

A PARTIAL DIFFERENTIAL EQUATION CONNECTED TO OPTION PRICING WITH STOCHASTIC VOLATILITY: REGULARITY RESULTS AND DISCRETIZATION

YVES ACHDOU, BRUNO FRANCHI, AND NICOLETTA TCHOU

ABSTRACT. This paper completes a previous work on a Black and Scholes equation with stochastic volatility. This is a degenerate parabolic equation, which gives the price of a European option as a function of the time, of the price of the underlying asset, and of the volatility, when the volatility is a function of a mean reverting Orstein–Uhlenbeck process, possibly correlated with the underlying asset. The analysis involves weighted Sobolev spaces. We give a characterization of the domain of the operator, which permits us to use results from the theory of semigroups. We then study a related model elliptic problem and propose a finite element method with a regular mesh with respect to the intrinsic metric associated with the degenerate operator. For the error estimate, we need to prove an approximation result.

1. INTRODUCTION

We consider a financial asset whose price is given by the stochastic differential equation

$$(1) \quad dX_t = \mu X_t dt + \sigma_t X_t dW_t,$$

where $\mu X_t dt$ is a drift term, (W_t) is a Brownian motion, and (σ_t) is the volatility. The simplest models (see [19] for a complete overview) take a constant volatility, but these models are generally too rough to match real prices. A more realistic model consists in assuming that (σ_t) is a function of a mean reverting Orstein–Uhlenbeck process:

$$(2) \quad \begin{aligned} \sigma_t &= f(Y_t), \\ dY_t &= \alpha(m - Y_t)dt + \beta d\hat{Z}_t, \end{aligned}$$

where α , m , and β are positive constants, and (\hat{Z}_t) is a linear combination of (W_t) and an independent Brownian motion (Z_t) :

$$(3) \quad \hat{Z}_t = \rho W_t + \sqrt{1 - \rho^2} Z_t.$$

The *correlation factor* ρ lies in $] -1, 1[$. In [9], the parameter α is called the *rate of mean reversion*, and $\frac{1}{\alpha}$ can be thought of as a *correlation time*. For convenience,

Received by the editor April 16, 2003 and, in revised form, March 3, 2004.

2000 *Mathematics Subject Classification*. Primary 35K65, 65M15, 65M60, 65N30.

Key words and phrases. Finance, degenerate parabolic operator, finite elements.

The second author was partially supported by University of Bologna, funds for selected research topics and by GNAMPA of INdAM, Italy, project “Analysis in metric spaces”.

we introduce the parameter ν

$$(4) \quad \nu^2 = \frac{\beta^2}{2\alpha}.$$

Consider a European derivative on this asset with expiration date T and payoff function $h(X_T)$. Its price at time t will depend on t , on the price of the underlying asset X_t , and on Y_t . We denote by $P(t, X_t, Y_t)$ the price of the derivative, and by $\tilde{r}(t)$ the interest rate. As explained in [9], by using the no arbitrage principle and the two-dimensional Itô's formula, it is possible to prove that P satisfies the partial differential equation

$$(5) \quad \frac{\partial P}{\partial t} + \frac{1}{2}f(y)^2x^2\frac{\partial^2 P}{\partial x^2} + \rho\beta xf(y)\frac{\partial^2 P}{\partial x\partial y} + \frac{1}{2}\beta^2\frac{\partial^2 P}{\partial y^2} + \tilde{r}(t)\left(x\frac{\partial P}{\partial x} - P\right) + \left(\alpha(m - y) - \beta\tilde{\Lambda}(t, x, y)\right)\frac{\partial P}{\partial y} = 0, \quad \begin{matrix} 0 \leq t < T, \\ x > 0, y \in \mathbb{R}, \end{matrix}$$

where

$$(6) \quad \tilde{\Lambda}(t, x, y) = \rho\frac{\mu - \tilde{r}(t)}{f(y)} + \sqrt{1 - \rho^2}\tilde{\gamma}(t, x, y),$$

and the function $\tilde{\gamma}(t, x, y)$ is used to model the contribution of the second source of randomness dZ_t to the market price of the volatility risk $\beta\tilde{\Lambda}(S, y, t)\frac{\partial P}{\partial y}$. Equation (5) is complemented with the terminal condition $P(T, x, y) = h(x)$. The function $\tilde{\gamma}(t, x, y)$ can be chosen arbitrarily. We will assume hereafter that $\tilde{\gamma}$ is a bounded function.

As in [9], we can group the differential operator in (5) as follows:

$$(7) \quad \underbrace{\frac{\partial P}{\partial t} + \frac{1}{2}f(y)^2x^2\frac{\partial^2 P}{\partial x^2} + \tilde{r}(t)\left(x\frac{\partial P}{\partial x} - P\right)}_{BS_{f(y)}} + \underbrace{\rho\beta xf(y)\frac{\partial^2 P}{\partial x\partial y}}_{\text{correlation}} + \underbrace{\frac{1}{2}\beta^2\frac{\partial^2 P}{\partial y^2} + \alpha(m - y)\frac{\partial P}{\partial y}}_{\text{Orstein-Uhlenbeck}} - \underbrace{\beta\tilde{\Lambda}(t, x, y)\frac{\partial P}{\partial y}}_{\text{premium}}.$$

In the premium term, the term $\frac{\mu - \tilde{r}(t)}{f(y)}$ is the *excess return to risk ratio*.

Remark 1. For options on stocks, typical values for m and \tilde{r} are 0.2 and 0.05, and generally $\alpha \gg 1$, while $\nu^2 = \frac{\beta^2}{2\alpha}$ (the variance of the invariant distribution of Y) is not too large; see [9].

There remains to choose the function f . In [18], Stein and Stein have considered the case when

$$(8) \quad f(y) = |y|.$$

In [1], a variational approach has been studied for (7), in both the uncorrelated case ($\rho = 0$) and the correlated case ($\rho \neq 0$). Here we will focus on the case when $\rho = 0$. Some of the results obtained in [1] will be recalled here. In particular, the following change of unknown function was employed

$$(9) \quad u(t, x, y) = P(T - t, x, y)e^{-(1-\eta)\frac{(y-m)^2}{2\nu^2}},$$

where η is a parameter such that $0 < \eta < 1$, because it can be seen very easily that if $\tilde{r} = 0$, $\tilde{\Lambda} = 0$, and $\rho = 0$, then the function $e^{\frac{(y-m)^2}{2\nu^2}}$ satisfies (5). If P is a bounded

solution of (5), the function u should decay rapidly when y goes to infinity: in [1] u was looked for in a weighted Sobolev space (weak solution), and it was proved that the function u obtained by this method has the correct behavior at infinity, so that the function $u(T - t, x, y)e^{(1-\eta)\frac{(y-m)^2}{2v^2}}$ is a bounded solution of (5).

With the notations $r(t) = \tilde{r}(T - t)$, $\gamma(t) = \tilde{\gamma}(T - t)$, and $\Lambda(t) = \tilde{\Lambda}(T - t)$ ($= \gamma(t)$ because $\rho = 0$), the new unknown u satisfies the degenerate parabolic partial differential equation

$$(10) \quad \frac{\partial u}{\partial t} - A_t u = 0$$

with

$$(11) \quad \begin{aligned} A_t(v) = & \frac{1}{2}y^2x^2\frac{\partial^2v}{\partial x^2} + \frac{1}{2}\beta^2\frac{\partial^2v}{\partial y^2} + r(t)x\frac{\partial v}{\partial x} + ((1 - 2\eta)\alpha(y - m) - \beta\gamma(t, x, y))\frac{\partial v}{\partial y} \\ & - \left(r(t) + 2\frac{\alpha^2}{\beta^2}\eta(1 - \eta)(y - m)^2 + 2(1 - \eta)\frac{\alpha}{\beta}(y - m)\gamma(t, x, y) - \alpha(1 - \eta) \right) v. \end{aligned}$$

The paper is organized as follows. Section 2 is devoted to a regularity result for the elliptic problem with operator A_t ; we give a characterization of the domain of A_t , and we prove in particular that it is independent of t . In the next section, we show that the evolution equation involves a perturbation of an analytic semigroup, so we have a smoothing effect. In the fourth section, we propose a finite element discretization for a related degenerate elliptic problem in a bounded domain, with a simpler operator retaining yet the essential characteristics of the operator A_t . The essential feature of this method is that the employed meshes are not regular (in the sense of [4]) with respect to the Euclidean metric. Error estimates are then obtained: a key ingredient for that is the study of a local regularization operator, done in Section 5.

2. THE ELLIPTIC PROBLEM

In the rest of the paper, we use the notation $Q = \mathbb{R}_+ \times \mathbb{R}$. For $t \geq 0$, let us denote by A_t the unbounded operator in $L^2(Q)$ defined as follows: let r be a continuous positive function, γ a bounded measurable function, and finally let α, β, η , and m be positive constants. We set

$$(12) \quad \begin{aligned} A_t &= A_{1,t} + A_{2,t}, \\ A_{1,t}v &= \frac{1}{2}y^2x^2\frac{\partial^2v}{\partial x^2} + \frac{1}{2}\beta^2\frac{\partial^2v}{\partial y^2} + r(t)x\frac{\partial v}{\partial x} - 2\frac{\alpha^2}{\beta^2}\eta(1 - \eta)y^2v + \alpha(1 - 2\eta)y\frac{\partial v}{\partial y}, \\ A_{2,t}v &= -((1 - 2\eta)m + \beta\gamma(t, x, y))\frac{\partial v}{\partial y} \\ & - \left(r + 2\frac{\alpha^2}{\beta^2}\eta(1 - \eta)(-2my + m^2) + 2(1 - \eta)\frac{\alpha}{\beta}(y - m)\gamma(t, x, y) - \alpha(1 - \eta) \right) v, \end{aligned}$$

with domain $D_t = \{v \in V : A_tv \in L^2(Q)\}$, where the weighted Sobolev space V is defined by

$$(13) \quad V = \left\{ v : \left(\sqrt{1 + y^2}v, \frac{\partial v}{\partial y}, xy\frac{\partial v}{\partial x} \right) \in (L^2(Q))^3 \right\}.$$

Endowed with the norm

$$(14) \quad \|v\|_V = \left(\int_Q (1+y^2)|v|^2 + \left| \frac{\partial v}{\partial y} \right|^2 + x^2 y^2 \left| \frac{\partial v}{\partial x} \right|^2 \right)^{\frac{1}{2}},$$

V is a Hilbert space. We call V' the dual of V . As it is proved in [1], V enjoys the following properties:

- (1) V is separable;
- (2) Denoting by $\mathcal{D}(Q)$ the space of smooth and compactly supported real-valued functions in Q , $\mathcal{D}(Q) \subset V$ and $\mathcal{D}(Q)$ is dense in V ;
- (3) V is dense in $L^2(Q)$;
- (4) $\|\cdot\|_V$ defined by

$$(15) \quad \|v\|_V = \left(\int_Q v^2 + \left| \frac{\partial v}{\partial y} \right|^2 + x^2 y^2 \left| \frac{\partial v}{\partial x} \right|^2 \right)^{\frac{1}{2}}$$

is a norm in V , equivalent to $\|\cdot\|$.

It is clear that $A_{2,t}$ is a bounded operator from V to $L^2(Q)$, so that we can write also

$$D_t = \{v \in V : A_{1,t}v \in L^2(Q)\}.$$

The following results can be found in [1]

Lemma 1. *The operator $v \rightarrow \beta x \frac{\partial v}{\partial x}$ is continuous from V into V' .*

Proposition 1. *Assume that r is a bounded function of time and that γ is bounded by a constant. Then the operators $-A_t$, $t \geq 0$, are bounded linear operators from V into V' with norm uniformly bounded by a constant $L > 0$ independent of t .*

2.1. Estimates on the resolvent. In this section, we deal with complex-valued functions. The notation $L^2(Q)$ is still used for complex-valued functions whose modulus is square integrable. Remember that the scalar product of complex-valued functions in $L^2(Q)$ is the sesquilinear form defined by $\langle u, v \rangle \equiv \int_Q u \bar{v}$. We can extend also the definition of V to complex-valued functions, so that all the results stated above are still valid.

Proposition 2. *Assume that r is a bounded function of time and that γ is bounded by a constant Γ . Assume that $\alpha > \beta$, then there exist two positive constants C_0 and c_0 independent of t and two constants $0 < \eta_1 < \frac{1}{2} < \eta_2 < 1$ such that, for $\eta_1 < \eta < \eta_2$ and for any $v \in V$,*

$$(16) \quad \Re(-\langle A_t v, v \rangle) \geq C_0 \|v\|_V^2 - c_0 \|v\|_{L^2(Q)}^2.$$

Proof. By density, it is enough to prove the assertion when $v \in \mathcal{D}(Q)$. We can write

$$\langle -A_t v, v \rangle = \sum_{i=1}^6 \mathcal{I}_i,$$

with

$$\begin{aligned} \mathcal{I}_1 &= -\frac{1}{2}\langle x^2y^2\frac{\partial^2v}{\partial x^2}, v \rangle, \quad \mathcal{I}_2 = -\frac{\beta^2}{2}\langle \frac{\partial^2v}{\partial y^2}, v \rangle, \quad \mathcal{I}_3 = -r(t)\langle x\frac{\partial v}{\partial x}, v \rangle, \\ \mathcal{I}_4 &= \int_Q (-\alpha(1-2\eta)(y-m) + \beta\gamma(t,x,y))\frac{\partial v}{\partial y}\bar{v}, \\ \mathcal{I}_5 &= \int_Q \left(r(t) + 2(1-\eta)\frac{\alpha}{\beta}(y-m)\gamma(t,x,y) - \alpha(1-\eta) \right) |v|^2, \\ \mathcal{I}_6 &= 2\frac{\alpha^2}{\beta^2}\eta(1-\eta) \int_Q (y-m)^2|v|^2. \end{aligned}$$

Integrating by part, we have

$$\begin{aligned} \Re(\mathcal{I}_1) &= \frac{1}{2} \int_Q x^2y^2 \left| \frac{\partial v}{\partial x} \right|^2 + \Re \left(\int_Q xy^2 \frac{\partial v}{\partial x} \bar{v} \right) = \frac{1}{2} \int_Q x^2y^2 \left| \frac{\partial v}{\partial x} \right|^2 - \frac{1}{2} \int_Q y^2|v|^2, \\ \mathcal{I}_2 &= \frac{\beta^2}{2} \int_Q \left| \frac{\partial v}{\partial y} \right|^2, \\ \Re(\mathcal{I}_3) &= \frac{r(t)}{2} \int_Q |v|^2, \\ \Re(\mathcal{I}_4) &= \frac{\alpha}{2}(1-2\eta) \int_Q |v|^2 + \beta \Re \left(\int_Q \gamma(t,x,y) \frac{\partial v}{\partial y} \bar{v} \right) \\ &\geq \frac{\alpha}{2}(1-2\eta) \int_Q |v|^2 - \frac{1}{2}\beta\Gamma(\zeta\|v\|_{L^2(Q)}^2 + \frac{1}{\zeta}\|\frac{\partial v}{\partial y}\|_{L^2(Q)}^2), \end{aligned}$$

where ζ is an arbitrary positive number. Similarly,

$$\begin{aligned} \mathcal{I}_5 &= (r(t) - \alpha(1-\eta))\|v\|_{L^2(Q)}^2 + 2(1-\eta)\frac{\alpha}{\beta} \int_Q (y-m)\gamma(t,x,y)|v|^2 \\ &\geq (r(t) - \alpha(1-\eta) - 2(1-\eta)\frac{\alpha}{\beta}\Gamma(|m| + \frac{1}{2}\zeta))\|v\|_{L^2(Q)}^2 - (1-\eta)\frac{\alpha\Gamma}{\zeta\beta}\|yv\|_{L^2(Q)}^2, \end{aligned}$$

and

$$\mathcal{I}_6 \geq 2\frac{\alpha^2}{\beta^2}\eta(1-\eta) \left(\|yv\|_{L^2(Q)}^2 + m^2\|v\|_{L^2(Q)}^2 - |m|(\zeta\|v\|_{L^2(Q)}^2 + \frac{1}{\zeta}\|yv\|_{L^2(Q)}^2) \right).$$

If $\alpha > \beta$, it is possible to choose C_0 , $0 < C_0 < \frac{1}{2}$ and η_1, η_2 , $0 < \eta_1 < \eta_2 < 1$, such that for any η with $\eta_1 < \eta < \eta_2$ and for ζ large enough

$$2\frac{\alpha^2}{\beta^2}\eta(1-\eta)\left(1 - \frac{|m|}{\zeta}\right) - \frac{1}{2} - (1-\eta)\frac{\alpha\Gamma}{\zeta\beta} > C_0,$$

and at the same time

$$\frac{\beta}{2}\left(\beta - \frac{\Gamma}{\zeta}\right) > C_0.$$

Then the estimate (16) follows by fixing ζ . □

Remark 2. The assumption $\alpha > \beta$ says that the rate of mean reversion should not be too small compared to the asymptotic variance of the volatility. This assumption is usually verified in practice; see Remark 1.

From Propositions 1 and 2, we deduce

Proposition 3. *With the assumptions of Proposition 2, for all $t \in [0, T]$, the resolvent $R(\lambda : A_t) = (\lambda I - A_t)^{-1}$ of A_t exists for $\Re(\lambda) \geq c_0$ (where c_0 is defined in Proposition 2), and there exists a constant M independent of $t \geq 0$ such that*

$$(17) \quad \|R(\lambda : A_t)\|_{\mathcal{L}(L^2(Q))} \leq \frac{M}{|\lambda - c_0| + 1}.$$

Proof. If $\Re(\lambda) \geq c_0$, the operator $\lambda I - A_t$ is bounded from V to V' and satisfies the coercivity estimate

$$(18) \quad \Re(\langle (\lambda I - A_t)v, v \rangle) \geq C_0 \|v\|_V^2 + \Re(\lambda - c_0) \|v\|_{L^2(Q)}^2.$$

Therefore, for all $f \in L^2(Q)$, there exists a unique function $u \equiv R(\lambda : A_t)f \in V$ such that $(\lambda I - A_t)u = f$ and hence

$$(19) \quad \lambda \|u\|_{L^2(Q)} - \langle A_t u, u \rangle = \int_Q f \bar{u}.$$

By taking the real part of (19), we get

$$(20) \quad (C_0 + \Re(\lambda) - c_0) \|u\|_{L^2(Q)} \leq \|f\|_{L^2(Q)},$$

and

$$(21) \quad C_0 \|u\|_V^2 \leq \|f\|_{L^2(Q)} \|u\|_{L^2(Q)}.$$

The imaginary part of (19) leads to

$$(22) \quad \begin{aligned} |\Im(\lambda)| \|u\|_{L^2(Q)}^2 &= |\Im(\int_Q f \bar{u} + \langle A_t u, u \rangle)| \leq \|f\|_{L^2(Q)} \|u\|_{L^2(Q)} + L \|u\|_{L^2(Q)}^2 \\ &\leq (1 + \frac{L}{C_0}) \|f\|_{L^2(Q)} \|u\|_{L^2(Q)}, \end{aligned}$$

where the last estimate is a consequence of (21). From (20) and (22), we deduce

$$(23) \quad (|\Im(\lambda)| + |\Re(\lambda) - c_0|) \|u\|_{L^2(Q)} \leq (2 + \frac{L}{C_0}) \|f\|_{L^2(Q)},$$

which together with (21) yields the desired result. □

2.2. The regularity result for the elliptic problem. In this section we go back to real-valued functions. Let us start with a density result.

Lemma 2. *The space $\mathcal{D}(Q)$ is dense in D_t with the graph norm.*

Proof. The proof is done in two steps: a truncation, then a convolution by a mollifier. We first perform the change of variables $Q \rightarrow \mathbb{R}^2, (x, y) \mapsto (\xi, y) = (\log(x), y)$. We introduce the weighting function $w(\xi, y) = \exp(\xi)$. By this change of variables, the space V is mapped to $\tilde{V} = \{v : \sqrt{1 + y^2}v, \frac{\partial v}{\partial y}, y \frac{\partial v}{\partial \xi} \in L_w^2(\mathbb{R}^2)\}$. Likewise, calling \tilde{A} the operator,

$$\tilde{A}v = \frac{1}{2}y^2(\frac{\partial^2 v}{\partial \xi^2} - \frac{\partial v}{\partial \xi}) + \frac{1}{2}\beta^2 \frac{\partial^2 v}{\partial y^2} - r(t) \frac{\partial v}{\partial \xi} + 2\frac{\alpha^2}{\beta^2}\eta(1 - \eta)y^2v - \alpha(1 - 2\eta)y \frac{\partial v}{\partial y},$$

the domain of the operator is mapped to

$$\tilde{D} = \{v \in \tilde{V} : \tilde{A}v \in L_w^2(\mathbb{R}^2)\},$$

and the desired result is equivalent to the density of $\mathcal{D}(\mathbb{R}^2)$ in \tilde{D} supplied with its graph norm.

Let $\phi_R : \mathbb{R}^2 \rightarrow [0, 1]$ be a smooth function such that

- $\phi_R(\xi, y) = 0$, if $\sqrt{\xi^2 + y^2} \geq 2R$,
- $\phi_R(\xi, y) = 1$, if $\sqrt{\xi^2 + y^2} \leq R$,
- $\|\nabla\phi_R\|_\infty \leq \frac{C}{R}$, $\|D^2\phi_R\|_\infty \leq \frac{C}{R^2}$.

One can check that for any $v \in \tilde{V}$, $\lim_{R \rightarrow \infty} \|v - \phi_R v\|_{\tilde{V}} = 0$. We are also interested in estimating $\|\tilde{A}v - \tilde{A}(\phi_R v)\|_{L^2_w(\mathbb{R}^2)}$ for $v \in \tilde{D}$. We have that

$$(24) \quad \begin{aligned} \tilde{A}(\phi_R v) - \tilde{A}v &= (\phi_R - 1)\tilde{A}v + y^2 \frac{\partial v}{\partial \xi} \frac{\partial \phi_R}{\partial \xi} + \beta^2 \frac{\partial v}{\partial y} \frac{\partial \phi_R}{\partial y} \\ &+ v \left(\frac{1}{2} y^2 \left(\frac{\partial^2 \phi_R}{\partial \xi^2} - \frac{\partial \phi_R}{\partial \xi} \right) + \frac{\beta^2}{2} \frac{\partial^2 \phi_R}{\partial y^2} + r(t) \frac{\partial \phi_R}{\partial \xi} - \alpha(1 - 2\eta)y \frac{\partial \phi_R}{\partial y} \right). \end{aligned}$$

It can be proved easily that the $L^2_w(\mathbb{R}^2)$ norm of each term in the right-hand side of (24) vanishes as $R \rightarrow \infty$. For the sake of brevity, we prove this result only for the term $y^2 \frac{\partial v}{\partial \xi} \frac{\partial \phi_R}{\partial \xi}$ (the proof for the other terms is similar). We have that $y \frac{\partial v}{\partial \xi} \in L^2_w(\mathbb{R}^2)$ and that $|y \frac{\partial \phi_R}{\partial \xi}|$ is a function bounded by C and with support in the ring $R \leq \sqrt{\xi^2 + y^2} \leq 2R$. The product $y^2 \frac{\partial v}{\partial \xi} \frac{\partial \phi_R}{\partial \xi}$ tends to zero in $L^2_w(\mathbb{R}^2)$ norm as R tends to ∞ by Lebesgue theorem.

So we have proved that the function in \tilde{D} with compact supports are dense in \tilde{D} .

The next part of the proof follows the arguments of Friedrichs [14]. We take a function v with compact support and such that $v \in \tilde{D}$, or equivalently, $v \in \hat{D}$,

$$\hat{D} = \left\{ v : v, \frac{\partial v}{\partial y}, y \frac{\partial v}{\partial \xi}, \hat{A}v \in L^2(\mathbb{R}^2) \right\},$$

with

$$\hat{A}v = y^2 \frac{\partial^2 v}{\partial \xi^2} + \beta^2 \frac{\partial^2 v}{\partial y^2} - 2r(t) \frac{\partial v}{\partial \xi},$$

because v has a compact support. The space \hat{D} is supplied with the norm $\|\cdot\|_{\hat{D}}$: $\|v\|_{\hat{D}}^2 = \|v\|_{L^2(\mathbb{R}^2)}^2 + \|\frac{\partial v}{\partial y}\|_{L^2(\mathbb{R}^2)}^2 + \|y \frac{\partial v}{\partial \xi}\|_{L^2(\mathbb{R}^2)}^2 + \|\hat{A}v\|_{L^2(\mathbb{R}^2)}^2$. We also introduce the space $\hat{V} = \{v : v, \frac{\partial v}{\partial y}, y \frac{\partial v}{\partial \xi} \in L^2(\mathbb{R}^2)\}$, supplied with the norm $\|v\|_{\hat{V}} = (\|v\|_{L^2(\mathbb{R}^2)}^2 + \|\frac{\partial v}{\partial y}\|_{L^2(\mathbb{R}^2)}^2 + \|y \frac{\partial v}{\partial \xi}\|_{L^2(\mathbb{R}^2)}^2)^{\frac{1}{2}}$. We wish to show that v can be approached in \hat{D} by a sequence of smooth functions with compact supports. For that we choose a smooth function $j : \mathbb{R} \rightarrow \mathbb{R}_+$, supported in $[-1, 1]$ and such that $\int_{\mathbb{R}} j = 1$. Then we call j_ϵ the function $j_\epsilon : \mathbb{R}^2 \rightarrow \mathbb{R}_+$, $j_\epsilon(\xi, y) = \frac{1}{\epsilon^2} j(\frac{\xi}{\epsilon}) j(\frac{y}{\epsilon})$. The function $v_\epsilon = j_\epsilon * v$ is a smooth function with compact support. We want to estimate $\|v - v_\epsilon\|_{\hat{D}}$. It is clear that $\|v - v_\epsilon\|_{L^2(\mathbb{R}^2)}$ tends to zero as ϵ tends to zero. Also, it is proved in the pioneering paper of Friedrichs [14] that $\lim_{\epsilon \rightarrow \infty} \|\frac{\partial}{\partial y}(v - v_\epsilon)\|_{\hat{V}} = 0$. This implies that $\mathcal{D}(\mathbb{R}^2)$ is dense in \hat{V} . It remains to prove that $\lim_{\epsilon \rightarrow \infty} \|\hat{A}(v - v_\epsilon)\|_{L^2(\mathbb{R}^2)} = 0$. Calling J_ϵ the integral operator, $J_\epsilon v = j_\epsilon * v$, we have that $\|\hat{A}v - \hat{A}J_\epsilon v\|_{L^2(\mathbb{R}^2)} \leq \|\hat{A}v - J_\epsilon \hat{A}v\|_{L^2(\mathbb{R}^2)} + \|\hat{A}J_\epsilon v - J_\epsilon \hat{A}v\|_{L^2(\mathbb{R}^2)}$. It is clear that $\lim_{\epsilon \rightarrow 0} \|\hat{A}v - J_\epsilon \hat{A}v\|_{L^2(\mathbb{R}^2)} = 0$. It remains to consider the term $\hat{A}J_\epsilon v - J_\epsilon \hat{A}v$. But $\hat{A}J_\epsilon v - J_\epsilon \hat{A}v = H_\epsilon v$, where

$$H_\epsilon v(\xi, y) = \int_{\mathbb{R}^2} h_\epsilon(\bar{\xi}, \bar{y}, \xi, y) v(\bar{\xi}, \bar{y}) d\bar{\xi} d\bar{y},$$

with $h_\epsilon(\bar{\xi}, \bar{y}, \xi, y) = (\bar{y}^2 - y^2) \frac{\partial^2}{\partial \bar{\xi}^2} g_\epsilon(\bar{\xi}, \bar{y}, \xi, y)$ and $g_\epsilon(\bar{\xi}, \bar{y}, \xi, y) = j_\epsilon(\xi - \bar{\xi}, y - \bar{y})$. The operator H_ϵ has three important properties:

- a) $h(\bar{\xi}, \bar{y}, \xi, y) = 0$, if $\max(|\xi - \bar{\xi}|, |y - \bar{y}|) > \epsilon$.
- b) $H_\epsilon 1 = 0$.
- c) There exists a positive constant C , such that $\forall v \in \hat{D}$, with compact support, $\|H_\epsilon v\|_{L^2(\mathbb{R}^2)} \leq C \|v\|_{\hat{V}}$.

Let us prove briefly the last property:

$$\begin{aligned} H_\epsilon v(\xi, y) &= \int_{\mathbb{R}^2} h_\epsilon(\bar{\xi}, \bar{y}, \xi, y) v(\bar{\xi}, \bar{y}) d\bar{\xi} d\bar{y} = \int_{\mathbb{R}^2} (\bar{y}^2 - y^2) \frac{\partial^2}{\partial \bar{\xi}^2} g_\epsilon(\bar{\xi}, \bar{y}, \xi, y) v(\bar{\xi}, \bar{y}) d\bar{\xi} d\bar{y} \\ &= - \int_{\mathbb{R}^2} (\bar{y} - y)^2 \frac{\partial^2 g_\epsilon}{\partial \bar{\xi}^2}(\bar{\xi}, \bar{y}, \xi, y) v(\bar{\xi}, \bar{y}) d\bar{\xi} d\bar{y} + 2 \int_{\mathbb{R}^2} \bar{y}(\bar{y} - y) \frac{\partial^2 g_\epsilon}{\partial \bar{\xi}^2}(\bar{\xi}, \bar{y}, \xi, y) v(\bar{\xi}, \bar{y}) d\bar{\xi} d\bar{y} \\ &= - \underbrace{\int_{\mathbb{R}^2} (\bar{y} - y)^2 \frac{\partial^2 g_\epsilon}{\partial \bar{\xi}^2}(\bar{\xi}, \bar{y}, \xi, y) v(\bar{\xi}, \bar{y}) d\bar{\xi} d\bar{y}}_I - 2 \underbrace{\int_{\mathbb{R}^2} (\bar{y} - y) \frac{\partial g_\epsilon}{\partial \bar{\xi}}(\bar{\xi}, \bar{y}, \xi, y) \bar{y} \frac{\partial v}{\partial \bar{\xi}}(\bar{\xi}, \bar{y}) d\bar{\xi} d\bar{y}}_{II} \end{aligned}$$

But $\frac{\partial}{\partial \bar{\xi}} g_\epsilon(\bar{\xi}, \bar{y}, \xi, y) = -(\frac{\partial}{\partial \bar{\xi}} j_\epsilon)(\bar{\xi} - \xi, \bar{y} - y)$, and it can be seen that there exists a constant C such that $\|y \frac{\partial}{\partial \bar{\xi}} j_\epsilon\|_{L^1(\mathbb{R}^2)} \leq C$, so Young's inequality gives $\|II\|_{L^2(\mathbb{R}^2)} \leq C \|y \frac{\partial}{\partial \bar{\xi}} v\|_{L^2(\mathbb{R}^2)}$. Likewise, $\frac{\partial^2}{\partial \bar{\xi}^2} g_\epsilon(\bar{\xi}, \bar{y}, \xi, y) = (\frac{\partial^2}{\partial \bar{\xi}^2} j_\epsilon)(\bar{\xi} - \xi, \bar{y} - y)$, and there exists a constant C such that $\|y^2 \frac{\partial^2}{\partial \bar{\xi}^2} j_\epsilon\|_{L^1(\mathbb{R}^2)} \leq C$, so we have $\|I\|_{L^2(\mathbb{R}^2)} \leq C \|v\|_{L^2(\mathbb{R}^2)}$. The property (c) is proved.

Let us prove that

$$(25) \quad \lim_{\epsilon \rightarrow 0} \|H_\epsilon v\|_{L^2(\mathbb{R}^2)} = 0.$$

For that, we use the density of $\mathcal{D}(\mathbb{R}^2)$ in \hat{V} , so, thanks to property (c), we only have to prove (25) for v in $\mathcal{D}(\mathbb{R}^2)$. Then $H_\epsilon v = \int_{\mathbb{R}^2} h_\epsilon(\bar{\xi}, \bar{y}, \xi, y) (v(\bar{\xi}, \bar{y}) - v(\xi, y)) d\bar{\xi} d\bar{y}$ thanks to property (b). Now (25) stems from the continuity of v . This concludes the proof. \square

Theorem 1. *If for all t , $r(t) > 0$, then the domain D_t of $A_{1,t}$ does not depend on t : $D_t = D$. Moreover, if there exists a constant $r_0 > 0$ such that $r(t) > r_0$ a.e., and if $\frac{\alpha^2}{\beta^2} > 2$, then for well-chosen values of η (in particular such that $2\frac{\alpha^2}{\beta^2}\eta(1-\eta) > 1$),*

$$(26) \quad D = \left\{ v \in V; y^2 x^2 \frac{\partial^2 v}{\partial x^2}, \frac{\partial^2 v}{\partial y^2}, yx \frac{\partial^2 v}{\partial x \partial y}, x \frac{\partial v}{\partial x}, y \frac{\partial v}{\partial y}, y^2 v \in L^2(Q) \right\}.$$

Proof. The proof is in two steps.

- (1) In the first step, we prove that

$$v \in D_t \Rightarrow x \frac{\partial v}{\partial x} \in L^2(Q).$$

This implies that $D_t = D$ does not depend on t .

- (2) In the second step, we prove that if $v \in D$ satisfies $x \frac{\partial v}{\partial x} \in L^2(Q)$, then

$$y^2 x^2 \frac{\partial^2 v}{\partial x^2}, \frac{\partial^2 v}{\partial y^2}, y \frac{\partial v}{\partial y}, y^2 v \in L^2(Q).$$

Let v be a smooth function with compact support in Q and call f the function $f = -A_{1,t} v$. We call $A_{1,1,t} v = r(t) x \frac{\partial v}{\partial x}$ and $Bv = A_{1,t} v - A_{1,1,t} v$. We have

$$\|Bv + A_{1,1,t} v\|_{L^2(Q)}^2 = \|f\|_{L^2(Q)}^2,$$

or

$$\|Bv\|_{L^2(Q)}^2 + \|A_{1,1,t}v\|_{L^2(Q)}^2 + 2 \int_Q BvA_{1,1,t}v = \|f\|_{L^2(Q)}^2.$$

Let us write in detail the last term in the left-hand side:

$$2 \int_Q BvA_{1,1,t}v = r(t) \left(\begin{array}{c} \underbrace{\int_Q y^2 x^3 \frac{\partial^2 v}{\partial x^2} \frac{\partial v}{\partial x}}_I + \beta^2 \underbrace{\int_Q x \frac{\partial^2 v}{\partial y^2} \frac{\partial v}{\partial x}}_{II} \\ -4 \frac{\alpha^2}{\beta^2} \eta(1-\eta) \underbrace{\int_Q y^2 xv \frac{\partial v}{\partial x}}_{III} + 2\alpha(1-2\eta) \underbrace{\int_Q yx \frac{\partial v}{\partial y} \frac{\partial v}{\partial x}}_{IV} \end{array} \right).$$

Let us look separately at each of the terms in the equality above.

$$\begin{aligned} |I| &= \left| \frac{1}{2} \int_Q y^2 x^3 \frac{\partial}{\partial x} \left(\frac{\partial v}{\partial x} \right)^2 \right| = \frac{3}{2} \int_Q y^2 x^2 \left(\frac{\partial v}{\partial x} \right)^2 \leq \frac{3}{2} \|v\|_V^2. \\ |II| &= \left| - \int_Q x \frac{\partial v}{\partial y} \frac{\partial^2 v}{\partial x \partial y} \right| = \frac{1}{2} \left| \int_Q x \frac{\partial}{\partial x} \left(\frac{\partial v}{\partial y} \right)^2 \right| = \frac{1}{2} \int_Q \left(\frac{\partial v}{\partial y} \right)^2 \leq \frac{1}{2} \|v\|_V^2. \\ |III| &= \left| \frac{1}{2} \int_Q y^2 x \frac{\partial v^2}{\partial x} \right| = \frac{1}{2} \int_Q y^2 v^2 \leq C \|v\|_V^2. \\ |IV| &= \left| \int_Q yx \frac{\partial v}{\partial y} \frac{\partial v}{\partial x} \right| \leq C \|v\|_V^2. \end{aligned}$$

Thus

$$(27) \quad \|Bv\|_{L^2(Q)}^2 + \|A_{1,1,t}v\|_{L^2(Q)}^2 \leq \|A_{1,t}v\|_{L^2(Q)}^2 + C \|v\|_V^2.$$

Since $\mathcal{D}(Q)$ is dense in D_t , the first step is achieved. We have proved that the domain of A_t does not depend on t .

Now we know that the function v satisfies

$$(28) \quad \begin{aligned} v &\in V, \\ Bv &= g \in L^2(Q), \\ A_{1,1,t}v &\in L^2(Q), \end{aligned}$$

and we have to write $\|Bv\|_{L^2(Q)}^2$ in detail. We have $Bv = \sum_{i=1}^4 B_i v$ with $B_1 v = \frac{1}{2} y^2 x^2 \frac{\partial^2 v}{\partial x^2}$, $B_2 v = \frac{1}{2} \beta^2 \frac{\partial^2 v}{\partial y^2}$, $B_3 v = 2 \frac{\alpha^2}{\beta^2} \eta(1-\eta) y^2 v$, and $B_4 v = \alpha(1-2\eta) y \frac{\partial v}{\partial y}$. Of course, we have

$$(29) \quad \sum_{i=1}^4 \|B_i v\|_{L^2(Q)}^2 + \sum_{i=1}^6 I_i = \|g\|_{L^2(Q)}^2,$$

with

$$(30) \quad \begin{aligned} I_1 &= \frac{1}{2} \beta^2 \int_Q x^2 y^2 \frac{\partial^2 v}{\partial y^2} \frac{\partial^2 v}{\partial x^2}, & I_2 &= -2 \frac{\alpha^2}{\beta^2} \eta(1-\eta) \int_Q x^2 y^4 v \frac{\partial^2 v}{\partial x^2}, \\ I_3 &= \alpha(1-2\eta) \int_Q x^2 y^3 \frac{\partial v}{\partial y} \frac{\partial^2 v}{\partial x^2}, & I_4 &= -2 \alpha^2 \eta(1-\eta) \int_Q y^2 v \frac{\partial^2 v}{\partial y^2}, \\ I_5 &= \alpha \beta^2 (1-2\eta) \int_Q y \frac{\partial v}{\partial y} \frac{\partial^2 v}{\partial y^2}, & I_6 &= -4 \frac{\alpha^3}{\beta^2} \eta(1-\eta) \int_Q y^3 v \frac{\partial v}{\partial y}. \end{aligned}$$

We have

$$\begin{aligned}
 (31) \quad I_1 &= -\frac{\beta^2}{2} \left(\int_Q 2xy^2 \frac{\partial^2 v}{\partial y^2} \frac{\partial v}{\partial x} + \int_Q x^2 y^2 \frac{\partial^3 v}{\partial x \partial y^2} \frac{\partial v}{\partial x} \right) \\
 &= \beta^2 \left(\frac{1}{2} \int_Q x^2 y^2 \left(\frac{\partial^2 v}{\partial y \partial x} \right)^2 + \int_Q xy^2 \frac{\partial v}{\partial y} \frac{\partial^2 v}{\partial y \partial x} + \int_Q x^2 y \frac{\partial v}{\partial x} \frac{\partial^2 v}{\partial y \partial x} + 2 \int_Q xy \frac{\partial v}{\partial y} \frac{\partial v}{\partial x} \right) \\
 &= \frac{\beta^2}{2} \left(\int_Q x^2 y^2 \left(\frac{\partial^2 v}{\partial y \partial x} \right)^2 - \int_Q y^2 \left(\frac{\partial v}{\partial y} \right)^2 - \int_Q x^2 \left(\frac{\partial v}{\partial x} \right)^2 + 4 \int_Q xy \frac{\partial v}{\partial y} \frac{\partial v}{\partial x} \right).
 \end{aligned}$$

But, from the assumption made on $r(t)$, we know that $x \frac{\partial v}{\partial x} \in L^2(Q)$, and therefore, we have proven that there exists a constant C such that

$$(32) \quad \left| I_1 - \frac{\beta^2}{2} \left(\int_Q x^2 y^2 \left(\frac{\partial^2 v}{\partial y \partial x} \right)^2 - \int_Q y^2 \left(\frac{\partial v}{\partial y} \right)^2 \right) \right| \leq C(\|A_{1,1,t}v\|_{L^2(Q)}^2 + \|v\|_V^2).$$

Switching to I_2 , we see that

$$\begin{aligned}
 (33) \quad I_2 &= 2 \frac{\alpha^2}{\beta^2} \eta(1-\eta) \left(\int_Q 2xy^4 v \frac{\partial v}{\partial x} + \int_Q x^2 y^4 \left(\frac{\partial v}{\partial x} \right)^2 \right) \\
 &= 2 \frac{\alpha^2}{\beta^2} \eta(1-\eta) \left(- \int_Q y^4 v^2 + \int_Q x^2 y^4 \left(\frac{\partial v}{\partial x} \right)^2 \right).
 \end{aligned}$$

For the next terms, we have

$$\begin{aligned}
 (34) \quad I_3 &= \alpha(1-2\eta) \left(- \int_Q 2xy^3 \frac{\partial v}{\partial y} \frac{\partial v}{\partial x} - \int_Q x^2 y^3 \frac{\partial^2 v}{\partial x \partial y} \frac{\partial v}{\partial x} \right) \\
 &= \alpha(1-2\eta) \left(-2 \int_Q xy^3 \frac{\partial v}{\partial y} \frac{\partial v}{\partial x} + \frac{3}{2} \int_Q x^2 y^2 \left(\frac{\partial v}{\partial x} \right)^2 \right),
 \end{aligned}$$

$$(35) \quad I_4 = -2\alpha^2 \eta(1-\eta) \int_Q y^2 v \frac{\partial^2 v}{\partial y^2} = 2\alpha^2 \eta(1-\eta) \left(\int_Q y^2 \left(\frac{\partial v}{\partial y} \right)^2 - \int_Q v^2 \right),$$

$$(36) \quad I_5 = -\frac{\alpha\beta^2(1-2\eta)}{2} \int_Q \left(\frac{\partial v}{\partial y} \right)^2,$$

$$(37) \quad I_6 = 6 \frac{\alpha^3}{\beta^2} \eta(1-\eta) \int_Q y^2 v^2.$$

Gathering (32)–(37), we obtain that

$$\begin{aligned}
 (38) \quad & \sum_{i=1}^4 \|B_i v\|_{L^2(Q)}^2 + \frac{\beta^2}{2} \int_Q x^2 y^2 \left(\frac{\partial^2 v}{\partial y \partial x} \right)^2 + \left(-\frac{\beta^2}{2} + 2\alpha^2 \eta(1-\eta) \right) \int_Q y^2 \left(\frac{\partial v}{\partial y} \right)^2 \\
 & + 2 \frac{\alpha^2}{\beta^2} \eta(1-\eta) \left(\int_Q x^2 y^4 \left(\frac{\partial v}{\partial x} \right)^2 - \int_Q y^4 v^2 \right) - 2\alpha(1-2\eta) \int_Q xy^3 \frac{\partial v}{\partial y} \frac{\partial v}{\partial x} \\
 & \leq \|g\|^2 + C(\|A_{1,1,t}v\|_{L^2(Q)}^2 + \|v\|_V^2).
 \end{aligned}$$

Expanding the first four terms, we get

$$\begin{aligned}
 & \frac{1}{4} \int_Q x^4 y^4 \left(\frac{\partial^2 v}{\partial x^2}\right)^2 + \frac{\beta^4}{4} \int_Q \left(\frac{\partial^2 v}{\partial y^2}\right)^2 + \frac{\beta^2}{2} \int_Q x^2 y^2 \left(\frac{\partial^2 v}{\partial y \partial x}\right)^2 \\
 & - 2\alpha(1 - 2\eta) \int_Q xy^3 \frac{\partial v}{\partial y} \frac{\partial v}{\partial x} + \left(\alpha^2(2\eta^2 - 2\eta + 1) - \frac{\beta^2}{2}\right) \int_Q y^2 \left(\frac{\partial v}{\partial y}\right)^2 \\
 (39) \quad & + 2\frac{\alpha^2}{\beta^2} \eta(1 - \eta) \int_Q x^2 y^4 \left(\frac{\partial v}{\partial x}\right)^2 \\
 & + \left(4\frac{\alpha^4}{\beta^4} \eta^2(1 - \eta)^2 - 2\frac{\alpha^2}{\beta^2} \eta(1 - \eta)\right) \int_Q y^4 u^2 \\
 & \leq \|g\|^2 + C(\|A_{1,1,t}v\|_{L^2(Q)}^2 + \|v\|_V^2),
 \end{aligned}$$

which implies, for any $\chi > 0$,

$$\begin{aligned}
 & \frac{1}{4} \int_Q x^4 y^4 \left(\frac{\partial^2 v}{\partial x^2}\right)^2 + \frac{\beta^4}{4} \int_Q \left(\frac{\partial^2 v}{\partial y^2}\right)^2 + \frac{\beta^2}{2} \int_Q x^2 y^2 \left(\frac{\partial^2 v}{\partial y \partial x}\right)^2 \\
 & + \left(\alpha^2(2\eta^2 - 2\eta + 1) - \frac{\beta^2}{2} - |1 - 2\eta| \frac{\alpha\beta}{\chi}\right) \int_Q y^2 \left(\frac{\partial v}{\partial y}\right)^2 \\
 & + \left(2\frac{\alpha^2}{\beta^2} \eta(1 - \eta) - |1 - 2\eta| \alpha \frac{\chi}{\beta}\right) \int_Q x^2 y^4 \left(\frac{\partial v}{\partial x}\right)^2 \\
 & + \left(4\frac{\alpha^4}{\beta^4} \eta^2(1 - \eta)^2 - 2\frac{\alpha^2}{\beta^2} \eta(1 - \eta)\right) \int_Q y^4 u^2 \\
 & \leq \|g\|^2 + C(\|A_{1,1,t}v\|_{L^2(Q)}^2 + \|v\|_V^2).
 \end{aligned}$$

The result will be proved if we can find η , $0 < \eta < 1$, and $\chi > 0$ such that

$$\begin{aligned}
 & \frac{\alpha^2}{\beta^2}(2\eta^2 - 2\eta + 1) - |1 - 2\eta| \frac{\alpha}{\beta} \frac{1}{\chi} - \frac{1}{2} > 0, \\
 & 2\frac{\alpha^2}{\beta^2} \eta(1 - \eta) - |1 - 2\eta| \frac{\alpha\beta}{\chi} > 0, \\
 & 4\frac{\alpha^4}{\beta^4} \eta^2(1 - \eta)^2 - 2\frac{\alpha^2}{\beta^2} \eta(1 - \eta) > 0.
 \end{aligned}$$

The last inequality is true as soon as $2\frac{\alpha^2}{\beta^2} \eta(1 - \eta) > 1$. Also, $2\eta^2 - 2\eta + 1 \geq \frac{1}{2}$ so we will look for η and χ such that

$$\begin{aligned}
 & 2\frac{\alpha^2}{\beta^2} \eta(1 - \eta) > 1, \\
 & \frac{\alpha^2}{\beta^2} - 2|1 - 2\eta| \frac{\alpha}{\beta} \frac{1}{\chi} - 1 > 0, \\
 & 2\frac{\alpha^2}{\beta^2} \eta(1 - \eta) - |1 - 2\eta| \frac{\alpha\beta}{\chi} > 0.
 \end{aligned}$$

If $\frac{\alpha^2}{\beta^2} > 2$, the inequalities above are satisfied for $\eta = \frac{1}{2}$ and χ arbitrary. □

Theorem 1 makes it possible to apply the theory of analytic semigroups. This is carried out in Section 3 below, where existence and regularity results for (10) are proved, in particular a smoothing property typical for parabolic problems but in this case in a weighted Sobolev norm which is adjusted to the degeneracy of the equation.

3. THE EVOLUTION EQUATION

We are interested in the evolution equation

$$\begin{aligned}
 (40) \quad & \frac{d}{dt}u - A_t u = 0, \quad 0 < t \leq T, \\
 & u(t = 0) = u_0.
 \end{aligned}$$

The following result can be found in [1].

Theorem 2. *Assume that $\alpha > \beta$ and that η has been chosen as in Proposition 2. Then there exists a unique u in $L^2(0, T; V) \cap C^0([0, T]; L^2(Q))$, with $\frac{\partial u}{\partial t} \in L^2(0, T; V')$ such that, in the sense of the distributions in time,*

$$(41) \quad \frac{d}{dt}(u, v) - \langle A_t u, v \rangle = 0, \quad \forall v \in V,$$

and

$$(42) \quad u(t = 0) = u_0.$$

The mapping $u_0 \mapsto u$ is continuous from $L^2(Q)$ to $L^2(0, T; V) \cap C^0([0, T]; L^2(Q))$.

We can obtain a more precise result with stronger assumptions

Theorem 3. *Assume that there exists ζ , where $0 < \zeta \leq 1$ such that γ belongs to $C^\zeta([0, T], L^\infty(Q))$ and r is a Hölder function of time with exponent ζ . Assume also that $r(t) > r_0$ for a positive constant r_0 and that $\frac{\alpha^2}{\beta^2} > 2$. Then for η chosen as in Proposition 2 and Theorem 1, if u_0 belongs to D defined by (26), then the solution of (41) and (42) given by Theorem 2 belongs also to $C^1((0, T); L^2(Q)) \cap C^0([0, T]; D)$ and satisfies the equation*

$$(43) \quad \frac{d}{dt}u - A_t u = 0,$$

for each $t \in [0, T]$.

Furthermore, for $u_0 \in L^2(Q)$, the weak solution of (41) and (42) given by Theorem 2 belongs also to $C^1((\tau, T); L^2(Q)) \cap C^0([\tau, T]; D)$, for all $\tau > 0$ and we have that $\|\frac{du}{dt}(t)\|_{L^2(Q)} + \|A_t u(t)\|_{L^2(Q)} \leq \frac{C}{\tau}$, for $t > 0$.

Proof. We have that

- the domain of A_t is dense in $L^2(Q)$ and independent on t by Theorem 1;
- for all $t \in [0, T]$, the resolvent $R(\lambda : A_t)$ of A_t exists for $\Re(\lambda) \geq c_0$ (where c_0 is defined in Proposition 2), and we have (17);
- there exists a constant L such that

$$(44) \quad \|(A_t - A_s)A_\tau^{-1}\|_{\mathcal{L}(L^2(Q))} \leq L|t - s|^\zeta.$$

With these three facts, we can apply a result of Kato (see [17] Theorem 5.6.1.) which yields the desired result. □

Remark 3. It is quite natural to wonder whether the results of Sections 2 and 3 can be extended to more general classes of differential operators. More precisely, we would like to better understand the structure properties of the operator A_t that make these results possible, in particular Gårding's inequality (16), as well as the second order estimates of Theorem 1 and the regularity results of Theorems 2 and 3. In fact it is straightforward matter to check that all our results hold when we assume that the coefficients α, β, η , and m depend on time (say, they are Hölder continuous in time in order to satisfy (44)), provided they satisfy the prescribed bounds of Proposition 2 and Theorem 1 uniformly for $t > 0$. Indeed, we just have to notice that all our estimates involve only space derivatives and integration with respect to space variables.

On the other hand, the situation changes if we are interested in having the coefficients depend on x and y . In this respect, it is worth noticing that the decomposition $A_t = A_{1,t} + A_{2,t}$ in (12) hides the fact that the elementary operators used to build A_t are the vector fields $X = xy \frac{\partial}{\partial x}$, $Y = \frac{\partial}{\partial y}$, and the multiplication

operator $Mv : v \rightarrow yv$. Denoting by X^* , Y^* , and M^* the formal adjoint operators of X , Y , and M , respectively, i.e., $X^* = -X - M$, $Y^* = -Y$, and $M^* = M$, an elementary computation shows that

$$\begin{aligned} x^2 y^2 \frac{\partial^2}{\partial x^2} &= -X^* X - 2M^* X, \\ x \frac{\partial}{\partial x} &= -Y^* X + X^* Y + M^* Y, \end{aligned}$$

and thus we can write $A_{1,t}$ as

$$\begin{aligned} (45) \quad A_{1,t} &= -\left(\frac{1}{2}X^* X + M^* X + \frac{1}{2}\beta^2 Y^* Y + r(t)(Y^* X - X^* Y - M^* Y)\right) \\ &\quad + 2\frac{\alpha^2}{\beta^2}\eta(1 - \eta)M^* M + \alpha(1 - 2\eta)M^* Y. \end{aligned}$$

To obtain a more symmetric structure, notice $M^* Y v = -Y^* M v - v$, so in (45), the term

$$(-r(t) + \alpha(1 - 2\eta))M^* Y$$

can be replaced by

$$\frac{1}{2}(-r(t) + \alpha(1 - 2\eta))M^* Y - \frac{1}{2}(-r(t) + \alpha(1 - 2\eta))Y^* M - \frac{1}{2}(-r(t) + \alpha(1 - 2\eta))Id,$$

where Id is the identity operator $Id(v) \equiv v$. Thus we can write $A_{1,t}$ in the *divergence form*

$$A_{1,t} = -(X^*, Y^*, M^*)\mathcal{A}(t) \begin{pmatrix} X \\ Y \\ M \end{pmatrix} + \frac{1}{2}(-r(t) + \alpha(1 - 2\eta))Id,$$

where

$$(46) \quad \mathcal{A}(t) = \begin{bmatrix} \frac{1}{2} & -r(t) & 0 \\ r(t) & \frac{1}{2}\beta^2 & -\frac{1}{2}(-r(t) + \alpha(1 - 2\eta)) \\ 1 & \frac{1}{2}(-r(t) + \alpha(1 - 2\eta)) & 2\frac{\alpha^2}{\beta^2}\eta(1 - \eta) \end{bmatrix}.$$

Notice that the symmetric part $\mathcal{A}^{\text{sym}}(t)$ of $\mathcal{A}(t)$ given by

$$(47) \quad \mathcal{A}^{\text{sym}}(t) = \begin{bmatrix} \frac{1}{2} & 0 & \frac{1}{2} \\ 0 & \frac{1}{2}\beta^2 & 0 \\ \frac{1}{2} & 0 & 2\frac{\alpha^2}{\beta^2}\eta(1 - \eta) \end{bmatrix}$$

is positive definite provided we have

$$2\alpha^2\eta(1 - \eta) - \frac{1}{2}\beta^2 > 0,$$

which is possible if $\alpha > \beta$ and $\eta_1 < \eta < \eta_2$, where η_1, η_2 are suitable positive constants, $0 < \eta_1 < \frac{1}{2} < \eta_2 < 1$.

On the other hand, the term $A_{2,t}$ in the decomposition (12) contains “lower order terms”. We emphasize the fact that here, the notion of “lower order term” must be understood with respect to the elementary operators X, Y , and M . Thus, for instance, a term like $c(x, y)v$ is a zero order term provided $c \in L^\infty$, and $c(x, y)yv$ and Yv are order one terms, but $\frac{\partial v}{\partial x}$ is not a first order term.

Looking back to Section 2 in this spirit, it is easy to see that Gårding’s inequality (16) holds for very general operators of the form

$$(48) \quad A_t = -(X^*, Y^*, M^*)\mathcal{A}(t, x, y) \begin{pmatrix} X \\ Y \\ M \end{pmatrix} + (\text{true}) \text{ lower order terms},$$

when $\mathcal{A}(t, x, y)$ is a bounded measurable 3×3 real matrix such that its symmetric part $\mathcal{A}^{\text{sym}}(t, x, y)$ is positive definite uniformly with respect to t, x , and y .

Concerning Theorem 1, we see that its proof relies heavily on iterate integration by parts with respect to space variables. But, if we allow the coefficients to depend on x, y , then these integration by parts yield new terms containing derivatives of the coefficients with respect to space variables. When dealing with elliptic equations with smooth coefficients in bounded regions, this is harmless, since these additional terms only contain lower order derivatives. Unfortunately, this is no more true when dealing with degenerate equations with polynomial coefficients in unbounded regions, as in the present case, since the terms containing the derivatives of the coefficients might fail to be “true lower order terms”. Thus, we cannot extend the a priori estimates of Theorem 1 to general operators of the form (48).

4. A FINITE ELEMENT DISCRETIZATION FOR A RELATED PROBLEM

4.1. A simpler elliptic problem. We shall hereafter address the question of the discretization by finite elements of equation (10). More precisely, we are going to focus on the question, How should one choose the mesh in order to cope with the degeneracy of the operator (11) on the axis $y = 0$ as well as the nonuniformity of the operator with respect to the variable x ? With this aim, we are going to simplify the equation and restrict the domain where this equation is posed, in order to retain only the difficulties due to the degeneracy of the operator.

Denote by X and Y the smooth vector fields in $\mathbb{R}^2 = \mathbb{R}_x \times \mathbb{R}_y$ defined by

$$(49) \quad X = xy\partial_x, \quad Y = \partial_y.$$

The simplifications are the following.

- We shall take for Ω a rectangle $(a, b) \times (-c, c)$ where a, b, c are three positive numbers, $0 < a < b$. With this choice, we avoid the issue of the discretization near $x = 0$, whose answer depends on available information on the regularity of the solution near $x = 0$ (for the equation (10), (11) it is shown in [1] that the solution is continuous up to $x = 0$, with bounded derivatives so the mesh does not need special refinement near $x = 0$). Nevertheless, our discretization scheme will take into account that a may be as close as necessary to zero, and hence all our error estimates will be explicit in terms of a . Up to appropriate scalings in the two variables, we can choose $\Omega = (m, 1) \times (-1, 1)$. We denote Γ_0 (respectively Γ_1) the horizontal (respectively vertical) part of $\partial\Omega$.
- We consider the boundary value problem in Ω :

$$(50) \quad \begin{aligned} -y^2x^2\frac{\partial^2u}{\partial x^2} - \frac{\partial^2u}{\partial y^2} + \Lambda u &= f \quad \text{in } \Omega, \\ n_yYu &= g_0 \quad \text{on } \Gamma_0, \\ n_xXu &= xyg_1 \quad \text{on } \Gamma_1. \end{aligned}$$

where n is the unit vector normal to $\partial\Omega$, pointing out of Ω , and Λ is a large enough positive number (see below).

If \mathcal{U} is an open subset of \mathbb{R}^2 , we denote by $S^{1,2}(\mathcal{U})$ the function space

$$(51) \quad S^{1,2}(\mathcal{U}) = \{u \in L^2(\mathcal{U}) : Xu, Yu \in L^2(\mathcal{U})\},$$

endowed with the natural norm

$$(52) \quad \|u\|_{S^{1,2}(\mathcal{U})}^2 = \|u\|_{L^2(\mathcal{U})}^2 + |u|_{S^{1,2}(\mathcal{U})}^2,$$

where the seminorm $|u|_{S^{1,2}(\mathcal{U})}$ is defined by

$$(53) \quad |u|_{S^{1,2}(\mathcal{U})}^2 = \|Xu\|_{L^2(\mathcal{U})}^2 + \|Yu\|_{L^2(\mathcal{U})}^2.$$

Analogously, we define

$$(54) \quad S^{2,2}(\mathcal{U}) = \{u \in S^{1,2}(\mathcal{U}) : X^2u, Y^2u, XYu, YXu \in L^2(\mathcal{U})\},$$

endowed again with the natural norm

$$(55) \quad \|u\|_{S^{2,2}(\mathcal{U})}^2 = \|u\|_{S^{1,2}(\mathcal{U})}^2 + |u|_{S^{2,2}(\mathcal{U})}^2,$$

where the seminorm $|u|_{S^{2,2}(\mathcal{U})}$ is defined by

$$(56) \quad |u|_{S^{2,2}(\mathcal{U})}^2 = \|X^2u\|_{L^2(\mathcal{U})}^2 + \|Y^2u\|_{L^2(\mathcal{U})}^2 + \|XYu\|_{L^2(\mathcal{U})}^2 + \|YXu\|_{L^2(\mathcal{U})}^2.$$

In addition, we define $S_0^{1,2}(\mathcal{U})$ as the closure of $\mathcal{D}(\Omega)$ in $S^{1,2}(\mathcal{U})$.

The equation in (50) can be written in divergence form

$$(57) \quad -\frac{\partial}{\partial x}(y^2x^2\frac{\partial u}{\partial x}) - \frac{\partial^2 u}{\partial y^2} + 2y^2x\frac{\partial u}{\partial x} + \Lambda u = f \quad \text{in } \Omega,$$

and admits a weak formulation in $S^{1,2}(\Omega)$: find $u \in S^{1,2}(\Omega)$ such that for all $v \in S^{1,2}(\Omega)$,

$$(58) \quad \int_{\Omega} XuXv + YuYv + 2yvXu + \Lambda uv = \int_{\Omega} fv + \int_{\Gamma_0} g_0v + \int_{\Gamma_1} xyg_1v,$$

which is well defined if $f \in L^2(\Omega)$, $g_0 \in L^2(\Gamma_0)$, and $g_1 \in L^2(\Gamma_1)$, because for a function $v \in S^{1,2}(\Omega)$, we can define the traces $v|_{\Gamma_0} \in L^2(\Gamma_0)$ and $(xyv)|_{\Gamma_1} \in L^2(\Gamma_1)$. For Λ large enough, the bilinear form $\mathcal{A}(u, v) = \int_{\Omega} XuXv + YuYv + 2yvXu + \Lambda uv$ is continuous and coercive in $S^{1,2}(\Omega)$, so the problem (50) has a unique solution, which depends continuously on the data. We shall hereafter suppose that Λ is large enough so that \mathcal{A} is coercive. This assumption is not restrictive, since the real problem we are interested in is parabolic.

In order to discretize (50), we need first to define an appropriate mesh.

4.2. The Carnot–Carathéodory metric. In the rest of this paper, the notation $a \approx b$ means that there exist two absolute positive constants c and C such that $cb \leq a \leq Cb$.

Let us recall now the following standard definition of the Carnot–Carathéodory metric associated with a family $X = (X_1, \dots, X_m)$ of Lipschitz continuous vector fields in \mathbb{R}^n (see, e.g., [8], [11], [16]).

Definition. We say that an absolutely continuous curve $\gamma : [0, T] \rightarrow \mathbb{R}^n$ is a subunit curve with respect to X if for any $\xi \in \mathbb{R}^n$

$$\langle \dot{\gamma}(t), \xi \rangle^2 \leq \sum_{j=1}^m \langle X_j(\gamma(t)), \xi \rangle^2,$$

for a.e. $t \in [0, T]$. If $P_1, P_2 \in \mathbb{R}^n$, we define

$$d(P_1, P_2) = \inf \{T > 0 : \text{there exists a subunit curve } \gamma, \\ \gamma : [0, T] \longrightarrow \mathbb{R}^n, \quad \gamma(0) = P_1, \quad \gamma(T) = P_2\}.$$

If the above set of curves is empty, we take $d(P_1, P_2) = \infty$.

If we choose now $X_1 = X, X_2 = Y$, it is well known that $d(P_1, P_2) < \infty$ for any couple of points P_1, P_2 belonging to the half-plane $\{(x, y); x > 0\}$. Thus, if $(a, b) \in \{(x, y); x > 0\}$ and $h > 0$, we denote by $B_c((a, b), h)$ the open ball of radius $h > 0$ centered at (a, b) for the Carnot–Carathéodory metric d associated with (X, Y) .

If a point (a, b) belongs to a compact subset K of $\{(x, y); x > 0\}$, then it is possible to provide a characterization of $B_c((a, b), h)$, provided h is sufficiently small (compared with $\inf_K x$ and $(\sup_K x)^{-1}$).

Since later in this paper this will be precisely the case of interest, take $K = [m, 1] \times [-1, 1]$, where $m > 0$, and assume $h < m$. Combining Theorem 2.3 in [10] and the previous Remark therein, we obtain the following property of metric balls centered in K .

Proposition 4. *If $(a, b) \in K$ and $h < m$, we take*

$$\Lambda_1^*((a, b), h) \equiv \max_{|y-b|<h} |y|a = (|b| + h)a, \quad \Lambda_2^*((a, b), h) \equiv 1, \\ F_i^*((a, b), h) \equiv h\Lambda_i^*((a, b), h), \quad i = 1, 2,$$

and

$$Q((a, b), h) \equiv (a - F_1^*((a, b), h), a + F_1^*((a, b), h)) \times (b - F_2^*((a, b), h), b + F_2^*((a, b), h)).$$

There exists an absolute constant $\theta > 1$ and a constant $h_0 = h_0(m)$ such that

$$Q((a, b), \frac{h}{\theta}) \subset B_c((a, b), h) \subset Q((a, b), \theta h),$$

for any $(a, b) \in K$ and $h < h_0$.

Proof. It is enough to choose in [10], Theorem 2.3, $\lambda_1(x, y) = |x||y|$ and $\lambda_2(x, y) = 1$. In the domain K , the Lipschitz constant of both λ_1 and λ_2 is clearly 1. \square

Remark 4. Since $F_1^*((a, b), 2h) \leq 4F_1^*((a, b), h)$, we can see directly that the distance d is doubling with respect to Lebesgue measure, i.e.,

$$(59) \quad |B_c((a, b), 2h)| \leq C|B_c((a, b), h)|, \quad (a, b) \in K, h < h_0,$$

where $|E|$ stands for the Lebesgue measure of $E \subset \mathbb{R}^2$, and C is an absolute constant, independent of m . This is a particular case of a well-known result (see e.g., [11] and [16]).

Remark 5. When we are “away from $y = 0$ ”, it will turn out to be easier to work with a slightly different family of equivalent rectangles. For any positive numbers h, a, b such that $bh \leq \frac{1}{2}$ and $b \geq 2h > 0$, we define $R_{a,b,h}$ as the rectangle

$$(60) \quad R_{a,b,h} = \{(x, y) \in \mathbb{R}_+^2, |x - a| \leq abh; |y - b| \leq h\}.$$

Since $h < b$, we have $abh \leq (b + h)ha \leq ab(2h)$, so that

$$(61) \quad \text{Int}(R_{a,b,h}) \subset Q((a, b), h) \subset R_{a,b,2h}.$$

4.3. Definition of the mesh. We want to construct a triangulation of the domain $\Omega = (m, 1) \times (-1, 1)$. With this aim, we choose an integer $N_x > 0$ and the real number $\lambda > 1$ (in the following N_x will be large and λ will be close to 1), defined by

$$(62) \quad m = \lambda^{-N_x} \Leftrightarrow \log(\lambda) = -\frac{\log(m)}{N_x}.$$

We denote by L_k the length of the interval $[\lambda^{-k-1}, \lambda^{-k}]$; i.e., we set

$$(63) \quad L_k \equiv \lambda^{-k-1}(\lambda - 1).$$

For a given constant c , we denote by h_λ the real number

$$(64) \quad h_\lambda \equiv c\sqrt{\lambda - 1}.$$

Note that $N_x \approx (\log m)h_\lambda^{-2}$.

For convenience, we choose a mesh constructed by assembling a Cartesian mesh of the domain $(m, 1) \times (-2h_\lambda, 2h_\lambda)$ and more general meshes in the domains $(m, 1) \times (-1, -2h_\lambda)$ and $(m, 1) \times (2h_\lambda, 1)$. For $k, j \in \mathbb{Z}$, let $S_{k,j}$ be the node in the plane

$$(65) \quad S_{k,j} \equiv (\lambda^{-k}, jh_\lambda).$$

We consider the set of nodes \mathcal{V}_1 of $[m, 1] \times [-2h_\lambda, 2h_\lambda]$ given by

$$(66) \quad \mathcal{V}_1 = \{S_{k,j}, k = 0, \dots, N_x, |j| \leq 2\}.$$

It is possible to construct a Cartesian grid covering $(m, 1) \times (-2h_\lambda, 2h_\lambda)$ whose vertices are the nodes of \mathcal{V}_1 . Then we divide each rectangle of the grid into two triangles by means of the diagonal straight line passing through its left-up vertex. Doing so, we construct a triangulation called \mathcal{T}_1 of $(m, 1) \times (-2h_\lambda, 2h_\lambda)$ which is not regular in the sense of [3, 4] with respect to the Euclidean metric. For any triangle t of this mesh, we introduce the size h_t :

$$(67) \quad h_t \equiv h_\lambda.$$

This size is of the same order as the diameter of the triangle t with respect to the intrinsic metric associated with the degenerate operator in (50) (see subsection 4.2 for the definition of the intrinsic metric). Indeed, if we take a point $S_{k,j}$ of this mesh, then

$$F_1^*(S_{k,j}, h_\lambda) = h_\lambda \lambda^{-k}(jh_\lambda + h_\lambda) \approx \lambda^{-k}h_\lambda^2 \approx L_{k-1},$$

and

$$F_2^*(S_{k,j}, h_\lambda) = h_\lambda,$$

so that the Carnot–Caratheodory distance of the contiguous points $S_{k-1,j}, S_{k-1,j+1}, S_{k-1,j+1}$ to $S_{k,j}$ is of the order of h_λ .

Next, for a positive constant $\mu > 1$, we consider a triangulation \mathcal{T}_2 of $(m, 1) \times (2h_\lambda, 1)$ satisfying the following assumptions.

- (1) The nodes of \mathcal{T}_2 and of \mathcal{T}_1 located on the straight line $y = 2h_\lambda$ coincide.
- (2) For each triangle $t \in \mathcal{T}_2$, we can choose two points $(x'_t, y'_t), (x''_t, y''_t)$ and two positive numbers h_t, ρ_t , such that $y'_t h_t \leq \frac{1}{2}, y'_t \geq 2h_t, y''_t \rho_t \leq \frac{1}{2}$, and $y''_t \geq 2\rho_t$, with

$$(68) \quad \begin{aligned} &R_{x'_t, y'_t, \rho_t} \subset t \subset R_{x''_t, y''_t, h_t}, \\ &h_t = \inf\{\tau > 0; \text{ there exists } (\xi, \eta) \text{ such that } \eta \geq 2\tau, 2\eta\tau \leq 1, \text{ and } t \subset R_{\xi, \eta, \tau}\}, \\ &\rho_t = \sup\{\tau > 0, \text{ there exists } (\xi, \eta) \text{ such that } \eta \geq 2\tau, 2\eta\tau \leq 1, \text{ and } R_{\xi, \eta, \tau} \subset t\}, \end{aligned}$$

and

$$\frac{h_t}{\rho_t} \leq \mu.$$

By Proposition 4 and Remark 5, this assumption says that the mesh is regular with respect to the intrinsic metric.

Note that (68) implies that

$$R_{x_t'', y_t'', \rho_t} \subset t \subset R_{x_t'', y_t'', 8\theta^2 \rho_t}.$$

Therefore, for simplicity, we are even going to make a slightly stronger assumption: we assume that there exist a point $(x_t, y_t) \in t$ and two positive numbers h_t, ρ_t such that $y_t h_t \leq \frac{1}{2}$, $y_t \geq 2h_t$, and

$$(69) \quad R_{x_t, y_t, \rho_t} \subset t \subset R_{x_t, y_t, h_t} \quad \text{with} \quad \frac{h_t}{\rho_t} \leq \mu.$$

We can think of (x_t, y_t) as of a generalized “center of gravity” of t .

For the triangulation \mathcal{T}_3 of $(m, 1) \times (-1, -2h_\lambda)$, we make the same assumptions as for \mathcal{T}_2 (all the definitions are modified by using the symmetry with respect to $y = 0$).

The triangulation \mathcal{T} is obtained as the union $\mathcal{T} = \mathcal{T}_1 \cup \mathcal{T}_2 \cup \mathcal{T}_3$.

Remark 6. We give an example of a Cartesian mesh satisfying the assumptions stated above. We define N_y as the integer part of $1/h_\lambda$. For $k, j \in \mathbb{Z}$, let $\tilde{S}_{k,j}$ be the node in the plane

$$(70) \quad \begin{aligned} \tilde{S}_{k,j} &\equiv S_{k,j} && \text{if } |j| \leq 2, \\ \tilde{S}_{k,j} &\equiv \left(\lambda^{-k}, \text{sign}(j) \frac{\sqrt{|j|+3}}{N_y} \right), && \text{if } |j| \geq 3. \end{aligned}$$

Note that if h_λ is small enough, then $\frac{\sqrt{6}}{N_y} > 2h_\lambda$. We consider the set of nodes \mathcal{V} of $\tilde{\Omega}$ given by

$$(71) \quad \mathcal{V} = \{ \tilde{S}_{k,j}, k = 0, \dots, N_x, |j| = 0, \dots, N_y^2 - 3 \}.$$

It is possible to construct a Cartesian grid covering Ω whose vertices are the nodes of \mathcal{V} . Then we divide each rectangle of the grid into two triangles by means of the diagonal straight line passing through its left-up vertex. Doing so, we construct a triangulation satisfying all the assumptions stated above.

Let M be the number of nodes of \mathcal{V} . We have

$$(72) \quad M = \#\mathcal{V} \approx N_y^2 N_x \approx N_x^2 \log m^{-1} \approx N_y^4 \log m.$$

Note that this Cartesian triangulation is close to the one proposed in [13] for the so-called Grushin operator. Note also that this triangulation is not quasi-uniform with respect to the intrinsic metric, because the triangles near the axis $y = 0$ have diameter (in the intrinsic metric) of order h_λ , whereas the triangles far from $y = 0$ have a diameter of order h_λ^2 .

Constructing a Cartesian grid which would be also quasi-uniform with respect to the intrinsic metric is not possible.

4.4. The discrete problem. We denote by V_λ the vector space of the continuous functions on Ω which are piecewise linear on the triangles of \mathcal{T} . We define $\mathcal{V} = \{S_1, \dots, S_M\}$ as the set of nodes of \mathcal{T} . The nodal basis of V_λ is $(\phi_i)_{i=1, \dots, M}$, where $\phi_i \in V_\lambda$ is such that $\phi_i(S_j) = \delta_{i,j}$.

The discrete version of (50) is find $u_\lambda \in V_\lambda$ such $\forall v_\lambda \in V_\lambda$,

$$(73) \quad \mathcal{A}(u_\lambda, v_\lambda) = \int_\Omega f v_\lambda + \int_{\Gamma_0} g_0 v_\lambda + \int_{\Gamma_1} x y g_1 v_\lambda.$$

It is now standard matter (see for example [3], [4], [2]) to prove that (73) has a unique solution, and that there exists a positive constant C such that

$$(74) \quad \|u - u_\lambda\|_{S^{1,2}(\Omega)} \leq C \inf_{v_\lambda \in V_\lambda} |u - v_\lambda|_{S^{1,2}(\Omega)}.$$

We are going to assume that $u \in S^{2,2}(\Omega)$ (this is precisely the regularity proved in Section 2) and we are going to compute the distance $\inf_{v_\lambda \in V_\lambda} |u - v_\lambda|_{S^{1,2}(\Omega)}$ of u to V_λ . For that, it is not possible to use the standard Lagrange interpolation, since a function $v \in S^{2,2}(\Omega)$ is not necessarily continuous on the line $y = 0$. Instead, we propose to use a local regularization operator due to Clément, [5].

5. A LOCAL REGULARIZATION OPERATOR

5.1. Preliminary results. Denote now by \tilde{X} and Y the smooth vector fields in $\mathbb{R}^2 = \mathbb{R}_x \times \mathbb{R}_y$ defined by

$$(75) \quad \tilde{X} = y \partial_x, \quad Y = \partial_y.$$

As above, if \mathcal{U} is an open subset of \mathbb{R}^2 , we denote by $\tilde{S}^{1,2}(\mathcal{U})$, $\tilde{S}_0^{1,2}(\mathcal{U})$, and $\tilde{S}^{2,2}(\mathcal{U})$ the function spaces defined as above replacing X by \tilde{X} . Related norms and seminorms are defined analogously.

Remark 7. An elementary computation shows that, if $\mathcal{U} \subset\subset (0, \infty) \times \mathbb{R}$, then $S^{i,2}(\mathcal{U})$ and $\tilde{S}^{i,2}(\mathcal{U})$ coincide as vector spaces, and the respective norms are equivalent for $i = 1, 2$.

In the rest of this paper, the following compactness result will play a key role.

Theorem 4. *Let \mathcal{U} be a bounded open rectangle in \mathbb{R}^2 . Then*

- i) $\tilde{S}^{1,2}(\mathcal{U})$ is compactly embedded in $L^2(\mathcal{U})$;
- ii) $\tilde{S}^{2,2}(\mathcal{U})$ is compactly embedded in $\tilde{S}^{1,2}(\mathcal{U})$.

Remark 8. Because of Remark 7, if $\mathcal{U} \Subset (0, \infty) \times \mathbb{R}$, then the above statement still holds when $\tilde{S}^{i,2}(\mathcal{U})$ is replaced by $S^{i,2}(\mathcal{U})$, $i = 1, 2$.

Remark 9. If we replace the vector field $\tilde{X} = y \partial_x$ by the vector field $(y + c) \partial_x$ with $c \in \mathbb{R}$, an elementary translation argument shows that Theorem 4 still holds for the new related function spaces.

Proof of Theorem 4. The proof consists of several steps, some of them having intrinsic interest. □

Proposition 5. *The space $C^\infty(\bar{\mathcal{U}})$ is dense in $\tilde{S}^{1,2}(\mathcal{U})$.*

Proof. The result is proved in [7] for smooth bounded domains. For the sake of completeness, we give here a full proof in the case of interest. Suppose now $\mathcal{U} = (a, b) \times (c_1, c_2)$. It will be enough to prove that

$$\mathbf{C}^\infty((a - \epsilon, b + \epsilon) \times (c_1, c_2)) \cap \tilde{S}^{1,2}((a - \epsilon, b + \epsilon) \times (c_1, c_2))$$

is dense in $\tilde{S}^{1,2}(\mathcal{U})$, and then to apply the density result of [7] to a smooth set $\tilde{\mathcal{U}}$, $\mathcal{U} \subset \tilde{\mathcal{U}} \subset (a - \epsilon, b + \epsilon) \times (c_1, c_2)$. Let $u \in \tilde{S}^{1,2}(\mathcal{U})$ be given; we want to approximate u in $\tilde{S}^{1,2}(\mathcal{U})$ by smooth functions defined in $(a - \epsilon, b + \epsilon) \times (c_1, c_2)$. Since by, e.g., [7] and [12], $\mathbf{C}^\infty(\mathcal{U}) \cap \tilde{S}^{1,2}(\mathcal{U})$ is dense in $\tilde{S}^{1,2}(\mathcal{U})$, we can assume without loss of generality that u is smooth. Moreover, we can always express u as a sum of two functions vanishing, respectively, in a neighborhood of $x = a$ and $x = b$. Therefore, without loss of generality, we may assume, for instance, that

$$u \equiv 0 \text{ in } (a', b) \times (c_1, c_2) \quad \text{with } a < a' < b.$$

In that case, u can be continued by zero on $(a', +\infty) \times (c_1, c_2)$. If $\epsilon > 0$ is small, we take

$$u_\epsilon(x, y) = u(x + \epsilon, y), \quad a - \epsilon < x < \infty, \quad c_1 < y < c_2.$$

Clearly, u_ϵ is smooth in $[a, +\infty) \times (c_1, c_2)$. If now v is any measurable function in $(a, +\infty) \times (c_1, c_2)$ continued by zero for $x < a$, then, if $(x, y) \in \mathcal{U}$, we have

$$|v(x + \epsilon, y) - v(x, y)| = |v(x + \epsilon, y)\mathbf{1}_{[a, +\infty)}(x + \epsilon) - v(x, y)\mathbf{1}_{[a, +\infty)}(x)|.$$

Indeed, if $x \in \mathcal{U}$, then both $x > a$ and $x + \epsilon > a$, so that $\mathbf{1}_{[a, +\infty)}(x + \epsilon) = \mathbf{1}_{[a, +\infty)}(x) = 1$. Thus, if $v \in L^2((a, +\infty) \times (c_1, c_2))$ (and hence $v\mathbf{1}_{[a, +\infty)} \in L^2(\mathbb{R} \times (c_1, c_2))$), then

$$(76) \quad \int_{\mathcal{U}} |v(x + \epsilon, y) - v(x, y)|^2 dx dy \leq \int_{\mathbb{R} \times (c_1, c_2)} |v(x + \epsilon, y)\mathbf{1}_{[a, +\infty)}(x + \epsilon) - v(x, y)\mathbf{1}_{[a, +\infty)}(x)|^2 dx dy \rightarrow 0$$

as $\epsilon \rightarrow 0$. Taking successively $v = u$, $v = y\partial_x u$, and then $v = \partial_y u$, we obtain that $u_\epsilon \rightarrow u$, $\tilde{X}u_\epsilon \rightarrow \tilde{X}u$, and $\tilde{Y}u_\epsilon \rightarrow \tilde{Y}u$ in $L^2(\mathcal{U})$, since $\tilde{Y}u_\epsilon(x, y) = (\tilde{Y}u)(x + \epsilon, y)$ and $\tilde{X}u_\epsilon(x, y) = (\tilde{X}u)(x + \epsilon, y)$. □

Proposition 6. *Let $a' \leq a$, $b' \geq b$, $c'_1 \leq c_1$, and $c'_2 \geq c_2$ be given. Then there exists a continuous extension operator \mathcal{T} from $\tilde{S}^{1,2}((a, b) \times (c_1, c_2))$ to $\tilde{S}^{1,2}((a', b') \times (c'_1, c'_2))$, i.e., a continuous linear map*

$$\mathcal{T} : \tilde{S}^{1,2}((a, b) \times (c_1, c_2)) \rightarrow \tilde{S}^{1,2}((a', b') \times (c'_1, c'_2)),$$

such that

$$\mathcal{T}u|_{(a, b) \times (c_1, c_2)} \equiv u$$

for any $u \in \tilde{S}^{1,2}((a, b) \times (c_1, c_2))$.

Proof. First, let us start by proving the existence of a continuous extension operator from $\tilde{S}^{1,2}((a, b) \times (c_1, c_2))$ to $\tilde{S}^{1,2}((a', b') \times (c_1, c_2))$. By a density argument and Proposition 5, it is enough to construct a continuous extension of a function $u \in \mathbf{C}^\infty(\mathcal{U})$. Let $\psi = \psi(x)$ be a smooth nonnegative function such that $0 \leq \psi \leq 1$, $\psi(x) \equiv 1$ in a neighborhood of $x = a$ and $\psi(x) \equiv 0$ for $x > (a + b)/2$. Since $u = u\psi + u(1 - \psi)$ and both $u\psi$ and $u(1 - \psi)$ belong to $\tilde{S}^{1,2}((a, b) \times (c_1, c_2))$, we can consider separately the case $a' < a$, $b' = b$, and the case $a' = a$, $b' > b$. Thus,

without loss of generality, we can assume $b' = b$, and we can assume u vanishing for $x > a + \epsilon$, so that u is continuous for $x \in [a, +\infty)$. We define the function v by

$$(77) \quad v(x, y) = \begin{cases} u(x, y) & \text{if } x > a; \\ u(2a - x, y) & \text{if } x < a. \end{cases}$$

Obviously

$$(78) \quad \frac{\partial v}{\partial y}(x, y) = \begin{cases} \frac{\partial u}{\partial y}(x, y) & \text{if } x > a; \\ \frac{\partial u}{\partial y}(2a - x, y) & \text{if } x < a. \end{cases}$$

On the other hand,

$$(79) \quad \frac{\partial v}{\partial x}(x, y) = \begin{cases} \frac{\partial u}{\partial x}(x, y) & \text{if } x > a; \\ -\frac{\partial u}{\partial x}(2a - x, y) & \text{if } x < a. \end{cases}$$

Thus, a straightforward computation shows that

$$v \in \tilde{S}^{1,2}((a', b) \times (c_1, c_2)) \quad \text{and} \quad \|v\|_{\tilde{S}^{1,2}((a', b) \times (c_1, c_2))} \leq C \|u\|_{\tilde{S}^{1,2}((a, b) \times (c_1, c_2))},$$

and hence our preliminary statement is proved.

We want now to build an extension operator from $(a', b') \times (c_1, c_2)$ to $(a', b') \times (c'_1, c'_2)$. Clearly, we can restrict ourselves to the cases $c_1 = 0$ or $c_2 = 0$, since otherwise the result follows from usual well-known extension theorems in classical Sobolev spaces, the space $\tilde{S}^{1,2}(\mathcal{U})$ being a classical Sobolev space away from $\{y = 0\}$.

Thus suppose now for instance that $c_1 = 0$, and set for sake of simplicity $c_2 \equiv c$. A continuous extension from $\tilde{S}^{1,2}((a', b') \times (0, c))$ to $\tilde{S}^{1,2}((a', b') \times (-c, c))$ can be obtained now again by reflection across $y = 0$. Finally, the extension from $\tilde{S}^{1,2}((a', b') \times (-c, c))$ to $\tilde{S}^{1,2}((a', b') \times (-c', c'))$ is immediate, again since away from $y = 0$, the space $\tilde{S}^{1,2}((a', b') \times (-c, c))$ is a usual Sobolev space. \square

End of the proof of Theorem 4. We can now prove Theorem 4 i), i.e., that the space $\tilde{S}^{1,2}(\mathcal{U})$ is compactly embedded in $L^2(\mathcal{U})$. Indeed, take $a' < a$, $b' > b$, $c'_1 < c_1$, $c'_2 > c_2$, and let \mathcal{T} be the continuous extension operator defined in Proposition 6. If $\psi \in \mathcal{D}((a', b') \times (c'_1, c'_2))$, $\psi \equiv 1$ in \mathcal{U} , then the map $u \rightarrow \psi \mathcal{T}u$ provides a continuous extension operator from $\tilde{S}^{1,2}(\mathcal{U})$ to $\tilde{S}_0^{1,2}((a', b') \times (c'_1, c'_2))$, and then the assertion follows by the compact embedding theorems for spaces of functions vanishing on the boundary. In fact, this result has been proved by several authors with different degrees of generality (different exponents, weighted measures and so on), but we refer to [6] as the first explicit formulation to our knowledge.

Let us now prove Theorem 4 ii), i.e., that the space $\tilde{S}^{2,2}(\mathcal{U})$ is compactly embedded in $\tilde{S}^{1,2}(\mathcal{U})$.

To this end, let $(u_k)_{k \in \mathbb{N}}$ be a bounded sequence in $\tilde{S}^{2,2}(\mathcal{U})$. Since $(u_k)_{k \in \mathbb{N}}$ is also bounded in $\tilde{S}^{1,2}(\mathcal{U})$, by i) there exists a subsequence (that we still denote by $(u_k)_{k \in \mathbb{N}}$) converging strongly to u in $L^2(\mathcal{U})$. On the other hand, the sequences $(Xu_k)_{k \in \mathbb{N}}$ and $(Yu_k)_{k \in \mathbb{N}}$ are also bounded in $\tilde{S}^{1,2}(\mathcal{U})$, so that we can assume they both converge strongly in $L^2(\mathcal{U})$ to v_1 and v_2 , respectively. It is easy to see that $v_1 = Xu$ and $v_2 = Yu$, and then the assertion is proved. \square

5.2. A linear mapping from $S^{2,2}(\Omega)$ to V_λ . A function $v \in S^{2,2}(\Omega)$ is continuous, except maybe on the axis $y = 0$. This is the reason why it is not possible to use the standard interpolation operator for approximating functions of $S^{2,2}(\Omega)$. Instead, we propose to use a local modification of this operator in the strip $y \in (-2h_\lambda, 2h_\lambda)$: there, we shall use a Clément's local regularization operator.

For $i \in \{1, \dots, M\}$, we denote by ω_i the support of the nodal basis function ϕ_i : $\omega_i = \text{supp}(\phi_i)$ and, for any $v \in L^2(\Omega)$, we denote by $P_i v$ the orthogonal projection of v on $\mathcal{P}_1(\omega_i)$, the space of affine polynomials on ω_i , i.e.,

$$(80) \quad \begin{cases} P_i v \in \mathcal{P}_1(\omega_i) \\ \int_{\omega_i} (v - P_i v) p \, dx dy = 0 \quad \text{for all } p \in \mathcal{P}_1(\omega_i). \end{cases}$$

For the sake of simplicity we write $S_i = (a_i, b_i)$, for $i = 1, \dots, M$. We define by \mathcal{I}_λ and \mathcal{J}_λ , respectively, the set of indices such that $|b_i| \leq h_\lambda$, respectively $|b_i| > h_\lambda$. We can now define the linear operator $\Pi = \Pi_\lambda : S^{2,2}(\Omega) \rightarrow V_\lambda$ by

$$(81) \quad \Pi v = \sum_{i \in \mathcal{I}_\lambda} (P_i v)(S_i) \phi_i + \sum_{i \in \mathcal{J}_\lambda} v(S_i) \phi_i,$$

so Π is a Clément's type operator in the strip $y \in [-h_\lambda, h_\lambda]$, and the standard interpolation operator in the strips $y \geq 2h_\lambda$ and $y \leq -2h_\lambda$.

The following lemma (see, e.g., [3], p. 126) will play an important role in the analysis of the approximation properties of Clément's operator.

Lemma 3 (Peetre–Tartar). *Let V, V_1, V_2, W be Banach spaces, and let $A_i \in \mathcal{L}(V, V_i)$ be continuous linear maps for $i = 1, 2$, the map A_1 being compact. Suppose there exists $c_0 > 0$ such that*

$$\|v\|_V \leq c_0 (\|A_1 v\|_{V_1} + \|A_2 v\|_{V_2}),$$

for any $v \in V$. In addition, let $L \in \mathcal{L}(V, W)$ be a continuous linear map such that

$$L|_{\ker A_2} \equiv 0.$$

Then there exists $C > 0$ such that

$$\|Lv\|_W \leq C \|A_2 v\|_{V_2},$$

for any $v \in V$.

Let $i \in \{1, \dots, M\}$ be fixed; in order to simplify what follows, we consider only indices i such that $b_i \geq 0$. For the case $b_i < 0$ the formulas below must be changed in an obvious manner by symmetry.

We want first to provide an estimate of the norm

$$\|v - P_i v\|_{L^2(\omega_i)},$$

and then of the seminorm

$$|v - P_i v|_{S^{1,2}(\omega_i)} \equiv \|X(v - P_i v)\|_{L^2(\omega_i)} + \|Y(v - P_i v)\|_{L^2(\omega_i)},$$

for $i \in \mathcal{I}_\lambda$. Similarly, for $i \in \mathcal{J}_\lambda$ and for a given triangle $t \in \mathcal{T}$ in ω_i , we call I_t the Lagrange interpolation on the space of affine functions on t , and we want to get an estimate for $\|v - I_t v\|_{L^2(t)}$ and then for the seminorm $|v - I_t v|_{S^{1,2}(t)} \equiv \|X(v - I_t v)\|_{L^2(t)} + \|Y(v - I_t v)\|_{L^2(t)}$.

- For $i \in \mathcal{I}_\lambda$, we have that $S_i = S_{k_i, j_i} \in \mathcal{V}$, with $j_i = 0$ or $j_i = 1$; let t_i be the triangle $t_i = (S_{k_i, j_i}, S_{k_i-1, j_i}, S_{k_i, j_i+1})$. Note that t_i is not necessarily contained in $\bar{\Omega}$. We call \hat{t} the reference triangle $\hat{t} = (\hat{S}_0, \hat{S}_1, \hat{S}_2)$, where $\hat{S}_0 = (0, 0)$, $\hat{S}_1 = (0, 1)$, $\hat{S}_2 = (1, 0)$, and we define the affine mapping B_i which maps \hat{S}_0 to S_i , \hat{S}_1 to S_{k_i-1, j_i} and \hat{S}_2 to S_{k_i, j_i+1} . We have $B_i = Z_i + T_i$ where

$$(82) \quad Z_i = S_i \quad \text{and} \quad T_i \equiv \begin{pmatrix} \alpha_i & 0 \\ 0 & \beta_i \end{pmatrix},$$

with

$$(83) \quad \alpha_i \equiv L_{k_i-1} \quad (\text{see (63)}) \quad \text{and} \quad \beta_i \equiv h_{t_i} = h_\lambda \quad (\text{see (64)}).$$

We define $\hat{\omega}_i$ by $\hat{\omega}_i = B_i^{-1}\omega_i$. It is easy to see that if $S_i \in \Omega$, $\hat{\omega}_i$ is the convex hull of $(-r, 0), (0, -1), (1, -1), (1, 0), (0, 1), (-r, 1)$, where $r = \frac{1}{\lambda}$.

Calling $\hat{v} \equiv v \circ B_i$, we denote by \hat{P}_i the projector on the space of affine polynomials in $L^2(\hat{\omega}_i)$, we know that

$$(84) \quad \widehat{P_i v} = \hat{P}_i \hat{v}.$$

Indeed, the linear map $p \rightarrow \hat{p}$ is an isomorphism of $\mathcal{P}_i(\omega_i)$ onto $\mathcal{P}_i(\hat{\omega}_i)$, since B_i is affine; hence $\widehat{P_i v}$ is an affine polynomial. In addition, if $\hat{p} \in \mathcal{P}_i(\hat{\omega}_i)$, then

$$\int_{\hat{\omega}_i} (\hat{v} - \widehat{P_i v}) \hat{p} \, d\xi d\eta = (\alpha_i \beta_i)^{-1} \int_{\omega_i} (v - P_i v) p \, dx dy = 0.$$

- For $i \in \mathcal{J}_\lambda$ and for a given triangle $t \in \mathcal{T}$ contained in ω_i , let (x_t, y_t) be the center of t (satisfying (69)). Let T_t be the linear map

$$(85) \quad T_t \equiv \begin{pmatrix} \alpha_t & 0 \\ 0 & \beta_t \end{pmatrix},$$

where

$$(86) \quad \alpha_t \equiv 2h_t x_t y_t \quad \text{and} \quad \beta_t \equiv 2h_t.$$

We call $B_t = Z_t + T_t$, where $Z_t = (x_t - h_t x_t y_t, y_t - h_t)$. Note that B_t maps the unit square to the rectangle $R(x_t, y_t, h_t)$. We define \hat{t} by $\hat{t} = B_t^{-1}t$. From the assumptions on the triangulation, \hat{t} is a shape regular triangle containing a ball of diameter r (in the Euclidean metric), where r depends only μ , and being contained in the unit cube.

Calling $\hat{v} \equiv v \circ B_t$, we denote by $I_{\hat{t}}$ the Lagrange interpolation operator on \hat{t} , we have

$$(87) \quad \widehat{I_t v} = I_{\hat{t}} \hat{v}.$$

In the rest of this paper we shall estimate:

- For $i \in \mathcal{I}_\lambda$, the error norms $\|v - P_i v\|_{L^2(\omega_i)}$ and $|v - P_i v|_{S^{1,2}(\omega_i)}$ by rescaling the polygonal domains ω_i through the map B_i^{-1} . We shall distinguish the case “near $y = 0$ ” (i.e., the case $b_i = h_\lambda$), where the model case is described by function spaces associated with the vector fields ∂_y and $(1 + y)\partial_x$, the coefficient $1 + y$ vanishing at the lower boundary of the rescaled cell, and finally the case “on $y = 0$ ” (i.e., the case $b_i = 0$), where the model case is described by function spaces associated with the vector fields ∂_y and $y\partial_x$, the coefficient y vanishing inside the rescaled cell.

- For $i \in \mathcal{J}_\lambda$ and for a given triangle $t \in \mathcal{T}$ contained in ω_i , the error norms $\|v - I_t v\|_{L^2(t)}$ and $|v - I_t v|_{S^{1,2}(t)}$ by rescaling the triangle t through the map B_t^{-1} . Here we will be able to use standard results on usual elliptic Sobolev spaces.

More precisely, we have the following lemma.

Lemma 4. *There exists a positive constant C such that:*

- for any $i \in \mathcal{J}_\lambda$ and for a given triangle $t \in \mathcal{T}$ contained in ω_i , for any $v \in S^{2,2}(t)$,

$$(88) \quad \begin{aligned} & h_t^{-2} \|v - I_t v\|_{L^2(t)}^2 + |v - I_t v|_{S^{1,2}(t)}^2 \\ & \leq C \alpha_t \beta_t^{-1} \left(\int_{\hat{t}} |\hat{v} - I_{\hat{t}} \hat{v}|^2 d\xi d\eta + \int_{\hat{t}} \left| \frac{\partial}{\partial \xi} (\hat{v} - I_{\hat{t}} \hat{v}) \right|^2 d\xi d\eta \right. \\ & \quad \left. + \int_{\hat{t}} \left| \frac{\partial}{\partial \eta} (\hat{v} - I_{\hat{t}} \hat{v}) \right|^2 d\xi d\eta \right), \end{aligned}$$

and note that in this case, $\alpha_t \beta_t^{-1} = x_t y_t$;

- for any $i \in \mathcal{I}_\lambda$ with $b_i = h_\lambda$, and for any $v \in S^{1,2}(\omega_i)$,

$$(89) \quad \begin{aligned} & h_{t_i}^{-2} \|v - P_i v\|_{L^2(\omega_i)}^2 + |v - P_i v|_{S^{1,2}(\omega_i)}^2 \\ & \leq C \alpha_i \beta_i^{-1} \left(\int_{\hat{\omega}_i} |\hat{v} - \hat{P}_i \hat{v}|^2 d\xi d\eta + \int_{\hat{\omega}_i} |(1 + \eta) \frac{\partial}{\partial \xi} (\hat{v} - \hat{P}_i \hat{v})|^2 d\xi d\eta \right. \\ & \quad \left. + \int_{\hat{\omega}_i} \left| \frac{\partial}{\partial \eta} (\hat{v} - \hat{P}_i \hat{v}) \right|^2 d\xi d\eta \right); \end{aligned}$$

- for any $i \in \mathcal{I}_\lambda$ with $b_i = 0$, and for any $v \in S^{1,2}(\omega_i)$,

$$(90) \quad \begin{aligned} & h_{t_i}^{-2} \|v - P_i v\|_{L^2(\omega_i)}^2 + |v - P_i v|_{S^{1,2}(\omega_i)}^2 \\ & \leq C \alpha_i \beta_i^{-1} \left(\int_{\hat{\omega}_i} |\hat{v} - \hat{P}_i \hat{v}|^2 d\xi d\eta + \int_{\hat{\omega}_i} |\eta \frac{\partial}{\partial \xi} (\hat{v} - \hat{P}_i \hat{v})|^2 d\xi d\eta \right. \\ & \quad \left. + \int_{\hat{\omega}_i} \left| \frac{\partial}{\partial \eta} (\hat{v} - \hat{P}_i \hat{v}) \right|^2 d\xi d\eta \right). \end{aligned}$$

Proof. Let us consider first $i \in \mathcal{J}_\lambda$, a given triangle $t \in \mathcal{T}$ contained in ω_i , and $v \in S^{2,2}(t)$. Thanks to (87), we have that

$$(91) \quad h_t^{-2} \int_t |v - I_t v|^2 dx dy \approx \beta_t^{-2} \int_t |v - I_t v|^2 dx dy = \alpha_t \beta_t^{-1} \int_{\hat{t}} |\hat{v} - I_{\hat{t}} \hat{v}|^2 d\xi d\eta.$$

Moreover (remember that $\widehat{\frac{\partial}{\partial x}} f = \alpha_t^{-1} \frac{\partial}{\partial \xi} \hat{f}$ and $\widehat{\frac{\partial}{\partial y}} f = \beta_t^{-1} \frac{\partial}{\partial \eta} \hat{f}$), and calling $z_{t,1}, z_{t,2}$ the coordinates of Z_t (where Z_t is defined just after (86)),

$$(92) \quad \begin{aligned} & \int_t |X(v - I_t v)|^2 dx dy = \int_t |xy \frac{\partial}{\partial x} (v - I_t v)|^2 dx dy \\ & = \alpha_t^{-1} \beta_t \int_{\hat{t}} |(z_{t,1} + \alpha_t \xi)(z_{t,2} + \beta_t \eta) \frac{\partial}{\partial \xi} (\hat{v} - I_{\hat{t}} \hat{v})|^2 d\xi d\eta \\ & = \alpha_t \beta_t^3 \int_{\hat{t}} \left| \left(\frac{z_{t,1}}{\alpha_t} + \xi \right) \left(\frac{z_{t,2}}{\beta_t} + \eta \right) \frac{\partial}{\partial \xi} (\hat{v} - I_{\hat{t}} \hat{v}) \right|^2 d\xi d\eta. \end{aligned}$$

But $\frac{z_{t,1}}{\alpha_t} = \frac{x_t}{\alpha_t} - \frac{1}{2} = \frac{1}{2y_t h_t} - \frac{1}{2}$ and $\beta_t = 2h_t$ so $\frac{z_{t,1}}{\alpha_t} + \xi \approx \frac{1}{y_t \beta_t}$, because $0 \leq \xi \leq 1$. Also, $0 < \frac{z_{t,2}}{\beta_t} + \eta \approx \frac{y_t}{\beta_t}$ because $-\frac{1}{2} \leq \frac{y_t}{\beta_t} - \left(\frac{z_{t,2}}{\beta_t} + \eta \right) = \frac{1}{2} - \eta \leq \frac{1}{2}$, and from the

assumptions on the mesh, $\frac{y_t}{\beta_t}$ is greater than 1, which yields $\frac{1}{2}\frac{y_t}{\beta_t} \leq \frac{x_t}{\alpha_t} - \frac{1}{2} \leq 2\frac{y_t}{\beta_t}$. Therefore, $0 \leq (\frac{z_{t,1}}{\alpha_t} + \xi)(\frac{z_{t,2}}{\beta_t} + \eta) \approx \frac{1}{\beta_t^2}$, and

$$(93) \quad \int_t |X(v - I_tv)|^2 dx dy \approx \alpha_t \beta_t^{-1} \int_{\hat{\omega}_t} \left| \frac{\partial}{\partial \xi} (\hat{v} - I_{\hat{t}} \hat{v}) \right|^2 d\xi d\eta.$$

Also

$$(94) \quad \int_t |Y(v - I_tv)|^2 dx dy = \int_t \left| \frac{\partial}{\partial y} (v - I_tv) \right|^2 dx dy = \alpha_i \beta_t^{-1} \int_{\hat{\omega}_t} \left| \frac{\partial}{\partial \eta} (\hat{v} - I_{\hat{t}} \hat{v}) \right|^2 d\xi d\eta,$$

Estimate (88) is obtained from (91), (93), (94).

Now consider $i \in \mathcal{I}_\lambda$ and $v \in S^{1,2}(\omega_i)$: thanks to (84),

$$(95) \quad h_{t_i}^{-2} \int_{\omega_i} |v - P_i v|^2 dx dy \approx \beta_i^{-2} \int_{\omega_i} |v - P_i v|^2 dx dy = \alpha_i \beta_i^{-1} \int_{\hat{\omega}_i} |\hat{v} - \hat{P}_i \hat{v}|^2 d\xi d\eta.$$

In addition, since $\widehat{\frac{\partial}{\partial x}} f = \alpha_i^{-1} \frac{\partial}{\partial \xi} \hat{f}$ and $\widehat{\frac{\partial}{\partial y}} f = \beta_i^{-1} \frac{\partial}{\partial \eta} \hat{f}$, and calling $z_{i,1}, z_{i,2}$ the coordinates of Z_i , we have

$$(96) \quad \begin{aligned} \int_{\omega_i} |X(v - P_i v)|^2 dx dy &= \int_{\omega_i} \left| xy \frac{\partial}{\partial x} (v - P_i v) \right|^2 dx dy \\ &= \alpha_i^{-1} \beta_i \int_{\hat{\omega}_i} |(z_{i,1} + \alpha_i \xi)(z_{i,2} + \beta_i \eta) \frac{\partial}{\partial \xi} (\hat{v} - \hat{P}_i \hat{v})|^2 d\xi d\eta \\ &= \alpha_i \beta_i^3 \int_{\hat{\omega}_i} \left| \left(\frac{z_{i,1}}{\alpha_i} + \xi \right) \left(\frac{z_{i,2}}{\beta_i} + \eta \right) \frac{\partial}{\partial \xi} (\hat{v} - \hat{P}_i \hat{v}) \right|^2 d\xi d\eta. \end{aligned}$$

Now, by definition, $\frac{z_{i,1}}{\alpha_i} = \frac{1}{\lambda-1} = \frac{1}{h_\lambda^2} = \frac{1}{\beta_i^2}$; therefore, in the limit $h_\lambda \rightarrow 0$, $\frac{z_{i,1}}{\alpha_i} + \xi \approx \frac{1}{\beta_i^2}$.

Concerning the term $\frac{z_{i,2}}{\beta_i} + \eta$, we have

- (1) if $b_i = 0$, $\frac{z_{i,2}}{\beta_i} + \eta = \eta$,
- (2) if $b_i = h_\lambda$, $\frac{z_{i,2}}{\beta_i} + \eta = 1 + \eta$.

This yields that

- (1) if $b_i = 0$, $(\frac{z_{i,1}}{\alpha_i} + \xi)(\frac{z_{i,2}}{\beta_i} + \eta) \approx \frac{\eta}{\beta_i^2}$,
- (2) if $b_i = h_\lambda$, $(\frac{z_{i,1}}{\alpha_i} + \xi)(\frac{z_{i,2}}{\beta_i} + \eta) \approx \frac{1+\eta}{\beta_i^2}$.

After simple calculations,

- (1) if $b_i = 0$,

$$(97) \quad \int_{\omega_i} |X(v - P_i v)|^2 dx dy \approx \alpha_i \beta_i^{-1} \int_{\hat{\omega}_i} \left| \eta \frac{\partial}{\partial \xi} (\hat{v} - \hat{P}_i \hat{v}) \right|^2 d\xi d\eta.$$

- (2) if $b_i = h_\lambda$,

$$(98) \quad \int_{\omega_i} |X(v - P_i v)|^2 dx dy \approx \alpha_i \beta_i^{-1} \int_{\hat{\omega}_i} \left| (1 + \eta) \frac{\partial}{\partial \xi} (\hat{v} - \hat{P}_i \hat{v}) \right|^2 d\xi d\eta.$$

Finally,

$$(99) \quad \begin{aligned} &\int_{\omega_i} |Y(v - P_i v)|^2 dx dy \\ &= \int_{\omega_i} \left| \frac{\partial}{\partial y} (v - P_i v) \right|^2 dx dy = \alpha_i \beta_i^{-1} \int_{\hat{\omega}_i} \left| \frac{\partial}{\partial \eta} (\hat{v} - \hat{P}_i \hat{v}) \right|^2 d\xi d\eta. \end{aligned}$$

The lemma is proved. □

To estimate the right-hand side of (88), we recall the following result.

Lemma 5. *There exists a positive constant C depending only on μ such that, for any $i \in \mathcal{J}_\lambda$, and for a given triangle $t \in \mathcal{T}$ contained in ω_i , we have*

$$(100) \quad \int_{\hat{t}} |\hat{v} - I_{\hat{t}} \hat{v}|^2 d\xi d\eta + \int_{\hat{t}} \left| \frac{\partial}{\partial \xi} (\hat{v} - I_{\hat{t}} \hat{v}) \right|^2 d\xi d\eta + \int_{\hat{t}} \left| \frac{\partial}{\partial \eta} (\hat{v} - I_{\hat{t}} \hat{v}) \right|^2 d\xi d\eta \leq C \left(\int_{\hat{t}} \left| \frac{\partial^2}{\partial \xi^2} \hat{v} \right|^2 d\xi d\eta + \int_{\hat{t}} \left| \frac{\partial^2}{\partial \xi \partial \eta} \hat{v} \right|^2 d\xi d\eta + \int_{\hat{t}} \left| \frac{\partial^2}{\partial \eta^2} \hat{v} \right|^2 d\xi d\eta \right)$$

for any $\hat{v} \in W^{2,2}(\hat{t})$ (the usual Sobolev space of order 2 in $L^2(\hat{t})$).

Proposition 7. *There exist positive constants C and $H > 0$ such that for all $i \in \mathcal{J}_\lambda$, for a given triangle $t \in \mathcal{T}$ contained in ω_i , and for all $v \in S^{2,2}(t)$,*

$$(101) \quad h_t^{-2} \|v - I_t v\|_{L^2(t)}^2 + \|v - I_t v\|_{S^{1,2}(t)}^2 \leq C h_t^2 \|v\|_{S^{2,2}(t)}^2,$$

if $h_t < H$.

Proof. To prove (101), we must estimate the terms in the right-hand side of (100) by integrals in t , repeating backward the arguments used to obtain (88).

Keeping in mind that for $(\xi, \eta) \in \hat{t}$, we have that $(z_{t,1} + \alpha_t \xi)(z_{t,2} + \beta_t \eta) \approx \frac{\alpha_t}{\beta_t}$, and that the Jacobian of T_t is $\alpha_t \beta_t$, we get

$$(102) \quad \int_{\hat{t}} \left| \frac{\partial^2}{\partial \xi^2} \hat{v} \right|^2 d\xi d\eta \approx \frac{\beta_t^4}{\alpha_t^4} \int_{\hat{t}} |(z_{t,1} + \alpha_t \xi)^2 (z_{t,2} + \beta_t \eta)^2 \frac{\partial^2}{\partial \xi^2} \hat{v}|^2 d\xi d\eta = \alpha_t^{-1} \beta_t^3 \int_t \left| x^2 y^2 \frac{\partial^2}{\partial x^2} v \right|^2 dx dy.$$

In addition, we have

$$(103) \quad \int_{\hat{t}} \left| \frac{\partial^2}{\partial \eta^2} \hat{v} \right|^2 d\xi d\eta = \alpha_t^{-1} \beta_t^3 \int_t \left| \frac{\partial^2}{\partial y^2} v \right|^2 dx dy.$$

Finally,

$$(104) \quad \int_{\hat{t}} \left| \frac{\partial^2}{\partial \xi \partial \eta} \hat{v} \right|^2 d\xi d\eta \approx \frac{\beta_t^2}{\alpha_t^2} \int_{\hat{t}} |(z_{t,1} + \alpha_t \xi)(z_{t,2} + \beta_t \eta) \frac{\partial^2}{\partial \xi \partial \eta} \hat{v}|^2 d\xi d\eta = \alpha_t^{-1} \beta_t^3 \int_t \left| xy \frac{\partial^2}{\partial x \partial y} v \right|^2 dx dy.$$

Now combining (88), (100), (102), (103), (104), and using the fact that $\beta_t = 2h_t$, we get that

$$(105) \quad h_t^{-2} \|v - I_t v\|_{L^2(t)}^2 + \|v - I_t v\|_{S^{1,2}(t)}^2 \leq C h_t^2 \left(\int_t \left| x^2 y^2 \frac{\partial^2}{\partial x^2} v \right|^2 dx dy + \int_t \left| \frac{\partial^2}{\partial y^2} v \right|^2 dx dy + \int_t \left| xy \frac{\partial^2}{\partial x \partial y} v \right|^2 dx dy \right).$$

We can write the above estimate better in terms of the vector fields X and Y as follows. Clearly, $\frac{\partial^2}{\partial y^2} = Y^2$ and $xy \frac{\partial^2}{\partial x \partial y} = XY$. On the other hand, notice that

$$X^2 v = xy \frac{\partial}{\partial x} \left(xy \frac{\partial v}{\partial x} \right) = x^2 y^2 \frac{\partial^2 v}{\partial x^2} + xy^2 \frac{\partial v}{\partial x},$$

so that

$$(106) \quad \int_t |x^2 y^2 \frac{\partial^2}{\partial x^2} v|^2 dx dy \leq 2 \|X^2 v\|_{L^2(t)}^2 + 2 \|Xv\|_{L^2(t)}^2.$$

Hence, if we choose in (105) $h_t < H$, with H such that

$$2C H^2 < \frac{1}{2},$$

we get eventually (101). □

We must now consider the case $i \in \mathcal{I}_\lambda$ and $b_i = h_\lambda$ in order to prove an estimate like (101) in this case. If we denote by q_i the smallest rectangle whose sides are aligned with the axes containing ω_i , it is much simpler to provide an estimate where in the right-hand side the norms over ω_i are replaced by norms over q_i , and this gives no trouble, since the sets q_i have bounded overlaps, as the sets ω_i do. Again as above, we set

$$q_i = B_i(\hat{q}_i).$$

If $\omega_i \subset \Omega$, and hence $q_i \subset \Omega$, then $\hat{q}_i \subset \mathbb{R}_{\xi, \eta}^2$ is the rectangle $(-r, 1) \times (-1, 1)$. The only difference in the case $b_i = h_\lambda$ is that (88) is replaced by (89), and then we have to prove the following lemma, analogous to Lemma 5.

Lemma 6. *There exists a positive constant C such that, for any $i \in \mathcal{I}_\lambda$ with $b_i = h_\lambda$,*

$$(107) \quad \begin{aligned} & \int_{\hat{\omega}_i} |\hat{v} - \hat{P}_i \hat{v}|^2 d\xi d\eta + \int_{\hat{\omega}_i} |(1 + \eta) \frac{\partial}{\partial \xi} (\hat{v} - \hat{P}_i \hat{v})|^2 d\xi d\eta + \int_{\hat{\omega}_i} \left| \frac{\partial}{\partial \eta} (\hat{v} - \hat{P}_i \hat{v}) \right|^2 d\xi d\eta \\ & \leq C \left(\int_{\hat{q}_i} |(1 + \eta)^2 \frac{\partial^2}{\partial \xi^2} \hat{v}|^2 d\xi d\eta + \int_{\hat{q}_i} |(1 + \eta) \frac{\partial^2}{\partial \xi \partial \eta} \hat{v}|^2 d\xi d\eta \right) \\ & \quad \left(+ \int_{\hat{q}_i} \left| \frac{\partial}{\partial \xi} \hat{v} \right|^2 d\xi d\eta + \int_{\hat{q}_i} \left| \frac{\partial^2}{\partial \eta^2} \hat{v} \right|^2 d\xi d\eta \right) \end{aligned}$$

for any $\hat{v} \in \tilde{S}^{2,2}(\hat{q}_i)$.

Proof. We notice first that the left-hand side of (107) is (the square of) the norm of $\hat{v} - \hat{P}_i \hat{v}$ in $\tilde{S}^{1,2}(\hat{\omega}_i)$, and that the right-hand side is equivalent to $\|\tilde{X}^2 v\|_{L^2(\hat{q}_i)}^2 + \|Y^2 v\|_{L^2(\hat{q}_i)}^2 + \|\tilde{X}Yv\|_{L^2(\hat{q}_i)}^2 + \|Y\tilde{X}v\|_{L^2(\hat{q}_i)}^2$, since $Y\tilde{X}v = \partial_\xi + \tilde{X}Yv$.

We define by \hat{R}_i the continuous linear operator from $\tilde{S}^{2,2}(\hat{q}_i)$ to $\tilde{S}^{1,2}(\hat{\omega}_i)$, which maps $v \in \tilde{S}^{2,2}(\hat{q}_i)$ to the affine polynomial $\hat{P}_i(v)|_{\hat{\omega}_i}$.

Thus, we can apply Lemma 3 with $V = \tilde{S}^{2,2}(\hat{q}_i)$, $V_1 = \tilde{S}^{1,2}(\hat{q}_i)$, $V_2 = L^2(\hat{q}_i)^4$, $W = \tilde{S}^{1,2}(\hat{\omega}_i)$, $A_1 = Id$, $A_2 v = (\tilde{X}^2 v, Y^2 v, \tilde{X}Yv, Y\tilde{X}v)$, $L = Id - \hat{R}_i$.

Keeping in mind Theorem 4 together with Remark 9, to verify the assumptions of Lemma 3, we have only to check that L vanishes identically on the kernel of A_2 , and hence it is enough to show that a function v belonging to the kernel of A_2 is an affine polynomial. Since $Y^2 v \equiv 0$, then $v(\xi, \eta) = \eta a(\xi) + b(\xi)$. On the other hand, $0 \equiv \tilde{X}Yv \equiv \tilde{X}a$, so that $a'(\xi) \equiv 0$ and hence $a(\xi) \equiv a_0$ in \hat{q}_i . Moreover, $0 \equiv Y\tilde{X}v - \tilde{X}Yv \equiv \partial_\xi v \equiv b'(\xi)$, so b is a constant too, and the assertion follows. □

Proposition 8. *There exist positive constants C and $\lambda_0 > 1$ such that for all $i = 1, \dots, M = \#\mathcal{V}$ with $b_i = h_\lambda$, and for all $v \in S^{2,2}(q_i)$,*

$$(108) \quad h_{t_i}^{-2} \|v - P_i v\|_{L^2(\omega_i)}^2 + \|v - P_i v\|_{S^{1,2}(\omega_i)}^2 \leq C h_{t_i}^2 \|v\|_{S^{2,2}(q_i)}^2$$

if $1 < \lambda < \lambda_0$.

Proof. The proof will be carried out by arguments akin to those of Proposition 7. By definition, $b_i = \beta_i$. Keeping in mind that $a_i + \alpha_i \xi \approx \alpha_i \beta_i^{-2}$ and (106), we get

$$(109) \quad \begin{aligned} & \int_{\hat{q}_i} |(1+\eta)^2 \frac{\partial^2}{\partial \xi^2} \hat{v}|^2 d\xi d\eta = \alpha_i^4 \int_{\hat{q}_i} |(1+\eta)^2 \widehat{\frac{\partial^2}{\partial x^2} v}|^2 d\xi d\eta \\ & \approx \alpha_i^4 \cdot \alpha_i^{-4} \beta_i^8 \cdot \beta_i^{-4} \int_{\hat{q}_i} (a_i + \alpha_i \xi)^4 (b_i + \beta_i \eta)^4 \widehat{\left| \frac{\partial^2}{\partial x^2} v \right|^2} d\xi d\eta \\ & = \alpha_i^{-1} \beta_i^3 \int_{q_i} |x^2 y^2 \frac{\partial^2}{\partial x^2} v|^2 dx dy \\ & \leq 2\alpha_i^{-1} \beta_i^3 \left(\|X^2 v\|_{L^2(q_i)}^2 + \|Xv\|_{L^2(q_i)}^2 \right). \end{aligned}$$

Analogously,

$$(110) \quad \begin{aligned} & \int_{\hat{q}_i} |(1+\eta) \frac{\partial^2}{\partial \xi \partial \eta} \hat{v}|^2 d\xi d\eta = \alpha_i^2 \beta_i^2 \int_{\hat{q}_i} |(1+\eta) \widehat{\frac{\partial^2}{\partial x \partial y} v}|^2 d\xi d\eta \\ & \approx \alpha_i^2 \beta_i^2 \cdot \alpha_i^{-2} \beta_i^4 \cdot \beta_i^{-2} \int_{\hat{q}_i} (a_i + \alpha_i \xi)^2 (b_i + \beta_i \eta)^2 \widehat{\left| \frac{\partial^2}{\partial x \partial y} v \right|^2} d\xi d\eta \\ & \approx \alpha_i^{-1} \beta_i^3 \int_{q_i} |xy \frac{\partial^2}{\partial x \partial y} v|^2 dx dy = \alpha_i^{-1} \beta_i^3 \|XYv\|_{L^2(q_i)}^2. \end{aligned}$$

Moreover, from the identity $x\partial_x = YX - XY$,

$$(111) \quad \begin{aligned} & \int_{\hat{q}_i} \left| \frac{\partial}{\partial \xi} \hat{v} \right|^2 d\xi d\eta = \alpha_i^2 \int_{\hat{q}_i} \left| \widehat{\frac{\partial}{\partial x} v} \right|^2 d\xi d\eta \\ & \approx \alpha_i^2 \cdot \alpha_i^{-2} \beta_i^4 \int_{\hat{q}_i} (a_i + \alpha_i \xi)^2 \widehat{\left| \frac{\partial}{\partial x} v \right|^2} d\xi d\eta \\ & \approx \alpha_i^{-1} \beta_i^3 \int_{q_i} \left| x \frac{\partial}{\partial x} v \right|^2 dx dy \\ & \leq 2\alpha_i^{-1} \beta_i^3 \left(\|YXv\|_{L^2(q_i)}^2 + \|XYv\|_{L^2(q_i)}^2 \right). \end{aligned}$$

Eventually, as in (103),

$$(112) \quad \int_{\hat{q}_i} \left| \frac{\partial^2}{\partial \eta^2} \hat{v} \right|^2 d\xi d\eta = \alpha_i^{-1} \beta_i^3 \int_{q_i} \left| \frac{\partial^2}{\partial y^2} v \right|^2 dx dy = \alpha_i^{-1} \beta_i^3 \|Y^2 v\|_{L^2(q_i)}^2.$$

Since $\beta_i = h_{t_i} = h_\lambda$, the conclusion follows by combining (89), (107), and the four estimates above. As in the conclusion of Proposition 7, we can get rid of the first order term $\|Xv\|_{L^2(q_i)}$ by taking $\lambda < \lambda_0$, λ_0 sufficiently small. \square

Finally, the same arguments used in the proofs of Lemma 6 and Proposition 8 yield the following corresponding estimates for the case $b_i = 0$, calling q_i the smallest rectangle with sides parallel to the axes and containing ω_i and $\hat{q}_i = B_i^{-1} q_i$.

Lemma 7. *There exists a positive constant C such that, for any $i = \mathcal{I}$ with $b_i = 0$,*

$$(113) \quad \int_{\hat{\omega}_i} |\hat{v} - \hat{P}_i \hat{v}|^2 d\xi d\eta + \int_{\hat{\omega}_i} \left| \eta \frac{\partial}{\partial \xi} (\hat{v} - \hat{P}_i \hat{v}) \right|^2 d\xi d\eta + \int_{\hat{\omega}_i} \left| \frac{\partial}{\partial \eta} (\hat{v} - \hat{P}_i \hat{v}) \right|^2 d\xi d\eta \\ \leq C \left(\int_{\hat{q}_i} \left| \eta^2 \frac{\partial^2}{\partial \xi^2} \hat{v} \right|^2 d\xi d\eta + \int_{\hat{q}_i} \left| \eta \frac{\partial^2}{\partial \xi \partial \eta} \hat{v} \right|^2 d\xi d\eta \right) \\ + \int_{\hat{q}_i} \left| \frac{\partial}{\partial \xi} \hat{v} \right|^2 d\xi d\eta + \int_{\hat{q}_i} \left| \frac{\partial^2}{\partial \eta^2} \hat{v} \right|^2 d\xi d\eta$$

for any $\hat{v} \in \tilde{S}^{2,2}(\hat{q}_i)$.

Hence we can derive the following a priori error estimate.

Proposition 9. *There exist positive constants C and $\lambda_0 > 1$ such that for all $i = 1, \dots, M = \#\mathcal{V}$ with $b_i = 0$, and for all $v \in S^{2,2}(q_i)$,*

$$(114) \quad h_{t_i}^{-2} \|v - P_i v\|_{L^2(\omega_i)}^2 + |v - P_i v|_{S^{1,2}(\omega_i)}^2$$

if $1 < \lambda < \lambda_0$.

We are able now to prove the main error estimates.

Theorem 5. *There exists a positive constant C such that, if $\max_{t \in \mathcal{T}} h_t$ is small enough, for all functions, $v \in S^{2,2}(\Omega)$,*

$$(115) \quad \|v - \Pi v\|_{S^{1,2}(\Omega)}^2 \leq C \max_{t \in \mathcal{T}} h_t^2 \left(\|X^2 v\|_{L^2(\Omega)}^2 + \|Y^2 v\|_{L^2(\Omega)}^2 + \|XY v\|_{L^2(\Omega)}^2 \right).$$

Proof. The first step is to estimate $\|v - \Pi v\|_{L^2(\Omega)}$. Let t be a triangle of \mathcal{T} , and let $S_{i_1}, S_{i_2}, S_{i_3}$ be its vertices. For $j = 1, 2, 3$, if $i_j \in \mathcal{I}_h$, we call Q_j the operator P_{i_j} . On the contrary, if $i_j \in \mathcal{J}_h$, we call Q_j the Lagrange interpolation operator obtained by assembling the operators $I_t, t \in \mathcal{T}, t \subset \omega_i$.

We have

$$\|v - \Pi v\|_{L^2(t)} = \|v - \sum_{j=1}^3 (Q_j v)(S_j) \phi_j\|_{L^2(t)} \\ \leq \|v - Q_1 v\|_{L^2(t)} + \|Q_1 v - \sum_{j=1}^3 (Q_j v)(S_j) \phi_j\|_{L^2(t)} \\ \leq \|v - Q_1 v\|_{L^2(t)} + \sum_{j=1}^3 \|(Q_1 v - Q_j v)(S_j) \phi_j\|_{L^2(t)} \\ \leq \|v - Q_1 v\|_{L^2(t)} + \sum_{j=1}^3 \|Q_1 v - Q_j v\|_{L^\infty(t)} \|\phi_j\|_{L^2(t)}.$$

We define q_j as

- $q_j = t$ if $i_j \in \mathcal{J}_\lambda$,
- q_j is the smallest rectangle with sides parallel to the axes containing ω_{i_j} in the opposite case.

It is clear from Propositions 7, 8, and 9 that $\|v - Q_1 v\|_{L^2(t)} \leq Ch_t^2 |v|_{S^{2,2}(q_1)}$.

Also $\|\phi_j\|_{L^2(t)} = \left(\frac{|t|}{6}\right)^{\frac{1}{2}}$, where $|t|$ is the Lebesgue measure of t . On the other hand, since $(Q_1 - Q_j)v$ is affine on t , we have the inverse inequality

$\|Q_1v - Q_jv\|_{L^\infty(t)} \leq C|t|^{-\frac{1}{2}}\|Q_1v - Q_jv\|_{L^2(t)}$, where C does not depend on t .
Therefore

$$\begin{aligned} \|Q_1v - Q_jv\|_{L^\infty(t)}\|\phi_j\|_{L^2(t)} &\leq C\|Q_1v - Q_jv\|_{L^2(t)} \\ &\leq C(\|Q_1v - v\|_{L^2(t)} + \|Q_jv - v\|_{L^2(t)}). \end{aligned}$$

Again by Propositions 7, 8, and 9 and using the fact that the triangles having vertices in the strip $|y| < 2h_\lambda$ have a size h_t of order h_λ , we have that

$$\sum_{j=1}^3 \|Q_1v - Q_jv\|_{L^\infty(t)}\|\phi_j\|_{L^2(t)} \leq Ch_t^2 \sum_{j=1}^3 |v|_{S^{2,2}(q_j)}$$

(the constant depends on μ), which yields finally the same estimate for $\|v - \Pi v\|_{L^2(t)}$.
Summing on all triangles, and keeping in mind that every triangle of \mathcal{T} is contained in a bounded number of subdomains q_j , we obtain that

$$(116) \quad \|v - \Pi v\|_{L^2(\Omega)} \leq C \max_{t \in \mathcal{T}} h_t^2 |v|_{S^{2,2}(\Omega)}.$$

In order to estimate $\|Y(v - \Pi v)\|_{L^2(\Omega)}$, we proceed in the same manner: we obtain exactly as above that

$$\|Y(v - \Pi v)\|_{L^2(t)} \leq \|Y(v - Q_1v)\|_{L^2(t)} + \sum_{j=1}^3 \|Q_1v - Q_jv\|_{L^\infty(t)}\|Y\phi_j\|_{L^2(t)}.$$

We know that $\|Y\phi_j\|_{L^2(t)} \leq \frac{1}{h_t}\|\phi_j\|_{L^2(t)}$, since h_t is the size of t (with respect to the Euclidean metric) in the y direction, so exactly as above, we obtain that

$$\|Y(v - \Pi v)\|_{L^2(t)} \leq \|Y(v - Q_1v)\|_{L^2(t)} + \frac{C}{h_t} \sum_{j=1}^3 \|v - Q_jv\|_{L^2(t)}.$$

By Propositions 7, 8, and 9 and using the fact that each triangle in the strip $|y| \leq 2h_\lambda$ has a size of order h_λ , we obtain that

$$\|Y(v - \Pi v)\|_{L^2(t)} \leq Ch_t \sum_{j=1}^3 |v|_{S^{2,2}(q_j)}.$$

By summing over the triangles, we obtain the global estimate

$$(117) \quad \|Y(v - \Pi v)\|_{L^2(\Omega)} \leq C \max_{t \in \mathcal{T}} h_t |v|_{S^{2,2}(\Omega)}.$$

In order to estimate $\|X(v - \Pi v)\|_{L^2(\Omega)}$, we obtain exactly as above that

$$\|X(v - \Pi v)\|_{L^2(t)} \leq \|X(v - Q_1v)\|_{L^2(t)} + \sum_{j=1}^3 \|Q_1v - Q_jv\|_{L^\infty(t)}\|X\phi_j\|_{L^2(t)},$$

and arguing as for (97), (98), and (93), we see that $\|X\phi_j\|_{L^2(t)}^2 \leq C\alpha_j\beta_j^{-1} \approx Ch_t^{-2}|t|$. Thus we can conclude as for (117) that

$$(118) \quad \|X(v - \Pi v)\|_{L^2(\Omega)} \leq C \max_{t \in \mathcal{T}} h_t |v|_{S^{2,2}(\Omega)}.$$

This achieves the proof. □

5.3. Consequence: Error estimate for the finite element method. We can now estimate the error in the finite element method.

Theorem 6. *Let u be the solution of problem (58), and assume that $u \in S^{2,2}(\Omega)$. There exists a positive constant C such that, if u_λ is the solution of the discrete problem (73), then*

$$(119) \quad \|u - u_\lambda\|_{S^{1,2}(\Omega)} \leq C \max_{t \in \mathcal{T}} h_t |u|_{S^{2,2}(\Omega)}.$$

Proof. Use (74) and Theorem 5. □

Remark 10. The a priori error estimate (119) is analogous to the one obtained for standard elliptic operators for conforming piecewise linear finite element with a regular family of meshes \mathcal{T}_h (with respect to the Euclidean metric), where h is the maximal diameter of the triangles of \mathcal{T}_h , i.e., $\|u - u_h\|_{H^1(\Omega)} \leq Ch|u|_{H^2(\Omega)}$.

Remark 11. If we consider other boundary value problems obtained by replacing the Neumann boundary conditions in (50) by Dirichlet or Robin conditions, the finite element method described above can be applied with standard modifications, and the error estimate (119) holds, as soon as the solution of the continuous problem is smooth enough.

Remark 12. For the parabolic problem obtained by replacing the first line of (50) by $\frac{\partial u}{\partial t} - y^2 x^2 \frac{\partial^2 u}{\partial x^2} - \frac{\partial^2 u}{\partial y^2} = f$, or by $\frac{\partial u}{\partial t} - A_t u = f$, with A_t given in (11), it is possible to use standard time integration schemes like Euler's implicit scheme or the Crank–Nicolson scheme coupled with the finite element method described above. The a priori error estimates are deduced in a standard manner from those obtained for the elliptic problem.

Remark 13. Concerning the Cauchy problem (40) from which we departed (the parabolic equation is posed in the unbounded domain Q), one strategy is to localize first the problem to a rectangle $(m_x, M_x) \times (0, M_y)$. One has to find suitable boundary conditions on the artificial boundaries. This is discussed qualitatively in [1], and more quantitatively by A.M. Matache and Christopher Schwab in [15]. Then one can use the finite element method described above for the problem in the truncated domain. Of course, the error comes from both the artificial boundary conditions and the discrete method. The error due to artificial boundary conditions is analyzed in [15].

ACKNOWLEDGMENT

The first author wishes to thank Professor Christopher Schwab for very helpful discussions.

REFERENCES

1. Y. Achdou and N. Tchou, *Variational analysis for the Black and Scholes equation with stochastic volatility*, M2AN Math. Model. Numer. Anal. **36** (2002), no. 3, 373–395. MR2003g:91050
2. S.C. Brenner and L. R. Scott, *The mathematical theory of finite element methods*, Texts in Applied Mathematics, vol. 15, Springer-Verlag, New York, 1994. MR95f:65001
3. P. G. Ciarlet, *The finite element method for elliptic problems*, North-Holland Publishing Co., Amsterdam, 1978, Studies in Mathematics and its Applications, Vol. 4. MR58:25001
4. ———, *Basic error estimates for elliptic problems*, Handb. Numer. Anal., II, North-Holland, Amsterdam, 1991, pp. 17–351.

5. Ph. Clément, *Approximation by finite element functions using local regularization*, Rev. Française Automat. Informat. Recherche Opérationnelle Sér. Rouge Anal. Numér. **9** (1975), no. R-2, 77–84. MR53:4569
6. M. Derridj and J.-P. Dias, *Le problème de Dirichlet pour une classe d'opérateurs non linéaires*, J. Math. Pures Appl. (9) **51** (1972), 219–230. MR548015
7. M. Derridj and C. Zuily, *Régularité C^∞ à la frontière, d'opérateurs dégénérés*, C. R. Acad. Sci. Paris Sér. A-B **271** (1970), A786–A788. MR43:2347
8. C. Fefferman and D. H. Phong, *Subelliptic eigenvalue problems*, Conference on harmonic analysis in honor of Antoni Zygmund, Vol. I, II (Chicago, Ill., 1981), Wadsworth Math. Ser., Wadsworth, Belmont, CA, 1983, pp. 590–606. MR86c:35112
9. J.-P. Fouque, G. Papanicolaou, and K. R. Sircar, *Derivatives in financial markets with stochastic volatility*, Cambridge University Press, Cambridge, 2000. MR2002g:91082
10. B. Franchi, *Weighted Sobolev-Poincaré inequalities and pointwise estimates for a class of degenerate elliptic equations*, Trans. Amer. Math. Soc. **327** (1991), no. 1, 125–158. MR91m:35095
11. B. Franchi and E. Lanconelli, *Hölder regularity theorem for a class of linear nonuniformly elliptic operators with measurable coefficients*, Ann. Scuola Norm. Sup. Pisa Cl. Sci. (4) **10** (1983), no. 4, 523–541. MR85k:35094
12. B. Franchi, R. Serapioni, and F. Serra Cassano, *Meyers-Serrin type theorems and relaxation of variational integrals depending on vector fields*, Houston J. Math. **22** (1996), no. 4, 859–890. MR98c:49037
13. B. Franchi and M. C. Tesi, *A finite element approximation for a class of degenerate elliptic equations*, Math. Comp. **69** (2000), no. 229, 41–63. MR2000i:65184
14. K. O. Friedrichs, *The identity of weak and strong extensions of differential operators*, Trans. Amer. Math. Soc. **55** (1944), 132–151. MR5:188b
15. N. Hilber, A.M. Matache, and C. Schwab, *Sparse wavelets methods for option pricing under stochastic volatility*, Seminar for Applied Mathematics, ETH Zurich, 2004
16. A. Nagel, E.M. Stein, and S. Wainger, *Balls and metrics defined by vector fields. I. Basic properties*, Acta Math. **155** (1985), no. 1-2, 103–147. MR86k:46049
17. A. Pazy, *Semigroups of linear operators and applications to partial differential equations*, Applied Mathematical Sciences, vol. 44, Springer-Verlag, New York, 1983. MR85g:47061
18. E. Stein and J. Stein, *Stock price distributions with stochastic volatility: an analytic approach*, The Review of Financial Studies **4** (1991), no. 4, 727–752.
19. P. Willmott, J. Dewynne, and J. Howison, *Option pricing: mathematical models and computations*, Oxford Financial Press, 1993.

UFR MATHÉMATIQUES, UNIVERSITÉ PARIS 7, 2 PLACE JUSSIEU, 75251 PARIS CEDEX 05, FRANCE;
AND LABORATOIRE J.L. LIONS, UNIVERSITÉ PARIS 6, 4 PLACE JUSSIEU, 75252 PARIS CEDEX 05,
FRANCE

E-mail address: `achdou@math.jussieu.fr`

DIPARTIMENTO DI MATEMATICA, UNIVERSITÀ DI BOLOGNA, PIAZZA DI PORTA S. DONATO, 5,
40126 BOLOGNA, ITALY

E-mail address: `bfranchi@dm.unibo.it`

IRMAR, UNIVERSITÉ DE RENNES 1, RENNES, FRANCE

E-mail address: `nicoletta.tchou@univ-rennes1.fr`