M2AN. MATHEMATICAL MODELLING AND NUMERICAL ANALYSIS - MODÉLISATION MATHÉMATIQUE ET ANALYSE NUMÉRIQUE

GABRIEL CALOZ

Approximation by finite element method of the model plasma problem

M2AN. Mathematical modelling and numerical analysis - Modélisation mathématique et analyse numérique, tome 25, nº 1 (1991), p. 49-65

<http://www.numdam.org/item?id=M2AN_1991__25_1_49_0>

MATHEMATICAL MODELLING AND NUMERICAL ANALYSIS
MODÉLISATION MATHÉMATIQUE ET ANALYSE NUMÉRIQUE

(Vol. 25, n° 1, 1991, p. 49 à 66)

APPROXIMATION BY FINITE ELEMENT METHOD OF THE MODEL PLASMA PROBLEM (*)

Gabriel CALOZ [1]

Communicated by F. BREZZI

Abstract. — *We analyze finite element approximations of the model plasma problem* $-\Delta w = \lambda (w-d)^+$ *in* Ω, $w = 0$ *on* $\partial\Omega$, $\lambda \int_\Omega (w-d)^+ \, dx = j$, *where* Ω *is a bounded domain in* $\mathbb{R}^2$ *with boundary* $\partial\Omega$ *and* j *is a given positive number ; the function* w *and the real* λ *are the unknowns,* d *is a parameter. We can show that the finite element approximation, in the case of numerical integration too, converges in the norms of* $H^1(\Omega)$ *and* $L^\infty(\Omega)$ *at the optimal rate.*

Résumé. — *Nous analysons une approximation par éléments finis du problème modèle des plasmas* $-\Delta w = \lambda (w-d)^+$ *dans* Ω, $w = 0$ *sur* $\partial\Omega$, $\lambda \int_\Omega (w-d)^+ \, dx = j$, *où* Ω *est un domaine borné de* $\mathbb{R}^2$ *de frontière* $\partial\Omega$ *et* j *un réel positif ; la fonction* w *et le réel* λ *sont les inconnues,* d *est un paramètre. Nous pouvons montrer que l'approximation par éléments finis proposée, y compris avec intégration numérique, a une convergence optimale dans les normes de* $H^1(\Omega)$ *et* $L^\infty(\Omega)$.

1. INTRODUCTION

Let Ω be a regular bounded domain in $\mathbb{R}^2$ with boundary $\partial\Omega$ and j be a real positive number. We define the mapping $S : \mathbb{R}^2 \times H^1(\Omega) \to \mathbb{R} \times H^1(\Omega)$ by

$$(1.1) \quad S(e, l, v) = \begin{pmatrix} l \int_\Omega (v-e)^+ \, dx - j \\ v - lT(v-e)^+ \end{pmatrix} \quad \forall (e, l, v) \in \mathbb{R}^2 \times H^1(\Omega),$$

(*) Received in January 1989.
This research was supported by the Swiss National Science Foundation.
[1] Department of Mathematics, University of Maryland, College Park, Maryland, MD 20742, U.S.A.

where the compact linear operator $T \in \mathcal{L}(L^2(\Omega)\,;\, H^1(\Omega))$ is given by : for a function $f \in L^2(\Omega)$, $Tf \equiv v \in H^1(\Omega)$ is the solution of

$$-\Delta v = f \text{ in } \Omega\,, \quad v = 0 \text{ on } \partial\Omega\,. \tag{1.2}$$

For ease of exposition, we assume that

$$(H)\ \Omega \text{ is a convex domain}\,. \tag{1.3}$$

Remark that our analysis can be carry out with some technical difficulties if we assume that Ω is such that $T \in \mathcal{L}(L^2(\Omega)\,;\, H^2(\Omega))$. We will study finite element approximations of

PROBLEM 1.1 : *For the value d of the parameter, find $\lambda \in \mathbb{R}$ and $w \in H^1(\Omega)$ such that*

$$S(d, \lambda, w) = 0\,.$$

Remark that Problem 1.1 is called the model plasma problem and has been studied first in [11] ; it can be related directly to the study of the static magnetohydrodynamic (MHD) equilibria of a plasma confined in a tokamak, see [11], [5], [2], for an account. Because the mapping $f : u \in H^1(\Omega) \mapsto f(u) \equiv u^+ \in L^2(\Omega)$ is not of class C^1 ($Df(u)$ does not exist if the measure of the set $\{x \in \Omega : u(x) = 0\}$ is different from zero), the mapping S is not continuously Fréchet differentiable.

Let $\lambda_1 < \lambda_2 \leq \lambda_3 \leq \ldots$, be the characteristic values of the operator T. The characteristic value λ_1 is positive, simple and the corresponding eigenvector ϕ_1 can be chosen positive such that

$$\phi_1 = \lambda_1 T\phi_1\,, \quad \lambda_1 \int_\Omega \phi_1\, dx = j\,. \tag{1.4}$$

Using a variant of the implicit function theorem given in [10], Caloz proved in [2] the following result.

THEOREM 1.2 : *There exist a positive number $\bar{d}$ and two C^1 mappings*

$$\lambda : d \in [0, \bar{d}] \mapsto \lambda(d) \in \mathbb{R} \quad \textit{and} \quad w : d \in [0, \bar{d}] \mapsto w(d) \in H^1(\Omega)$$

such that

$$\operatorname{meas}\, \{x \in \Omega : w(d)(x) = d\} = 0\,, \tag{1.5}$$

$$S(d, \lambda(d), w(d)) = 0\,, \quad \forall d \in [0, \bar{d}]\,, \tag{1.6}$$

$$\lambda(0) = \lambda_1\,, \quad w(0) = \phi_1\,; \quad \lambda(\bar{d}) = \lambda_2\,. \tag{1.7}$$

Moreover $\lambda(\,.\,)$ *is increasing and the branch* $d \in [0, \bar{d}] \mapsto (\lambda(d), w(d)) \in \mathbb{R} \times H^1(\Omega)$ *is regular, i.e.* $D_{l,v}\, S(d, \lambda(d), w(d))$ *exists and is regular for all* $d \in [0, \bar{d}]$.

The main goal of this paper is to analyze the gap between the solution branch given in Theorem 1.2 and an approximation of this branch by a finite element method. We will focus our study on the approximation of Problem 1.1 with piecewise linear elements, including numerical integration, and obtain rates of convergence.

The layout of the paper is as follows. In Section 2, we state all the theoretical results concerning the approximation of Problem 1.1. In Section 3, we present in an abstract framework a result on the approximation of the regular solutions of nonlinear problems stated in [4]. Finally Section 4 is devoted to the proofs.

Let us mention that the error estimates at optimal rates were already proved in [2] but without numerical integration. Moreover here we follow [4] where a simpler, alternate proof has been developed. In that way, we can improve the results obtained by Kikuchi *et al.* [8] and the ones by Barrett and Elliott [1], when numerical integration is introduced.

2. THEORETICAL RESULTS

Throughout the paper we adopt the following standard notation. Let X, Z be two Banach spaces with the norms $\|\,.\,\|_X$, $\|\,.\,\|_Z$; when there is no ambiguity we omit the subscript. $\mathcal{L}(X; Z)$ denotes the space of continuous linear operators from X to Z. The norm in the product space $\mathbb{R} \times X$ is defined by $\|\,.\,\|_{\mathbb{R}\times X} = |\,.\,| + \|\,.\,\|_X$. The $W^{m,p}(\Omega)$ are the Sobolev spaces on Ω, with the norm $\|\,.\,\|_{m,p,\Omega}$ and the seminorm $|\,.\,|_{m,p,\Omega}$. When p is equal to 2, $W^{m,p}(\Omega)$ is a Hilbert space denoted by $H^m(\Omega)$ with the scalar product $(\,.\,,\,.\,)_{m,\Omega}$ inducing the norm $\|\,.\,\|_{m,\Omega} \equiv \|\,.\,\|_{m,2,\Omega}$. We set $H_0^1(\Omega) = \{v \in H^1(\Omega) : v = 0 \text{ on } \partial\Omega\}$, $a(\,.\,,\,.\,)$ denotes the Dirichlet form on $H_0^1(\Omega)$. $C^0(\bar{\Omega})$ is the space of continuous functions on $\bar{\Omega}$.

Let $\{\mathcal{T}_h\}_{0<h<1}$ be a regular family of triangulations of Ω, (see [3] for instance), where h is the maximum of the diameters of the triangles K in $\mathcal{T}_h$. We set $\bar{\Omega}_h = \bigcup_{K \in \mathcal{T}_h} K$ and suppose that the vertices on the boundary of $\bar{\Omega}_h$ are also on the boundary $\partial\Omega$. With the hypothesis (1.3), we have $\Omega_h \subset \Omega$, Ω_h being the interior of $\bar{\Omega}_h$. We define the finite element space

$$(2.1) \qquad V_h = \left\{ v_h \in C^0(\bar{\Omega}) : v_h(x) = 0 \text{ if } x \notin \Omega_h,\ v_h|_K \in \mathcal{P}_1 \quad \forall K \in \mathcal{T}_h \right\},$$

where $\mathscr{P}_1$ denotes the space of polynomials of degree 1. With this space V_h, we associate the forms

$$(2.2) \qquad a(u, v) = \int_\Omega \nabla u \, \nabla v \, dx \,,$$

$$(2.3) \qquad (f, v)_h = \int_\Omega f v \, dx \,,$$

$$(2.4) \qquad (f, v)_{\hat{h}} = \sum_{K \in \mathcal{T}_h} \left[\frac{s_K}{3} \sum_{i=1}^{3} f(a_{iK}) \, v(a_{iK}) \right],$$

where s_K is the measure of $K \in \mathcal{T}_h$ and a_{iK}, $i = 1, 2, 3$, its vertices. We define now the operators $T_h : L^2(\Omega) \to V_h$, $\hat{T}_h : C^0(\bar{\Omega}) \mapsto V_h$ by

$$(2.5) \qquad \text{for } f \in L^2(\Omega)\,, \quad a(T_h f, v) = (f, v)_h \quad \forall v \in V_h\,,$$

$$(2.6) \qquad \text{for } f \in C^0(\bar{\Omega})\,, \quad a(\hat{T}_h f, v) = (f, v)_{\hat{h}} \quad \forall v \in V_h\,,$$

and the mappings $S_h : \mathbb{R}^2 \times V_h \to \mathbb{R} \times V_h$, $\hat{S}_h : \mathbb{R}^2 \times V_h \to \mathbb{R} \times V_h$ by

$$(2.7) \quad S_h(e, l, v) = \begin{pmatrix} l((v - e)^+, 1)_h - j \\ v - l T_h (v - e)^+ \end{pmatrix},$$

$$(2.8) \quad \hat{S}_h(e, l, v) = \begin{pmatrix} l((v - e)^+, 1)_{\hat{h}} - j \\ v - l \hat{T}_h (v - e)^+ \end{pmatrix}, \quad \forall (e, l, v) \in \mathbb{R}^2 \times V_h\,.$$

Finite element approximations of Problem 1.1 can be stated in the following way.

PROBLEM 2.1 : *For the value d of the parameter, find $\lambda_h \in \mathbb{R}$ and $w_h \in V_h$ such that*

$$S_h(d, \lambda_h, w_h) = 0\,.$$

PROBLEM 2.2 : *For the value d of the parameter, find $\lambda_h \in \mathbb{R}$ and $w_h \in V_h$ such that*

$$\hat{S}_h(d, \lambda_h, w_h) = 0\,.$$

In Section 4, we will prove the following results.

PROPOSITION 2.3 : *Let $\bar{d}$, $\lambda(\,.\,)$ and $w(\,.\,)$ be the positive real number and the functions given in Theorem 1.2. There exist two positive constants α, h_0 and for $h \leqslant h_0$ two continuous mappings*

$$\lambda_h : d \in [0, \bar{d}] \mapsto \lambda_h(d) \in \mathbb{R} \quad \text{and} \quad w_h : d \in [0, \bar{d}] \mapsto w_h(d) \in V_h$$

such that for $d \in [0, \bar{d}]$, $h \leqslant h_0$

$$(2.9) \quad (S_h(d, \lambda, w) = 0,\ (\lambda, w) \in B((\tilde{\lambda}_h(d), \tilde{w}_h(d)), \alpha)) \Leftrightarrow \Leftrightarrow (\lambda = \lambda_h(d),\ w = w_h(d)),$$

where

$$(\tilde{\lambda}_h(d), \tilde{w}_h(d)) = (\lambda(d), \lambda(d)\, T_h(w(d) - d)^+)$$

and $B((\tilde{\lambda}_h(d), \tilde{w}_h(d)), \alpha)$ *is the closed ball in* $\mathbb{R} \times H^1(\Omega)$ *with radius* α *and centre* $(\tilde{\lambda}_h(d), \tilde{w}_h(d))$. *Moreover there exists a constant* c *such that*

$$(2.10) \qquad |\lambda(d) - \lambda_h(d)| \leqslant ch^2,$$

$$(2.11) \qquad \|w(d) - w_h(d)\|_{0,\Omega} + h\,\|w(d) - w_h(d)\|_{1,\Omega} \leqslant ch^2;$$

for all $\varepsilon > 0$ *there exists a constant* $c(\varepsilon)$ *with*

$$(2.12) \qquad \|w(d) - w_h(d)\|_{0,\infty,\Omega} \leqslant c(\varepsilon)\, h^{2-\varepsilon},$$

and if $w(d)$ *is in* $W^{2,\infty}(\Omega)$

$$(2.13) \qquad \|w(d) - w_h(d)\|_{0,\infty,\Omega} \leqslant ch^2\, |\ln h|.$$

For $h \leqslant h_0$ and $d \in [0, \bar{d}]$, let us define the sets

$$(2.14) \quad \Delta_h = \{x \in \Omega : (w(d)(x) > d \text{ and } w_h(d)(x) \leqslant d) \text{ or } (w(d)(x) \leqslant d \text{ and } w_h(d)(x) > d)\},$$

$$(2.15) \quad \Gamma_h = \{x \in \Omega : w_h(d)(x) = d\}.$$

PROPOSITION 2.4 : *Let* $d \in [0, \bar{d}]$ *be fixed. For all* $\varepsilon > 0$, *there exists a constant* $c(\varepsilon) > 0$ *such that for all* $h \leqslant h_0$

$$(2.16) \qquad \text{meas}\,(\Delta_h) \leqslant c(\varepsilon)\, h^{2-\varepsilon},$$

$$(2.17) \qquad \text{meas}\,(\Gamma_h) = 0,$$

$$(2.18) \qquad \text{dist}\,(\Gamma, \Gamma_h) \leqslant c(\varepsilon)\, h^{2-\varepsilon},$$

where Γ *is the set* $\{x \in \Omega : w(d)(x) = d\}$.

We have similar results with the mapping $\hat{S}_h$.

PROPOSITION 2.5 : *Let* $\bar{d}$ *be given in Theorem* 1.2 *and* $\lambda_h(\,.\,)$, $w_h(\,.\,)$ *with* $h \leqslant h_0$ *in Proposition* 2.3. *There exist two positive constants* $\hat{\alpha}$, $\hat{h}_0$, *and for*

$h \leq \hat{h}_0$ two continuous mappings

$$\hat{\lambda}_h : d \in [0, \bar{d}] \mapsto \hat{\lambda}_h(d) \in \mathbb{R} \quad \text{and} \quad \hat{w}_h : d \in [0, \bar{d}] \mapsto \hat{w}_h(d) \in V_h$$

such that for $d \in [0, \bar{d}]$, $h \leq \hat{h}_0$

$$(2.19) \quad (\hat{S}_h(d, \lambda, w) = 0,\ (\lambda, w) \in B((\lambda_h(d), w_h(d)), \hat{\alpha}) \cap V_h) \Leftrightarrow \Leftrightarrow (\lambda = \hat{\lambda}_h(d),\ w = \hat{w}_h(d)),$$

where $B((\lambda_h(d), w_h(d)), \hat{\alpha})$ is the closed ball in $\mathbb{R} \times C^0(\bar{\Omega})$ with radius $\hat{\alpha}$ and centre $(\lambda_h(d), w_h(d))$. Moreover there exists a constant c such that

$$(2.20) \quad |\lambda(d) - \hat{\lambda}_h(d)| + \|w(d) - \hat{w}_h(d)\|_{0,\Omega} \leq ch^2 |\ln h|,$$

$$(2.21) \quad \|w(d) - \hat{w}_h(d)\|_{1,\Omega} \leq ch;$$

for all $\varepsilon > 0$ there exists a constant $c(\varepsilon)$ with

$$(2.22) \quad \|w(d) - \hat{w}_h(d)\|_{0,\infty,\Omega} \leq c(\varepsilon) h^{2-\varepsilon},$$

and if $w(d)$ is in $W^{2,\infty}(\Omega)$

$$(2.23) \quad \|w(d) - \hat{w}_h(d)\|_{0,\infty,\Omega} \leq ch^2 |\ln h|.$$

The analogue of Proposition 2.4 with $\hat{w}_h$ remains still valid.

Here, we have limited our study to the approximation with piecewise linear elements. When we take piecewise degree 2 elements, we can improve the estimates (2.10)-(2.12), (2.16) and (2.18) as long as $w(d)$ is in $H^3(\Omega)$. If Ω is a polygonal domain, this is the case only if Ω is a rectangle or a triangle with angles inferior to $\frac{\pi}{2}$ (see [7]). If Ω is a regular domain (for example of class C^4), then $w(d)$ is in $W^{3,p}(\Omega)$ for $p \in [1, \infty)$ (see [11]) : this result is optimal as $(w(d) - d)^+$ is at best in $W^{1,p}(\Omega)$ if d is greater than 0. Then taking isoparametric elements of degree 2, we can improve the estimates in the case without numerical integration along the same lines.

3. APPROXIMATION OF REGULAR SOLUTIONS OF NONLINEAR PROBLEMS

We follow the method developed in [4] and applied in [6]. The key point in the proof of Proposition 2.3, is the use of a result on the approximation of regular solutions of nonlinear problems. In this section, we give an overlook of the method in a general framework.

We consider X, Z two Banach spaces and F a continuous map, $F : (l, x) \in \mathbb{R} \times X \mapsto F(l, x) \in Z$. We assume that the problem

$$F(l, x) = 0 , \tag{3.1}$$

has a regular solution branch $\{(l, x(l)) : l \in L\}$ with $L \subset \mathbb{R}$ compact. This means that at every pair $(l, x(l)) \in L \times X$ solution of (3.1), F is strongly partially differentiable with respect to x (i.e. for all $\varepsilon > 0$ there are $\delta > 0$ and a neighborhood N of l such that $\|x_1 - x(l)\|_X < \delta$, $\|x_2 - x(l)\|_X < \delta$, $e \in N$ imply

$$\|F(e, x_1) - F(e, x_2) - D_x F(l, x(l))(x_1 - x_2)\|_Z \leq \varepsilon \|x_1 - x_2\|_X ,$$

cf. [9]) and $D_x F(l, x(l))$ is an isomorphism.

We introduce a positive parameter $0 < h < 1$ which will tend to zero, closed subspaces $X_h \subset X$, $Z_h \subset Z$ and a family of continuous maps, $F_h : \mathbb{R} \times X_h \to Z_h$. We consider the problem to find $(l, x_h) \in \mathbb{R} \times X_h$ such that

$$F_h(l, x_h) = 0 . \tag{3.2}$$

With each regular point $(l, x(l))$ of the solution branch, we associate a point $\tilde{x}_h(l) \in X_h$ such that $F_h(l, \tilde{x}_h(l))$ is small and the difference $x(l) - \tilde{x}_h(l)$ can be easily studied. We assume that for $l \in L$, $0 < h \leq \tilde{h}$ for some $\tilde{h} > 0$,

(3.3) there exists $B(l, h) \in \mathscr{L}(X_h ; Z_h)$ which is an isomorphism,

(3.4) there exists a function $\alpha \in \mathbb{R}^+ \mapsto L_h^l(\alpha) \in \mathbb{R}^+$ such that

$$\|F_h(l, w) - F_h(l, v) - B(l, h)(w - v)\|_Z \leq L_h^l(\alpha) \|w - v\|_X \qquad \forall v, w \in B_h^l(\alpha) \cap X_h ,$$

where the ball $B_h^l(\alpha) = \left\{x \in X : \|x - \tilde{x}_h(l)\|_X \leq \alpha\right\}$. Let us introduce the following notations

$$\varepsilon_h(l) = \|F_h(l, \tilde{x}_h(l))\|_Z , \tag{3.5}$$

$$\gamma_h(l) = \|B(l, h)^{-1}\|_{\mathscr{L}(Z_h ; X_h)} . \tag{3.6}$$

PROPOSITION 3.1 : *We suppose that the mapping $\tilde{x}_h : l \in L \mapsto \tilde{x}_h(l) \in X_h$ is continuous and*

$$\sup_{0 < h \leq \tilde{h}} \left(\max_{l \in L} \gamma_h(l)\right) = \gamma \quad \textit{for some} \quad \tilde{h} > 0 , \tag{3.7}$$

$$\lim_{h \to 0} \left(\max_{l \in L} \varepsilon_h(l) \right) = 0 . \tag{3.8}$$

Moreover we assume (3.3), (3.4) *and in addition for all* $\zeta > 0$, *there are* α_ζ, h_ζ *with*

$$\max_{l \in L} L_h^l(\alpha) \leqslant \zeta \quad \forall h < h_\zeta , \quad \alpha < \alpha_\zeta . \tag{3.9}$$

Then there exist $h_0 > 0$, $\alpha_0 > 0$ *and a continuous map* $x_h : l \in L \mapsto x_h(l) \in X_h$ *such that for all* $h \leqslant h_0$, $l \in L$

$$(F_h(l, x) = 0, x \in B_h^l(\alpha_0) \cap X_h) \Leftrightarrow (x = x_h(l)) . \tag{3.10}$$

Moreover we have the estimate

$$\| x_h(l) - x \|_X \leqslant 2 \gamma \| F_h(l, x) \|_Z \quad \forall x \in B_h^l(\alpha_0) \cap X_h . \tag{3.11}$$

The proof of Proposition 3.1 can be obtained along the same lines as in [4] with minor modifications.

4. PROOFS

Throughout this section, we shall refer to the notation introduced in the previous one. Before starting with the proof of Proposition 2.3, let us make some comments. When we handle with the mapping S_h, we can simplify somehow the theory in Section 3, because then it suffices to consider F_h defined on $\mathbb{R} \times X$ with values in Z. This will be no more the case with the mapping $\hat{S}_h$. As we need to differentiate the mappings S and S_h, the following result will be helpful.

LEMMA 4.1 : *Let* $q > 2$ *be an integer,* Y *be the space* $L^q(\Omega)$ *or* $C^0(\bar{\Omega})$, ψ *be a function in* Y *and* γ *be a real number. We set*

$$\Omega_p = \{ x \in \Omega : \psi(x) > \gamma \} , \quad \Gamma = \{ x \in \Omega : \psi(x) = \gamma \}$$

and χ *the characteristic function of* Ω_p. *We assume the* $\mathbb{R}^2$*-measure of* Γ *is zero.*

Then for all $\zeta > 0$, *there exist positive constants* e_0 *and* δ_0 *such that for all* e, e^* *with* $|e - \gamma| \leqslant e_0$, $|e^* - \gamma| \leqslant e_0$, *for all* u, $u^* \in Y$ *with* $\| u - \psi \|_Y \leqslant \delta_0$, $\| u^* - \psi \|_Y \leqslant \delta_0$, *we have*

$$\| (u-e)^+ - (u^* - e^*)^+ + (e - e^*)\chi - \chi(u - u^*)\|_{0,\Omega} \leqslant$$
$$\leqslant \zeta(|e - e^*| + \|u - u^*\|_Y) .$$

The proof of Lemma 4.1 is given in [2] when $Y = L^q(\Omega)$; when $Y = C^0(\bar{\Omega})$, it is an easy variant.

We consider the particular case $X = \mathbb{R} \times H^1(\Omega)$, $Z = \mathbb{R} \times H^1(\Omega)$, $l = d$, $x = (\lambda, w)$ and $F = S$. From Theorem 1.2, we know that Problem 1.1 has a regular solution branch $\{(d, \lambda(d), w(d)) : d \in [0, \bar{d}]\}$. S is strongly partially differentiable with respect to (l, v) and referring to Lemma 4.1, we have

$$\text{(4.1)} \quad D_{l,v} S(d, \lambda(d), w(d))(\delta, u) \equiv B(d)(\delta, u) =$$
$$= \begin{pmatrix} \lambda(d) \int_\Omega \chi_d u \, dx + \delta \int_\Omega (w(d) - d)^+ \, dx \\ u - \lambda(d) T\chi_d u - \delta T(w(d) - d)^+ \end{pmatrix}, \quad \forall (\delta, u) \in \mathbb{R} \times H^1(\Omega)$$

which is an isomorphism, (see [2] for details); here χ_d denotes the characteristic function of the set

$$\text{(4.2)} \qquad \Omega_p(d) = \{x \in \Omega : w(d)(x) > d\} .$$

Proof of Proposition 2.3: The relation (2.9) will be an immediate consequence of Proposition 3.1 and the error estimates will be obtained from (3.11) and the classical error estimates for approximation of linear problems.

Let us consider with the parameter h, $0 < h < 1$, the spaces $X_h = \mathbb{R} \times H^1(\Omega)$, $Z_h = \mathbb{R} \times H^1(\Omega)$ and the mapping $F_h = S_h$ defined in (2.7). To each regular point $(d, \lambda(d), w(d))$ of the solution branch, we associate the point $(\tilde{\lambda}_h(d), \tilde{w}_h(d)) = (\lambda(d), \lambda(d) T_h(w(d) - d)^+)$. For $d \in [0, \bar{d}]$ and $h \in (0, 1)$, we define $B(d, h) \in \mathscr{L}(\mathbb{R} \times H^1(\Omega);$ $\mathbb{R} \times H^1(\Omega))$ by $B(d, h) = B(d)$ with $B(d)$ given in (4.1).

Now to apply Proposition 3.1, we have to check (3.3), (3.4), (3.7), (3.8) and (3.9). The hypothesis (3.3) is verified since we know from Theorem 1.2 that $B(d)$ is an isomorphism, while (3.7) is an immediate consequence of the fact that $B(d, h)$ is independent of h. The continuity of the mapping S and the fact, $\lim_{h \to 0} \| T - T_h \|_{\mathscr{L}(H^1(\Omega), H^1(\Omega))} = 0$, imply (3.8). Let us check now (3.4) and (3.9). Using Lemma 4.1 and the result,

$$\lim_{h \to 0} \| T - T_h \|_{\mathscr{L}(L^2(\Omega); H^1(\Omega))} = 0 ,$$

we easily check that for $d \in [0, \bar{d}]$ and for all real $\zeta > 0$, there are positive numbers $\bar{h}$, $\bar{\xi}$ and $\bar{\delta}$ such that for all $h \leq \bar{h}$, for all l, l^* with $|l - \lambda(d)| \leq \bar{\xi}$, $|l^* - \lambda(d)| \leq \bar{\xi}$ and for all u, $u^* \in H^1(\Omega)$ with $\|u - w(d)\|_{1,\Omega} \leq \bar{\delta}$, $\|u^* - w(d)\|_{1,\Omega} \leq \bar{\delta}$, we have

$$
(4.3) \quad \|S_h(d, l, u) - S_h(d, l^*, u^*) - B(d)(l - l^*, u - u^*)\| \leq \leq \zeta \{ |l - l^*| + \|u - u^*\|_{1,\Omega} \}.
$$

A detailed proof of (4.3) is given in [2]. As

$$
\lim_{h \to 0} \|\tilde{w}_h(d) - w(d)\|_{1,\Omega} = 0 ,
$$

we can choose $h_\zeta \leq \bar{h}$ such that (4.3) is true for all $h \leq h_\zeta$ and all (l, u), (l^*, u^*) in $B_h^d(\alpha_\zeta) =$

$$
= \left\{ (l, u) \in \mathbb{R} \times H^1(\Omega) : \left| l - \tilde{\lambda}_h(d) \right| + \|u - \tilde{w}_h(d)\|_{1,\Omega} \leq \alpha_\zeta \right\},
$$

with $\alpha_\zeta = \min \left(\bar{\xi}, \dfrac{\bar{\delta}}{2} \right)$. With this and the continuity of $\tilde{w}_h$, we can easily construct a function $L_h^d(\alpha)$ satisfying (3.4) and (3.9). The equivalence (3.10) implies (2.9). Hereafter c denotes a generic constant independent of h and d. In our case, Inequality (3.11) reads

$$
(4.4) \qquad |\lambda(d) - \lambda_h(d)| + \|\tilde{w}_h(d) - w_h(d)\|_{1,\Omega} \leq c \|S_h(d, \lambda(d), \tilde{w}_h(d))\|.
$$

Remark that we can develop the right-hand side of (4.4) in the following way

$$
\|S_h(d, \lambda(d), \tilde{w}_h(d))\| = \lambda(d) \left| \int_\Omega [(\tilde{w}_h(d) - d)^+ - (w(d) - d)^+] \, dx \right| + \lambda(d) \|T_h[(w(d) - d)^+ - (\tilde{w}_h(d) - d)^+]\|_{1,\Omega} .
$$

Using the fact that $\|T_h\|_{\mathscr{L}(L^2(\Omega); H^1(\Omega))}$ is bounded independently of h, we obtain

$$
(4.5) \quad \|S_h(d, \lambda(d), \tilde{w}_h(d))\| \leq c \|(\tilde{w}_h(d) - d)^+ - (w(d) - d)^+\|_{0,\Omega} \leq c \|\tilde{w}_h(d) - w(d)\|_{0,\Omega} .
$$

Introducing the classical error estimates for linear problems (*cf.* [3], [4] for instance), (4.5) implies

$$\| S_h(d, \lambda(d), \tilde{w}_h(d)) \| \leqslant ch^2 . \tag{4.6}$$

Finally the estimates (2.10)-(2.13) are a consequence of (4.6) and the classical error estimates for linear problems. □

We skip the proof of Proposition 2.4, which can be found in [2] or [1]. To prove Proposition 2.5, it does not suffice to adapt the proof of Proposition 2.3, because the mapping $\hat{S}_h$ must be defined on continuous functions, while S_h could be defined on $\mathbb{R} \times H^1(\Omega)$. Moreover the choice of the point $(\tilde{\lambda}_h(d), \tilde{w}_h(d))$ cannot be $(\lambda(d), \lambda(d)\, \hat{T}_h(w(d) - d)^+)$ because $(w(d) - d)^+$ is in $W^{1,p}(\Omega)$, $1 \leqslant p < \infty$, but not in $W^{2,1}(\Omega)$, which is needed to obtain optimal estimates.

We consider now the case $X = \mathbb{R} \times C^0(\bar{\Omega})$, $Z = \mathbb{R} \times C^0(\bar{\Omega})$. As the operator $T \in \mathscr{L}(L^2(\Omega) ; C^0(\bar{\Omega}))$ is compact, the operator $B(d) \in \mathscr{L}(\mathbb{R} \times C^0(\bar{\Omega}) ; \mathbb{R} \times C^0(\bar{\Omega}))$ is an isomorphism and $\{(d, \lambda(d), w(d)) : d \in [0, \bar{d}]\}$ is a regular solution branch of the problem to find for the value d of the parameter, $(\lambda, w) \in \mathbb{R} \times C^0(\bar{\Omega})$ such that $S(d, \lambda, w) = 0$.

Proof of Proposition 2.5 : Hereafter c denotes a generic constant independent of h and d; to specify a constant, we use c indexed by an integer. Let us consider with the parameter h, $0 < h < 1$, the spaces $X_h = \mathbb{R} \times V_h$, $Z_h = \mathbb{R} \times V_h$, remember that here V_h is provided with the norm $\|.\|_{0, \infty, \Omega}$, and the mapping $F_h = \hat{S}_h$ defined in (2.8). To each regular solution $(d, \lambda(d), w(d))$ we associate the point

$$(\tilde{\lambda}_h(d), \tilde{w}_h(d)) = (\lambda_h(d), w_h(d)) ,$$

where the functions $\lambda_h(\,.\,)$, $w_h(\,.\,)$ are the ones given in Proposition 2.3. For $d \in [0, \bar{d}]$ and $h \in (0, 1)$, we define the operator $B(d, h) \in \mathscr{L}(\mathbb{R} \times V_h ; \mathbb{R} \times V_h)$ by

$$B(d, h)(\delta, u) = \tag{4.7}$$

$$= \begin{pmatrix} \lambda_h(d) \displaystyle\int_\Omega \chi_{hd}\, u\, dx + \delta \int_\Omega (w_h(d) - d)^+ \, dx \\ u - \lambda_h(d)\, T_h\, \chi_{hd}\, u - \delta T_h(w_h(d) - d)^+ \end{pmatrix} \quad \forall (\delta, u) \in \mathbb{R} \times V_h ;$$

here χ_{hd} denotes the characteristic function of the set

$$\Omega_p(h, d) = \{x \in \Omega : w_h(d)(x) > d\} . \tag{4.8}$$

Remark that with Lemma 4.1 and the equality (2.17), we show that for h small enough $B(d, h) = D_{l,v} S_h(d, \lambda_h(d), w_h(d))$. To apply Proposition 3.1, we shall check first (3.3) and (3.7), secondly (3.8) and thirdly (3.4) and (3.9).

1° Let us show that there exists $\tilde{h} > 0$ such that for $d \in [0, \bar{d}]$, $0 < h \leqslant \tilde{h}$, $B(d, h)$ is an isomorphism and (3.7) is satisfied. Remark that we can consider $B(d, h) \in \mathscr{L}(\mathbb{R} \times C^0(\bar{\Omega}) ; \mathbb{R} \times C^0(\bar{\Omega}))$ and write

$$B(d, h) = B(d)(I + A(h, d)) \tag{4.9}$$

with $A(h, d) \in \mathscr{L}(\mathbb{R} \times C^0(\bar{\Omega}) ; \mathbb{R} \times C^0(\bar{\Omega}))$. Using the estimates (2.10), (2.11), (2.12), (2.16) and the property $\lim\limits_{h \to 0} \|T - T_h\|_{\mathscr{L}(L^2(\Omega) ; C^0(\bar{\Omega}))} = 0$, we easily show that $\lim\limits_{h \to 0} \|A(d, h)\| = 0$. Let us recall that $B(d) \in \mathscr{L}(\mathbb{R} \times C^0(\bar{\Omega}) ; \mathbb{R} \times C^0(\bar{\Omega}))$ is an isomorphism. There is a $\tilde{h} > 0$ such that for all $0 < h \leqslant \tilde{h}$, $B(d, h)$ is an isomorphism on $\mathbb{R} \times C^0(\bar{\Omega})$ and the norm of $A(d, h)$ is bounded by $1/2$; in particular, we have

$$\|B(d, h)^{-1}\| \leqslant \|(I + A(d, h))^{-1}\| \, \|B(d)^{-1}\| \leqslant 2 \|B(d)^{-1}\| , \tag{4.10}$$

where $\|.\|$ denotes the norm in $\mathscr{L}(\mathbb{R} \times C^0(\bar{\Omega}) ; \mathbb{R} \times C^0(\bar{\Omega}))$. As a consequence, Assumptions (3.3) and (3.7) are verified.

2° We check now (3.8). By the definition (3.5) of ε_h, we have

$$\begin{aligned} \varepsilon_h(d) &= \left\| \hat{S}_h(d, \lambda_h(d), w_h(d)) \right\|_{\mathbb{R} \times C^0(\bar{\Omega})} \\ &= \lambda_h(d) \left| ((w_h(d) - d)^+, 1)_{\hat{h}} - ((w_h(d) - d)^+, 1)_h \right| \\ &\quad + \lambda_h(d) \left\| T_h(w_h(d) - d)^+ - \hat{T}_h(w_h(d) - d)^+ \right\|_{0, \infty, \Omega} . \end{aligned} \tag{4.11}$$

We analyze separately both terms of the right-hand side of (4.11). Let us consider the partition of $\mathcal{T}_h = \mathcal{T}_{1h} \cup \mathcal{T}_{2h} \cup \mathcal{T}_{3h}$ given by

$$\begin{aligned} \mathcal{T}_{1h} &= \{K \in \mathcal{T}_h : w_h(d) - d \geqslant 0 \text{ in } K\}, \\ \mathcal{T}_{2h} &= \{K \in \mathcal{T}_h : w_h(d) - d \text{ changes sign in } K\}, \\ \mathcal{T}_{3h} &= \{K \in \mathcal{T}_h : w_h(d) - d < 0 \text{ in } K\}. \end{aligned}$$

Using the maximum principle, the fact that $\{x \in \Omega : w(d)(x) = d\}$ is a Jordan curve (these arguments are developed in [2], Lemmas 4.6 and 4.7) and the estimate (2.12), we can show the existence of $\bar{h}$ and a constant c such that for $h \leqslant \bar{h}$,

$$\text{(4.12)} \quad \text{meas}\left(\bigcup_{K \in \mathcal{T}_{2h}} K \equiv \Omega_{2h}\right) \leqslant ch\,, \quad \| w_h(d) - d \|_{0,\infty,\Omega_{2h}} \leqslant ch\,.$$

With the partition of $\mathcal{T}_h$ given above, we can develop

$$\text{(4.13)} \quad |((w_h(d) - d)^+, 1)_{\hat{h}} - ((w_h(d) - d)^+, 1)_h|$$
$$\leqslant \sum_{K \in \mathcal{T}_{1h}} \left| \frac{s_K}{3} \sum_{i=1}^{3} (w_h(d) - d)(a_{iK}) - \int_K (w_h(d) - d)\, dx \right|$$
$$+ \sum_{K \in \mathcal{T}_{2h}} \left| \frac{s_K}{3} \sum_{i=1}^{3} (w_h(d) - d)^+ (a_{iK}) + \int_K (w_h(d) - d)^+ \, dx \right| \leqslant ch^2,$$

when we use (4.12) and the fact that numerical integration is exact for polynomials of degree 1. To estimate the last term in (4.11), we introduce some notations ; if we denote by

$$u_h = T_h(w_h(d) - d)^+\,, \quad \hat{u}_h = \hat{T}_h(w_h(d) - d)^+\,,$$

then the functions u_h, $\hat{u}_h$ satisfy

$$\text{(4.14)} \qquad \int_\Omega \nabla u_h \, \nabla v_h \, dx = ((w_h(d) - d)^+, v_h)_h\,,$$

$$\text{(4.15)} \qquad \int_\Omega \nabla \hat{u}_h \, \nabla v_h \, dx = ((w_h(d) - d)^+, v_h)_{\hat{h}}\,, \quad \forall v_h \in V_h\,.$$

Substracting (4.15) from (4.14), we obtain

$$\text{(4.16)} \qquad \int_\Omega \nabla (u_h - \hat{u}_h) \, \nabla v_h \, dx = E((w_h(d) - d)^+, v_h)\,,$$

where we have used the notation

$$E((w_h(d) - d)^+, v_h) = ((w_h(d) - d)^+, v_h)_h - ((w_h(d) - d)^+, v_h)_h\,.$$

We plug $v_h = u_h - \hat{u}_h$ in (4.16) and obtain

$$\text{(4.17)} \qquad |u_h - \hat{u}_h|^2_{1,\Omega} = E((w_h(d) - d)^+, u_h - \hat{u}_h)\,.$$

If $E_K(.,.)$ denotes the same error as $E(.,.)$ but on the triangle K, then we can write

$$\text{(4.18)} \quad E((w_h(d) - d)^+, u_h - \hat{u}_h) = \sum_{K \in \mathcal{T}_{1h}} E_K(w_h(d) - d, u_h - \hat{u}_h)$$
$$+ \sum_{K \in \mathcal{T}_{2h}} E_K((w_h(d) - d)^+, u_h - \hat{u}_h)\,.$$

Since the integration rule is exact for polynomials of degree 1 and $w_h(d) - d$, $u_h - \hat{u}_h$ are piecewise polynomials of degree 1, we have

$$(4.19) \qquad \sum_{K \in \mathcal{C}_{1h}} E_K(w_h(d) - d, u_h - \hat{u}_h) \leqslant ch^2 \| w_h(d) - d \|_{1,\Omega} \| u_h - \hat{u}_h \|_{1,\Omega} .$$

Applying the inverse inequality

$$(4.20) \qquad \| v_h \|_{0,\infty,\Omega} \leqslant c \, |\ln h|^{1/2} \| v_h \|_{1,\Omega} ,$$

and (4.12) we can bound

$$(4.21) \qquad \sum_{K \in \mathcal{C}_{2h}} E_K((w_h(d) - d)^+, u_h - \hat{u}_h) \leqslant ch^2 |\ln h|^{1/2} \| u_h - \hat{u}_h \|_{1,\Omega} .$$

So from (4.21) and (4.19), a bound of (4.17) is obtained

$$(4.22) \qquad \| u_h - \hat{u}_h \|_{1,\Omega} \leqslant ch^2 |\ln h|^{1/2} .$$

Finally with (4.13), (4.20) and (4.22), we can overestimate (4.11) in the following way

$$(4.23) \qquad \varepsilon_h(d) \leqslant ch^2 |\ln h| .$$

As a consequence, the assumption (3.8) is verified.

3° Now we have to check (3.4) and (3.9). To this end, we shall show that for $d \in [0, \bar{d}]$ and for all $\zeta > 0$, there are h_ζ, α_ζ with

$$(4.24) \qquad \| \hat{S}_h(d, l, u) - \hat{S}_h(d, l^*, u^*) - B(d, h)(l - l^*, u - u^*) \|_{\mathbb{R} \times C^0(\bar{\Omega})} \leqslant \\ \leqslant \zeta \{ |l - l^*| + \| u - u^* \|_{0,\infty,\Omega} \},$$

for $h \leqslant h_\zeta$ and (l, u), $(l^*, u^*) \in B_h^d(\alpha_\zeta)$. Let $\zeta > 0$ be fixed. In the same way we have proved (4.3), we can show there are positive numbers h_1, ξ_1 and δ_1 such that for all $h \leqslant h_1$, for all l, l^* with $|l - \lambda(d)| \leqslant \xi_1$, $|l^* - \lambda(d)| \leqslant \xi_1$ and for all u, $u^* \in C^0(\bar{\Omega})$ with $\| u - w(d) \|_{0,\infty,\Omega} \leqslant \delta_1$, $\| u^* - w(d) \|_{0,\infty,\Omega} \leqslant \delta_1$ we have

$$(4.25) \qquad \| S_h(d, l, u) - S_h(d, l^*, u^*) - B(d, h)(l - l^*, u - u^*) \|_{\mathbb{R} \times C^0(\bar{\Omega})} \leqslant \\ \leqslant \frac{\zeta}{2} \{ |l - l^*| + \| u - u^* \|_{0,\infty,\Omega} \} .$$

We analyze now

(4.26)

$$\left\| \hat{S}_h(d, l, u) - \hat{S}_h(d, l^*, u^*) - S_h(d, l, u) + S_h(d, l^*, u^*) \right\|_{\mathbb{R} \times C^0(\bar{\Omega})} =$$

$$= \left| E(l(u-d)^+ - l^*(u^*-d)^+, 1) \right| + \left\| (T_h - \hat{T}_h)(l(u-d)^+ - l^*(u^*-d)^+) \right\|_{0,\infty,\Omega},$$

where $E(.,.)$ has been defined in (4.16). Before developing both terms of the right-hand side of (4.26), let us recall a useful result. For $\varepsilon > 0$, let us define the set $A(\varepsilon) = \{x \in \Omega : |w(d)(x) - d| \leqslant \varepsilon\}$; then there exists a constant c_1 such that

$$\text{meas } A(\varepsilon) \leqslant c_1 \varepsilon . \tag{4.27}$$

This is a consequence of the maximum principle and the fact that $\{x \in \Omega : w(d)(x) = d\}$ is a Jordan curve, (*cf.* [2] for details). We consider now $\varepsilon_0 > 0$, which will be expressed later, and associate a partition of $\mathcal{T}_h = \mathcal{T}_{1h} \cup \mathcal{T}_{2h} \cup \mathcal{T}_{3h}$ given by

$$\mathcal{T}_{1h} = \left\{ K \in \mathcal{T}_h : w(d) - d \geqslant \frac{\varepsilon_0}{2} \text{ in } K \right\},$$

$$\mathcal{T}_{3h} = \left\{ K \in \mathcal{T}_h : w(d) - d \leqslant -\frac{\varepsilon_0}{2} \text{ in } K \right\},$$

$$\mathcal{T}_{2h} = \{ K \in \mathcal{T}_h : K \notin \mathcal{T}_{1h} \cup \mathcal{T}_{3h} \}.$$

Let us choose now $h_2 \leqslant h_1$ such that for $h \leqslant h_2$

$$\bigcup_{K \in \mathcal{T}_{2h}} K \subset A(\varepsilon_0) . \tag{4.28}$$

We choose $\delta_2 = \min\left(\delta_1, \frac{\varepsilon_0}{2}\right)$ and consider l, l^* with $|l - \lambda(d)| \leqslant \xi_1$, $|l^* - \lambda(d)| \leqslant \xi_1$, $u, u^* \in V_h$ with

$$\|u - w(d)\|_{0,\infty,\Omega} \leqslant \delta_2 , \quad \|u^* - w(d)\|_{0,\infty,\Omega} \leqslant \delta_2 ,$$

then

$$\begin{aligned} (4.29) \quad & |E(l(u-d)^+ - l^*(u^*-d)^+, 1)| \leqslant \\ & \leqslant |l - l^*| \, |E((u-d)^+, 1)| + |l^*| \, |E((u-d)^+ - (u^*-d)^+, 1)| \\ & \leqslant |l - l^*| \, 2 c_1 \varepsilon_0 \|u - d\|_{0,\infty,\Omega} + |l^*| \, 2 c_1 \varepsilon_0 \|u - u^*\|_{0,\infty,\Omega} . \end{aligned}$$

On the other hand, the last term in (4.26) is developed in the following way

$$(4.30) \quad \left\|(T_h - \hat{T}_h)(l(u-d)^+ - l^*(u^*-d)^+)\right\|_{0,\infty,\Omega} \leq$$
$$\leq |l - l^*| \left\|(T_h - \hat{T}_h)(u-d)^+\right\|_{0,\infty,\Omega} + |l^*| \left\|(T_h - \hat{T}_h)((u-d)^+ - (u^*-d)^+)\right\|_{0,\infty,\Omega}.$$

Let us study the two terms of (4.30); hereafter f represents either $(u-d)^+$ or $(u-d)^+ - (u^*-d)^+$. We introduce the Galerkin approximation $G_h(.,.)$ in V_h to the Green function $G(.,.)$; we have for $x \in \Omega$

$$\int_\Omega \nabla G_h(x,y)\, \nabla v_h(y)\, dy = v_h(x) \quad \forall v_h \in V_h.$$

The following standard results will be useful,

$$(4.31) \qquad \|G_h(x,.)\|_{0,\Omega} \leq c, \quad \|G_h(x,.)\|_{1,\Omega} \leq c\,|\ln h|^{1/2}, \quad \text{for} \quad x \in \Omega.$$

Then with this function G_h, we have

$$\left\|(T_h - \hat{T}_h)\, f\right\|_{0,\infty,\Omega} = \max_{x \in \Omega} |E(f, G_h(x,.))|.$$

For $v_h \in V_h$ we develop likewise in (4.18)

$$|E(f, v_h)| \leq \left|\sum_{K \in \mathcal{T}_{1h}} E_K(f, v_h)\right| + \left|\sum_{K \in \mathcal{T}_{2h}} E_K(f, v_h)\right|$$
$$\leq ch^2 \|f\|_{1,\Omega}\, \|v_h\|_{1,\Omega} + 2(c_1\, \varepsilon_0)^{1/2} \|f\|_{0,\infty,\Omega}\, \|v_h\|_{0,\Omega},$$

which implies if v_h is $G_h(x,.)$

$$(4.32) \quad \left\|(T_h - \hat{T}_h)\, f\right\|_{0,\infty,\Omega} \leq c_2 \{h^2 |\ln h|^{1/2} + 2(c_1\, \varepsilon_0)^{1/2}\}\, \|f\|_{0,\infty,\Omega}.$$

We can express now ε_0 precisely; it has to be chosen according to (4.29), (4.30) and (4.32) such that

$$2\, c_1\, \varepsilon_0 \{\|w(d) - d\|_{0,\infty,\Omega} + \delta_1\} \leq \frac{\zeta}{4},$$

$$2\, c_1\, \varepsilon_0 \{\lambda(d) + \xi_1\} \leq \frac{\zeta}{4},$$

$$c_2 \{\varepsilon_0^2 |\ln \varepsilon_0|^{1/2} + 2(c_1\, \varepsilon_0)^{1/2}\} \{\|w(d) - d\|_{0,\infty,\Omega} + \delta_1\} \leq \frac{\zeta}{4},$$