SOME REMARKS ON APPROXIMATION BY FINITE ELEMENT METHODS

GIUSEPPE GEYMONAT

*Laboratoire de Mécanique et Technologie, E.N.S. de Cachan
(C.N.R.S.) Université Paris VI, 61, Avenue du Président Wilson
F-94230 Cachan, France*

PETER OSWALD

*Technische Universität Dresden, Mommsenstrasse 13
DDR-8027 Dresden, German Democratic Republic*

Introduction

The aim of these lectures is to briefly introduce to some aspects of the use of finite element methods in approximation of solutions of linear and nonlinear problems. Special emphasis is put on the connections with some aspects of modern approximation theory: for this we carefully analyze the question of the optimal asymptotic approximation rate. This can be done in terms of n -width for general data and we see how FEM are optimal for regular solutions (see Sect. 1 and Sect. 2.2). But this also leads naturally to nonlinear approximation methods for particular problems: see Sect. 3. As we shall see in Sects. 5 and 6, the error estimates for the approximation of solution branches of nonlinear equations can be obtained with the aid of the known estimates for the corresponding linear problems. These results raise the question of the possibility of applying nonlinear approximation methods also for solution branches of nonlinear equations.

The text is mainly based on some lectures given by the authors in April-May 1986 at the XXVII Semester on Approximation and Function Spaces at the Banach Center. The authors warmly thank Professor Z. Ciesielski for the stimulating and calm atmosphere he created during their stay at the Banach Center.

1. Internal approximation for linear equations: abstract results

1.1. Let V be a real Hilbert space equipped with the norm $\|\cdot\|$; we denote by V' its strong dual equipped with the norm $\|\cdot\|_*$ and by $(\cdot, \cdot)$ the duality pairing.

Let $a(u, v)$ be a continuous and coercive bilinear form on $V \times V$, i.e.:

(i) (continuity) there exists $M > 0$ such that for all $u, v \in V$

$$|a(u, v)| \leq M \|u\| \|v\|,$$

(ii) (coerciveness) there exists $\alpha > 0$ such that for all $u \in V$

$$a(u, u) \geq \alpha \|u\|^2.$$

The Lax–Milgram theorem ensures the existence and uniqueness of solution of the following problem:

(P) For a given $f \in V'$ find $u \in V$ such that

$$a(u, v) = (f, v) \quad \text{for all } v \in V.$$

Moreover, the solution u satisfies

$$(1.1) \quad \|u\| \leq \frac{M}{\alpha} \|f\|_*.$$

1.2. EXAMPLE 1. Let Ω be a bounded open polygonal set in $\mathbb{R}^2$, $V = W_2^1(\Omega)$, $V' = W_2^{-1}(\Omega)$ and

$$a(u, v) = \int_{\Omega} \left\{ \frac{\partial u}{\partial x_1} \frac{\partial v}{\partial x_1} + \frac{\partial u}{\partial x_2} \frac{\partial v}{\partial x_2} \right\} dx;$$

then (i) and (ii) are satisfied thanks to the Poincaré inequality and the solution of the problem (P) is the weak solution of the Dirichlet problem

$$(1.2) \quad \begin{cases} -\Delta u = f & \text{in } \Omega, \\ u = 0 & \text{on } \partial\Omega. \end{cases}$$

Let us remark that if $f \in W_p^l(\Omega)$, $1 < p < +\infty$, l integer ≥ 0 , then the Sobolev embedding theorem ensures that $f \in W_2^{-1}(\Omega)$ (here and in the following we shall freely use the main properties of Sobolev spaces; see e.g. [1], [23], ...).

For our applications of the theory, the so-called *regularity results* are of great importance. A typical such result is the following:

Let $f \in W_p^l(\Omega)$, $1 < p < +\infty$, l integer ≥ 0 , and suppose f satisfies suitable compatibility conditions on $\partial\Omega$. Then $u \in W_p^{l+2}(\Omega) \cap W_2^1(\Omega)$.

In particular, if f has compact support $\subset \Omega$ then such compatibility conditions are fulfilled.

In Section 3 we shall state a more precise regularity result following Grisvard [10].

1.3. Let V_N be a fixed subspace of V of finite dimension N . Then the approximate problem

(P_{V_N}) For a given $f \in V'$ find $u_N \in V_N$ such that

$$a(u_N, v) = (f, v) \quad \text{for all } v \in V_N$$

has a unique solution. Clearly, the problem (P_{V_N}) is equivalent to an N -dimensional linear system.

A first comparison of the solutions of the problems (P) and (P_{V_N}) is given by the following result.

THEOREM 1.1 (J. Céa). *One has*

$$(1.3) \quad \|u - u_N\| \leq \frac{M}{\alpha} \inf_{v \in V_N} \|u - v\|.$$

If moreover $u \in Z$, a Banach space continuously embedded in V , then

$$(1.4) \quad \|u - u_N\| \leq \frac{M}{\alpha} \|u\|_Z \sup_{\substack{z \in Z \\ \|z\|_Z \leq 1}} \inf_{v \in V_N} \|z - v\|.$$

Proof. Comparing (P) and (P_{V_N}) one has

$$a(u - u_N, z) = 0 \quad \text{for all } z \in V_N$$

and so for all $v \in V_N$, $a(u - u_N, u - u_N) = a(u - u_N, u - v)$. Then (1.3) follows from the properties (i) and (ii) of a and (1.4) is clear. ■

In the sequel we shall use the notation

$$e(V_N; Z, V) := \sup_{\substack{z \in Z \\ \|z\|_Z \leq 1}} \inf_{v \in V_N} \|z - v\|.$$

Let us remark that, usually, the condition $u \in Z$ follows from a certain regularity result for the problem (P):

(RR) *If $f \in W$, a Banach space continuously embedded in V' , then $u \in Z$ and there exists $c_1 > 0$ such that $\|u\|_Z \leq c_1 \|f\|_W$.*

Since

$$\inf_{v \in V_N} \|z - v\| = d(z, V_N) = \|z - P_{V_N} z\|$$

where P_{V_N} is the orthogonal projection on V_N (and $d(z, V_N)$ denotes the distance in V from z to V_N) we are reduced to a regularity problem

independent of V_N and to an approximation problem independent of the form $a(u, v)$. This last result simplifies many considerations; however, it also implies a loss of information.

If we now consider a sequence of subspaces $\{V_N\}$, $N \in \mathbb{N}$, with $\dim V_N = N$, we find that the *approximation error* $\varepsilon(u, V_N) := \|u - u_N\|$ will converge to zero as $N \rightarrow \infty$ for all $f \in V'$ if and only if the following *completeness condition* holds:

(CC) For all $w \in V$ one has $\lim_{N \rightarrow \infty} d(w, V_N) = 0$.

A natural and vague question is to look for *quantitative asymptotic error estimates* depending on the choice of V_N and of f ; in order to clarify the statement of the question it may be useful to study a very particular example.

1.4. EXAMPLE 2. Let be $V = L_2(0, 1) = V'$ and set

$$a(u, v) = \int_0^1 u(x) v(x) dx.$$

Let $\Phi = \{\varphi_n\}_{n=1}^\infty$ be any complete orthonormal system on $[0, 1]$ and consider the subspaces $V_N = \text{span}\{\varphi_1, \dots, \varphi_N\}$ generated by Φ . Clearly the solutions of the problems (P) and (P_{V_N}) are

$$u = f \quad \text{and} \quad u_N = S_N^\Phi f = \sum_{k=1}^N a_k \varphi_k$$

where $a_k = (f, \varphi_k)$ are the Fourier coefficients of f with respect to Φ . Moreover,

$$(1.5) \quad \|u - u_N\|^2 = \sum_{k=N+1}^\infty a_k^2 = o(1), \quad N \rightarrow \infty.$$

This relation shows that the only information that $u \in V$ does not automatically involve any quantitative asymptotic error estimate.

However, the situation changes if we make additional assumptions in terms of the smoothness of u . For instance, let us consider the Haar system $H = \{\chi_n\}_{n=1}^\infty$ given by $\chi_1(x) = 1$, $x \in [0, 1]$, and

$$(1.6) \quad \chi_n(x) = \begin{cases} 0, & x \notin \left[\frac{l-1}{2^k}, \frac{l}{2^k}\right], \\ 2^{k/2}, & x \in \left(\frac{l-1}{2^k}, \frac{2l-1}{2^{k+1}}\right), \\ -2^{k/2}, & x \in \left(\frac{2l-1}{2^{k+1}}, \frac{l}{2^k}\right), \end{cases} \quad \text{for } n = 2^k + l \geq 2,$$

where $k = 0, 1, \dots$, $l = 1, \dots, 2^k$. In this case, the subspaces V_N consist of

piecewise constant functions with respect to a sequence of special dyadic partitions of $[0, 1]$, and for $1 \leq p < +\infty$ and $f \in L_p(0, 1)$ we have (cf. [11])

$$(1.7) \quad \|f - S_N^H f\|_{L_p} \leq C\omega(1/N, f)_p, \quad N = 1, 2, \dots$$

where

$$\omega(t, f)_p := \sup_{0 \leq h \leq t} \left(\int_0^{1-h} |f(x+h) - f(x)|^p dx \right)^{1/p}, \quad 0 \leq t \leq 1,$$

denotes the L_p -modulus of continuity of $f(x)$.

Thus, if f is absolutely continuous and $f' \in L_2(0, 1)$ (i.e. $f = u \in W_2^1(0, 1)$) we obtain

$$(1.8) \quad \|u - u_N\| \leq CN^{-1} \|u'\|, \quad N = 1, 2, \dots,$$

which represents the sharp $O(N^{-1})$ asymptotic error estimate for the chosen class $W_2^1(0, 1)$. Changing the class we obtain other estimates. For instance, from the usual embedding theorems and (1.7) it easily follows that

$$(1.9) \quad \|u - u_N\| \leq CN^{-(3/2-1/p)} \|u'\|_{L_p}, \quad N = 1, 2, \dots,$$

for any $f = u \in W_p^1(0, 1)$, $1 \leq p < 2$. This estimate is sharp with respect to the whole class $W_p^1(0, 1)$. Now, let us change our point of view in order to, at least partially, improve the above result. Instead of $\|u - u_N\|$ we consider some kind of nonlinear best approximation $\inf_{V_N = V(n_1, \dots, n_N)} \|u - u_N\|$ where the

infimum is taken over the family of all N -dimensional subspaces

$$V(n_1, \dots, n_N) = \text{span} \{\chi_{n_1}, \dots, \chi_{n_N}\}, \quad 1 \leq n_1 < n_2 < \dots < n_N.$$

With this notation, one can prove that

$$(1.10) \quad \inf_{V_N = V(n_1, \dots, n_N)} \|u - u_N\| \leq CN^{-1} \|u'\|_{L_p}, \quad N = 1, 2, \dots,$$

for arbitrary $f = u \in W_p^1(0, 1)$, $1 \leq p < 2$. Thus, choosing an appropriate $V_N = V(n_1, \dots, n_N)$ which depends in general on u we get an improvement of (1.9): instead of $O(N^{-3/2+1/p})$ we now have $O(N^{-1})$.

Though not constructive, the result indicates that strategies of choosing individual subspaces, from a given class, for each u would be useful for constructing asymptotically improved approximations. These problems are connected with approximation by splines with free nodes and more generally with the choice of good triangulations in finite element methods as we shall see in Sect. 3.

1.5. Let there be given a sequence of subspaces $\{V_N\}$, $N \in \mathbb{N}$, of finite dimension (for simplicity we assume $\dim V_N = N$) and satisfying (CC), and let us suppose that an asymptotic error estimate has been found:

$$(1.11) \quad \|u - u_N\| \leq \sigma(V_N) \|u\|_Z.$$

Then it is natural to ask if the result obtained is, at least up to some multiplicative constant, asymptotically optimal. A first answer to this question can be obtained in terms of Kolmogorov N -widths; indeed, from (1.11) it follows immediately that $e(V_N; Z, V) \leq \sigma(V_N)$ and so if $\mathcal{G}_N$ denotes the family of all subspaces of V of dimension N one obtains the lower bound

$$\inf_{V_N \in \mathcal{G}_N} e(V_N; Z, V) \leq \sigma(V_N).$$

But the N -width of the unit ball of Z in V is

$$(1.12) \quad \begin{aligned} d_N(Z, V) &:= \inf_{V_N \in \mathcal{G}_N} e(V_N; Z, V) \\ &= \inf_{V_N \in \mathcal{G}_N} \sup_{\substack{z \in Z \\ \|z\|_Z \leq 1}} \inf_{v \in V_N} \|z - v\| \end{aligned}$$

and so we have proved the following result:

PROPOSITION 1.1. *If for $u \in Z$ one has $\|u - u_N\| \leq \sigma(V_N) \|u\|_Z$ then*

$$(1.13) \quad d_N(Z, V) \leq \sigma(V_N).$$

Since the precise asymptotic behaviour of the N -width is known for Sobolev spaces (see e.g. Pinkus [16]), one can see if (1.11) is asymptotically optimal up to a multiplicative constant; in Example 2, e.g., (1.8) and (1.9) are sharp since they are asymptotically optimal. However, as is shown by (1.10) in the case of Example 2, it is also possible to work with subspaces V_N adapted to u and, at least in some cases, improve the previous asymptotic error estimates. More precisely, for a given family $\mathcal{V}_N \subset \mathcal{G}_N$ of vector spaces of dimension N we define

$$(1.14) \quad \eta(z; \mathcal{V}_N) := \inf_{V_N \in \mathcal{V}_N} \inf_{v \in V_N} \|z - v\|,$$

$$(1.15) \quad d_N(Z, V; \mathcal{V}_N) := \sup_{\substack{z \in Z \\ \|z\|_Z \leq 1}} \eta(z; \mathcal{V}_N).$$

The previous definitions and (1.10) suggest that starting from (1.11) one can look for a suitable family $\mathcal{V}_N \subset \mathcal{G}_N$ such that $\inf_{V_N \in \mathcal{V}_N} \|u - u_N\|$ gives asymptotically improved error estimates. In Section 3 we shall apply this nonlinear approximation method to the finite element method for the Dirichlet problem of Example 1.

1.6. Some complements. 1. Let H be a real Hilbert space such that $V \subset H$ with continuous embedding and dense image. Then identifying H with its dual H' we have $V \subset H \simeq H' \subset V'$ with H continuously and densely embedded in V' ; moreover, we can without ambiguity always denote by $(\cdot, \cdot)$ the scalar product in H and the duality pairing of V and V' . The norm of H

will be denoted by $|\cdot|$ and there exists $k \geq 1$ such that

$$k^{-1} \|u\|_* \leq |u| \leq k \|u\|.$$

The bilinear and continuous form $a(u, v)$ defines obviously a linear and continuous operator $A: V \rightarrow V'$ by the variational equality

$$(1.16) \quad (Au, v) = a(u, v) \quad \text{for all } v \in V.$$

We can also define on the Hilbert space H an unbounded linear operator still denoted by A and called the *Lax-Milgram operator* associated to the form a . For this, set

$$D(A) := \{v \in V; \text{ there exists } k_v > 0 \text{ with } |a(v, w)| \leq k_v |w| \text{ for all } w \in V\}.$$

Then if $u \in D(A)$ the real-valued linear map $Au: v \rightarrow a(u, v)$ is continuous on V equipped with the H -norm and hence can be uniquely extended to all H to define an element of $H' \simeq H$. If moreover the form $a(u, v)$ is coercive (i.e. satisfies (ii)) then $D(A)$ is dense in V and in H , and $D(A)$ equipped with the graph norm is a Hilbert space and A is an isomorphism from $D(A)$ onto H .

When a is coercive we can also define an operator $T \in \mathcal{L}(V', V)$ by $T = A^{-1}$, i.e.

$$(1.17) \quad a(Tf, v) = (f, v) \quad \text{for all } v \in V \text{ and all } f \in V'.$$

The regularity result (RR) can then be written:

(RR) *If $f \in W \subset V'$ then $u = Tf \in Z$ and there exists $c_1 > 0$ such that*

$$\|Tf\|_Z \leq c_1 \|f\|_W.$$

Define now $a^*(u, v) := a(v, u)$ for $u, v \in V$ and A^* by $(A^*u, v) = a^*(u, v) = a(v, u)$; then A^* has the same properties as A . In particular, if a is coercive then $D(A^*)$ is dense in V and we can define $T^* \in \mathcal{L}(V', V)$ by

$$(1.18) \quad a(v, T^*f) = (f, v) \quad \text{for all } v \in V \text{ and all } f \in V'.$$

If a is symmetric and coercive, as in Example 1, where $H = L_2(\Omega)$, then $A = A^*$ and $T = T^* = A^{-1}$.

2. In many applications one also looks for error estimates in H . This can be done by the following duality argument.

THEOREM 1.2 (J. P. Aubin–J. Nitsche). *With the previous assumptions and notation*

$$(1.19) \quad |u - u_N| \leq M \|u - u_N\| \sup_{\substack{g \in H \\ |g| \leq 1}} \inf_{z \in V_N} \|T^*g - z\|.$$

Proof. For a given $g \in H$ let $\varphi \in V$ be the unique solution of

$$(1.20) \quad a^*(\varphi, v) = (g, v) \quad \text{for all } v \in V.$$

Recalling that $a(u - u_N, z) = 0$ for all $z \in V_N$ (see the proof of Theorem 1.1) and taking $v = u - u_N$ in (1.20) we obtain for all $z \in V_N$

$$a(u - u_N, \varphi - z) = (g, u - u_N)$$

which gives, since $\varphi = T^*g$,

$$|(g, u - u_N)| \leq M \|u - u_N\| \|T^*g - z\|.$$

Then (1.19) follows immediately. ■

Let us remark that (1.19) implies

$$\|u - u_N\| \leq C \|u - u_N\| e(V_N; D(A^*), V)$$

and all the considerations developed in Section 1.5 can be repeated for $\|u - u_N\|$.

2. Elementary theory of straight finite elements

2.1. We recall the basic definitions of the finite element method essentially following Ciarlet [5], where one can also find an extensive bibliography.

Let Ω be a polygonal domain in $\mathbb{R}^2$. A finite family $\mathcal{T}$ of closed triangles is a *triangulation* of Ω if $\bar{\Omega} = \bigcup_{K \in \mathcal{T}} K$ and the intersection of two triangles is either empty or reduced to a common vertex or to a common side (in this last case the triangles are adjacent). A *Lagrange finite element of type (k)* or of *degree k* ≥ 1 is a triple $\{K, \Sigma_K, P_k\}$ consisting of a nondegenerate triangle K , of the set P_k of all polynomials of degree $\leq k$ and of a certain set $\Sigma_K = \{a_1, \dots, a_d\}$ of points of K such that $d = \dim P_k$ and every polynomial $p \in P_k$ is uniquely determined by Lagrange interpolation from its values at the points of Σ_K (see the examples below). The points of Σ_K are called the *degrees of freedom* or the *nodes of K* and we say that Σ_K is P_k -*unisolvent*; since $\dim P_k = \binom{2+k}{k}$ one must have $d = \binom{2+k}{k}$.

EXAMPLE. Let K be a nondegenerate triangle with vertices a_1, a_2, a_3 . Courant's triangle or triangle of type (1) is obtained by taking $k = 1$ and $\Sigma_K = \{a_1, a_2, a_3\}$ (see Fig. 1, a). A triangle of type (2) is obtained by taking $k = 2$ and $\Sigma_K = \{a_1, a_2, a_3, a_4, a_5, a_6\}$ where a_4, a_5, a_6 are the midpoints of the sides (see Fig. 1, b). A triangle of type (3) is obtained by taking $k = 3$ and $\Sigma_K = \{a_1, a_2, a_3, a_4, a_5, a_6, a_7, a_8, a_9, a_{10}\}$ where a_4, a_5 (resp. a_6, a_7 and a_8, a_9) divide the side $a_1 a_2$ (resp. $a_2 a_3$ and $a_3 a_1$) into three equal parts and a_{10} is the barycenter of K (see Fig. 1, c).

Two Lagrange finite elements of type (k) , $\{K, \Sigma_K, P_k\}$ and $\{L, \Sigma_L, P_k\}$, are affine-equivalent since there exists a unique invertible affine mapping F of $\mathbb{R}^2$ such that $F(K) = L$ and $a_i \rightarrow F(a_i)$ defines a bijection of Σ_K onto Σ_L .

Fig. 1

For a given triangulation $\mathcal{T}$ and a given choice of Lagrange finite elements of degree k , all the elements can thus be taken affine-equivalent to a fixed reference element $\{\hat{K}, \hat{\Sigma}_{\hat{K}}, P_k\}$.

For a given triangulation $\mathcal{T}$ and a given choice of Lagrange finite elements of degree k , we define the finite element space

$$V_{\mathcal{T},k} = \{v \in W_2^1(\Omega); v|_K \in P_k \text{ for every } K \in \mathcal{T}\} \subset C^0(\bar{\Omega}).$$

Let us remark that the construction of a basis of $V_{\mathcal{T},k}$ is easy; indeed, for every node b_j of the triangulation we have to take the function whose value is 1 at b_j and 0 at all other nodes. It is important to remark that the functions of the basis have small support, since they are identically zero on each triangle K such that $b_j \notin K$.

With the notation of Theorem 1.1, if $Z \subset C^0(\bar{\Omega})$ one can define the $V_{\mathcal{T},k}$ -interpolate of $z \in Z$, i.e. the function $z_I \in V_{\mathcal{T},k}$ such that $z_I(b_j) = z(b_j)$ for all the nodes b_j of the triangulation $\mathcal{T}$. One then obtains

$$(2.1) \quad e(V_{\mathcal{T},k}; Z, V) \leq \sup_{\substack{z \in Z \\ \|z\|_Z \leq 1}} \|z - z_I\|.$$

2.2. In order to be more specific, let us consider the application of the finite element method to Example 1 of Section 1 (for a more general situation see Ciarlet [5]).

Let us define for $1 \leq p \leq +\infty$, $m \geq 0$ integer, the seminorms

$$(2.2) \quad |\varphi|_{m,p,K} = \left(\sum_{r+s=m} \int_K |\partial^m \varphi / \partial x_1^r \partial x_2^s|^p dx \right)^{1/p};$$

then since $V = W_2^1(\Omega)$ we have

$$\|z - z_I\|^2 = \sum_{K \in \mathcal{T}} \{ |z - z_I|_{1,2,K}^2 + |z - z_I|_{0,2,K}^2 \}$$

and the problem of finding error estimates is reduced to the evaluation of the local interpolation error on each element $K \in \mathcal{T}$. This can be done at first for the reference element $\{\hat{K}, \hat{\Sigma}_{\hat{K}}, P_k\}$ and then it only remains to take into

account the distortion effect of the invertible affine map from $\hat{K}$ to K . The following is a particular case of a general theorem of Ciarlet–Raviart.

THEOREM 2.1. *Let $\{K, \Sigma_K, P_k\}$ be a Lagrange finite element of degree $k \geq 1$. Let $1 \leq p, q \leq +\infty$ and let k, m be integers ≥ 0 that*

$$W_p^{k+1}(K) \subset C^0(K) \cap W_q^m(K).$$

Then there exists a constant $C(\hat{K}, \hat{\Sigma}_{\hat{K}}, P_k)$ such that for all $z \in W_p^{k+1}(K)$

$$(2.3) \quad |z - z_I|_{m,q,K} \leq C(\hat{K}, \hat{\Sigma}_{\hat{K}}, P_k) (\text{meas } K)^{1/q - 1/p} \frac{h^{k+1-m}}{(\sin \theta)^m} |z|_{k+1,p,K}$$

where $h = h(K)$ is the diameter of K and $\theta = \theta(K)$ is the smallest angle of K .

From these local estimates it is an easy matter to pass to global ones for a sequence of triangulations $\{\mathcal{T}_n\}_{n \in \mathbb{N}}$ satisfying the following regularity assumption:

(2.4) *There exists $\theta_* > 0$ such that for all $n \in \mathbb{N}$ and all $K \in \mathcal{T}_n$ one has $\theta(K) \geq \theta_*$.*

We can now give the precise error estimates for the approximation of the Dirichlet problem of Example 1 of Section 1 with Lagrange finite elements of degree k . For this let $V_{\mathcal{T}_n,k}$ be the finite element space for $\mathcal{T}_n$ and set $\dot{V}_{\mathcal{T}_n,k} = V_{\mathcal{T}_n,k} \cap \dot{W}_2^1(\Omega)$; since $\dot{V}_{\mathcal{T}_n,k}$ is included in $V = \dot{W}_2^1(\Omega)$ we use a conforming finite element method. Let us always denote by u_N the solution of the approximate problem (P_{V_N}) with $V_N = V_{\mathcal{T}_n,k}$. Then we can state the fundamental theorem.

THEOREM 2.2. *Let $\{\mathcal{T}_n\}_{n \in \mathbb{N}}$ be a sequence of triangulations satisfying (2.4), all formed by Lagrange finite elements of degree $k \geq 1$, and suppose that $h_n := \max_{K \in \mathcal{T}_n} h(K)$ satisfies $h_n \rightarrow 0$ as $n \rightarrow \infty$. If $u \in W_p^{l+2}(\Omega) \cap \dot{W}_2^1(\Omega)$ for some $l \geq 0$ and $1 < p < +\infty$, then the following error estimate holds:*

$$(2.5) \quad \|u - u_N\| \leq C h_n^{\min(k, l+2-2/p^*)} \|u\|_{W_p^{l+2}(\Omega)}$$

with a constant C independent of $u, u_N, \mathcal{T}_n$ and h_n , where $p^* = \min(p, 2)$.

This theorem needs some comments. First, the regularity assumption $u \in W_p^{l+2}(\Omega) \cap \dot{W}_2^1(\Omega)$ will be satisfied only if the data $f \in W_p^l(\Omega)$ and fulfills suitable compatibility conditions on $\partial\Omega$. As we shall see in Section 3, u has in general a singular part.

A second comment concerns the optimality of the asymptotic error bounds with respect to the N -width:

$$(2.6) \quad d_N(W_p^{l+2}(\Omega) \cap \dot{W}_2^1(\Omega), \dot{W}_2^1(\Omega)) \sim N^{-l/2-1+1/p^*} \quad \text{as } N \rightarrow \infty.$$

For this we must first find a relation between $N = \dim V_{\mathcal{T}_n, k}$ and h_n ; obviously, there exists $c > 0$ such that $\dim V_{\mathcal{T}_n, k} \geq ch_n^{-2}$, but the reverse inequality is not always fulfilled. In any case we have the following result.

PROPOSITION 2.1. *Let the assumptions of Theorem 2.2 be satisfied. Suppose also that the triangulations $\mathcal{T}_n$ are such that*

$$(2.7) \quad N = \dim V_{\mathcal{T}_n, k} \sim h_n^{-2} \quad \text{as } n \rightarrow \infty.$$

Then the error estimate (2.5) is asymptotically optimal if and only if $\min(k, l+2-2/p^) = l+2-2/p^*$.*

From this proposition it follows that the choice of the degree of the finite elements, at least for the asymptotic error estimates, is dictated by the regularity of the solution u and hence by the properties of f . Obviously this is only an abstract point of view, since we have not taken into account the computational problems connected with the effective solution of the problem (P_{V_N}) ; for a review of them see [7] and the references therein. In a more recent paper, Werschulz [25] has stated related results by using an information-theoretic approach.

3. Nonlinear approximation with finite elements and its applications

3.1. As in Section 2 let Ω be a polygonal domain in $\mathbb{R}^2$, $\mathcal{T}$ a triangulation of Ω consisting of affine-equivalent Lagrange elements of degree k , $V_{\mathcal{T}, k}$ the corresponding finite element space and $\dot{V}_{\mathcal{T}, k} = V_{\mathcal{T}, k} \cap \dot{W}_2^1(\Omega)$.

We need some further function spaces. Let $0 < p < \infty$ and define for $f \in L_p(\Omega)$ its m th order modulus of smoothness by

$$(3.1) \quad \omega_m(t, f)_p = \sup_{0 \leq |h| \leq t} \left(\int_{\Omega_{m, h}} |\Delta_h^m f(x)|^p dx \right)^{1/p}, \quad t > 0,$$

where

$$\Delta_h^m f(x) = \Delta_h^1 (\Delta_h^{m-1} f)(x), \quad \Delta_h^1 f(x) = f(x+h) - f(x)$$

and

$$\Omega_{m, h} = \{x \in \Omega; [x, x+mh] \subset \Omega\}.$$

The Nikol'skii classes $A_p^{s, m}(\Omega)$, $0 < s \leq m$,

$$(3.2) \quad A_p^{s, m}(\Omega) := \{f \in L_p(\Omega); \|f\|_{A_p^{s, m}(\Omega)} = |f|_{0, p, \Omega} + \sup_{t > 0} t^{-s} \omega_m(t, f)_p < +\infty\}$$

are usually defined and studied for $1 \leq p < \infty$ (cf. [12], [23]); here they are

also considered for $p < 1$. In that case they represent quasi-Banach spaces which still preserve some nice properties (but not all!) of the classical spaces $A_p^{s,m}(\Omega)$ with $p \geq 1$. For the connection with general function spaces of Besov–Sobolev type we refer to [24]. The definition (3.2) based on the notion of the modulus of smoothness $\omega_m(t, f)_p$ given here is more convenient for applications in approximation theory.

THEOREM 3.1. *Suppose that $1 \leq q < \infty$, $l = 0, 1$, $0 < p < q$ and $k+1 > l + 2(1/p - 1/q)$ for some integer $k \geq 1$. Then, for any $f \in A_p^{k+1, k+1}(\Omega)$ and $n \geq n_0(\Omega)$, there exists a triangulation $\mathcal{T}_n$ and an element $v_n \in V_{\mathcal{T}_n, k}$ such that:*

$$(3.3) \quad \|f - v_n\|_{W_q^l(\Omega)} \leq C n^{-(k+1-l)/2} \|f\|_{A_p^{k+1, k+1}(\Omega)},$$

$$(3.4) \quad \mathcal{T}_n \text{ contains } \leq n \text{ triangles; } h_n \leq C n^{-1/2},$$

$$(3.5) \quad \mathcal{T}_n \text{ is regular (cf. (2.4)) with } \theta_* = \theta_*(\Omega) > 0.$$

If, in addition, $k+1 > 2/p$ and $f \in \dot{W}_2^1(\Omega)$ then we also have $v_n \in \dot{V}_{\mathcal{T}_n, k}$.

For $\Omega = [0, 1]^2$, Theorem 3.1 is actually contained in [14, Section 3] as a very particular case. The proof of Theorem 3.1 and of more general results will be given in a forthcoming paper [15]. Here we sketch the proof in the case of linear elements ($k = 1$) and $p \geq 1$ only.

First, starting from some initial triangulation $\mathcal{T}^{(0)}$ of Ω we construct, by the standard subdivision procedure (cf. Fig. 2), a sequence of *dyadic* triangulations $\mathcal{T}^{(0)}, \mathcal{T}^{(1)}, \mathcal{T}^{(2)}, \dots$ and the corresponding finite element spaces $V^{(0)} \subset V^{(1)} \subset V^{(2)} \dots$ where $V^{(r)} = V_{\mathcal{T}^{(r)}, 1}$ satisfies $N^{(r)} = \dim V^{(r)} \sim 2^{2r}$.

Fig. 2. The first 3 dyadic triangulations $\mathcal{T}^{(r)}$ for an L-shaped domain

By using the standard approximation techniques for finite elements (e.g. interpolants (if $k+1 > 2/p$) or quasiinterpolants) we obtain a sequence $v^{(r)} \in V^{(r)}$ satisfying

$$(3.6) \quad |f - v^{(r)}|_{0,p,\Omega} \leq C \omega_{k+1}(2^{-r}, f)_p \leq C 2^{-r(k+1)} \|f\|_{A_p^{k+1, k+1}(\Omega)}.$$

(Let us briefly mention that in this formulation, in terms of the modulus of smoothness, it is possible to generalize (3.6) also to the case of $p < 1$ while for $p \geq 1$ it is essentially equivalent to the usual seminorm estimates (2.3)).

Now, fix some integer $s = 0, 1, \dots$ and for $r > s$ consider the representations

$$\tilde{v}^{(r)} := v^{(r)} - v^{(r-1)} = \sum_{i=1}^{N^{(r)}} \tilde{a}_{r,i} w_i^{(r)}$$

in terms of the standard basis functions in $V^{(r)}$. Defining a new set of coefficients

$$(3.7) \quad \tilde{a}_{r,i}^{(s)} = \begin{cases} \tilde{a}_{r,i} & \text{if } |\tilde{a}_{r,i}| > 2^{\alpha r + \beta s} \|f\|_{\Lambda_p^{k+1, k+1}(\Omega)}, \\ 0 & \text{otherwise,} \end{cases}$$

where $\alpha = -k - 1 + (2 + \varepsilon)/p$, $\beta = -(2 + \varepsilon)/p$ with some $\varepsilon > 0$, we set

$$(3.8) \quad \bar{v}_s = v^{(s)} + \sum_{r=s+1}^{\infty} \tilde{v}_s^{(r)}, \quad \tilde{v}_s^{(r)} = \sum_{i=1}^{N^{(r)}} \tilde{a}_{r,i}^{(s)} w_i^{(r)}.$$

Roughly speaking, $\bar{v}_s$ is obtained from $v^{(s)}$ by adding the basis functions $\tilde{a}_{r,i} w_i^{(r)}$ with large coefficients from the decomposition $f = v^{(s)} + \sum_{r=s+1}^{\infty} \tilde{v}^{(r)}$. Using the straightforward relation

$$(3.9) \quad |v|_{0, \nu, \Omega} \sim (2^{-2r} \sum_{i=1}^{N^{(r)}} |a_{r,i}|^{\nu})^{1/\nu}$$

$$\text{for all } v = \sum_{i=1}^{N^{(r)}} a_{r,i} w_i^{(r)} \in V^{(r)}, \quad 0 < \nu < \infty,$$

from (3.6)–(3.8) we easily get

$$M_s^{(r)} := \text{card} \{i; \tilde{a}_{r,i}^{(s)} \neq 0\} \leq C 2^{-\alpha r} 2^{(2+\varepsilon)s}, \quad r > s.$$

This shows that the number of additional summands in (3.8) is bounded by

$$M_s := \text{card} \{(r, i); \tilde{a}_{r,i}^{(s)} \neq 0, r > s\} = \sum_{r>s} M_s^{(r)} \leq C 2^{2s}.$$

Thus the representation (3.8) of $\bar{v}_s$ implies that the triangulation $\bar{\mathcal{T}}_s$ corresponding to $\bar{v}_s$ represents a suitable refinement of $\mathcal{T}^{(s)}$ induced by $O(2^{2s})$ smaller triangles contained in the support of some $w_i^{(r)}$ with $\tilde{a}_{r,i}^{(s)} \neq 0$, $r > s$. After some additional considerations it turns out that for some $n \sim 2^{2s}$ the triangulation $\mathcal{T}_n = \bar{\mathcal{T}}_s$ satisfies the conditions (3.4), (3.5).

Concerning (3.3), we first obtain by (3.6)–(3.9) the estimate for $r > s$

$$|\tilde{v}^{(r)} - \bar{v}_s^{(r)}|_{0, q, \Omega} \leq C 2^{(\alpha r + \beta s)(1 - p/q)} \cdot 2^{-(k+1)r p/q} \|f\|_{\Lambda_p^{k+1, k+1}(\Omega)}$$

which immediately gives the case $l = 0$. Indeed, setting $v_n = \bar{v}_s$, we find

$$\begin{aligned} |f - v_n|_{0,q,\Omega} &= |f - \bar{v}_s|_{0,q,\Omega} \leq \sum_{r>s} |\tilde{v}^{(r)} - \bar{v}_s^{(r)}|_{0,q,\Omega} \\ &\leq C \|f\|_{\Lambda_p^{k+1,k+1}(\Omega)} 2^{\beta s(1-p/q)} \sum_{r>s} 2^{(\alpha(1-p/q)-(k+1)p/q)r} \\ &\leq C 2^{-(k+1)s} \|f\|_{\Lambda_p^{k+1,k+1}(\Omega)} \leq C n^{-(k+1)/2} \|f\|_{\Lambda_p^{k+1,k+1}(\Omega)}, \end{aligned}$$

since $\alpha(1-p/q)-(k+1)p/q = -(k+1)+(2+\varepsilon)(1/p-1/q) < 0$ with a proper choice of $\varepsilon > 0$.

For $l = 1$ we additionally have to apply to $v := \tilde{v}^{(r)} - \bar{v}_s^{(r)} \in V^{(r)}$ some inequality of Bernstein type (or inverse inequality, cf. [5, Theorem 3.2.6]):

$$(3.10) \quad \|v\|_{W_q^1(\Omega)} \leq C 2^r |v|_{0,q,\Omega}, \quad \text{for all } v \in V^{(r)}.$$

The modifications needed for the zero boundary values are obvious (since for $k+1 > 2/p$ one has $\Lambda_p^{k+1,k+1}(\Omega) \subset C^0(\bar{\Omega})$ the usual interpolants do the job).

For the other cases, especially the harder proof $p < 1$, we refer the reader to [14], [15].

Remark. The proof as sketched above is, in some sense, a pure existence result, i.e. it does not immediately lead to an effective construction of the triangulations $\mathcal{T}_n$ which, in general, really depend on the given f . However, in some important cases, e.g. for functions with typical singularities such as given in Section 3.2, where it is reasonable to apply Theorem 3.1 (cf. Section 3.3) the optimal triangulation can be explicitly determined. For the one-dimensional case, see [3].

3.2 Let us now return to the Dirichlet problem (1.2) in a polygonal domain Ω . We denote by $S_1, S_2, \dots, S_M$ ($= S_0$) the vertices, by $\omega_1, \omega_2, \dots, \omega_M$ ($\omega_j \leq 2\pi$) the interior angles and by $\Gamma_1, \Gamma_2, \dots, \Gamma_M$ ($\Gamma_j = S_{j-1}S_j$) the sides of $\partial\Omega$ (see Fig. 3 for an example).

Fig. 3

Let $\varphi_j(x)$, $j = 0, 1, \dots, M$, be a resolution of unity with respect to Ω , for simplicity defined by

$$(3.11) \quad \varphi_j(x) = \varphi(\text{dist}(P, S_j)) \quad \text{where } P := x \in \Omega, j = 1, \dots, M,$$

$$\varphi_0(x) = 1 - \sum_{j=1}^M \varphi_j(x),$$

where $\varphi(r) \in C^\infty([0, \infty[)$ satisfies $\varphi(r) = 1$ for $r \in [0, r_0/2]$, $\varphi(r) = 0$ for $r > r_0$ and $r_0 > 0$ is chosen so small that

$$\text{supp } \varphi_j(x) \cap \text{supp } \varphi_i(x) = \emptyset \quad \text{for } i \neq j, i, j = 1, \dots, M.$$

In order to describe the singularities near the vertices S_j in the solution of (1.2), following Grisvard [10] we introduce the functions

$$(3.12) \quad \sigma_{ij}(x) = \varphi_j(r) \begin{cases} r^{\alpha_i} \sin \alpha_i \theta & \text{if } \alpha_i := i\pi/\omega_j \text{ is noninteger,} \\ r^{\alpha_i} (\log r \cdot \sin \alpha_i \theta + \theta \cos \alpha_i \theta) & \text{if } \alpha_i \text{ is an integer,} \end{cases}$$

where $r = \text{dist}(P, S_j)$ and θ are the local polar coordinates at S_j ($\theta = 0$ if $x \in \Gamma_{j+1}$). With this notation we have the following regularity result (see Grisvard [10]).

THEOREM 3.2 (P. Grisvard). *Let $1 < q < \infty$, $l = 0, 1, 2, \dots$. Suppose that all the numbers $(l+2-2/q)\omega_j/\pi$ are noninteger and put $\kappa_j = [(l+2-2/q)\omega_j/\pi]$, $j = 1, 2, \dots, M$. Then, for any $f \in W_q^l(\Omega)$, the weak solution $u \in \dot{W}_2^1(\Omega)$ of (1.2) has a unique representation*

$$(3.13) \quad u(x) = u_{\text{reg}}(x) + u_{\text{sing}}(x)$$

where

$$(3.14) \quad u_{\text{reg}} \in W_q^{l+2}(\Omega), \quad \|u_{\text{reg}}\|_{W_q^{l+2}(\Omega)} \leq C \|f\|_{W_q^l(\Omega)},$$

$$(3.15) \quad u_{\text{sing}}(x) = \sum_{j=1}^M \sum_{i=1}^{\kappa_j} c_{ij} \sigma_{ij}(x), \quad |c_{ij}| \leq C \|f\|_{W_q^l(\Omega)}.$$

EXAMPLE 1. In the case of the L-shaped domain of Fig. 3, we have at most one singularity for $l = 0$:

$$(3.16) \quad u_{\text{sing}}(x) = \begin{cases} c_{11} \varphi_1(r) r^{2/3} \sin(2\theta/3), & q > 3/2, \\ 0, & q < 3/2, \end{cases}$$

while for $l \geq 1$ the expressions become (in dependence on q) more complicated. E.g., for $l = 2$, $q = 2$:

$$(3.17) \quad u_{\text{sing}}(x) = \varphi_1(x) [c_{11} r^{2/3} \sin(2\theta/3) + c_{21} r^{4/3} \sin(4\theta/3) \\ + c_{31} r^2 (\log r \cdot \sin 2\theta + \theta \cos 2\theta) \\ + c_{41} r^{8/3} \sin(8\theta/3)] + \sum_{j=2}^6 c_{1j} \sigma_{1j}(x)$$

where the functions $\sigma_{1j}(x)$, $j = 2, \dots, 6$, coincide (in local coordinates) with the indicated $\sigma_{31}(x)$.

We need the following easy result.

PROPOSITION 3.1. *Let $\omega = \max \omega_j$ and $l = 0, 1, 2, \dots$. Then, under the assumptions of Theorem 3.2, the function u_{sing} belongs to $\Lambda_p^{l+2, l+2}(\Omega)$ for $0 < p < 2/(l+2 - \pi/\omega)$ (note that $l+2 > \pi/\omega$), and*

$$(3.18) \quad \|u_{\text{sing}}\|_{\Lambda_p^{l+2, l+2}(\Omega)} \leq C \|f\|_{W_q^l(\Omega)}.$$

Proof. Actually, in view of (3.11), (3.12) and Theorem 3.2 (cf. also the definition of $\Lambda_p^{m,m}(\Omega)$) it is sufficient to prove that functions of the type

$$g(x) = \varphi(r) r^\alpha (\psi_1(\theta) \log r + \psi_2(\theta)), \quad x \in \Omega_\omega = \{r > 0, 0 < \theta < \omega\},$$

belong to $\Lambda_p^{m,m}(\Omega)$ for $m > \alpha$ and $0 < p < 2/(m - \alpha)$. Here, $\varphi(r)$ is defined after (3.11), and $\psi_1(\theta), \psi_2(\theta) \in C^\infty([0, \omega])$. For this we check the modulus of smoothness $\omega_m(t, g)_p$. Let $h \in \mathbb{R}^2$ be arbitrarily chosen, and define

$$E_1 := \{x \in \Omega_\omega; r \leq 2m|h|\}, \quad E_2 := \{x \in \Omega_\omega; r \geq r_0 + m|h|\}, \\ E_3 := \Omega_\omega \setminus (E_1 \cup E_2).$$

Clearly

$$\Delta_h^m g(x) = 0 \quad \text{for } x \in E_2.$$

Thus, for $0 < p < 2/(m - \alpha)$,

$$\begin{aligned} \int_{E_1 \cup E_2} |\Delta_h^m g(x)|^p dx &= \int_{E_1} |\Delta_h^m g(x)|^p dx \leq C \int_{r \leq 3m|h|} |g(x)|^p dx \\ &\leq C \int_0^{\min(r_0, 3m|h|)} r^{1+\alpha p} |\log r|^p dr \leq C |h|^{mp}. \end{aligned}$$

On the other hand, $E_3 \neq \emptyset$ iff $|h| < r_0/m$. In this case, by a straightforward estimation of the partial derivatives of $g(x)$ up to order m , we get

$$\begin{aligned} |\Delta_h^m g(x)| &\leq C |h|^m \max_{y \in [x, x+mh]} |(\partial/\partial h)^m g(y)| \leq C |h|^m r^{\alpha-m} (1 + |\log r|), \\ \int_{E_3} |\Delta_h^m g(x)|^p dx &\leq C |h|^{mp} \int_{2m|h|}^{2r_0} r^{1+(\alpha-m)p} (1 + |\log r|)^p dr \leq C |h|^{mp}. \end{aligned}$$

Putting together the estimates we see that $\omega_m(t, g)_p \leq Ct^m$. ■

3.3. The basic idea for improving the asymptotic error estimates for the finite element approximation of the Dirichlet problem (1.2) in a polygonal domain Ω is to deal first with the singular part u_{sing} by Theorem 3.1. For this, we fix $q = 2$, $l = k - 1$, we choose $p > 0$ with $2/(k+1) < p < 2/(k+1 - \pi/\omega)$ and we take into account Proposition 3.1 with $l = k - 1$. This leads to a

sequence of triangulations $\mathcal{T}_n$ and the corresponding N -dimensional subspaces $V_N = \dot{V}_{\mathcal{T}_n, k} \subset \dot{W}_2^1(\Omega)$ such that

$$\inf_{v \in V_N} \|u_{\text{sing}} - v\| = O(N^{-k/2}) \quad \text{as } N \rightarrow \infty.$$

On the other hand, if $f \in W_2^{k+1}(\Omega)$ then the regular part $u_{\text{reg}} \in W_2^{k+1}(\Omega)$ can be approximated as

$$\inf_{v \in V_N} \|u_{\text{reg}} - v\| = O(N^{-k/2}) \quad \text{as } N \rightarrow \infty$$

with respect to the same sequence of subspaces. This follows from the regularity properties of the triangulations $\mathcal{T}_n$ (cf. (3.4), (3.5) and the results of Section 2). Clearly, this argument can be applied to other norms and regularity assumptions. To be more precise let us first consider the case $k = 1$ (linear elements).

THEOREM 3.3. *Let Ω be a polygonal domain with interior angles $\omega_1, \dots, \omega_M$. Then there exists a sequence $\{\mathcal{T}_n, n \geq n_0(\Omega)\}$ of triangulations such that the following properties hold:*

(3.19) *The finite element spaces $V_N = \dot{V}_{\mathcal{T}_n, 1}$ have dimension N .*

(3.20) *$\mathcal{T}_n$ is regular and $h_n \sim N^{-1/2}$ as $N \rightarrow \infty$.*

(3.21) *For $f \in L_2(\Omega)$ the finite element approximations $u_N \in V_N$ of the solution $u \in \dot{W}_2^1(\Omega)$ of (1.2) satisfy*

$$N^{-1/2} |u - u_N|_{1,2,\Omega} + |u - u_N|_{0,2,\Omega} \leq CN^{-1} \|f\|_{0,2,\Omega}.$$

Proof. Theorem 3.2 with $l = 0$ and $q = 2$ yields

$$u(x) = u_{\text{reg}}(x) + u_{\text{sing}}(x)$$

where

$$\|u_{\text{reg}}\|_{W_2^2(\Omega)} \leq C \|f\|_{L_2(\Omega)}$$

and

$$(3.22) \quad u_{\text{sing}}(x) = \sum_{j: \omega_j > \pi} c_j \sigma_j(x), \quad |c_j| \leq C \|f\|_{L_2(\Omega)}.$$

The singular functions $\sigma_j(x)$ ($\equiv \sigma_{1j}(x)$) are of the form

$$(3.23) \quad \sigma_j(x) = \varphi_j(r) r^{\alpha_j} \sin \alpha_j \theta, \quad \alpha_j = \pi/\omega_j \in (1/2, 1)$$

(here r, θ are local polar coordinates at the corresponding S_j ; see (3.12)). Thus, from Proposition 3.1 we see that $u_{\text{sing}} \in A_p^{2,2}(\Omega) \cap C^0(\bar{\Omega})$ for $p < 4/3$. Now, for fixed $p \in (1, 4/3)$, $q = 2$ and $l = 0, 1$, Theorem 3.1 ensures (with a proper choice of $n \sim N$) the existence of triangulations $\mathcal{T}_n$ and subspaces V_N

satisfying (3.19), (3.20) and

$$(3.24) \quad \|u_{\text{sing}} - v_N\|_{W_2^l(\Omega)} \leq CN^{-(2-l)/2} \|u_{\text{sing}}\|_{A_p^{2,2}} \leq CN^{-(2-l)/2} \|f\|_{L_2(\Omega)}.$$

It should be mentioned that the triangulation $\mathcal{T}_n$ can, actually, be chosen independent of the concrete values c_j in the representation of u_{sing} . For this, we have only to take into account that the functions $\sigma_j(x)$ have nonintersecting supports, and look at the proof of Theorem 3.1.

Since $u_{\text{reg}} \in W_2^2(\Omega)$ the proof of (3.21) can be finished by the standard results of the preceding section and by remarking that

$$\inf_{v \in V_N} \|u - v\| \leq \|u_{\text{sing}} - v_N\| + \inf_{v \in V_N} \|u_{\text{reg}} - v\|.$$

Remarks. 1. Clearly, in this result it is actually not necessary to have Theorem 3.1: any *ad hoc* chosen triangulation which realizes the estimate (3.24) for the concrete u_{sing} given in (3.22), (3.23) yields the same result (cf. e.g. [10; 8.4.1] or [13, Ch. 6]).

2. Usually, error estimates for the finite element approximation of (1.2) in polygonal domains are given in terms of weighted norms and mesh refinements around the corners (see e.g. [20], [18], [19], [2], ...). For instance, G. Raugel [18], [19] has obtained the same optimal error estimates on triangulations such that: (α) if a triangle K has no vertices common with $S_1, \dots, S_M$ then its diameter $h(K)$ satisfies $h(K) \leq CN^{-1/2}$, and (β) if K has vertex S_j then

$$h(K) \leq CN^{-(1/2)((l+1)/(l+1-\kappa_j\pi/\omega_j))}.$$

In order to compare this result with (3.20) one has to take $l = 0$.

3.4. Let us finish this section by two further examples.

EXAMPLE 2. Let Ω be the L-shaped domain of Fig. 3, and consider the case $k = 3$ (cubic elements). The description of the corresponding singular part u_{sing} of the solution $u \in W_2^1(\Omega)$ of (1.2) with $f \in W_2^2(\Omega)$ has been given in (3.17). Before applying Theorem 3.1 we define a certain modification of u_{sing} by including the terms $c_{ij} \varphi(r) r^2 \theta \cos \theta$ into the regular part. The u'_{sing} so defined belongs to $A_p^{4,4}(\Omega) \cap C^0(\bar{\Omega})$ if $p < 3/5$ thanks to Proposition 3.1. Now, Theorem 3.1, when applied to some $1/2 < p < 3/5$, $q = 2$, $l = 1$, yields the existence of triangulations $\mathcal{T}_n$ and finite element spaces $V_N = \dot{V}_{\mathcal{T}_n,3}$ of cubic elements such that for $n \geq n_0(\Omega)$

$$\inf_{v \in V_N} \|u'_{\text{sing}} - v\| \leq CN^{-3/2} \|u'_{\text{sing}}\|_{A_p^{4,4}} \leq CN^{-3/2} \|f\|_{W_2^2(\Omega)}.$$

Now, proceeding with the modified regular part u'_{reg} as in Theorem 3.3 we obtain

$$(3.25) \quad \|u - u_N\| \leq CN^{-3/2} \|f\|_{W_2^2(\Omega)} \quad \text{as } N \rightarrow \infty.$$

It should be mentioned that even in this case the triangulations $\mathcal{T}_n$ which realize (3.25) could be defined independent of the coefficients c_{ij} in (3.17) (although this does not immediately follow from Theorem 3.1). Furthermore, it is surprising that strange function spaces (here $\Lambda_p^{m,m}(\Omega)$ with $p < 1$) could have nice applications! Let us also remark that also in this case G. Raugel [18], [19] has obtained the same error estimate (3.25).

EXAMPLE 3. Let us suppose that $f \in L_q(\Omega)$ for some $1 < q < 2$. For simplicity, let us consider once more the L-shaped Ω of Fig. 3 and linear elements ($k = 1$). Using the regularity result given in Theorem 3.2 ($q \neq 3/2$, cf. (3.16)), we see that in any case $u \in \Lambda_p^{2,2}(\Omega) \cap C^0(\bar{\Omega})$ for $1 < p < \min(q, 3/2)$ (use Proposition 3.1 and the trivial embeddings $W_q^2(\Omega) \subset \Lambda_q^{2,2}(\Omega) \subset \Lambda_p^{2,2}(\Omega)$, $p < q$). Thus, Theorem 3.1 directly leads to a sequence of triangulations $\mathcal{T}_n$ and N -dimensional finite element spaces $V_N = \hat{V}_{\mathcal{T}_n,1}$ with the estimate ($q \neq 3/2$)

$$(3.26) \quad \|u - u_N\| \leq CN^{-1/2} \|u\|_{\Lambda_p^{2,2}} \leq CN^{-1/2} \|f\|_{L_q(\Omega)}.$$

Here, it is important to point out the dependence of the sequence $\mathcal{T}_n$ on the u (or f), i.e. on the *individual* right-hand side in (1.2). Let us recall that for any sequence $\mathcal{T}_n$ or V_N fixed for all $f \in L_q(\Omega)$ according to Proposition 2.1 the optimal asymptotic error estimate is bounded from below by $cN^{-(1-1/q)}$. Thus, (3.26) shows a further situation where, in principle, the asymptotic error can be significantly decreased by choosing triangulations better adapted to the solution. However, it is an open problem to give an effective construction of such optimal triangulations.

4. Internal approximation of nonlinear equations: preliminaries

4.1. With the notation of Section 1, let W be a reflexive Banach space such that $H \subset W \subset V'$ with continuous embeddings. We assume, moreover that

$$(4.1) \quad \text{The canonical injection of } W \text{ into } V' \text{ is compact.}$$

This implies that also the canonical injection of V into H is compact.

For some interval $I \subset \mathbf{R}$ let $G: V \times I \rightarrow W$ be a C^p mapping ($p \geq 2$) and consider the nonlinear problem

(NLP) Find pairs $(u, \lambda) \in V \times I$ which are solutions of

$$(4.2) \quad a(u, v) + (G(u, \lambda), v) = 0 \quad \text{for all } v \in V.$$

If we define an operator T by (1.17), then an equivalent form of (NLP) consists in finding pairs $(u, \lambda) \in V \times I$ which are solutions of

$$(4.3) \quad F(u, \lambda) := u + TG(u, \lambda) = 0.$$

Obviously, (4.1) implies that T is compact from W to V ; moreover, if the regularity condition (RR) is satisfied then $u \in Z$.

4.2. Let now V_N be a fixed subspace of V of finite dimension N . Then we can consider the nonlinear approximate problem

(NLP $_{V_N}$) Find pairs $(u_N, \lambda) \in V_N \times I$ such that

$$(4.4) \quad a(u_N, v) + (G(u_N, \lambda), v) = 0 \quad \text{for all } v \in V_N.$$

If we define an operator $T_N \in \mathcal{L}(V', V_N)$ by

$$(4.5) \quad a(T_N f, v) = (f, v) \quad \text{for all } f \in V' \text{ and all } v \in V_N,$$

then an equivalent form of (NLP $_{V_N}$) consists in finding pairs $(u, \lambda) \in V_N \times I$ which are solutions of

$$(4.6) \quad F_N(u_N, \lambda) := u_N + T_N G(u_N, \lambda) = 0.$$

If the completeness condition (CC) holds, then we obtain, since $T \in \mathcal{L}(W, V)$ is compact,

$$(4.7) \quad \lim_{N \rightarrow \infty} \|T - T_N\|_{\mathcal{L}(W, V)} = 0.$$

It is not difficult to find, for a fixed λ , suitable conditions on the operator $F(u, \lambda)$ such that $\|u - u_N\| \rightarrow 0$ as $N \rightarrow \infty$, and also such that an estimate of type

$$\|u - u_N\| \leq Cd(u, V_N)$$

holds; see e.g. Ciarlet [5] and the bibliography therein.

More difficult (and interesting) problems are connected with the simultaneous approximation of a set of solutions. We shall give some results in this direction mainly following the ideas of Brezzi–Rappaz–Raviart [4].

4.3. EXAMPLE 1. Let Ω be a bounded open polygonal set of $\mathbf{R}^2$. We consider the classical nonlinear eigenvalue problem of finding $(u, \lambda) \in \dot{W}_2^1(\Omega) \times \mathbf{R}$ such that

$$(4.8) \quad \begin{cases} -\Delta u - \lambda u + u^3 = 0 & \text{in } \Omega, \\ u = 0 & \text{on } \partial\Omega. \end{cases}$$

Since $G(u, \lambda): (u, \lambda) \rightarrow -\lambda u + u^3$ is a C^∞ mapping from $\dot{W}_2^1(\Omega) \times \mathbf{R}$ into $L_p(\Omega)$ for all $p \geq 2$ and the canonical embedding of $L_p(\Omega)$ into $W_2^{-1}(\Omega)$ is compact,

we are in the framework of the theory previously developed. In order to establish the regularity of the solution u , we may proceed in the following way. A first use of Theorem 3.2 with $f = -G(u, \lambda)$ implies that $u = u_{\text{reg}} + u_{\text{sing}}$ with $u_{\text{reg}} \in W_p^2(\Omega)$ for all $p \geq 2$ and u_{sing} given by (3.15). But then a careful inspection of (3.12) implies that at least $G(u, \lambda) \in W_p^{1/2+2/p}(\Omega)$, and a second use of Theorem 3.2 yields that any solution u of (4.8) has a decomposition $u = u_{\text{reg}} + u_{\text{sing}}$ with $u_{\text{reg}} \in W_p^{5/2+2/p}(\Omega)$. Clearly, for any concrete domain Ω the smoothness of u_{reg} might be further improved.

EXAMPLE 2. Let Ω be a polygonal domain in $\mathbb{R}^2$. We consider the nonlinear eigenvalue problem of finding $(u, \lambda) \in \dot{W}_2^1(\Omega) \times \mathbb{R}$ such that

$$(4.9) \quad \begin{cases} -\Delta u + \alpha(x) \cdot \text{grad}(u^2) - \lambda u = 0 & \text{in } \Omega, \\ u = 0 & \text{on } \partial\Omega, \end{cases}$$

where $\alpha(x)$ is a C^∞ vector field on $\bar{\Omega}$. Thanks to the Sobolev embedding theorem, if $u \in W_2^1(\Omega)$ then $uu_{x_1}, uu_{x_2} \in L_q(\Omega)$ for all q with $1 < q < 2$; moreover, the embedding of $L_q(\Omega)$ into $W_2^{-1}(\Omega)$ is compact for $1 < q < 2$. Therefore

$$G(u, \lambda): (u, \lambda) \rightarrow \alpha(x) \cdot \text{grad}(u^2) - \lambda u$$

is a C^∞ mapping from $\dot{W}_2^1(\Omega) \times \mathbb{R}$ into $L_q(\Omega)$ for all $1 < q < 2$ and we are in the framework of the theory previously developed. For the regularity condition we can argue as in the previous example. If (u, λ) is a solution of (4.9) then $u \in \dot{W}_2^1(\Omega)$ satisfies $-\Delta u = f$ with $f \in L_q(\Omega)$, $1 < q < 2$, and so $u = u_{\text{reg}} + u_{\text{sing}}$ with $u_{\text{reg}} \in W_q^2(\Omega)$ and u_{sing} given by (3.15). But this regularity can be hopefully improved for a particular domain Ω . For instance, in the case of the L-shaped domain Ω of Fig. 3, u_{sing} is given by (3.16) and thus the most singular part of u^2 is indeed given by $(u_{\text{sing}})^2 \sim r^{4/3} \sin^2(2\theta/3) \in W_p^2(\Omega)$ for $p < 3$. Therefore in this case, if (u, λ) is a solution of (4.9) then indeed $G(u, \lambda) \in W_s^1(\Omega)$ for all $s < 3$. In order to proceed further we must now consider the singular functions corresponding to all corners of the L-shaped domain Ω .

EXAMPLE 3. Let us consider the problem

$$(4.10) \quad \begin{cases} -u'' + (u')^2 - \lambda u = 0 & \text{in } I = (0, 1), \\ u(0) = u(1) = 0. \end{cases}$$

Besides the trivial solution $u = 0$, we can find for $\lambda = 8$ the solution $u(x) = 2x(x-1)$. Here

$$G(u, \lambda): (u, \lambda) \rightarrow -\lambda u + (u')^2$$

is a C^∞ mapping from $W_2^1(I)$ into $L_1(I)$ whose canonical embedding into $W_2^{-1}(I)$ is compact. Thus we are in the framework of the previous theory. The regularity of the problem is obvious.

5. Approximation of a branch of nonsingular solutions

With the assumptions and notation of Section 4, we say that $\mathcal{B} = \{(u(\lambda), \lambda); \lambda \in \Lambda\}$ with Λ a compact subinterval of I is a *branch of nonsingular solutions of (NLP)* if the following properties are satisfied:

$$(5.1) \quad F(u(\lambda), \lambda) = 0 \text{ for all } \lambda \in \Lambda.$$

$$(5.2) \quad D_u F(u(\lambda), \lambda) \in \mathcal{L}(V, V) \text{ is an isomorphism for all } \lambda \in \Lambda.$$

$$(5.3) \quad \lambda \rightarrow u(\lambda) \text{ is a continuous function from } \Lambda \text{ to } V.$$

Indeed, by the implicit function theorem it follows that $\lambda \rightarrow u(\lambda)$ is a C^p function.

It is natural to look for a branch $\mathcal{B}_N$ of nonsingular solutions of (NLP_{V_N}) . The following theorem of Brezzi–Rappaz–Raviart [4, Part I] answers the question; see also Descloux–Rappaz–Scholz [6] for similar error estimates in other norms.

THEOREM 5.1. *Assume that $D^p G$ is bounded on all bounded subsets of $V \times \Lambda$. Then there exists $\alpha_0 > 0$ and, for all $N \geq N_0$ large enough, a unique C^p function*

$$\lambda \in \Lambda \rightarrow u_N(\lambda) \in V_N$$

such that for all $\lambda \in \Lambda$

$$(5.4) \quad F_N(u_N(\lambda), \lambda) = 0,$$

$$(5.5) \quad \|u_N(\lambda) - u(\lambda)\| \leq \alpha_0.$$

Furthermore, for some constants K_r , $r = 0, \dots, p-1$, independent of $N \geq N_0$ and of $\lambda \in \Lambda$, the following estimates hold:

$$(5.6) \quad \|u_N(\lambda) - u(\lambda)\| \leq K_0 \{ \|u(\lambda) - P_{V_N} u(\lambda)\| + \|(T_N - T)G(u(\lambda), \lambda)\| \},$$

$$(5.7) \quad \|u_N^{(r)}(\lambda) - u^{(r)}(\lambda)\|$$

$$\leq K_r \sum_{k=0}^r \{ \|u^{(k)}(\lambda) - P_{V_N} u^{(k)}(\lambda)\| + \|(T_N - T)G^{(k)}(u(\lambda), \lambda)\| \},$$

$$r = 1, \dots, p-1,$$

where $G^{(k)}(u(\lambda), \lambda)$ denotes the (total) derivative with respect to λ of the function $G(u(\lambda), \lambda)$.

The result obtained in Theorem 5.1 is fundamental because it reduces the problem of error estimates to a regularity problem for the linear operator T and to an approximation problem. Indeed, with the notation of Section 1, if the regularity property (RR) holds for the problem (P) then $G(u(\lambda), \lambda) \in W$

implies $u(\lambda) = -TG(u(\lambda), \lambda) \in Z$. Therefore also for this type of nonlinear problem we can look for asymptotic optimal error estimates as has been done for linear problems in Section 1.

A more difficult problem seems to be to find asymptotically improved error estimates for a suitable family $\mathcal{V}_N$ as has been done in Section 3 for the linear Dirichlet problem. Indeed, such a family $\mathcal{V}_N$ has to work for all the points of the branch $\mathcal{B}$. For instance, for the nonlinear eigenvalue problem (4.9) (Example 2 of Section 4) the situation is particularly complicated because the asymptotic improved error estimate (3.26) (Example 3 of Section 3) has been obtained for a family $\mathcal{V}_N$ depending on u .

The proof of Theorem 5.1 is based on a suitable version of the implicit function theorem.

6. Approximation near a bifurcation point

6.1. Thanks to (4.1) for fixed (u_0, λ_0) the map

$$D_u F(u_0, \lambda_0) = I + TD_u G(u_0, \lambda_0) \in \mathcal{L}(V, V)$$

is a compact perturbation of the identity and so, when singular, it satisfies

$$(6.1) \quad \begin{cases} \dim \text{Ker } D_u F(u_0, \lambda_0) = \text{codim Im } D_u F(u_0, \lambda_0), \\ \text{Range } D_u F(u_0, \lambda_0) \text{ is closed in } V. \end{cases}$$

We shall say that (u_0, λ_0) is a *bifurcation point* (resp. a *limit point*) of F , of multiplicity $m \geq 1$ if:

$$(6.2) \quad \begin{aligned} & \text{(i)} \quad F(u_0, \lambda_0) = 0. \\ & \text{(ii)} \quad D_u F(u_0, \lambda_0) \text{ is singular and } \dim \text{Ker } D_u F(u_0, \lambda_0) = m. \\ & \text{(iii)} \quad D_\lambda F(u_0, \lambda_0) \in \text{Range } D_u F(u_0, \lambda_0) \\ & \quad \text{(resp. } D_\lambda F(u_0, \lambda_0) \notin \text{Range } D_u F(u_0, \lambda_0)). \end{aligned}$$

Let us first recall the classical Lyapunov–Schmidt procedure used to solve (4.2) in a neighborhood of (u_0, λ_0) .

Thanks to (6.2) (ii) and the classical Riesz–Schauder theory of compact operators, there exist m linearly independent eigenvectors $\varphi_i \in V$, $i = 1, \dots, m$, such that

$$(6.3) \quad V_1 := \text{Ker } D_u F(u_0, \lambda_0) = \text{span} \{ \varphi_1, \dots, \varphi_m \}.$$

It is an easy matter to check that

$$(6.4) \quad V_1 = \{ \varphi \in V; a(\varphi, v) + (D_u G(u_0, \lambda_0) \varphi, v) = 0 \text{ for all } v \in V \}.$$

Moreover, if we set

$$\begin{aligned} K &:= \text{Ker}(I + T^*(D_u G(u_0, \lambda_0))^*) \\ &= \{\psi^* \in V; a(v, \psi^*) + (D_u G(u_0, \lambda_0)v, \psi^*) = 0 \text{ for all } v \in V\}, \end{aligned}$$

it is possible to choose $\psi_1^*, \dots, \psi_m^* \in V$ such that

$$a(\varphi_i, \psi_j^*) = \delta_{ij}, \quad i, j = 1, \dots, m,$$

$$K = \text{span}\{\psi_1^*, \dots, \psi_m^*\},$$

$$V_2 := \text{Range } D_u F(u_0, \lambda_0) = \{v \in V; a(v, \psi_i^*) = 0, i = 1, \dots, m\}.$$

Finally, we have

$$(6.5) \quad V = V_1 \oplus V_2.$$

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

$$Qv = v - \sum_{i=1}^m a(v, \psi_i^*) \varphi_i.$$

The equation (4.3) is then obviously equivalent to the system

$$(6.6) \quad \begin{cases} QF(u, \lambda) = 0, \\ a(u, \psi_i^*) + (G(u, \lambda), \psi_i^*) = 0, \quad i = 1, \dots, m. \end{cases}$$

Let us set

$$(6.7) \quad \begin{cases} u = u_0 + \sum_{i=1}^m \alpha_i \varphi_i + v & \text{with } v \in V_2, \\ \lambda = \lambda_0 + \xi, \end{cases}$$

and let us define the C^p map

$$(6.8) \quad \begin{cases} \mathcal{F}: V_2 \times \mathbb{R}^m \times \mathbb{R} \rightarrow V_2, \\ (v, \alpha, \xi) \rightarrow \mathcal{F}(v, \alpha, \xi) := QF(u_0 + \sum_{i=1}^m \alpha_i \varphi_i + v, \lambda_0 + \xi). \end{cases}$$

From (6.3) and (6.5) it follows that

$$(6.9) \quad D_v \mathcal{F}(0, 0, 0) = D_u F(u_0, \lambda_0)|_{V_2} \text{ is an isomorphism of } V_2.$$

Hence from the implicit function theorem we deduce that there exist $\alpha_0, \xi_0 > 0$ and a unique map of class C^p

$$v: B(\alpha_0) \times]-\xi_0, \xi_0[\rightarrow V_2,$$

where

$$B(\alpha_0) = \{\alpha = (\alpha_1, \dots, \alpha_m) \in \mathbb{R}^m; |\alpha_i| < \alpha_0, i = 1, \dots, m\},$$

such that

$$(6.10) \quad \begin{cases} \mathcal{F}(v(\alpha, \xi), \alpha, \xi) = 0, \\ v(0, 0, 0) = 0. \end{cases}$$

With this procedure, called the Lyapunov–Schmidt reduction, solving the equation (4.3) in a neighborhood of (u_0, λ_0) leads us to the *bifurcation equation*

$$(6.11) \quad f(\alpha, \xi) = 0$$

in a neighborhood of the origin in $\mathbb{R}^m \times \mathbb{R}$ where the components of the function $f: \mathbb{R}^m \times \mathbb{R} \rightarrow \mathbb{R}^m$ of class C^p are defined on $B(\alpha_0) \times]-\xi_0, \xi_0[$ by

$$(6.12) \quad f_i(\alpha, \xi) := a(u_0, \psi_i^*) + \alpha_i + a(v(\alpha, \xi), \psi_i^*) \\ + \left(G(u_0 + \sum_{j=1}^m \alpha_j \varphi_j + v(\alpha, \xi), \lambda_0 + \xi), \psi_i^* \right), \quad i = 1, \dots, m.$$

By elementary calculations we obtain

$$(6.13) \quad f(0, 0) = \frac{\partial f}{\partial \alpha_k}(0, 0) = 0, \quad k = 1, \dots, m,$$

$$(6.14) \quad \frac{\partial f_i}{\partial \xi}(0, 0) = (D_\lambda G(u_0, \lambda_0), \psi_i^*), \quad i = 1, \dots, m,$$

and so (u_0, λ_0) is a bifurcation point of multiplicity $m \geq 1$ if and only if

$$(6.15) \quad (D_\lambda G(u_0, \lambda_0), \psi_i^*) = 0, \quad \text{for all } i = 1, \dots, m.$$

6.2. The approximate equation (4.6) in $V_N \times I$ can also equivalently be solved in $V \times I$; indeed, if a pair $(u, \lambda) \in V \times I$ satisfies (4.6) it follows that $u = -T_N G(u, \lambda) \in V_N$. Therefore (4.6) is equivalent to the system

$$(6.16) \quad \begin{cases} QF_N(u_N, \lambda) = 0, \\ a(u_N + T_N G(u_N, \lambda), \psi_i^*) = 0, \quad i = 1, \dots, m. \end{cases}$$

Let us set

$$(6.17) \quad \begin{cases} u_N = u_0 + \sum_{i=1}^m \alpha_i \varphi_i + v_N & \text{with } v_N \in V_2, \\ \lambda = \lambda_0 + \xi, \end{cases}$$

and let us define the C^p map

$$\begin{cases} \mathcal{F}_N: V_2 \times \mathbb{R}^m \times \mathbb{R} \rightarrow V_2, \\ (v_N, \alpha, \xi) \rightarrow \mathcal{F}_N(v_N, \alpha, \xi) := QF_N(u_0 + \sum_{i=1}^m \alpha_i \varphi_i + v_N, \lambda_0 + \xi). \end{cases}$$

Following Brezzi–Rappaz–Raviart [4; Part II] we can prove

THEOREM 6.1. *Assume in addition that the mapping $D^p G$ is bounded on each bounded subset of $V \times R$. Then there exist four positive constants $\alpha_0, \xi_0, \beta, N_0$ (all independent of N) and for each $N \geq N_0$, a unique C^p mapping*

$$v_N: B(\alpha_0) \times]-\xi_0, \xi_0[\rightarrow V_2$$

such that for all $(\alpha, \xi) \in B(\alpha_0) \times]-\xi_0, \xi_0[$

$$(6.18) \quad \mathcal{F}_N(v_N(\alpha, \xi), \alpha, \xi) = 0,$$

$$\|v_N(\alpha, \xi) - v(\alpha, \xi)\| \leq \beta.$$

Moreover, there exists a constant $k > 0$ independent of N such that for all $\alpha, \alpha^* \in B(\alpha_0)$ and all $\xi, \xi^* \in]-\xi_0, \xi_0[$ the following error estimate holds:

$$(6.19) \quad \|v_N(\alpha^*, \xi^*) - v(\alpha, \xi)\| \leq k \{ \|\alpha^* - \alpha\|_{R^m} + |\xi^* - \xi| + \|(T - T_N)J(\alpha, \xi)\| \}$$

where J is a C^p mapping of $B(\alpha_0) \times]-\xi_0, \xi_0[$ into W defined by

$$(6.20) \quad J(\alpha, \xi) := G(u_0 + \sum_{j=1}^m \alpha_j \varphi_j + v(\alpha, \xi), \lambda_0 + \xi).$$

Analogous error estimates hold also for the derivatives $D^r v_N$ for $1 \leq r \leq p-1$.

Thanks to this result solving equation (4.6) in a neighborhood of (u_0, λ_0) is reduced to solving the approximate bifurcation equation

$$(6.21) \quad f_N(\alpha, \xi) = 0$$

in a neighborhood of the origin in $R^m \times R$ where the components of the function $f_N: R^m \times R \rightarrow R^m$ of class C^p are defined on $B(\alpha_0) \times]-\xi_0, \xi_0[$ by

$$(6.22) \quad f_{iN}(\alpha, \xi) := a(u_0, \psi_i^*) + \alpha_i + a(v_N(\alpha, \xi), \psi_i^*) \\ + a(T_N G(u_0 + \sum_{j=1}^m \alpha_j \varphi_j + v_N(\alpha, \xi), \lambda_0 + \xi), \psi_i^*), \quad i = 1, \dots, m.$$

Now in order to analyze the approximation of the solutions of (4.3) by those of (4.6) it remains to compare the solutions of the bifurcation equation (6.11) with those of the approximate bifurcation equation (6.21). Such an analysis is indeed the most interesting part since it strongly depends on the local structure of the set of solutions of (6.11) near $(0, 0)$. It has been developed by Brezzi–Rappaz–Raviart [4] in the case $m = 1$, Rappaz–Raugel [17] for $m \geq 1$, and Raugel [21] gave a very careful description of the situation for $m = 2$ in the presence of symmetries. The presentation of these interesting results would, however, take us too far. We would only remark that they are of the same kind as those obtained when studying the small

regular perturbations of the type $F(u, \lambda; \varepsilon) = u + TG(u, \lambda; \varepsilon)$. In the recent years such perturbations have been carefully studied with the help of deep instruments of algebra and geometry; see e.g. Golubitsky–Schaeffer [9]. Also if the numerical approximation is not a regular perturbation, in most cases one can prove that indeed it behaves as a regular perturbation: see e.g. Geymonat–Raugel [8], Raugel [20].

References

- [1] R. A. Adams, *Sobolev Spaces*, Academic Press, N. Y., 1975.
- [2] I. Babuška, R. B. Kellog and J. Pitkäranta, *Direct and inverse error estimates for finite elements with mesh refinements*, Numer. Math. 33 (1979), 447–471.
- [3] C. de Boor, *A Practical Guide to Splines*, Springer Verlag, 1978.
- [4] F. Brezzi, J. Rappaz and P. A. Raviart, *Finite dimensional approximation of non-linear problems*. Part I: *Branches of nonsingular solutions*, Numer. Math. 36 (1980), 1–25; Part II: *Limit points*, ibid. 37 (1981), 1–28; Part III: *Simple bifurcation points*, ibid. 38 (1981), 1–30.
- [5] P. G. Ciarlet, *The Finite Element Method for Elliptic Problems*, North-Holland, 1978.
- [6] J. Descloux, J. Rappaz and R. Scholz, *On the rate of convergence for the approximation of nonlinear problems*, Math. Comp. 45 (1985), 51–64.
- [7] G. Geymonat, *Alcuni aspetti del metodo degli elementi finiti*, Bollettino U.M.I. 13-A (1976), 242–272.
- [8] G. Geymonat and G. Raugel, *Finite dimensional approximation of some bifurcation problems in presence of symmetries*, in: *Numerical Methods for Bifurcation Problems* (T. Küpper, H. D. Mittelman, H. Weber, Editors), Birkhäuser Verlag, 1984, 369–384.
- [9] M. Golubitsky and D. G. Schaeffer, *Singularities and Groups in Bifurcation Theory*, Volume I, Springer-Verlag, 1985.
- [10] P. Grisvard, *Elliptic Problems in Non-Smooth Domains*, Pitman, 1985.
- [11] B. S. Kashin and A. A. Saakyan, *Orthogonal Series*, Nauka, Moscow 1984 (in Russian).
- [12] S. M. Nikol'skii, *Approximation of Functions of Several Variables and Imbedding Theorems*, Springer-Verlag, 1975.
- [13] L. A. Oganessian and L. A. Rykhovets, *Variational-difference methods of solving elliptic equations*, Izdat. Akad. Nauk Armyan. SSR, Erevan' 1979 (in Russian).
- [14] P. Oswald, *On the degree of nonlinear spline approximation in Besov–Sobolev spaces*, J. Approx. Theory, to appear.
- [15] —, *On function spaces related to finite element approximation theory*, Analysis Anwendungen (to appear).
- [16] A. Pinkus, *n-Widths in Approximation Theory*, Springer-Verlag, 1985.
- [17] J. Rappaz and G. Raugel, *Finite-dimensional approximation of bifurcation problems at a multiple eigenvalue*, Rapport Interne n° 71, 1981, Centre de Mathématiques Appliquées, École Polytechnique, Palaiseau, France.
- [18] G. Raugel, *Résolution numérique de problèmes elliptiques dans des domaines avec coins*, Thèse de Troisième Cycle, 1978, Université de Rennes, Rennes, France.
- [19] —, *Résolution numérique par une méthode d'éléments finis du problème de Dirichlet pour le laplacien dans un polygone*, C. R. Acad. Sci. Paris 286 (1978), Série A, 791–794.
- [20] —, *Approximation numérique de problèmes non linéaires*, Thèse d'Etat, 1984, Université de Rennes–Beaulieu, Rennes, France.

- [21] G. Raugel, *Finite dimensional approximation of bifurcation problems in presence of symmetries*, Numer. Math. 48 (1986), 137–198.
- [22] A. H. Schatz and L. B. Wahlbin, *Maximum norm estimates in the finite element method on plane polygonal domains*, Part I, Math. Comp. 32 (1978), 73–109; Part 2, *Refinements*, ibid., 33 (1979), 465–492.
- [23] H. Triebel, *Interpolation Theory, Function Spaces, Differential Operators*, North-Holland, 1978.
- [24] —, *Theory of Function Spaces*, Geest & Portig, Leipzig 1983.
- [25] A. G. Werschulz, *Optimal error properties of finite element methods for second order elliptic Dirichlet problems*, Math. Comp. 38 (1982), 401–413.

*Presented to the Semester
Approximation and Function Spaces
February 27–May 27, 1986*