Solutions*, Numer Math 36, 1980, 1-25

[2] F BREZZI, J RAPPAZ and P A RAVIART, *Finite Dimensional Approximation of Nonlinear Problems, Part II Limit Points*, Numer Math , 37, 1981, 1-28

[3] F BREZZI, J RAPPAZ and P A RAVIART, *Finite Dimensional Approximation of Nonlinear Problems, Part III Simple Bifurcation Points*, Numer Math , 38, 1981, 1-30

[4] P G CIARLET, *The Finite Element Method for Elliptic Problems*, Studies in Mathematics and its Applications, Vol 4, North-Holland, 1978

[5] D W DECKER and H B KELLER, *Multiple Limit Point Bifurcation*, Report of Math Dept , California Institute of Technology, 1979

[6] J DESCLOUX, M LUSKIN and J RAPPAZ, *Approximation of the Spectrum of Closed Operators The Determination of Normal Modes of a Rotating Basin*, Math Comp , 36, 1981, 137-154

[7] J DESCLOUX, N NASSIF and J RAPPAZ, *On Spectral Approximation*, R A I R O Numerical Analysis, 12, 1978, 97-119

[8] J DESCLOUX and J RAPPAZ, *On Numerical Approximation of Solution Branches of Nonlinear Equations*, Rapport, Departement de Mathematiques, EPFL, 1981

[9] J DESCLOUX and J RAPPAZ, *Numerical Approximation of a Nonlinear Sturm-Liouville Problem on an Infinite Interval*, to appear in the Proceedings of Equadiff V, Teubner

[10] R J MAGNUS, *On the Local Structure of the Zero-Set of a Banach Space Valued Mapping*, J Functional Analysis, 22, 1976, 58-72

[11] J -C PAUMIER, *Une methode numerique pour le calcul des points de retournement Application a un probleme aux limites non-lineaire* II *Analyse numerique d'un probleme aux limites non-lineaire*, Numer Math , 37, 1981, 445-454

[12] J RAPPAZ and G RAUGEL, *Finite Dimensional Approximation of Bifurcation Problems at a Multiple Eigenvalue*, Rapport no 71, Centre de Mathematiques appliquees, Ecole Polytechnique, Palaiseau, 1981

[13] J RAPPAZ and G RAUGEL, *Approximation of Double Bifurcation Points for Nonlinear Eigenvalue Problems*, MAFELAP 1981 (Ed J Whiteman), Academic Press (to appear)

[14] D H SATTINGER, *Group Theoric Methods in Bifurcation Theory*, Lecture Notes in Mathematics 762, Springer-Verlag, Berlin, 1979