Numer. Math. 69: 213–231 (1994)

Numerische Mathematik

Consistency, stability, a priori and a posteriori errors for Petrov-Galerkin methods applied to nonlinear problems

J. Pousin*, J. Rappaz

Department of Mathematics, EPF-Lausanne, Switzerland

Received June 10, 1992 / Revised version received February 28, 1994

Summary. In an abstract framework we present a formalism which specifies the notions of consistency and stability of Petrov-Galerkin methods used to approximate nonlinear problems which are, in many practical situations, strongly nonlinear elliptic problems. This formalism gives rise to a priori and a posteriori error estimates which can be used for the refinement of the mesh in adaptive finite element methods applied to elliptic nonlinear problems. This theory is illustrated with the example: $-\operatorname{div}(k(u)\nabla u) + \mathbf{c} \cdot \nabla u = f$ in a two dimensional domain Ω with Dirichlet boundary conditions.

Mathematics Subject Classification (1991): 35K55, 65N15, 65N30

1. Introduction

Let X and Y be two reflexive Banach spaces and let us consider a C^1-mapping $F : X \to Y'$ where Y' is the dual space of Y, the duality pairing is denoted by $\langle\cdot|\cdot\rangle$. We are interested in approximations of an element $u \in X$ satisfying

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

or equivalently

$$\langle F(u)|v\rangle = 0 \ , \quad \forall\, v \in Y \ . \tag{1.2}$$

In order to build approximations u_h of u, we use a Petrov-Galerkin method on (1.2), that is to say we choose finite dimensional subspaces $X_h \subset X$ and $Y_h \subset Y$ with $\dim X_h = \dim Y_h$, and we find $u_h \in X_h$ satisfying

$$\langle F(u_h)|v_h\rangle = 0 \ , \quad \forall\, v_h \in Y_h \ . \tag{1.3}$$

* *Present address:* Mathematical Modelling and Scientific Computing Laboratory URA CNRS 740, National Institut of Applied Sciences in Lyon, 20, Av. A. Einstein, F-69621 Villeurbanne

In this paper we assume we have a solution u of (1.1) such that the Fréchet derivative $DF(u)$ at the point u is an isomorphism from X onto Y'. Then, under consistency and stability conditions which are essentially linked to approximation properties of X by X_h, of Y by Y_h and to discrete "inf-sup" conditions on the bilinear form $b(\varphi,\psi) \equiv \langle DF(u)\varphi|\psi\rangle$, we prove that Problem (1.3) has a unique solution u_h in a neighborhood of u and we establish a priori and a posteriori error estimates for $(u-u_h)$ in the norm of X. Let us point out that the main term in the a posteriori estimate is the residual value $\|F(u_h)\|_{Y'}$ in the Y'-norm. In concrete situations, the knowledge of this value is important since it allows to minimize the error by means of adaptive techniques (see for instance Johnson [11], Johnson and Hansboo [12], Baranger-Elamri [3], Verfürth [15], Picasso [13]).

The main results presented in this paper have been announced in Pousin-Rappaz [14]. For proving them, we begin to build a C^1-mapping $F_h : X \to Y'$ such that if $w \in X$ is a solution of

$$F_h(w) = 0 \,, \tag{1.4}$$

then w belongs to X_h and $u_h = w$ is a solution of (1.3) and conversely. Next we compare the mappings F and F_h in a similar way as in Crouzeix-Rappaz [5] for obtaining the existence and the error estimates for the approximate problem (1.3). Nevertheless, in contrast to the method previously employed in Crouzeix-Rappaz [5], the new approach presented here, does not require to invert the principal part of the operator when treating approximations of nonlinear elliptic problems. Consequently our theory permits to consider strongly nonlinear problems.

An outline of the paper is as follows. In Sect. 2 we define the functional framework in which the problem is set and we give some abstract theorems of convergence and error estimates when we compare Problem (1.4) to Problem (1.1) under reasonable assumptions. Section 3 is devoted to the particular case where the nonlinear problem (1.1) is approximated by Petrov-Galerkin methods (1.3). We give a priori and a posteriori error estimates. In Sect. 4 we consider a nonlinear stationary heat problem with a convection term. Under appropriate hypotheses we show how the abstract results of the previous sections can be applied to the finite element approximation of this problem, and we give a posteriori error estimates by means of local indicators which are used in an adaptive code solving this problem (see Picasso [13]).

2. Abstract results

Let X and Z be two Banach spaces the norm of which are respectively denoted by $\|\cdot\|_X$ and by $\|\cdot\|_Z$. If $\mathscr{L}(X,Z)$ is the Banach space of all continuous linear operators from X into Z, we denote by $\|T\|_{XZ} = \sup_{u\in X, \|u\|_X=1} \|Tu\|_Z$ the norm of $T \in \mathscr{L}(X,Z)$. If $G : X \to Z$ is a C^1-mapping from X into Z and if $u \in X$, we denote by $DG(u)$ the Fréchet derivative of G at the point u.

We begin to establish a result which is similar to one of them we find in Girault-Raviart [9].

Theorem 1. *Let $G : X \to Z$ be a C^1-mapping from X into Z and let w be an element of X. We assume*

(2.1) *(i) $DG(w)$ is an isomorphism from X onto Z,* (2.2) *(ii)* $\|DG(w)^{-1}\|_{ZX}\|G(w)\|_Z \leq \delta/2$

where $\delta > 0$ *is such that*

(2.3) $$\sup_{x\in X, \|w-x\|_X \leq \delta} \|DG(w) - DG(x)\|_{XZ} \leq (2\|DG(w)^{-1}\|_{ZX})^{-1} \ .$$

Then there exists a unique $v \in X$ *satisfying*

(2.4) $G(v) = 0$ *and* $\|w - v\|_X \leq \delta$.

Moreover we have the estimate:

(2.5) $$\|w - v\|_X \leq 2\|DG(w)^{-1}\|_{ZX}\|G(w)\|_Z \ .$$

Proof. Let $\delta > 0$ be given by (2.3). If we define the mapping

$$L(x) = x - DG(w)^{-1}G(x) \ , \quad x \in X \ ,$$

we have for $x, y \in X$ such that $\|w - x\|_X \leq \delta$, $\|w - y\|_X \leq \delta$:

$$\begin{aligned}\|L(x) - L(y)\|_X &= \|DG(w)^{-1}\int_0^1 (DG(w) - DG(sx + (1-s)y))(x-y)ds\|_X \\ &\leq \|DG(w)^{-1}\|_{ZX}(2\|DG(w)^{-1}\|_{ZX})^{-1}\|x-y\|_X \\ &\leq \frac{1}{2}\|x-y\|_X \ .\end{aligned}$$

By using the above inequality with $y = w$ together with (2.2) we have:

$$\begin{aligned}\|w - L(x)\|_X &\leq \|w - L(w)\|_X + \|L(w) - L(x)\|_X \\ &\leq \|DG(w)^{-1}\|_{ZX}\|G(w)\|_Z + \frac{1}{2}\|w-x\|_X \leq \delta \ . \qquad (2.6)\end{aligned}$$

We conclude that L is a contracting mapping from the ball centered at w and with radius δ into itself. Consequently there exists a unique fixed point v of L in this ball, i.e. $v = L(v)$, and (2.4) is proved. Estimate (2.5) is a direct consequence of (2.6) with $x = v$. □

A consequence of Theorem 1 is

Corollary 1. *Let* $G : X \to Z$ *be a* C^2*-mapping such that its second Fréchet derivative is bounded on all bounded subsets of* X*. We assume there exist a sequence* $\{u_n\}_{n=1}^{\infty} \subset X$ *and constants* C, M*, independent of* n*, satisfying:*

(i) $\lim_{n\to\infty} \|G(u_n)\|_Z = 0$

(ii) $\|u_n\|_X \leq C$, $n = 1, 2, 3, \ldots$,

(iii) $DG(u_n)$ *is an isomorphism from* X *onto* Z *and* $\|DG(u_n)^{-1}\|_{ZX} \leq M$.

Then there exists $u \in X$ *such that* $G(u) = 0$.

Proof. Since the second derivative of G is bounded on all bounded subsets of X and by using (ii), we can find $\delta > 0$ such that $\sup_{x\in X, \|u_n - x\|_X \leq \delta} \|DG(u_n) - DG(x)\|_{ZX} \leq$

$1/(2M)$. By using (i) and (iii) there exists N such that $\|DG(u_n)^{-1}\|_{ZX}\|G(u_n)\|_Z \leq \delta/2$ for all $n > N$. Theorem 1 allows to conclude. □

Let us remark that Corollary 1 applied to concrete nonlinear elliptic problems gives, in several situations, a numerical proof of existence in the same way as Bamberger [2].

Now we consider an element $u \in X$ and a family $\{F_h\}_{0<h\leq 1}$ of C^1-mappings from X into Z parametrized by h. In concrete examples, u will be a zero of a mapping $F : X \to Z$ and F_h will be an approximation of F.

We assume that the family $\{F_h\}_{0<h\leq 1}$ and the point $u \in X$ are satisfying:

(H1) there exists a constant $\varepsilon_0 > 0$ and for all $h \in (0,1]$ there exists a positive number L_h such that

$$\|DF_h(u) - DF_h(v)\|_{XZ} \leq L_h\|u - v\|_X \text{ for all } v \in X \text{ with } \|u - v\|_X \leq \varepsilon_0\ ;$$

(H2) $\lim_{h\to 0}(1 + L_h)\|F_h(u)\|_Z = 0$;

(H3) $DF_h(u)$ is an isomorphism from X onto Z for all $h \in (0,1]$ and there exists a constant M, independent of h, such that

$$\|DF_h(u)^{-1}\|_{ZX} \leq M\ , \quad \forall\, h \in (0,1]\ .$$

Let us note that if u is a zero of a mapping F which is approximated by F_h, Hypothesis (H3) is a stability assumption and Hypothesis (H2) is a consistency assumption. In fact if L_h is bounded with respect to h (which is realized in a lot of applications) then it is sufficient to assume $\lim_{h\to 0}\|F_h(u)\|_Z = 0$ which is the standard consistency hypothesis if F_h is an approximation of a mapping F.

Now we establish the two main results of this section.

Theorem 2. *Assume that the family of mappings $\{F_h\}_{0<h\leq 1}$ satisfies Hypotheses (H1), (H2) and (H3). Then there exist $h_0 > 0$ and $\delta_0 > 0$ such that for all $h \in (0, h_0]$ there is a unique $u_h \in X$ satisfying*

$$F_h(u_h) = 0 \quad \textit{and} \quad \|u - u_h\|_X \leq \delta_0/(1 + L_h)\ . \tag{2.7}$$

Moreover, for $h \leq h_0$ we have the estimate:

$$\|u - u_h\|_X \leq 2\|DF_h(u)^{-1}\|_{ZX}\|F_h(u)\|_Z\ . \tag{2.8}$$

Proof. We set $\delta_0 = \min(\varepsilon_0, \frac{1}{2M})$ and we define $\delta_h = \delta_0/(1 + L_h)$. By using (H1) we verify that, for all $v \in X$ such that $\|u - v\|_X \leq \delta_h$, we have:

$$\|DF_h(u) - DF_h(v)\|_{XZ} \leq \frac{1}{2M}\ . \tag{2.9}$$

It follows, by using (H3), that for all $h \in (0,1]$ and $\|u - v\|_X \leq \delta_h$:

$$\|DF_h(u) - DF_h(v)\|_{XZ} \leq \frac{1}{2\|DF_h(u)^{-1}\|_{ZX}}\ . \tag{2.10}$$

Now using Hypothesis (H2), there exists $h_0 > 0$ such that for $h \leq h_0$

$$(1 + L_h)\|F_h(u)\|_Z \leq \delta_0/(2M)\ . \tag{2.11}$$

Since Hypothesis (H3) implies

$$\|DF_h(u)^{-1}\|_{ZX}\|F_h(u)\|_Z \le M\|F_h(u)\|_Z \ ,$$

it is easy to verify by using (2.11) that, for $h \le h_0$:

$$\|DF_h(u)^{-1}\|_{ZX}\|F_h(u)\|_Z \le \delta_h/2 \ . \tag{2.12}$$

Theorem 2 is a direct consequence of (2.10), (2.12) and of Theorem 1 with $G = F_h$ and $w = u$. □

Remark 1. If L_h is bounded, we can see that, in a fixed neighborhood of u, there exists a unique zero u_h of F_h if h is small enough.

Let us observe that, under Hypotheses of consistency and stability (H2) and (H3), we have

$$\|u - u_h\|_X \le 2M\|F_h(u)\|_Z \ . \tag{2.13}$$

Consequently u_h converges to u when h tends to zero with the same order as the one of the consistency. Inequality (2.13) is an a priori error estimate since it depends on u. □

Now, if u is a zero of a C^1-mapping $F : X \to Z$, we have the following a posteriori error estimate which does not depend on u.

Theorem 3. *Assume that the family of mappings $\{F_h\}_{0<h\le 1}$ satisfies (H1), (H2), (H3). Moreover we assume that we have a C^1-mapping $F : X \to Z$ such that:*

(H4) $F(u) = 0$ and $DF(u)$ is an isomorphism from X onto Z.

If h_0, δ_0 and u_h are given by Theorem 2, there exists $h_1 \le h_0$ such that for all $h \le h_1$ we have:

$$\|u - u_h\|_X \le 2\|DF(u_h)^{-1}\|_{ZX}\|F(u_h)\|_Z \ . \tag{2.14}$$

Proof. Since F is a C^1-mapping, there exists $\delta > 0$ such that

$$\sup_{v\in X, \|u-v\|_X \le 2\delta} \|DF(u) - DF(v)\|_{XZ} \le \frac{1}{8\|DF(u)^{-1}\|_{ZX}} \ . \tag{2.15}$$

A consequence of Theorem 2 is that u_h converges to u when h tends to zero and so there exists $\tilde{h}_1 \le h_0$ such that

$$\|u - u_h\|_X \le \delta \ , \quad \forall\ h \le \tilde{h}_1 \ . \tag{2.16}$$

Now, if $h \le \tilde{h}_1$ and $x \in X$ satisfies $\|u_h - x\|_X \le \delta$, then $\|u - x\|_X \le 2\delta$ and by using (2.15) we obtain:

$$\sup_{x\in X, \|u_h-x\|_X \le \delta} \|DF(u_h) - DF(x)\| \le \frac{1}{4\|DF(u)^{-1}\|_{ZX}} \ . \tag{2.17}$$

Besides, there exists $h_1 \le \tilde{h}_1$ such that for all $h \le h_1$ we have:

(2.18) $$\|DF(u_h)^{-1}\|_{ZX} \leq 2\|DF(u)^{-1}\|_{ZX}$$

and

(2.19) $$\|DF(u_h)^{-1}\|_{ZX}\|F(u_h)\|_Z \leq \delta/2 \ .$$

By using Relations (2.17), (2.18) and (2.19) we can conclude from Theorem 1 with $G = F$ and $w = u_h$: there is a unique $v \in X$ satisfying

$$F(v) = 0 \quad \text{and} \quad \|v - u_h\|_X \leq \delta \ .$$

Moreover the following estimates hold:

(2.20) $$\|u_h - v\|_X \leq 2\|DF(u_h)^{-1}\|_{ZX}\|F(u_h)\|_Z \ .$$

Inequality (2.14) is a direct consequence of Inequality (2.20) if we prove that $v = u$. In fact we have $\|u - v\|_X \leq 2\delta$ and $F(u) = F(v) = 0$. From the identity

$$(u - v) = DF(u)^{-1} \int_0^1 (DF(u) - DF(su + (1-s)v))(u - v)ds$$

and from Inequality (2.15) we deduce $\|u - v\|_X = 0$. $\square$

3. Petrov-Galerkin methods

Let X and Y be two reflexive real Banach spaces equipped respectively with the norms $\|\cdot\|_X$ and $\|\cdot\|_Y$ and we denote by $\langle\cdot|\cdot\rangle$ the duality pairing between Y' and Y.

In all the following, $F : X \to Y'$ is a C^1-mapping defined on X (or eventually on an open subset of X) with values in Y'. We call the exact problem, the following one: find $u \in X$ satisfying

(3.1) $$F(u) = 0 \ ,$$

or equivalently

(3.2) $$\langle F(u)|v\rangle = 0 \ , \quad \forall\, v \in Y \ .$$

If $\{X_h\}_{0<h\leq 1}$ and $\{Y_h\}_{0<h\leq 1}$ are respectively two families of finite dimensional subspaces of X and Y, we shall say that the problem of finding $u_h \in X_h$ such that

(3.3) $$\langle F(u_h)|v_h\rangle = 0 \ , \quad \forall\, v_h \in Y_h$$

is a Petrov-Galerkin approximation of Problem (3.2).

In this section, we build, under suitable assumptions, a family $\{F_h\}_{0<h\leq 1}$ of C^1-mapping from X into Y' which allows to obtain the solutions of (3.3) as the zeros of F_h. Next using the results established in Sect. 2, we get error estimates for the Petrov-Galerkin approximation of the exact problem.

We start by assuming there is a solution $u \in X$ of Problem (3.1), i.e. $F(u) = 0$, and we define the continuous bilinear form $b : X \times Y \to \mathbb{R}$ by the following relation:

(3.4) $$b(x, y) = \langle DF(u)x|y\rangle \ , \quad \forall\, x \in X \ , \quad \forall\, y \in Y \ .$$

We denote by $\|b\|$ the norm of b, i.e. $\|b\| = \sup_{\substack{y\in Y,\|y\|_Y=1 \\ x\in X,\|x\|_X=1}} b(x,y) = \|DF(u)\|_{XY'}$ and we assume the "inf-sup" conditions on b hold (see [1] for instance), i.e.

$$\text{(3.5)} \qquad \inf_{x\in X,\|x\|_X=1} \sup_{y\in Y,\|y\|_Y=1} b(x,y) = \beta > 0 \ ,$$

(H5)

$$\text{(3.6)} \qquad \sup_{x\in X,\|x\|_X=1} b(x,y) > 0 \ , \quad \forall\, y \in Y \ , \ y \neq 0 \ .$$

Hypothesis (H5) is equivalent to assume that $DF(u)$ is an isomorphism from X onto Y' since Y is a reflexive space (Hypothesis (H4) with $Z = Y'$) ; the norm of its inverse is given by

$$\text{(3.7)} \qquad \|DF(u)^{-1}\|_{Y'X} = \beta^{-1} \ .$$

Concerning the spaces X_h, Y_h and the form b, we assume the "discrete inf-sup" conditions hold:

$$\text{(3.8)} \qquad \inf_{x\in X_h,\|x\|_X=1} \sup_{y\in Y_h,\|y\|_Y=1} b(x,y) = \beta_h > 0 \ , \quad \text{and}$$

(H6)

$$\text{(3.9)} \qquad \dim X_h = \dim Y_h \ ,$$

where β_h, for $h \in (0,1]$, is a positive constant which could possibly tend to zero when h tends to zero. For this reason we assume in addition the following approximation property:

$$\text{(3.10)} \qquad \lim_{h\to 0} \beta_h^{-2} \min_{x_h\in X_h} \|u - x_h\|_X = 0. \qquad \text{(H7)}$$

Let us remark that $\beta_h \le \|b\|$ and if β_h is bounded from below, then (3.10) is a consequence of the standard hypothesis for the approximation of X by X_h, i.e. $\lim_{h\to 0} \min_{x_h\in X_h} \|x - x_h\|_X = 0$, $\forall\, x \in X$.

If β_h is not bounded from below we shall see, in concrete applications related to elliptic problems, that u must be sufficiently regular to obtain (3.10).

Now we give the main result of this section.

Theorem 4. *We suppose Hypotheses (H5), (H6) and (H7) are fulfilled. Moreover we assume that DF is Lipschitzian at u, i.e.*

$$\exists \varepsilon_0 > 0 \text{ and } L \text{ such that for all } v \in X,\ \|u - v\|_X \le \varepsilon_0\text{:}$$

(H8)

$$\text{(3.11)} \qquad \|DF(u) - DF(v)\|_{XY'} \le L\|u - v\|_X \ .$$

Then there exist $h_0 > 0$ and $\eta_0 > 0$ such that for all $h \in (0, h_0]$ there is a unique $u_h \in X_h$ satisfying

$$\text{(3.12)} \qquad \langle F(u_h)|v_h\rangle = 0 \ , \quad \forall\, v_h \in Y_h \quad \textit{and} \quad \|u - u_h\|_X \le \eta_0\beta_h \ .$$

Moreover we have the error estimates:

$$\text{(3.13)} \qquad \|u - u_h\|_X \le \frac{2\|b\|}{\beta}\left(1 + \frac{\|b\|}{\beta_h}\right) \min_{\chi_h\in X_h} \|u - \chi_h\|_X \ ,$$

$$\text{(3.14)} \qquad \|u - u_h\|_X \le \frac{4}{\beta}\|F(u_h)\|_{Y'} \ , \quad \forall\, h \le h_0.$$

Before to prove Theorem 4, we build a mapping $F_h : X \to Y'$, the zeros of which will be the solutions of (3.3).

It is well known (see [1] for instance) that Hypotheses (H6), that is to say (3.8), (3.9), imply the existence of two projectors

$$\Pi_h^1 : X \to X_h \quad \text{and} \quad \Pi_h^2 : Y \to Y_h \ ,$$

defined by

$$b(x - \Pi_h^1 x, y_h) = 0 \ , \quad \forall\, y_h \in Y_h \ , \quad \forall\, x \in X \ , \tag{3.15}$$

$$b(x_h, y - \Pi_h^2 y) = 0 \ , \quad \forall\, x_h \in X_h \ , \quad \forall\, y \in Y \ . \tag{3.16}$$

It is easy to verify that

$$\|\Pi_h^1\|_{XX} \le \frac{\|b\|}{\beta_h} \ . \tag{3.17}$$

In order to give a bound for $\|\Pi_h^2\|_{YY}$, we use Relations (3.4), (3.7), (3.15), (3.16); we write for $y \in Y$:

$$\begin{aligned} \|\Pi_h^2 y\|_Y &= \sup_{\varphi \in Y', \|\varphi\|_{Y'}=1} < \varphi | \Pi_h^2 y \rangle = \sup_{\varphi \in Y', \|\varphi\|_{Y'}=1} b(DF(u)^{-1}\varphi, \Pi_h^2 y) \\ &= \sup_{\varphi \in Y', \|\varphi\|_{Y'}=1} b(\Pi_h^1 DF(u)^{-1}\varphi, y) \le \|b\| \cdot \|y\|_Y \cdot \|\Pi_h^1\|_{XX} \cdot \beta^{-1} \ . \end{aligned}$$

Finally, we use (3.17) in order to obtain

$$\|\Pi_h^2\|_{YY} \le \frac{\|b\|^2}{\beta \cdot \beta_h} \ . \tag{3.18}$$

Now, for all $h \in (0, 1]$, we build $F_h : X \to Y'$ in the following way:

$$\langle F_h(x) | y \rangle \stackrel{\text{def}}{=} \langle F(x) | \Pi_h^2 y \rangle + b(x, y - \Pi_h^2 y) \ , \quad \forall\, x \in X \ , \quad \forall\, y \in Y \ . \tag{3.19}$$

Let us notice that F_h depends on the solution u of (3.1) through the bilinear form b. We have

Lemma 1. *If $u_h \in X_h$ is a solution of Problem (3.3), then u_h is such that $F_h(u_h) = 0$. Conversely, if $\mu \in X$ is such that $F_h(\mu) = 0$, then $\mu \in X_h$ and $u_h = \mu$ is a solution of Problem (3.3).*

Proof. It is easy to show that all the solutions of Problem (3.3) are zeros of the mapping F_h. Conversely, let $\mu \in X$ be a zero of the mapping F_h, i.e.

$$\langle F(\mu) | \Pi_h^2 y \rangle + b(\mu, y - \Pi_h^2 y) = 0 \ , \quad \forall\, y \in Y \ . \tag{3.20}$$

If w belongs to Y and if we choose $y = w - \Pi_h^2 w$ in (3.20), we obtain

$$b(\mu, w - \Pi_h^2 w) = 0 \ , \quad \forall\, w \in Y \ . \tag{3.21}$$

By using definitions (3.15), (3.16) of projectors Π_h^1 and Π_h^2, we can see that (3.21) implies

(3.22) $$b(\mu - \Pi_h^1\mu, w) = 0 \ , \forall \ w \in Y \ .$$

Hypothesis (H5) together with (3.22) imply $\mu = \Pi_h^1\mu \in X_h$.

By setting $y = v_h \in Y_h$ in (3.20), we obtain

$$\langle F(\mu)|v_h\rangle = 0 \ , \quad \forall \ v_h \in Y_h \ ,$$

and consequently $u_h = \mu$ is a solution of Problem (3.3). □

Now we are in position to derive the proof of our main result as a consequence of Theorems 2 and 3 thanks to Lemma 1.

Proof of Theorem 4. According to Lemma 1 we can deal with the family $\{F_h\}_{0<h\leq 1}$ defined by (3.19) instead of considering Problem (3.3). Let us check that the Hypotheses (H1), (H2), (H3) and (H4) are fulfilled by the family $\{F_h\}$ in order to apply Theorems 2 and 3. From the definition (3.19) of F_h we easily deduce that for: $v, \varphi \in X$, $y \in Y$, we have:

$$\langle (DF_h(u) - DF_h(v))\varphi|y\rangle = \langle (DF(u) - DF(v))\varphi|\Pi_h^2 y\rangle \ .$$

Thanks to (3.18), (3.11) and to inequalities $\beta \leq \|b\|$, $\beta_h \leq \|b\|$, it follows that for $v \in X$, $\|u - v\|_X \leq \varepsilon_0$:

$$\|DF_h(u) - DF_h(v)\|_{XY'} \leq \frac{\|b\|^2}{\beta \cdot \beta_h}\|DF(u) - DF(v)\|_{XY'}$$

(3.23) $$\leq L\frac{\|b\|^2}{\beta \cdot \beta_h}\|u - v\|_X \leq \left[(L+1)\frac{\|b\|^2}{\beta \cdot \beta_h} - 1\right]\|u - v\|_X \ .$$

So, we have proven Hypothesis (H1) of Sect. 2 with $Z = Y'$ and

(3.24) $$L_h = (L+1)\frac{\|b\|^2}{\beta \cdot \beta_h} - 1 \ .$$

In order to verify Hypothesis (H2) of Sect. 2, we use (3.19), (3.2), (3.15), (3.16) to obtain:

$$\begin{aligned}\|F_h(u)\|_{Y'} &= \sup_{v\in Y, \|v\|_Y = 1} \langle F_h(u)|v\rangle \\ &= \sup_{v\in Y, \|v\|_Y = 1} b\left(u, v - \Pi_h^2 v\right) \\ &= \sup_{v\in Y, \|v\|_Y = 1} b\left(u - \Pi_h^1 u, v\right) \\ &\leq \|b\|\|u - \Pi_h^1 u\|_X \ .\end{aligned}$$

Considering the inequality

$$\|u - \Pi_h^1 u\|_X \leq \|u - \chi_h\|_X + \|\Pi_h^1(\chi_h - u)\|_X \ , \quad \forall \ \chi_h \in X_h$$

and Relation (3.17), we conclude that:

(3.25) $$\|F_h(u)\|_{Y'} \leq \|b\|\left(1 + \frac{\|b\|}{\beta_h}\right)\min_{\chi_h\in X_h}\|u - \chi_h\|_X \ .$$

Inequality (3.25) together with (3.24) and the relation $\beta_h \leq \|b\|$ imply:

$$(3.26) \qquad (1 + L_h)\|F_h(u)\|_{Y'} \leq \frac{2(1+L)\|b\|^4}{\beta \cdot \beta_h^2} \min_{\chi_h \in X_h} \|u - \chi_h\|_X \;,$$

thus Hypothesis (H2) with $Z = Y'$ is a consequence of (H7).

Hypothesis (H3) of Sect. 2 is a direct consequence of (H5), or equivalently of (3.7), because we can easily verify that our choice of F_h and of form b implies $DF_h(u) = DF(u)$.

Theorem 2 with $Z = Y'$ can be applied in this particular situation. We obtain the existence of $\delta_0 > 0$ and $h_0 > 0$ such that for all $h \leq h_0$, there is a unique $u_h \in X_h$ satisfying $F_h(u_h) = 0$ and $\|u - u_h\|_X \leq \frac{\delta_0}{1+L_h}$. By setting $\eta_0 = \frac{\delta_0 \beta}{(L+1)\|b\|^2}$ we obtain $\|u - u_h\|_X \leq \eta_0 \beta_h$. By Lemma 1, the element u_h belongs to X_h and is solution to Problem (3.3). A priori error estimate (3.13) is a direct consequence of (2.8), (3.7), (3.25) together with $DF_h(u) = DF(u)$. A posteriori error estimate (3.14) is derived from Inequality (2.14) and (3.7) if we prove that

$$(3.27) \qquad \|DF(u_h)^{-1}\|_{Y'X} \leq 2\|DF(u)^{-1}\|_{Y'X}$$

for $h \leq h_0$ even if it means to take h_0 smaller. In fact Relations (3.13), (3.10) and $\beta_h \leq \|b\|$, imply that u_h converges to u when h tends to zero thus (3.27) is true. □

Let us terminate this section with two remarks.

Remark 2. According to Remark 1 of Sect. 2, we can see that, if β_h is bounded from below, then, in a fixed neighborhood of u in X, there exists a unique solution u_h of the Petrov-Galerkin approximation of the exact Problem (3.1).

Remark 3. If $X = Y$ and if the bilinear form $b(\cdot,\cdot)$ defined in (3.4) is coercive, then the constant β in equality (3.14) can be evaluated.

4. An example and practical considerations

In this section, we investigate a stationary heat problem with convection. We show how the formalism previously developed applies to this problem and to its numerical approximation.

Let $\Omega = (0,1) \times (0,1) \subset \mathbb{R}^2$ be the unit square, the boundary of which is denoted by $\partial\Omega$. The notations $W^{m,p}(\Omega)$ will denote the usual standard Sobolev's spaces of functions equipped with the norms $\|\cdot\|_{m,p,\Omega}$ and the semi-norms $|\cdot|_{m,p,\Omega}$. For $p = 2$, we will write $H^m(\Omega)$ for $W^{m,2}(\Omega)$ and we miss out the index p in the norms and semi-norms, that is to say $\|\cdot\|_{m,\Omega} = \|\cdot\|_{m,2,\Omega}$ and $|\cdot|_{m,\Omega} = |\cdot|_{m,2,\Omega}$. Finally, $W_0^{1,p}(\Omega)$ denotes the subspace $\{\varphi \in W^{1,p}(\Omega),\ \varphi = 0 \text{ on } \partial\Omega\}$, and if q is such that $\frac{1}{p} + \frac{1}{q} = 1$, then $W^{-1,q}(\Omega)$ is the dual space of $W_0^{1,p}(\Omega)$.

The problem we want to discuss, is to find $u \in H_0^1(\Omega)$ such that:

$$(4.1) \qquad -\text{div}\,(k(u)\nabla u) + \mathbf{c} \cdot \nabla u = f \;, \quad \text{in } \Omega \;,$$

where the functions $f \in L^\infty(\Omega)$, $\mathbf{c} \in [C^0(\overline{\Omega})]^2$ are given, and $k \in C^2(\mathbb{R})$ is a positive mapping satisfying:

$$k(s) \geq \alpha > 0\ , \quad \forall\, s \in \mathbb{R}\ , \tag{4.2}$$

$$|k^{(\ell)}(s)| \leq \gamma_\ell\ , \quad \forall\, s \in \mathbb{R}\ , \quad \forall\, \ell = 0, 1, 2, \tag{4.3}$$

(H9)

where α, γ_0, γ_1, and γ_2 are positive constants and $k^{(\ell)}$ stands for the ℓ^{th} derivative of k.

If we define the function K by $K(s) = \int_0^s k(t)dt$, using the property (4.2) we deduce that K is increasing, thus G the inverse of K exists, i.e. $K(G(s)) = s$, $\forall\, s \in \mathbb{R}$, and G is a C^2 mapping which first derivative g is defined by $G'(s) = g(s) = k(G(s))^{-1}$. Moreover, estimate (4.2) implies that g is bounded. Set $U = K(u)$, then $\nabla K(u) = k(u)\nabla u$ and Problem (4.1) is equivalent to find $U \in H_0^1(\Omega)$ such that:

$$-\Delta U + g\,(U)\,\mathbf{c} \cdot \nabla U = f\ , \quad \text{in } \Omega\ . \tag{4.4}$$

Since the function g is bounded, Problem (4.4) is equivalent to find $U \in H_0^1(\Omega)$ such that:

$$\int_\Omega \nabla U \cdot \nabla \varphi dx + \int_\Omega g(U)\,\mathbf{c} \cdot \nabla U \varphi dx = \int_\Omega f \varphi dx\ , \quad \forall\, \varphi \in H_0^1(\Omega)\ . \tag{4.5}$$

Theorem 5. *Assume Hypothesis (H9) holds and that* $\mathbf{c} \in [C^1(\overline{\Omega})]^2$ *with* $\operatorname{div}\mathbf{c} = 0$. *Then Problem (4.5) has at least one solution* $U \in H_0^1(\Omega)$. *Moreover,* $U \in W^{2,p}(\Omega)$ *for all* p, $2 \leq p < +\infty$, *and* $u = G(U) \in W^{2,p}(\Omega) \cap H_0^1(\Omega)$ *for all* p, $2 \leq p < +\infty$ *is a solution to Problem (4.1).*

Proof. We know that $-\Delta$ is an isophormism from $H_0^1(\Omega)$ onto $H^{-1}(\Omega)$, then denoting by T its inverse, we easily verify that (4.5) is equivalent to:

$$U = T(f - g(U)\,\mathbf{c} \cdot \nabla U)\ . \tag{4.6}$$

The mapping $U \to g(U)\,\mathbf{c}\cdot\nabla U$ is continuous and compact from $H_0^1(\Omega)$ into $H^{-1}(\Omega)$, so by setting $S(U) = T(f - g(U)\,\mathbf{c} \cdot \nabla U)$, we have that S is a compact continuous mapping from $H_0^1(\Omega)$ into itself.

It is not difficult to check that problem (4.5) is equivalent to find a fixed point of the operator S. It is well known (see for example Gilbarg-Trudinger [8]) that S has at least one fixed point if there exists a constant C such that for all solutions of $V = \lambda S(V)$ with $\lambda \in [0, 1]$, we have $\|V\|_{1,\Omega} \leq C$. Consider V satisfying $V = \lambda S(V)$ for $\lambda \in [0, 1]$ that is to say: $V \in H_0^1(\Omega)$ and verifies:

$$\int_\Omega \nabla V \cdot \nabla \psi dx + \lambda \int_\Omega g(V)\,\mathbf{c} \cdot \nabla V \psi dx = \lambda \int_\Omega f \psi dx\ \forall\, \psi \in H_0^1(\Omega)\ . \tag{4.7}$$

If we prove that $\int_\Omega g(V)\,\mathbf{c} \cdot \nabla V V dx = 0$, then Equality (4.7) with $\psi = V$ combined with Poincaré and Schwarz inequalities provide a bound for $\|V\|_{1,\Omega}$ independent of λ.

Let M be defined by $M(s) = \int_0^s g(t)tdt$. We verify that $M(V) \in W_0^{1,1}(\Omega)$ if $V \in H_0^1(\Omega)$ since g is bounded, and we have $\nabla M(V) = g(V)V\nabla V$. It follows that

$$\int_\Omega g(V)\,\mathbf{c} \cdot \nabla V V dx = \int_\Omega \mathbf{c} \cdot \nabla M(V) dx = -\int_\Omega \operatorname{div}\mathbf{c} M(V) dx = 0\ ,$$

which proves the existence of U solution of (4.5). Since the function g is bounded, $g(U)\,\mathbf{c} \cdot \nabla U$ belongs to $L^2(\Omega)$; the elliptic regularity of the laplacian operator in a square implies $U \in H^2(\Omega)$ (see for example Grisvard [10]). The embedding of $H^2(\Omega)$ into $W^{1,p}(\Omega)$ for all $p, 2 \leq p < +\infty$ and one more time the elliptic regularity of the laplacian operator lead to $U \in W^{2,p}(\Omega) \cap H_0^1(\Omega)$ for all $p, 2 \leq p < +\infty$.

It is easy to show that $u = G(U)$ is a solution of (4.1) and $u \in W^{2,p}(\Omega) \cap H_0^1(\Omega)$ for all p, $2 \leq p < +\infty$. □

Remark 4. In fact, by using a maximum principle, we can generalize Theorem 5 as follows: If $\mathbf{c} \in [C^0(\overline{\Omega}]^2$, Problem (4.1) has always a solution $u \in W^{2,p}(\Omega)$. Moreover, if $\mathbf{c} \in [C^1(\overline{\Omega})]^2$ is such that $\operatorname{div} \mathbf{c} = 0$, then the solution is unique.

In all the following, we will deal with Problem (4.1) instead of Problem (4.4) or Problem (4.5). In fact, for theoretical reasons, it is easier to control Problem (4.4) because the nonlinearity does not take place in the principal part of the operator. However, it is Problem (4.1) which is discretized in practice, when seeking for a numerical approximation of u. It is the reason why we are interested in Problem (4.1).

In all the following, $u \in W^{2,p}(\Omega) \cap H_0^1(\Omega)$, $2 \leq p < +\infty$ will be a solution of Problem (4.1) and F will be the mapping defined by:

$$F(v) = -\operatorname{div}(k(v)\nabla v) + \mathbf{c} \cdot \nabla v - f \ , \quad \forall\, v \in H_0^1(\Omega) \ .$$

Clearly, F is well defined as a mapping from $H_0^1(\Omega)$ into its dual $H^{-1}(\Omega)$ when hypothesis (H9) holds, and we have $F(u) = 0$. Unfortunately this mapping is not C^1. To overcome this difficulty we introduce the standard Sobolev's spaces $X = W_0^{1,p}(\Omega)$ with $p > 2$, $Y = W_0^{1,q}(\Omega)$ with $\frac{1}{p} + \frac{1}{q} = 1$, and their dual spaces $X' = W^{-1,q}(\Omega)$, $Y' = W^{-1,p}(\Omega)$. Clearly $W_0^{1,p}(\Omega) \hookrightarrow C^0(\overline{\Omega})$ and we have:

Theorem 6. *Under Hypothesis (H9), the mapping F, considered from X into Y' is a C^1-mapping. Moreover there exists c_0 such that if $\|\mathbf{c}\|_{0,\infty,\Omega} \leq c_0$, then $DF(u)$ is an isomorphism from X onto Y'.*

Proof. If we denote by $\langle\cdot|\cdot\rangle$ the duality pairing between Y' and Y, we have, for all $\psi \in Y$ and $v \in X$:

$$\langle F(v)|\psi\rangle = \int_\Omega k(v)\nabla v \cdot \nabla\psi dx + \int_\Omega \mathbf{c} \cdot \nabla v \psi dx - \int_\Omega f\psi dx \ . \tag{4.8}$$

We easily check that F is C^1 and that for all $v, w \in X$, $\psi \in Y$ we have:

$$\begin{aligned} \langle DF(v)w|\psi\rangle &= \int_\Omega k(v)\nabla w \cdot \nabla\psi dx + \int_\Omega k'(v)w\nabla v \cdot \nabla\psi dx + \int_\Omega \mathbf{c} \cdot \nabla w\psi dx \\ &= \int_\Omega \nabla(k(v)w) \cdot \nabla\psi dx + \int_\Omega \mathbf{c} \cdot \nabla w\psi dx \ . \end{aligned} \tag{4.9}$$

Now we show that $DF(u)$ is an injective operator if $\mathbf{c}$ is small enough. If $w \in X$ is such that $DF(u)w = 0$, we have:

$$\int_\Omega \nabla(k(u)w) \cdot \nabla\psi dx + \int_\Omega \mathbf{c} \cdot \nabla w\psi dx = 0 \ , \quad \forall\, \psi \in Y \ . \tag{4.10}$$

By setting $\omega = k(u)w$, we can choose $\psi = \omega$ in (4.10), then using Hölder's inequality we get:

(4.11) $$|\omega|^2_{1,\Omega} \le \|\mathbf{c}\|_{0,\infty,\Omega}\|\omega\|_{0,\Omega}\left\|\nabla\frac{\omega}{k(u)}\right\|_{0,\Omega} .$$

Poincaré's inequality and Hypothesis (H9) give the existence of a constant d (independent of ω) such that

$$\|\omega\|_{0,\Omega}\left\|\nabla\frac{\omega}{k(u)}\right\|_{0,\Omega} \le d|\omega|^2_{1,\Omega} .$$

It follows that for $d\|\mathbf{c}\|_{0,\infty,\Omega} < 1$ we have $\omega = 0$ and consequently $w = 0$, this means that $DF(u)$ is injective when $\|\mathbf{c}\|_{0,\infty,\Omega}$ is small enough.

Combining the characterization of $W^{-1,p}(\Omega)$ given in [10], p. 17, the result of regularity given in Dautray-Lions [7], p. 538, and a symmetry method for the square (see [7], p. 652), we can prove that the laplacian operator is an isomorphism from $W^{1,p}_0(\Omega)$ onto $W^{-1,p}(\Omega)$, the inverse of which is denoted by T (see also item (5.5) of Dau [6], p. 241). From relation (4.9) it follows:

$$T\ DF(u)w = k(u)w + T(\mathbf{c}\cdot\nabla w) .$$

Since the mapping $R : \varphi \in W^{1,p}_0(\Omega) \to R\varphi = \varphi/k(u) \in W^{1,p}_0(\Omega)$ is an isomorphism, the inverse of which is the multiplying operator by $k(u)$, we have:

(4.12) $$T\ DF(u)R\varphi = \varphi + T(\mathbf{c}\cdot\nabla R\varphi) .$$

Since the operator T is compact from $L^p(\Omega)$ into $W^{1,p}_0(\Omega)$ (due to the elliptic regularity of the laplacian operator and the compact embedding of $W^{2,p}(\Omega)$ into $W^{1,p}(\Omega)$), then $T\ DF(u)R$ is a Fredholm's operator with index zero from $W^{1,p}_0(\Omega)$ onto $W^{-1,p}(\Omega)$ when $DF(u)$ is injective.

Since R is an isomorphism of $W^{1,p}_0(\Omega)$ and since T is an isomorphism from $W^{-1,p}(\Omega)$ onto $W^{1,p}_0(\Omega)$, it follows that $DF(u)$ is an isomorphism from $W^{1,p}_0(\Omega)$ onto $W^{-1,p}(\Omega)$ when $DF(u)$ is injective. □

Remark 5. In fact, in Theorem 6, it is not necessary to assume that **c** is small in the C^0-norm. By using a maximum principle, we can prove that if $\mathbf{c} \in C^1(\overline{\Omega})$ is such that $\operatorname{div}\mathbf{c} = 0$, then $DF(u)$ is an isomorphism from X onto Y'.

Remark 6. By considering (4.9), it is easy to see that $DF(u)$ admits a continuous extension from $H^1_0(\Omega)$ into the dual space $H^{-1}(\Omega)$. Moreover, using Lax-Milgram theorem, it is standard to define the continuous one-to-one linear operator $T_2 : H^{-1}(\Omega) \to H^1_0(\Omega)$ as the inverse of the laplacian operator. The operator T_2 is compact from $L^2(\Omega)$ into $H^1_0(\Omega)$. It follows that the operator $T_2DF(u)R$, defined in the proof of Theorem 6, is a Fredholm's operator with index zero from $H^1_0(\Omega)$ into itself. The density of $W^{1,p}_0(\Omega)$ into $H^1_0(\Omega)$ implies that the range of $T_2DF(u)R$ considered in $H^1_0(\Omega)$ is all the space $H^1_0(\Omega)$. Thus, $T_2DF(u)R$ is an isomorphism of $H^1_0(\Omega)$ and consequently $DF(u)$ is an isomorphism from $H^1_0(\Omega)$ onto $H^{-1}(\Omega)$. We conclude that

(4.13) $$\inf_{\substack{v\in H^1_0(\Omega)\\ |v|_{1,\Omega}=1}} \sup_{\substack{w\in H^1_0(\Omega)\\ |w|_{1,\Omega}=1}} \langle DF(u)v|w\rangle = \gamma > 0 . \qquad \square$$

Now we can use the abstract framework of Sect. 3 for treating an approximation of Problem (4.1) by finite element method. To do this, we assume in the following that Problem (4.1) has a solution u such that $DF(u)$ is an isomorphism from X onto Y' (this hypothesis is not void because it is the case when $\mathbf{c} \in [C^1(\overline{\Omega})]^2$ is such that div $\mathbf{c}$ = 0 or $\mathbf{c}\|_{0,\infty,\Omega}$ is small enough (Theorems 5 and 6)). As in Sect. 3 we set $b(w,\psi) = \langle DF(u)w|\psi\rangle$ and we have (see (4.9)):

$$b(w,\psi) = \int_\Omega \nabla(k(u)w)\nabla\psi dx + \int_\Omega \mathbf{c}\cdot\nabla w\psi dx \ .$$

Naturally Hypothesis (H5) of Sect. 3 is true with $\beta = \|DF(u)^{-1}\|^{-1}_{Y'X}$.

Let now $\mathscr{C}_h$ be a quasi-uniform regular triangulation of $\overline{\Omega}$ (see [4] for the definition) and let V_h be the finite element subspace defined by

$$V_h = \{\varphi \in C^0(\overline{\Omega}) : \varphi/K \text{ is a polynomial of degree } \leq 1 \ , \ \forall\, K \in \mathscr{C}_h \ ; \varphi = 0 \text{ on } \partial\Omega\} \ .$$

Clearly $V_h \subset W_0^{1,s}(\Omega)$, $s \geq 1$, and we choose $X_h = Y_h = V_h$ as finite dimensional subspaces of X and Y respectively.

It follows that a finite element aproximation of Problem (4.1) consists on finding $u_h \in V_h$ satisfying:

$$\int_\Omega k(u_h)\nabla u_h\cdot\nabla v_h dx + \int_\Omega \mathbf{c}\cdot\nabla u_h v_h dx = \int_\Omega f v_h dx \ , \quad \forall\, v_h \in V_h \ . \tag{4.14}$$

Note that, in practice, we compute these terms by using numerical integration. For the sake of simplicity, we analyze the approximate problem (4.14) without numerical integration. In the following, we verify Hypotheses (H6) and (H7) of Sect. 3.

Theorem 7. *Under Hypothesis (H9) and if $F(u) = 0$ and $DF(u)$ is an isomorphism from $W_0^{1,p}(\Omega)$ onto $W^{-1,p}(\Omega)$ with $p > 2$, then we have*

$$\inf_{\substack{v_h\in V_h\\ |v_h|_{1,p,\Omega}=1}} \ \sup_{\substack{w_h\in V_h\\ |w_h|_{1,q,\Omega}=1}} b(v_h,w_h) \geq \xi h^{(p-2)/p} \tag{4.15}$$

where ξ is a positive constant independent of $\mathscr{C}_h$, and h is the maximum of diameters of triangles $K \in \mathscr{C}_h$.

Proof. By considering that $DF(u)$ is an isomorphism from $W_0^{1,p}(\Omega)$ onto $W^{-1,p}(\Omega)$ and by taking into account Remark 6, we state that:

(4.16) $\qquad DF(u)$ is an isomorphism from $H_0^1(\Omega)$ onto $H^{-1}(\Omega)$.

We now define the operator $T_h : H^{-1}(\Omega) \to V_h$ by the following relation:

$$\int_\Omega \nabla(T_h g)\cdot\nabla\varphi_h dx = \int_\Omega g\varphi_h dx \ , \quad \forall\, \varphi_h \in V_h \ , \quad g \in H^{-1}(\Omega) \ ,$$

where the second integral means the duality pairing between $H^{-1}(\Omega)$ and $H_0^1(\Omega)$. If T is the inverse of the laplacian operator with homogeneous boundary conditions and if R is the mapping defined by $R\varphi = \varphi/k(u)$, then we have seen in Remark 6 that the operator $TDF(u)R$ is an isomorphism from $H_0^1(\Omega)$ onto itself. Moreover we have seen that

(4.17) $$T\ DF(u)R\varphi = \varphi + T(\mathbf{c}\cdot\nabla R\varphi)\ ,\quad \forall\ \varphi \in H_0^1(\Omega)\ ,$$

and we can verify that

(4.18) $$T_h DF(u)R\varphi_h = \varphi_h + T_h(\mathbf{c}\cdot\nabla R\varphi_h)\ ,\quad \forall\ \varphi_h \in V_h\ .$$

It is well known that the hypothesis about the triangulation implies

(4.19) $$\lim_{h\to 0} \|T - T_h\|_{L^2(\Omega)H_0^1(\Omega)} = 0\ .$$

From relations (4.16), (4.17), (4.18) and (4.19), we conclude that $(I + T_h(\mathbf{c}\cdot\nabla R))$ is an isomorphism of $H_0^1(\Omega)$ with uniformely bounded inverse with respect to h, and there exists a positive constant $\gamma > 0$ such that:

(4.20) $$|T_h DF(u)R\varphi_h|_{1,\Omega} \geq \gamma|\varphi_h|_{1,\Omega}\ ,\quad \forall\ \varphi_h \in V_h\ .$$

For v_h and $w_h \in V_h$ we obtain:

$$b(Rv_h, w_h) = \langle DF(u)Rv_h|w_h\rangle = \int_\Omega \nabla(T_h DF(u)Rv_h)\cdot\nabla w_h dx\ .$$

By taking the supremum on w_h and by using (4.20) we have:

(4.21) $$\sup_{\substack{w_h\in V_h\\ |w_h|_{1,\Omega}=1}} b(Rv_h, w_h) = |T_h DF(u)Rv_h|_{1,\Omega} \geq \gamma|v_h|_{1,\Omega}\ .$$

By using Hypothesis (H9) on k, the regularity of the triangulation and the regularity of the solution u (see the arguments of Theorem 5), we can easily prove, by standard calculations on the reference triangle, that the multiplying operator R^{-1} satisfies the following properties:

(4.22) $$\lim_{h\to 0}\ \max_{\substack{\chi_h\in V_h\\ |\chi_h|_{1,\Omega}=1}} |R^{-1}\chi_h - r_h R^{-1}\chi_h|_{1,\Omega} = 0$$

where r_h is the interpolation operator on V_h.

Relations (4.21) and (4.22) lead to the existence of a positive constant $\tilde{\gamma}$ (independent of h) such that for $h \leq h_0$ small enough we have:

(4.23) $$\inf_{\substack{v_h\in V_h\\ |v_h|_{1,\Omega}=1}}\ \sup_{\substack{w_h\in V_h\\ |w_h|_{1,\Omega}=1}} b(v_h, w_h) \geq \tilde{\gamma}\ .$$

Taking into account the following inverse inequality (see [4], p. 140):

$$C_1 h^{(p-2)/p}|v_h|_{1,p,\Omega} \leq |v_h|_{1,\Omega}\ ,\quad \forall\ v_h \in V_h\ ,$$

where C_1 is a positive constant independent of h, we deduce (4.15) from (4.23). □

Remark 7. After this article has been submitted, the result of Theorem 7 has been improved. For two dimensional convex polygonal domains Ω, Inf-Sup conditions (4.15) can be replaced by:

$$\inf_{\substack{v_h \in V_h \\ |v_h|_{1,p\Omega}=1}} \sup_{\substack{w_h \in V_h \\ |w_h|_{1,q,\Omega}=1}} b(v_h, w_h) \geq \xi ,$$

where ξ is a positive constant independent of $\mathscr{C}_h$ and where $p > 2$.

Theorem 7 implies Hypothesis (H6) of Sect. 3 with $\beta_h = \xi h^{(p-2)/p}$. On the other hand, we have the following interpolation inequality if $u \in W^{2,p}(\Omega)$:

$$|u - r_h u|_{1,p,\Omega} \leq ch\|u\|_{2,p,\Omega} \ . \tag{4.24}$$

It follows that Hypothesis (H7) of Sect. 3 is satisfied for $p < 4$. Moreover we verify from relation (4.9) that DF is Lipschitzian in u and the hypotheses of Thereom 4 hold. Consequently, we obtain the main result:

Theorem 8. *We assume that Hypothesis (H9) holds, that the triangulation $\mathscr{C}_h$ is quasi-uniform, that the solution u of Problem (4.1) belongs to $W^{2,p}(\Omega)$ with $p \in]2,4[$ and that $DF(u)$ is an isomorphism from $W_0^{1,p}(\Omega)$ onto $W^{-1,p}(\Omega)$ (see Theorems 5 and 6 which prove these hypotheses are not void).*

Then there exist $h_0 > 0$ and $\eta_0 > 0$ such that for $h \leq h_0$ there exists a unique $u_h \in V_h$ solution of Problem (4.14) satisfying $|u - u_h|_{1,p,\Omega} \leq \eta_0 h^{(p-2)/p}$. Moreover we have the error estimates:

$$|u - u_h|_{1,p,\Omega} \leq ch^{2/p} \tag{4.25}$$

and

$$|u - u_h|_{1,p,\Omega} \leq c \left(\sum_{K \in \mathscr{C}_h} \eta(K)^p \right)^{1/p} \tag{4.26}$$

where c is a constant independent of the triangulation $\mathscr{C}_h$ and $\eta(K)$ is the local estimator given by

$$\eta(K) = h_k\| -\operatorname{div}(k(u_h)\nabla u_h) + \mathbf{c} \cdot \nabla u_h - f\|_{0,p,K} + h_K^{(2-p)/p} \sum_{i=1}^{3} h_{s_i}^{(p-1)/p} \left\| \left[k(u_h) \frac{\partial u_h}{\partial_n} \right] \right\|_{0,p,S_i} ;$$

here h_K is the diameter of $K \in \mathscr{C}_h$, h_{s_i} are the lengths of edges S_i of the triangle K, $1 \leq i \leq 3$, and $[v]$ denotes the jump of v across the considered edge (we adopt the convention $\frac{\partial u}{\partial n} = 0$ outside the domain Ω).

Proof. Theorem 8 is a consequence of Theorems 4 and 8 if we prove the two following error estimates:

$$\frac{1}{\beta_h} \min_{v_h \in V_h} |u - v_h|_{1,p,\Omega} \leq ch^{2/p} \tag{4.27}$$

and

$$\|F(u_h)\|_{Y'} \leq c \left(\sum_{K \in \mathscr{C}_h} \eta(K)^p \right)^{1/p} \ . \tag{4.28}$$

Inequality (4.27) is a direct consequence of (4.24) and $\beta_h = \xi h^{(p-2)/p}$.

Let us show how we can compute the residual norm $\|F(u_h)\|_{Y'}$ (see for instance [3] or [14]).

For $v \in Y$, $v_h \in Y_h$ we have (using (4.14)):

$$\begin{aligned}
&\langle F(u_h)|v\rangle = \langle F(u_h)|v - v_h\rangle \\
&= \sum_{K\in\mathscr{C}_h} \left\{ \int_K k(u_h)\nabla u_h \cdot \nabla(v - v_h)dx + \int_K \mathbf{c}\cdot\nabla u_h(v - v_h)dx - \int_K f(v - v_h)dx \right\} \\
&= \sum_{K\in\mathscr{C}_h} \left\{ \int_K (-\mathrm{div}\,(k(u_h)\nabla u_h) + \mathbf{c}\cdot\nabla u_h - f)(v - v_h)dx \right. \\
&\quad \left. + \int_{\partial K} k(u_h)\frac{\partial u_h}{\partial n_K}(v - v_h)ds \right\} ,
\end{aligned}$$

where $\frac{\partial u_h}{\partial n_K}$ is the exterior normal derivative of u_h on the boundary ∂K of K.

By applying Hölder's inequalities, we obtain:

$$\begin{aligned}
\langle F(u_h)|v\rangle \leq \sum_{K\in\mathscr{C}_h} & \left\{ (\| - \mathrm{div}\,(k(u_h)\nabla u_h) + \mathbf{c}\cdot\nabla u_h - f\|_{0,p,K}\|v - v_h\|_{0,q,K} \right. \\
& \left. + \sum_{i=1}^{3} \left\| \left[k(u_h)\frac{\partial u_h}{\partial n} \right] \right\|_{0,p,S_i} \|v - v_h\|_{0,q,S_i} \right\}
\end{aligned}$$

where, here $[\frac{\partial u_h}{\partial n}]$ is the jump of a normal derivative of u with the convention $\frac{\partial u}{\partial n} = 0$ outside the domain $\overline{\Omega}$.

It is known (see [3] for instance) that if $v \in Y$ and if $r_h v$ is the Clement's interpolate of v on V_h, we have:

$$\begin{aligned}
&\|v - r_h v\|_{0,q,K} \leq ch_K \sum_{K'\in S_K} \|v\|_{1,q,K'}; S_K = \{\cup K/K \cap K' = \phi\} \\
&\|v - r_h v\|_{0,q,S_i} \leq ch_{S_i}^{1/q} h_K^{1-2/p} \sum_{K'\in S_K} \|v\|_{1,q,K'},\ i = 1,2,3 ,
\end{aligned}$$

where, here c is independent of the triangle K. It follows that

$$\begin{aligned}
\langle F(u_h)|v\rangle &\leq c \sum_{K\in\mathscr{C}_h} \left\{ h_K\| - \mathrm{div}\,(k(u_h)\nabla u_h) + \mathbf{c}\cdot\nabla u_h - f\|_{0,p,K} \right. \\
&\quad \left. + \sum_{i=1}^{3} \left\| \left[k(u_h)\frac{\partial u_h}{\partial n} \right] \right\|_{0,p,S_i} h_{S_i}^{1/q} h_K^{1-2/q} \right\} \sum_{K'\in S_K} \|v\|_{1,q,K'} \\
&\leq c_1 \left(\sum_{K\in\mathscr{C}_h} \left\{ h_K\| - \mathrm{div}\,(k(u_h)\nabla u_h) + \mathbf{c}\cdot\nabla u_h - f\|_{0,p,K} \right. \right. \\
&\quad \left. \left. + \sum_{i=1}^{3} \left\| \left[k(u_h)\frac{\partial u_h}{\partial n} \right] \right\|_{0,p,S_i} h_{S_i}^{1/q} h_K^{1-2/q} \right\}^p \right)^{1/p} |v|_{1,q,\Omega} . \qquad \square
\end{aligned}$$

Finally let us show how to obtain a priori and a posteriori error estimates in H^1-norm similar to estimates (4.25) and (4.26) with $p = 2$. To do this we use the following arguments for obtaining a posteriori error estimates:

(i) For $p \in]2,4[$ we have $\lim_{h\to 0} |u-u_h|_{1,p,\Omega} = 0$ (see (4.25));

(ii) for $p \in]2,\infty[$ and $w \in W_0^{1,p}(\Omega)$, the mapping $DF(w) \in \mathscr{L}(W_0^{1,p}(\Omega), W^{-1,p}(\Omega))$ admits a continuous extension $DF(w) \in \mathscr{L}(H_0^1(\Omega), H^{-1}(\Omega))$; moreover the function $w \in W_0^{1,p}(\Omega) \to DF(w) \in \mathscr{L}(H_0^1(\Omega), H^{-1}(\Omega))$ is continuous;

(iii) since $F(u) = 0$ we can write

$$u - u_h = DF(u)^{-1} \int_0^1 (DF(u) - DF(su + (1-s)u_h))(u-u_h)ds - DF(u)^{-1}F(u_h) \ .$$

It is clear that items (i)–(iii) give rise to a posteriori error estimate of type

$$|u-u_h|_{1,\Omega} \leq C\|F(u_h)\|_{-1,\Omega}$$

where

$$C = 2\|DF(u)^{-1}\|_{\mathscr{L}(H^{-1}(\Omega),H_0^1(\Omega))} \text{ if } h \leq h_0 \text{ is small enough.}$$

Such a posteriori error estimates have been used in adaptive mesh refinement techniques applied to a 2D-regularized Stefan problem (see [13]) and yield excellent results.

Remark 8. According to Remark 7, estimate (4.25) can be improved and becomes:

$$|u-u_h|_{1,p,\Omega} \leq ch \ .$$

References

1. Babuska, I., Aziz, I.K. (1972): Survey lecture on the mathematical foundations of the finite element method with applications to partial differential equations. Academic Press, New York, London
2. Bamberg, A. (1989): Définition d'une méthodologie de démonstration numérique d'existence de solutions d'équations elliptiques semi-linéaires. In: Glowinski, R., ed., Comp. Math. Appl. Sciences Eng., INRIA
3. Baranger, J., Elamri, H. (1991): A posteriori estimates for the adaptive computation of quasi non newtonian flows. M^2AN **25**, 31–48
4. Ciarlet, P.G. (1978): The finite element method for elliptic problems. In: Lions, J.L., Paponicolaou, G., Rockafellar, R.T., eds., North Holland
5. Crouzeix, M., Rappaz, J. (1990): On numerical approximation in bifurcation theory. Masson, Paris
6. Dauge, M. (1992): Neumann and mixed problems on curvilinear polyhedra. Integral equations and operator theory, vol. 15, No. 2
7. Dautray, R., Lions, J.L. (1987): Analyse mathématique et calcul numérique pour les sciences et les techniques. T.2, Masson
8. Gilbarg, D., Trudinger, N.S. (1983): Elliptic partial differential equations of second order. A series of comprehensive studies in Mathematics No. 224. Springer
9. Girault, V., Raviart, P.A. (1982): An analysis of upwind schemes for the Navier Stokes equations. SIAM J. Numer. Anal. **19**, 312–333
10. Grisvard, P. (1985) Elliptic problems in nonsmooth domains. Monographs and Studies in Mathematics, No. 24. Pitman
11. Johnson, C. (1990): Adaptive finite element methods for diffusion and convection problems. Proceedings of the workshop on reliability in Comp. Mach. Lakeway, Austin Texas, September 1990. Comp. Meth. Appl. Mech. Ing. **82**

12. Johnson, C., Hansbo, P. (1990): Adaptive streamline diffusion methods for compressible flow using conservation variables. Comp. Meth. Appl. Mech. Eng. Proceedings from Symposium on large scale computations in fluid dynamics. Minnesota, Super-Computer Institute
13. Picasso, M. (1992): Simulation numérique de traitements de surface par laser. Thèse EPFL, no. 1011
14. Pousin, J., Rappaz, J. (1991): Consistance, stabilité, erreurs a priori et a posteriori pour des problèmes non linéaires. C.R. Acad. Sci. Paris, T. 312, Série I, pp. 699–703
15. Verfuehrt, R. (1989): A posterio error estimates for the Stokes equations. Numer. Math. **15**, 309–325