# Convergence of a numerical scheme for a nonlinear oblique derivative boundary value problem

Florian Méhats

MIP, Laboratoire CNRS (UMR 5640), Université Paul Sabatier,
118, route de Narbonne, 31062 Toulouse Cedex 04, France

mehats@mip.ups-tlse.fr

**Abstract**

We present here a discretization of a nonlinear oblique derivative boundary value problem for the heat equation in dimension two. This finite difference scheme takes advantages of the structure of the boundary condition, which can be reinterpreted as a Burgers equation in the space variables. This enables to obtain an energy estimate and to prove the convergence of the scheme. We also provide some numerical simulations of this problem and a numerical study of the stability of the scheme, which appears to be in good agreement with the theory.

# 1 Introduction: the continuous problem

This paper deals with the discretization and the numerical simulation of a nonlinear oblique boundary condition coupled with the heat equation. Let $\Omega$ be a two-dimensional domain of boundary $\Gamma$. If we denote respectively by $\tau$ and $\nu$ the local tangential and inner normal vectors on $\gamma$, this boundary condition writes

$$\partial_\nu U = KU\partial_\tau U, \tag{1.1}$$

where $K$ is a given positive constant.

This problem arises in the modelling of the plasma opening switch [8] and has been studied mathematically in [9], in the case $\Omega = \{(z,x) \in \mathbb{R}^2 \ : \ x > 0\}$ and $\Gamma = \{(z,x) \in \mathbb{R}^2 \ : \ x = 0\}$. The physical problem modelled by this system is the diffusion of a magnetic field inside a plasma in contact with perfect conductors. In this context a fast propagation of the magnetic field can be observed near the contact surface. In [9] a qualitative description of this fast propagation has been achieved and in [1] an analogy with the porous medium equation has been proved. The motivation of the numerical work presented here was to complete these studies by a quantitative description of the fast propagation.

More precisely, let $\Omega$ be the rectangle

$$\Omega = \{0 < z < L, \quad 0 < x < l\}.$$

The four parts of its boundary $\Gamma$ are labelled as follows:

$$\Gamma = \Gamma_1 \cup \Gamma_2 \cup \Gamma_3 \cup \Gamma_4,$$

$$\Gamma_1 = \{0 \le z \le L, \quad x = 0\}, \quad \Gamma_2 = \{z = 0, \quad 0 \le x \le l\},$$
$$\Gamma_3 = \{0 \le z \le L, \quad x = l\}, \quad \Gamma_4 = \{z = L, \quad 0 \le x \le l\}.$$

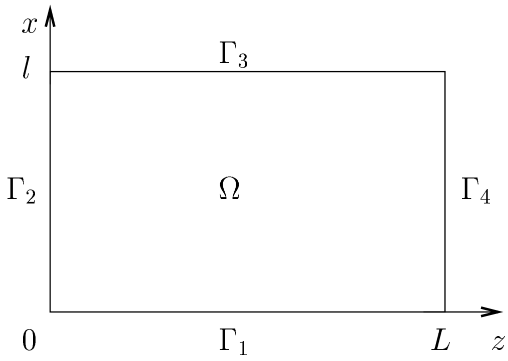


Figure 1 : Calculation domain

Consider the following scalar problem on $\Omega$:

$$\begin{cases} \partial_t U - \Delta U = 0 & (\Omega) \quad (1.2) \\ \partial_x U = KU\partial_z U & (\Gamma_1) \quad (1.3) \\ U = 1 & (\Gamma_2) \quad (1.4) \\ \partial_x U = 0 & (\Gamma_3) \quad (1.5) \\ U = 0 & (\Gamma_4) \quad (1.6) \\ U(0,z,x) = U^0(z,x) & (\Omega) \quad (1.7) \end{cases}$$

The non homogenous Dirichlet boundary condition (1.4) represents a source of magnetic field on $\Gamma_2$; there is no source of magnetic field on $\Gamma_4$ (this is modelled by an homogeneous Dirichlet condition) and the Neumann condition on $\Gamma_3$ is an open boundary condition.

This paper is organized as follows. In Section 2, we propose a discretization of the Cauchy problem (1.2)–(1.7) by a finite difference scheme, remarking that the nonlinear condition (1.3) writes as a Burgers equation in the variables $x$ and $z$. Our numerical scheme is a coupling between a standard five points scheme for the heat equation and the Godunov method for this conservation law.

This choice enables to show that the scheme preserves the discrete maximum principle and is stable in $\mathrm{H}^1$, under some constraints on the gridsteps and on the initial data. These estimates are obtained in Section 3. Hence we deduce the convergence of the scheme and obtain a weak solution for the continuous problem (1.2)–(1.7), in Section 4. This existence result is completed by a proof of uniqueness.

Finally, in Section 5, we give some numerical results, which highlight the fast propagation of the field at the boundary. In this section, we also give a numerical study of the stability of the scheme, which is in agreement with the constraints on the gridsteps found in Section 2.

Let us now make precise the functional framework of this work. We obtain the existence and uniqueness of weak solutions in $\mathrm{L}^2((0,T),\mathrm{H}^1(\Omega)) \cap \mathrm{L}^\infty((0,T)\times\Omega)$. The study of a better regularity goes beyond the scope of this paper. The case when $\Omega$ is the half-plane was extensively studied in [9]: in this case the solution is $C^\infty$. Moreover, if the domain $\Omega$ is bounded and regular then the results of [2, 10] apply and indicate that the solution is regular.

The most natural weak formulation of (1.2)–(1.7) is

$$\frac{d}{dt}\int_\Omega U\,\varphi + \int_\Omega \nabla U.\nabla\varphi + \frac{K}{2}\int_{\Gamma_1} \partial_z(U^2)\,\varphi = 0, \tag{1.8}$$

where $\varphi$ is in the space $V = \{\varphi \in \mathrm{H}^1(\Omega)$ such that $v$ vanishes on $\Gamma_2$ and $\Gamma_4\}$ and the boundary integral can be understood in the duality sense

$$\int_{\Gamma_1} \partial_z(U^2)\,\varphi = \langle \partial_z(U^2), \varphi\rangle_{\mathrm{H}^{-1/2}(\Gamma_1),\mathrm{H}^{1/2}(\Gamma_1)}$$

(recall that the direction $z$ is parallel to $\Gamma_1$). It is convenient to transform this nonlinear boundary term into a *volume* integral, as it is suggested in [7] for linear oblique derivative boundary value problems. Since boundary conditions are not symmetric on $\Gamma_1$ and $\Gamma_3$, we introduce the following auxiliary function:

$$\theta(x) = 1 - \frac{x}{l}.$$

Hence the boundary term can be written –at least for regular functions–

$$K\int_{\Gamma_1} U\partial_z U\,\varphi = K\int_\Omega U\,(\partial_x U\,\partial_z\varphi - \partial_z U\,\partial_x\varphi)\,\theta + \frac{K}{l}\int_\Omega U\,\partial_z U\,\varphi.$$

It gives a second weak formulation for (1.2)–(1.7):

$$\frac{d}{dt}\int_\Omega U\,\varphi + \int_\Omega \nabla U.\nabla\varphi + K\int_\Omega U\,(\partial_x U\,\partial_z\varphi - \partial_z U\,\partial_x\varphi)\,\theta + \frac{K}{l}\int_\Omega U\,\partial_z U\,\varphi = 0. \tag{1.9}$$

The advantage of this formulation is that it contains no boundary integral so it is easier to deal with.

# 2 The finite difference scheme

## 2.1 Notations

Let us introduce a regular grid on the domain $[0, T] \times \overline{\Omega}$, of time step $\Delta t$ and space steps $\Delta z$ and $\Delta x$:

$$\Delta t = \frac{T}{N_t}, \quad \Delta z = \frac{L}{N_z}, \quad \Delta x = \frac{l}{N_x}.$$

For $n \in \mathbb{N}$ we set $t^n = n\Delta t$. The domain $\overline{\Omega}$ is subdivised by the grid formed by the $(N_z + 1) \times (N_x + 1)$ gridpoints, which are labelled by $X_{i,j}$, for $0 \leq i \leq N_z$ and $0 \leq j \leq N_x$, and have the coordinates

$$(X_{i,j}) = (i\Delta z,\, j\Delta x)\,. \tag{2.1}$$

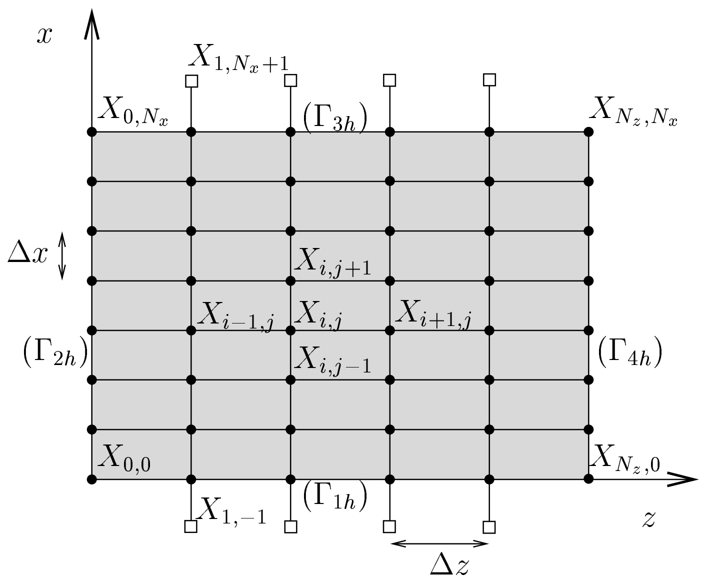


Figure 2 : The grid

The unknowns of the discrete problem are the values $U^n_{i,j}$, for $0 \leq i \leq N_z$, $0 \leq j \leq N_x$, $n \geq 0$. They approximate the gridfunction $U(X_{i,j}, t^n)$, where $U$ is the solution of the continuous problem (1.2)–(1.7).

In order to treat the Neumann and oblique type boundary conditions, we introduce some fictitious points outside the domain $\overline{\Omega}$ (the white squares on Figure 2). These are the points

$$(X_{i,-1})_{0<i<N_z}, \quad (X_{i,N_x+1})_{0<i<N_z}.$$

whose coordonates are also given by (2.1). The parameter of discretization is denoted by $h = (\Delta z, \Delta x, \Delta t)$ and we introduce the following subset of $\mathbb{N}^2$:

$$\Omega_h = \{(i,j) : \; 0 \le i \le N_z, \quad 0 \le j \le N_x\},$$

$$\Gamma_h = \Gamma_{1h} \cup \Gamma_{2h} \cup \Gamma_{3h} \cup \Gamma_{4h}, \quad \text{with} \left\{ \begin{array}{l} \Gamma_{1h} = \{0 \le i \le N_z, \quad j = 0\}, \\ \Gamma_{2h} = \{i = 0, \quad 0 \le j \le N_x\}, \\ \Gamma_{3h} = \{0 \le i \le N_z, \quad j = N_x\}, \\ \Gamma_{4h} = \{i = N_z, \quad 0 \le j \le N_x\}, \end{array} \right.$$

$$\Gamma_h^* = \Gamma_{1h}^* \cup \Gamma_{3h}^*, \quad \text{with} \left\{ \begin{array}{l} \Gamma_{1h}^* = \{1 \le i \le N_z - 1, \quad j = -1\}, \\ \Gamma_{3h}^* = \{1 \le i \le N_z - 1, \quad j = N_x + 1\}. \end{array} \right.$$

We have clearly

$$\left\{ \begin{array}{l} (i,j) \in \Omega_h \Longleftrightarrow X_{i,j} \in \overline{\Omega}, \\ (i,j) \in \Gamma_{\alpha h} \Longleftrightarrow X_{i,j} \in \Gamma_\alpha \quad \text{for } \alpha = 1,2,3 \text{ or } 4, \\ (i,j) \in \; \Gamma_h^* \Longrightarrow X_{i,j} \in \mathbb{R}^2 \setminus \overline{\Omega}. \end{array} \right.$$

Finally, if $(n,i,j) \in \mathbb{N}^3$, $k \in \mathbb{Z}$ and $U_{i,j}^n$ is a gridfunction, we denote the discrete derivatives by

$$D_{k,0}\, U_{i,j}^n = \frac{U_{i+k,j}^n - U_{i,j}^n}{k \Delta z}, \qquad D_{0,k}\, U_{i,j}^n = \frac{U_{i,j+k}^n - U_{i,j}^n}{k \Delta x}, \qquad D^{(k)}\, U_{i,j}^n = \frac{U_{i,j}^{n+k} - U_{i,j}^n}{k \Delta t}.$$

## 2.2 Discretization of the continuous problem

Throughout this paper we shall assume that the initial data belongs to $\mathrm{L}^2(\Omega)$ and verifies

$$0 \le U^0 \le 1 \text{ almost everywhere in } \Omega. \tag{2.2}$$

Let $U_h^0$ be a sequence of continuous functions which also verify (2.2) and tend to $U_0$ in $\mathrm{L}^2(\Omega)$ as $h \to 0$. For $(i,j) \in \Omega_h$ we set $U_{i,j}^0 = U_h^0(X_{i,j})$.

We now describe our finite difference scheme for (1.2)–(1.7). Assume that $U_{i,j}^n$ is given for a fixed $n \ge 0$ and for all $(i,j) \in \Omega_h$.
$a$) We first treat the boundary conditions (1.3) and (1.5), setting $U_{i,j}^n$ on the fictitious gridpoints $\Gamma_h^*$. Consider the nonlinear condition (1.3). This equation can be seen as a Burgers equation where $-x$ plays the role of the time. Between the points of $\Gamma_{1h}$ and $\Gamma_{1h}^*$, we will thus apply a scheme adapted to this conservation law, following the Godunov method (see for instance [3, 5]): we set

$$U_{i,-1}^n = U_{i,0}^n - \frac{K}{2} \frac{\Delta x}{\Delta z} \left[ \left(U_{i,0}^n\right)^2 - \left(U_{i-1,0}^n\right)^2 \right], \quad \text{for } 1 \le i \le N_z - 1. \tag{2.3}$$

The Neumann condition (1.5) can be discretized simply by

$$U_{i,\,N_x+1}^n = U_{i,\,N_x}^n, \quad \text{for } 1 \le i \le N_z - 1,$$

$b$) Next we can define the discrete solution $U_{i,j}^{n+1}$ for $(i,j) \in \Omega_h$. The Dirichlet conditions (1.4) and (1.6) fix its values on $\Gamma_{2h}$ and $\Gamma_{4h}$:

$$U_{0,j}^{n+1} = 1, \quad \text{for } 0 \le j \le N_x \qquad ; \qquad U_{N_z,j}^{n+1} = 0, \quad \text{for } 0 \le j \le N_x.$$

Finally we discretize the heat equation by the five points explicit scheme. For $(i,j) \in \Omega_h \setminus \Gamma_{2h} \setminus \Gamma_{4h}$, we set

$$\frac{U_{i,j}^{n+1} - U_{i,j}^n}{\Delta t} = \frac{U_{i,j+1}^n + U_{i,j-1}^n - 2U_{i,j}^n}{(\Delta x)^2} + \frac{U_{i+1,j}^n + U_{i-1,j}^n - 2U_{i,j}^n}{(\Delta z)^2}. \tag{2.4}$$

Using the discrete derivative operators $D_{k,0}$ and $D_{0,k}$ introduced in the previous subsection, this finite difference scheme can be rewritten in a more compact way:

$$\left\{ \begin{array}{ll} a) & \text{for } (i,j) \in \Gamma_{1h} \text{ and } 1 \le i \le N_z - 1, \quad D_{0,-1}\, U_{i,j}^n = \dfrac{K}{2} D_{-1,0}\left[(U_{i,j}^n)^2\right], \qquad (2.5) \\[2ex] & \text{for } (i,j) \in \Gamma_{3h} \text{ and } 1 \le i \le N_z - 1, \quad D_{0,+1}\, U_{i,j}^n = 0, \qquad (2.6) \\[2ex] b) & \text{for } (i,j) \in \Gamma_{2h}, \qquad U_{i,j}^{n+1} = 1, \qquad (2.7) \\[2ex] & \text{for } (i,j) \in \Gamma_{4h}, \qquad U_{i,j}^{n+1} = 0, \qquad (2.8) \\[2ex] & \text{for } (i,j) \in \Omega_h \setminus \Gamma_{2h} \setminus \Gamma_{4h}, \\ & U_{i,j}^{n+1} = U_{i,j}^n + \Delta t\left[D_{0,-1}\, D_{0,+1}\, U_{i,j}^n + D_{-1,0}\, D_{+1,0}\, U_{i,j}^n\right]. \qquad (2.9) \end{array} \right.$$

The standard results [3, 5] on these schemes enable to deduce the following lemma, stated here without proof:

**Lemma 2.1** *The scheme* (2.5)–(2.9) *is an approximation of Problem* (1.2)–(1.7) *which is consistent and accurate of order 1.*

# 3 Stability of the scheme

In this section, we shall obtain some estimates which will be sufficient to prove the convergence of the scheme.

## 3.1 The discrete maximum principle

The first estimates are obtained thanks to the maximum principle:

**Lemma 3.1** *If the initial data verifies* (2.2) *and if the time and space steps verify*

$$\left(1 + \frac{1}{2}\max\left(0,\, K\frac{\Delta x}{\Delta z} - 1\right)\right)\frac{\Delta t}{\Delta x^2} + \frac{\Delta t}{\Delta z^2} \le \frac{1}{2}, \tag{3.1}$$

*then for all* $(n,i,j) \in \mathbb{N} \times \Omega_h$ *we have*

$$0 \le U^n_{i,j} \le 1. \tag{3.2}$$

**Proof.** If the $U^n_{i,j}$ are bounded by 1, the CFL condition for the Burgers equation (1.3) writes

$$K\frac{\Delta x}{\Delta z} \le 1. \tag{CFL}$$

In a first step, we assume that (CFL) is verified. Then (3.1) becomes

$$\frac{\Delta t}{\Delta x^2} + \frac{\Delta t}{\Delta z^2} \le \frac{1}{2} \tag{3.3}$$

and the proof of the lemma is immediate, since each step of the scheme preserves the discrete maximum principle. One can for instance refer to [3] for the Godunov scheme and to [4] for the 5-points finite difference scheme for the Laplacian. Remark indeed that (3.3) is the classical $\mathrm{L}^\infty$-stability condition for this scheme.

Therefore we only treat the case where (CFL) is not verified. In this case the boundary condition on $\Gamma_{1h}$ is $\mathrm{L}^\infty$-unstable. Nevertheless this lemma says that the scheme will be stabilized thanks to the equation inside the domain, under Condition (3.1) which is stronger than (3.3). We denote

$$\lambda = K\frac{\Delta x}{\Delta z}, \qquad \mu_x = \frac{\Delta t}{\Delta x^2}, \qquad \mu_z = \frac{\Delta t}{\Delta z^2}.$$

With these notations, we have $\lambda > 1$ and (3.1) reads

$$(1+\lambda)\mu_x + 2\mu_z \le 1. \tag{3.4}$$

Consider a solution which verifies (3.2) at step $n$. Since (3.1) implies (3.3), we already have $0 \le U^{n+1}_{i,j} \le 1$ for $(i,j) \in \Omega_h \setminus \Gamma_{1h}$. It remains to show this formula for $(i,j)$ on the part of the boundary $\Gamma_{1h}$. Equation (2.3) reads

$$U^n_{i,-1} = U^n_{i,0} - \frac{\lambda}{2}\left(U^n_{i,0} - U^n_{i-1,0}\right)\left(U^n_{i,0} + U^n_{i-1,0}\right), \quad \text{for } 0 < i < N_z.$$

Then we plug it into (2.4). For $0 < i < N_z$ we obtain

$$\begin{aligned} U^{n+1}_{i,0} = {} & \left(1 - 2\mu_z - \mu_x - \lambda\,\frac{U^n_{i,0} + U^n_{i-1,0}}{2}\,\mu_x\right) U^n_{i,0} \\ & + \left(\mu_z + \lambda\,\frac{U^n_{i,0} + U^n_{i-1,0}}{2}\,\mu_x\right) U^n_{i-1,0} + \mu_z\, U^n_{i+1,0} + \mu_x\, U^n_{i,1}\,. \end{aligned} \tag{3.5}$$

To conclude, it suffices to check that the second-hand side is a convex combination of $U^n_{i,0}$, $U^n_{i-1,0}$, $U^n_{i+1,0}$ and $U^n_{i,1}$. By (3.2) we have

$$1-2\mu_z-(1+\lambda)\mu_x \le 1-2\mu_z-\mu_x-\lambda\,\frac{U^n_{i,0}+U^n_{i-1,0}}{2}\,\mu_x \quad \text{and} \quad 0 \le \mu_z+\lambda\,\frac{U^n_{i,0}+U^n_{i-1,0}}{2}\,\mu_x.$$

Thus, under Condition (3.4), the four coefficients in the right-hand side of (3.5) are nonnegative. One can conclude the proof since their sum is equal to 1. $\square$

**Lemma 3.2** *Assume that the initial data verifies* (2.2) *and that* (3.1) *is fulfilled. If moreover there exists a real number $A \geq 0$ independent of the grid such that*

$$\forall (i,j) \in \Omega \setminus \Gamma_{2h}, \qquad D_{-1,0}\, U^0_{i,j} \leq A, \tag{3.6}$$

*then we have*

$$\forall (n,i,j) \in \mathbb{N} \times (\Omega \setminus \Gamma_{2h}), \qquad D_{-1,0}\, U^n_{i,j} \leq A. \tag{3.7}$$

One can remark that this condition (3.7) is formally analogous to the standard condition for entropic shocks of the Burgers equation. This inequality will be used twice in this paper: for the energy estimate (3.16) and for the proof of uniqueness in Theorem 4.3.

**Proof.** Let

$$W^n_{i,j} = D_{-1,0}\, U^n_{i,j} = \frac{U^n_{i,j} - U^n_{i-1,j}}{\Delta z},$$

for $(i,j) \in \Omega_h \setminus \Gamma_{2h}$ and for $(i,j) \in \Gamma^*_{1h} \cup \Gamma^*_{3h}$ such that $i \geq 2$. Assumption (3.6) implies $W^n_{i,j} \leq A$.

It is straightforward to check that the $W^n_{i,j}$ verify the following scheme (with the same notations $\lambda$, $\mu_z$ and $\mu_x$ as in the proof of Lemma 3.1):

$$\begin{cases} a) \quad \text{for } 2 \leq i \leq N_z - 1, \qquad W^n_{i,\,N_x+1} = W^n_{i,\,N_x} \\ \qquad W^n_{i,\,-1} = W^n_{i,0} - \dfrac{\lambda}{2}\left[ (U^n_{i,0} + U^n_{i-1,0})\, W^n_{i,0} - (U^n_{i-1,0} + U^n_{i-2,0})\, W^n_{i-1,0} \right], \\ b) \quad \text{for } 0 \leq j \leq N_x, \qquad W^{n+1}_{1,j} = \dfrac{U^{n+1}_{1,j} - 1}{\Delta z}, \qquad W^{n+1}_{N_z,j} = -\dfrac{U^{n+1}_{N_z-1,j}}{\Delta z}, \\ \quad\;\; \text{for } 2 \leq i \leq N_z - 1, \quad 0 \leq j \leq N_x, \\ \qquad W^{n+1}_{i,j} = (1 - 2\mu_z - 2\mu_x) W^n_{i,j} + \mu_z (W^n_{i-1,j} + W^n_{i+1,j}) + \mu_x (W^n_{i,j-1} + W^n_{i,j+1}). \end{cases}$$

For $0 \leq j \leq N_x$ (3.2) directly gives

$$W^{n+1}_{1,j} \leq 0 \qquad \text{and} \qquad W^{n+1}_{N_z,j} \leq 0.$$

Next for $2 \leq i \leq N_z - 1$ and $1 \leq j \leq N_x$, the five-points laplacian and the discrete Neumann condition on $\Gamma_{3h}$ imply $W^{n+1}_{i,j} \leq A$. Therefore, as in the proof of Lemma 3.1, the main difficulty consists in proving this inequality for $(W^{n+1}_{i,0})_{i \geq 2}$, *i.e.* on the boundary $\Gamma_{1h}$. This element is defined by

$$W^{n+1}_{i,0} = (1 - 2\mu_z - 2\mu_x) W^n_{i,0} + \mu_z W^n_{i-1,0} + \mu_z\, W^n_{i+1,0} + \mu_x\, W^n_{i,1} + \mu_x\, W^n_{i,-1}. \tag{3.8}$$

Recall that by assumption we have

$$W_{i,0}^n \le A, \quad W_{i-1,0}^n \le A, \quad W_{i+1,0}^n \le A, \quad W_{i,1}^n \le A. \tag{3.9}$$

Besides for $2 \le i \le N_z - 1$ the discrete boundary condition on $\Gamma_{1,h}^*$ gives

$$W_{i,-1}^n = (1-\lambda_1) W_{i,0}^n + \lambda_2 W_{i-1,0}^n, \tag{3.10}$$

with

$$\lambda_1 = \frac{\lambda}{2}(U_{i,0}^n + U_{i-1,0}^n) \quad \text{and} \quad \lambda_2 = \frac{\lambda}{2}(U_{i-1,0}^n + U_{i-2,0}^n).$$

From (3.2) and (3.1) we get

$$0 \le \lambda_1 \le 1, \quad 0 \le \lambda_2 \le \lambda, \quad 2\mu_z + (1+\lambda_1)\mu_x \le 1, \quad 2\mu_z + (1+\lambda_2)\mu_x \le 1. \tag{3.11}$$

Plugging (3.10) into (3.8) gives

$$W_{i,0}^{n+1} = (1 - 2\mu_z - (1+\lambda_1)\,\mu_x)\, W_{i,0}^n + (\mu_z + \lambda_2\,\mu_x)\, W_{i-1,0}^n + \mu_z\, W_{i+1,0}^n + \mu_x\, W_{i,1}^n\,. \tag{3.12}$$

In order to show that $W_{i,0}^{n+1} \le A$, different cases have to be distinguished:

• If $W_{i-1,0}^n \ge 0$ and $W_{i,0}^n \ge 0$. Since $W_{i,0}^n$, $W_{i-1,0}^n$, $W_{i+1,0}^n$ and $W_{i,1}^n$ are bounded by $A$ (by (3.9)) and their coefficients in (3.12) are nonnegative (by (3.11)), we get

$$W_{i,0}^{n+1} \;\le\; (1 + (\lambda_2 - \lambda_1)\,\mu_x))\, A \;=\; (1 - \frac{\lambda}{2}\,\Delta z\,(W_{i,0}^n + W_{i-1,0}^n)\,\mu_x)\, A \;\le\; A,$$

where we used $W_{i-1,0}^n + W_{i,0}^n \ge 0$.

• If $W_{i-1,0}^n \le 0$ and $W_{i,0}^n \ge 0$. From (3.10) and (3.11) we deduce

$$W_{i,-1}^n \le (1-\lambda_1) W_{i,0}^n \le W_{i,0}^n \le A.$$

Then (3.8) and (3.9) give directly the result.

• If $W_{i,0}^n \le 0$. By (3.11) the coefficient of $W_{i,0}^n$ in (3.12) is nonnegative. Thus, thanks to (3.11),

$$\begin{aligned} W_{i,0}^{n+1} &\le (\mu_z + \lambda_2\,\mu_x)\, W_{i-1,0}^n + \mu_z\, W_{i+1,0}^n + \mu_x\, W_{i,1}^n \\ &\le (2\mu_z + (1+\lambda_2)\mu_x)\, A \le A. \end{aligned}$$

□

## 3.2 H¹ estimate

Before proving the energy estimate with the assumptions of Lemma 3.2, we state -without proof- two formulae of discrete integration by parts:

**Lemma 3.3** *Let $(\alpha_{i,j})$ and $(\beta_{i,j})$ be two families respectively indexed on $\Omega_h \cup \Gamma_h^*$ and on $\Omega_h$. Then we have*

$$\sum_{\Omega_h \backslash \Gamma_{2h}} (D_{-1,0}\, \alpha_{i,j})\ \beta_{i,j} = -\sum_{\Omega_h \backslash \Gamma_{4h}} \alpha_{i,j}\ (D_{1,0}\, \beta_{i,j}) - \frac{1}{\Delta z} \sum_{\Gamma_{2h}} \alpha_{i,j}\, \beta_{i,j} + \frac{1}{\Delta z} \sum_{\Gamma_{4h}} \alpha_{i,j}\, \beta_{i,j}\ , \tag{3.13}$$

$$\begin{aligned} \sum_{\Omega_h \backslash \Gamma_{2h} \backslash \Gamma_{4h}} (D_{0,-1}\, \alpha_{i,j})\ \beta_{i,j} = & - \sum_{\Omega_h \backslash \Gamma_{2h} \backslash \Gamma_{3h} \backslash \Gamma_{4h}} \alpha_{i,j}\ (D_{0,1}\, \beta_{i,j}) \\ & - \frac{1}{\Delta x} \sum_{i=1}^{N_z - 1} \alpha_{i,-1}\, \beta_{i,0} + \frac{1}{\Delta x} \sum_{i=1}^{N_z-1} \alpha_{i,N_x}\, \beta_{i,N_x}\, . \end{aligned} \tag{3.14}$$

In the introduction of this paper, we defined $V$ as the set of the functions of $\mathrm{H}^1(\Omega)$ which vanish on $\Gamma_2$ and $\Gamma_4$. The semi-norm $|\cdot|_{\mathrm{H}^1}$ is a norm of this space and will be denoted by $\|\cdot\|_V$. We also denote its dual space by $V^*$. Let us now introduce the discrete analogous of classical functional spaces. Define firstly the linear space $V_h$ of dimension $N_z \times (N_x + 1)$:

$$V_h = \left\{ (v_{i,j})_{(i,j)\in\Omega_h} \text{ such that } v_{i,j} = 0 \text{ for } (i,j) \in \Gamma_{2h} \cup \Gamma_{4h} \right\}$$

Two norms are defined on this space, corresponding respectively to the $\mathrm{L}^2(\Omega)$ and $V$ norms.

$$\left\|\alpha_{i,j}^n\right\|_{\mathrm{L}_h^2} = \left( \Delta z\, \Delta x \sum_{\Omega_h} \left(\alpha_{i,j}^n\right)^2 \right)^{1/2},$$

$$\left\|\alpha_{i,j}^n\right\|_{V_h} = \left[ \Delta z\, \Delta x \sum_{\Omega_h \backslash \Gamma_{3h}} \left(D_{0,1}\, \alpha_{i,j}^n\right)^2 + \Delta z\, \Delta x \sum_{\Omega_h \backslash \Gamma_{4h}} \left(D_{1,0}\, \alpha_{i,j}^n\right)^2 \right]^{1/2}.$$

We denote by $\|\cdot\|_{V_h^*}$ the dual norm of $\|\cdot\|_{V_h}$ with respect to the scalar product

$$\left(\alpha_{i,j}\, ,\, \beta_{i,j}\right)_h = \Delta z\, \Delta x \sum_{\Omega_h} \alpha_{i,j}\, \beta_{i,j}.$$

Consider a sequence of elements of $V_h$, denoted by $\alpha_{i,j}^n$. One finally defines the analogous of the $\mathrm{L}^\infty((0,t^N), \mathrm{L}^2(\Omega))$, $\mathrm{L}^2((0,t^N), V)$ and $\mathrm{L}^2((0,t^N), V^*)$ norms:

$$\left\|\alpha_{i,j}^n\right\|_{\mathrm{L}_h^\infty\left((0,N),\mathrm{L}_h^2\right)} = \max_{0\le n\le N} \left( \left\|\alpha_{i,j}^n\right\|_{\mathrm{L}_h^2} \right), \qquad \left\|\alpha_{i,j}^n\right\|_{\mathrm{L}_h^2((0,N),V_h)} = \left( \Delta t \sum_{n=0}^{N} \left\|\alpha_{i,j}^n\right\|_{V_h}^2 \right)^{1/2},$$

$$\left\|\alpha_{i,j}^n\right\|_{\mathrm{L}_h^2\left((0,N),V_h^*\right)} = \left( \Delta t \sum_{n=0}^{N} \left\|\alpha_{i,j}^n\right\|_{V_h^*}^2 \right)^{1/2}.$$

In the sequel of this paper, we denote by $C$ a constant independent of $h = (\Delta t,\ \Delta z,\ \Delta x)$.

**Proposition 3.4** *With the assumptions of Lemma 3.2, if*

$$\frac{\Delta t}{\Delta x^2} < \frac{1}{4}, \qquad \frac{\Delta t}{\Delta z^2} < \frac{1}{4}, \tag{3.15}$$

*then for all $N$ we have the following estimates:*

$$\left\| U_{i,j}^n \right\|_{\mathrm{L}_h^\infty\left((0,N),\mathrm{L}_h^2\right)} + \left\| U_{i,j}^n \right\|_{\mathrm{L}_h^2((0,N),V_h)} \leq C, \tag{3.16}$$

$$\left\| D^{(1)} U_{i,j}^n \right\|_{\mathrm{L}_h^2((0,N),V_h^*)} \leq C. \tag{3.17}$$

**Proof.** Let us start by proving (3.16). Setting

$$u_{i,j}^n = U_{i,j}^n - \psi_{i,j}, \quad \text{with} \quad \psi_{i,j} = 1 - \frac{i}{N_z},$$

we remark that it is equivalent to prove the estimate (3.16) for the $u_{i,j}^n$ instead of the $U_{i,j}^n$. We have $u_{i,j}^n \in V_h$ and

$$\left\{ \begin{array}{lll} a) & \text{for } (i,j) \in \Gamma_{1h} \text{ and } 0 < i < N_z, & D_{0,-1}\, u_{i,j}^n = \dfrac{K}{2} D_{-1,0} \left[ (U_{i,j}^n)^2 \right], \quad (3.18) \\ & \text{for } (i,j) \in \Gamma_{3h} \text{ and } 0 < i < N_z, & D_{0,1}\, u_{i,j}^n = 0, \quad (3.19) \\ b) & \text{for } (i,j) \in \Gamma_{2h} \cup \Gamma_{4h}, & u_{i,j}^{n+1} = 0, \quad (3.20) \\ & \text{for } (i,j) \in \Omega_h \setminus \Gamma_{2h} \setminus \Gamma_{4h}, & \\ & \quad u_{i,j}^{n+1} = u_{i,j}^n + \Delta t \left[ D_{0,-1}\, D_{0,+1}\, u_{i,j}^n + D_{-1,0}\, D_{+1,0}\, u_{i,j}^n \right]. & (3.21) \end{array} \right.$$

Take the square of (3.21) and sum it for $(i,j) \in \Omega_h \setminus \Gamma_{2h} \setminus \Gamma_{4h}$:

$$\sum_{\Omega_h \setminus \Gamma_{2h} \setminus \Gamma_{4h}} \left( u_{i,j}^{n+1} \right)^2 = \sum_{\Omega_h \setminus \Gamma_{2h} \setminus \Gamma_{4h}} \left( u_{i,j}^n \right)^2 + 2\Delta t\, \mathcal{I}_1 + 2\Delta t\, \mathcal{I}_2 + \Delta t^2\, \mathcal{R}, \tag{3.22}$$

where $\mathcal{I}_1 = \sum_{\Omega_h \setminus \Gamma_{2h} \setminus \Gamma_{4h}} \left[ D_{0,-1} D_{0,1}\, u_{i,j}^n \right] u_{i,j}^n, \qquad \mathcal{I}_2 = \sum_{\Omega_h \setminus \Gamma_{2h} \setminus \Gamma_{4h}} \left[ D_{-1,0} D_{1,0}\, u_{i,j}^n \right] u_{i,j}^n$

and $\mathcal{R} = \sum_{\Omega_h \setminus \Gamma_{2h} \setminus \Gamma_{4h}} \left( D_{0,-1} D_{0,1}\, u_{i,j}^n + D_{-1,0} D_{1,0}\, u_{i,j}^n \right)^2.$

Remark that the Dirichlet conditions (3.20) on $\Gamma_{2h}$ and $\Gamma_{4h}$ imply

$$\sum_{\Gamma_{2h} \cup \Gamma_{4h}} \left( u_{i,j}^{n+1} \right)^2 = \sum_{\Gamma_{2h} \cup \Gamma_{4h}} \left( u_{i,j}^n \right)^2 = 0,$$

thus we have

$$\sum_{\Omega_h} \left( u_{i,j}^{n+1} \right)^2 = \sum_{\Omega_h} \left( u_{i,j}^n \right)^2 + 2\Delta t\, \mathcal{I}_1 + 2\Delta t\, \mathcal{I}_2 + \Delta t^2\, \mathcal{R}. \tag{3.23}$$

Lemma 3.3 enables to calculate $\mathcal{I}_1$ and $\mathcal{I}_2$ thanks to discrete integrations by parts, using the discrete boundary conditions (3.18)–(3.20):

$$\mathcal{I}_1 = -\sum_{\Omega_h \backslash \Gamma_{3h}} \left(D_{0,1}\, u_{i,j}^n\right)^2 - \frac{K}{2\Delta x} \sum_{\Gamma_{1h}} \left[D_{-1,0}\left(\left(U_{i,j}^n\right)^2\right)\right] u_{i,j}^n$$

and

$$\mathcal{I}_2 = -\sum_{\Omega_h \backslash \Gamma_{4h}} \left(D_{1,0}\, U_{i,j}^n\right)^2$$

(in this formulae, all the terms make sense if we set for instance $U_{-1,0} = 1$ and $U_{-1,N_z} = 0$). Denote

$$\mu = 4 \max\left(\frac{\Delta t}{\Delta x^2}, \frac{\Delta t}{\Delta z^2}\right).$$

Thanks to (3.15) we have $\mu < 1$. Besides $\mathcal{R}$ can be estimated as follows, using (3.18)–(3.20):

$$\begin{aligned}
\mathcal{R} &= \sum_{\Omega_h \backslash \Gamma_{2h} \backslash \Gamma_{4h}} \left(\frac{1}{\Delta x} D_{0,1}\, u_{i,j}^n - \frac{1}{\Delta x} D_{0,1}\, u_{i,j-1}^n + \frac{1}{\Delta z} D_{1,0}\, u_{i,j}^n - \frac{1}{\Delta z} D_{1,0}\, u_{i-1,j}^n\right)^2 \\
&\leq \frac{8}{\Delta x^2} \sum_{\Omega_h \backslash \Gamma_{3h}} \left(D_{0,1}\, u_{i,j}^n\right)^2 + \frac{8}{\Delta z^2} \sum_{\Omega_h \backslash \Gamma_{4h}} \left(D_{1,0}\, u_{i,j}^n\right)^2 + \frac{4}{\Delta x^2} \sum_{i=1}^{N_z - 1} \left(D_{0,-1}\, u_{i,0}^n\right)^2 \\
&\leq \frac{2\mu}{\Delta t\, \Delta x\, \Delta z} \left\|u_{i,j}^n\right\|_{V_h}^2 + \frac{K^2}{\Delta x^2} \sum_{\Gamma_{1h}} \left[D_{-1,0}\left(\left(U_{i,j}^n\right)^2\right)\right]^2.
\end{aligned}$$

Therefore (3.23) becomes

$$\begin{aligned}
&\left\|u_{i,j}^{n+1}\right\|_{\mathrm{L}_h^2}^2 - \left\|u_{i,j}^n\right\|_{\mathrm{L}_h^2}^2 + 2(1-\mu)\Delta t \left\|u_{i,j}^n\right\|_{V_h}^2 \leq \\
&\quad \leq -K\, \Delta t\, \Delta z \sum_{\Gamma_{1h}} \left[D_{-1,0}\left(\left(U_{i,j}^n\right)^2\right)\right] u_{i,j}^n + \frac{K^2 \Delta t^2 \Delta z}{\Delta x} \sum_{\Gamma_{1h}} \left[D_{-1,0}\left(\left(U_{i,j}^n\right)^2\right)\right]^2.
\end{aligned}$$

It remains to estimate the two terms in the right-hand side of this inequality. These crucial estimates can be obtained thanks to (3.2) and (3.7). Recall that we have:

$$0 \leq U_{i,j}^n \leq 1, \qquad -1 \leq u_{i,j}^n \leq 1, \qquad D_{1,0}\, U_{i,j}^n \leq A.$$

Let us start with the first term. If for $0 \leq i \leq N_z$ we set $\Phi_i = \left(U_{i-1,0}^n + U_{i,0}^n\right) u_{i,0}^n$, then we have $0 \leq 2 - \Phi_i \leq 4$ and

$$\begin{aligned}
-K\, \Delta t\, \Delta z \sum_{\Gamma_{1h}} \left[D_{-1,0}\left(\left(U_{i,j}^n\right)^2\right)\right] u_{i,j}^n &= -K\, \Delta t\, \Delta z \sum_{\Gamma_{1h}} \left[D_{-1,0}\left(U_{i,j}^n\right)\right] \Phi_i \\
&= K\, \Delta t\, \Delta z \sum_{\Gamma_{1h}} \left[D_{-1,0}\left(U_{i,j}^n\right)\right] (2 - \Phi_i) - 2K\, \Delta t\, \Delta z \sum_{\Gamma_{1h}} \left[D_{-1,0}\left(U_{i,j}^n\right)\right] \\
&\leq 4\, A\, K\, \Delta t\, \Delta z\, N_z + 2K\, \Delta t\, \Delta z \frac{\left(U_{0,0}^n - U_{N_z,0}^n\right)}{\Delta z} \leq C \Delta t.
\end{aligned}$$

The other boundary term can be treated similarly, setting

$$\tilde{\Phi}_i = \left(U_{i,j}^n + U_{i-1,j}^n\right) \left[\left((U_{i,j}^n\right)^2 - \left(U_{i-1,j}^n\right)^2\right].$$

Indeed we have $0 \leq \tilde{\Phi}_i + 2 \leq 4$ and this term is written

$$\begin{aligned}
\frac{K^2 \Delta t^2 \Delta z}{\Delta x} \sum_{\Gamma_{1h}} \left[D_{-1,0}\left(\left(U_{i,j}^n\right)^2\right)\right]^2 &= \frac{K^2 \Delta t^2}{\Delta x} \sum_{\Gamma_{1h}} \left[D_{-1,0}\left(U_{i,j}^n\right)\right] \tilde{\Phi}_i \\
&= \frac{K^2 \Delta t^2}{\Delta x} \sum_{\Gamma_{1h}} \left[D_{-1,0}\left(U_{i,j}^n\right)\right] (\tilde{\Phi}_i + 2) - 2\frac{K^2 \Delta t^2}{\Delta x} \sum_{\Gamma_{1h}} \left[D_{-1,0}\left(U_{i,j}^n\right)\right] \\
&\leq C \frac{\Delta t^2}{\Delta x \Delta z} \leq C\mu \Delta t.
\end{aligned}$$

Finally we get

$$\left\|u_{i,j}^{n+1}\right\|_{\mathrm{L}_h^2}^2 - \left\|u_{i,j}^n\right\|_{\mathrm{L}_h^2}^2 + 2(1-\mu)\Delta t \left\|u_{i,j}^n\right\|_{V_h}^2 \leq C\Delta t,$$

which yields (3.16), after a sum on $n = 0 \dots N_t - 1$.

To complete the proof of the theorem, it remains to show (3.17). The proof of this estimate is based on a discrete variational formulation analogous to (1.9). Let $(\varphi_{i,j}) \in V_h$. We multiply (2.9) by $\varphi_{i,j}$ and sum on $\Omega_h \setminus \Gamma_{2h} \setminus \Gamma_{4h}$. After some discrete integration by parts it gives a first variational formulation (remark that $\varphi_{i,j}$ vanishes on $\Gamma_{2h} \cup \Gamma_{4h}$):

$$\sum_{\Omega_h} \left(D^{(1)} U_{i,j}^n\right) \varphi_{i,j} = -\sum_{\Omega_h \setminus \Gamma_{3h}} \left(D_{0,1}\, U_{i,j}^n\right) \left(D_{0,1}\, \varphi_{i,j}\right) - \sum_{\Omega_h \setminus \Gamma_{4h}} \left(D_{1,0}\, U_{i,j}^n\right) \left(D_{1,0}\, \varphi_{i,j}\right) - \frac{K}{2} \mathcal{T}, \tag{3.24}$$

where $\mathcal{T}$ is the following discrete integral on $\Gamma_{1h}$:

$$\mathcal{T} = -\frac{1}{\Delta x} \sum_{\Gamma_{1h}} \left[D_{-1,0}\left(\left(U_{i,0}^n\right)^2\right)\right] \varphi_{i,0}.$$

We shall now transform (3.24) into a "volumic" formulation, *i.e.* without any sum on boundaries. Let

$$\theta_{i,j} = \theta\left(X_{i,j}\right) = 1 - \frac{j}{N_x}.$$

We have $\theta_{i,j} = 1$ on $\Gamma_{1h}$ and $\theta_{i,j} = 0$ on $\Gamma_{3h}$. Hence, applying some discrete integrations by parts, we compute

$$\begin{aligned}
\mathcal{T} &= -\frac{1}{\Delta x} \sum_{\Gamma_{1h}} \left[D_{-1,0}\left(\left(U_{i,0}^n\right)^2\right)\right] \varphi_{i,0}\, \theta_{i,0} \\
&= \sum_{\Omega_h \setminus \Gamma_{3h}} \left[D_{-1,0}\left(\left(U_{i,j}^n\right)^2\right)\right] \left[D_{0,1}\left(\varphi_{i,j}\, \theta_{i,j}\right)\right] + \sum_{\Omega_h \setminus \Gamma_{1h}} \left[D_{0,-1} D_{-1,0}\left(\left(U_{i,j}^n\right)^2\right)\right] \varphi_{i,j}\, \theta_{i,j} \\
&= \sum_{\Omega_h \setminus \Gamma_{3h}} \left[D_{-1,0}\left(\left(U_{i,j}^n\right)^2\right)\right] \left[D_{0,1}\left(\varphi_{i,j}\, \theta_{i,j}\right)\right] - \sum_{\Omega_h \setminus \Gamma_{1h} \setminus \Gamma_{4h}} \left[D_{0,-1}\left(\left(U_{i,j}^n\right)^2\right)\right] \left[D_{1,0}\left(\varphi_{i,j}\, \theta_{i,j}\right)\right].
\end{aligned}$$

This enables to write the discrete variational formulation of Problem (1.2)–(1.7):

$$\begin{aligned}\sum_{\Omega_h}\left(D^{(1)}U^n_{i,j}\right)\varphi_{i,j} = & -\sum_{\Omega_h\setminus\Gamma_{3h}}\left(D_{0,1}\,U^n_{i,j}\right)\left(D_{0,1}\,\varphi_{i,j}\right) -\sum_{\Omega_h\setminus\Gamma_{4h}}\left(D_{1,0}\,U^n_{i,j}\right)\left(D_{1,0}\,\varphi_{i,j}\right)\\ & +\frac{K}{2}\sum_{\Omega_h\setminus\Gamma_{3h}}\left[D_{-1,0}\left(\left(U^n_{i,j}\right)^2\right)\right]\left[D_{0,1}\left(\varphi_{i,j}\,\theta_{i,j}\right)\right]\\ & -\frac{K}{2}\sum_{\Omega_h\setminus\Gamma_{1h}\setminus\Gamma_{4h}}\left[D_{0,-1}\left(\left(U^n_{i,j}\right)^2\right)\right]\left[D_{1,0}\left(\varphi_{i,j}\,\theta_{i,j}\right)\right].\end{aligned} \tag{3.25}$$

Remark now that we have

$$D_{-1,0}\left(\left(U^n_{i,j}\right)^2\right) = \left(U^n_{i,j}+U^n_{i-1,j}\right)D_{-1,0}\,U^n_{i,j},$$

$$D_{0,-1}\left(\left(U^n_{i,j}\right)^2\right) = \left(U^n_{i,j}+U^n_{i,j-1}\right)D_{0,-1}\,U^n_{i,j}.$$

Hence from (3.2) we deduce

$$\left\|\left(U^n_{i,j}\right)^2\right\|_{V_h} \le 2\left\|U^n_{i,j}\right\|_{V_h}.$$

Besides we have also

$$\left\|\varphi_{i,j}\,\theta_{i,j}\right\|_{V_h} \le C\left\|\varphi_{i,j}\right\|_{V_h}.$$

Therefore the Cauchy-Schwarz inequality applied to the right-hand side of (3.25) gives

$$\Delta z\,\Delta x\left|\sum_{\Omega_h}\left(D^{(1)}U^n_{i,j}\right)\varphi_{i,j}\right| \le C\left\|U^n_{i,j}\right\|_{V_h}\ \left\|\varphi_{i,j}\right\|_{V_h},$$

which implies

$$\left\|D^{(1)}U^n_{i,j}\right\|_{V_h^*} \le C\left\|U^n_{i,j}\right\|_{V_h}.$$

The proof is complete thanks to (3.16). □

We now reorganize the different terms of the discrete variational formulation in order to compute all the two-dimensional sums on the set $\tilde{\Omega}_h = \Omega_h\setminus\Gamma_{3h}\setminus\Gamma_{4h}$, which contains exactly $N_z \times N_x$ points. The following weak formulation will be used for the convergence proof:

$$\begin{aligned}\sum_{\tilde{\Omega}_h}\left(D^{(1)}U^n_{i,j}\right)\varphi_{i,j} = & -\sum_{\tilde{\Omega}_h}\left(D_{0,1}\,U^n_{i,j}\right)\left(D_{0,1}\,\varphi_{i,j}\right) - \sum_{\tilde{\Omega}_h}\left(D_{1,0}\,U^n_{i,j}\right)\left(D_{1,0}\,\varphi_{i,j}\right)\\ & +\frac{K}{2}\sum_{\tilde{\Omega}_h}\left[D_{1,0}\left(\left(U^n_{i,j}\right)^2\right)\right]\left[D_{0,1}\left(\varphi_{i+1,j}\,\theta_{i+1,j}\right)\right]\\ & -\frac{K}{2}\sum_{\tilde{\Omega}_h}\left[D_{0,1}\left(\left(U^n_{i,j}\right)^2\right)\right]\left[D_{1,0}\left(\varphi_{i,j+1}\,\theta_{i,j+1}\right)\right]\\ & +\mathcal{R}_n+\mathcal{S}_n,\end{aligned} \tag{3.26}$$

where the two terms of the remainder are

$$\mathcal{R}_n = -\sum_{\Gamma_{3h}} \left(D^{(1)} U^n_{i,j}\right) \varphi_{i,j}\,, \qquad \mathcal{S}_n = -\sum_{\Gamma_{3h},\, i<N_z} \left(D_{1,0}\, U^n_{i,j}\right) \left(D_{1,0}\, \varphi_{i,j}\right).$$

# 4 Convergence to the unique weak solution

The aim of this part is to let $h$ tend to 0 in (3.26) and to show that the limit problem is (1.9). The method is standard and consists in reinterpreting the sums in (3.26) as integral of interpolating functions.

## 4.1 Interpolation Operators

Let us first introduce some one-dimensional interpolation operators. Consider a sequence $u_i$ defined for $i = 0 \dots N_z$ and the gridpoints $z_i = i\Delta z$. We define the piecewise constant interpolation operator $\mathbb{P}^0_z$ along the direction $z$ by

$$\begin{cases} \forall i < N_z,\ \forall z \in [z_i, z_{i+1}[, \quad \left(\mathbb{P}^0_z\, u_i\right)(z) = u_i \\ \left(\mathbb{P}^0_z\, u_i\right)(L) = u_{N_z} \end{cases}$$

and the piecewise $\mathrm{C}^1$ interpolation operator $\mathbb{P}^1_z$ by

$$\forall i < N_z, \quad \forall z \in [z_i, z_{i+1}], \quad \left(\mathbb{P}^1_z\, u_i\right)(z) = u_i \left(\frac{z_{i+1} - z}{\Delta z}\right) + u_{i+1} \left(\frac{z - z_i}{\Delta z}\right).$$

We similarly define $\mathbb{P}^0_x$, $\mathbb{P}^1_x$, $\mathbb{P}^0_t$ and $\mathbb{P}^1_t$.

The domain $\Omega$ being a rectangle, we may construct several three-dimensional (one in time and two in space) interpolation operators as tensor products of these one-dimensional operators. Let $U^n_{i,j}$ be defined for $n \geq 0$ and $(i,j) \in \Omega_h$. We define a piecewise constant interpolation of this sequence on $[0,T] \times \Omega$ by

$$\mathbb{P}^0 U^n_{i,j} = \left(\mathbb{P}^0_t \otimes \mathbb{P}^0_z \otimes \mathbb{P}^0_x\right) U^n_{i,j}$$

and a piecewise $\mathrm{C}^1$ function by

$$\mathbb{P}^1 U^n_{i,j} = \left(\mathbb{P}^1_t \otimes \mathbb{P}^1_z \otimes \mathbb{P}^1_x\right) U^n_{i,j}.$$

We also define a third interpolation operator –discontinuous in time but continuous in space– by

$$\mathbb{P}^{0,1} U^n_{i,j} = \left(\mathbb{P}^0_t \otimes \mathbb{P}^1_z \otimes \mathbb{P}^1_x\right) U^n_{i,j}.$$

The following technical lemma, given without proof, states the properties of these operators that we shall use to pass to the limit as $h \to 0$:

**Lemma 4.1** *(a) Let $u_{i,j}$ and $v_{i,j}$ be two sequences defined for $(i,j) \in \Omega_h$. Then we have*

$$\Delta z\, \Delta x \sum_{\tilde{\Omega}_h} u_{i,j}\, v_{i,j} = \int_\Omega \left(\mathbb{P}^0\, u_{i,j}\right) \left(\mathbb{P}^0\, v_{i,j}\right), \tag{4.1}$$

*(b) The operators $\mathbb{P}^1$, $\mathbb{P}^0$ and $\mathbb{P}^{0,1}$ are bounded*

$$\text{from } \mathrm{L}^2_h((0,N), \mathrm{L}^2_h) \text{ to } \mathrm{L}^2((0,t^N), \mathrm{L}^2(\Omega)),$$
$$\text{from } \mathrm{L}^\infty_h((0,N), \mathrm{L}^2_h) \text{ to } \mathrm{L}^\infty((0,t^N), \mathrm{L}^2(\Omega))$$
$$\text{and from } \mathrm{L}^2_h((0,N), V_h^*) \text{ to } \mathrm{L}^2((0,t^N), V^*).$$

*(c) The operator $\mathbb{P}^1$ is bounded from $\mathrm{L}^2_h((0,N), V_h)$ to $\mathrm{L}^2((0,t^N), V)$.*

*(d) Let $u^n_{i,j}$ be bounded in $\mathrm{L}^2_h\left((0,N), \mathrm{L}^2_h\right)$. Then we have*

$$\left\|\mathbb{P}^1\, u^n_{i,j} - \mathbb{P}^0\, u^n_{i,j}\right\|_{\mathrm{L}^2((0,t^N),\mathrm{L}^2(\Omega))} = \mathcal{O}(h). \tag{4.2}$$

*(e) Let $u^n_{i,j}$ be bounded in $\mathrm{L}^2_h\left((0,N), V_h^*\right)$. Then we have*

$$\left\|\mathbb{P}^1\, u^n_{i,j} - \mathbb{P}^{0,1}\, u^n_{i,j}\right\|_{\mathrm{L}^2((0,t^N),V^*(\Omega))} + \left\|\mathbb{P}^1\, u^n_{i,j} - \mathbb{P}^0\, u^n_{i,j}\right\|_{\mathrm{L}^2((0,t^N),V^*(\Omega))} = \mathcal{O}(h). \tag{4.3}$$

*(f) Let $u^n_{i,j}$ be bounded in $\mathrm{L}^2_h\left((0,N), V_h\right)$. Then we have*

$$\begin{aligned} &\left\|\partial_z \left(\mathbb{P}^1\, u^n_{i,j}\right) - \mathbb{P}^0\, \left(D_{1,0} u^n_{i,j}\right)\right\|_{\mathrm{L}^2((0,t^N),\mathrm{L}^2(\Omega))} \\ &\quad + \left\|\partial_x \left(\mathbb{P}^1\, u^n_{i,j}\right) - \mathbb{P}^0\, \left(D_{0,1} u^n_{i,j}\right)\right\|_{\mathrm{L}^2((0,t^N),\mathrm{L}^2(\Omega))} = \mathcal{O}(h). \end{aligned} \tag{4.4}$$

## 4.2 The convergence result

The following theorem is the main result of this paper. It states the convergence of our finite difference scheme and provides a constructive existence proof for Problem (1.2)–(1.7). It will be completed by a uniqueness result in next section (Theorem 4.3).

**Theorem 4.2** *Let $U^0$ be an initial data in $\mathrm{L}^2(\Omega)$ satisfying $0 \le U^0 \le 1$. Assume that there exists a real number $A \ge 0$ such that*

$$\partial_z U^0 \le A \qquad \text{a. e. on } \Omega \tag{4.5}$$

*and consider the sequence $U^n_{i,j}$ constructed by the scheme (2.5)–(2.9). If the gridsteps satisfy (3.1) and (3.15) then for all $T > 0$ the sequence $\left(\mathbb{P}^1\, U^n_{i,j}\right)$ converges as $h \to 0$ to a weak solution $U$ of (1.2)–(1.7). This convergence holds in the $\mathrm{L}^2((0,T), \mathrm{L}^2(\Omega))$ norm and almost everywhere on $(0,T) \times \Omega$. Moreover this solution verifies*

$$U \in \mathrm{L}^2\left((0,T), \mathrm{H}^1(\Omega)\right) \cap \mathrm{C}^0\left([0,T], \mathrm{L}^2(\Omega)\right), \tag{4.6}$$

$$0 \le U \le 1 \qquad \text{and} \qquad \partial_z U \le A \qquad \text{a.e. on } (0,T) \times \Omega. \tag{4.7}$$

**Proof.** Let $U_h = \mathbb{P}^1 U^n_{i,j}$. Remark that $\partial_t U_h = \mathbb{P}^{0,1}\left(D^{(1)} U^n_{i,j}\right)$. Hence by Proposition 3.4 and Lemma 4.1 we have

$$\begin{cases} U_h \text{ is bounded in } \mathrm{L}^2((0,T),V)\cap \mathrm{L}^\infty((0,T),\mathrm{L}^2(\Omega)), \\ \partial_t U_h \text{ is bounded in } \mathrm{L}^2((0,T),V^*). \end{cases}$$

This enables us to apply a standard compactness result [6]. After extraction of a subsequence we have

$$\begin{cases} U_h \to U & \text{in } \mathrm{L}^2((0,T),\mathrm{L}^2(\Omega)) \text{ strong}, \\ U_h \rightharpoonup U & \text{in } \mathrm{L}^2((0,T),\mathrm{H}^1(\Omega)) \text{ weak and in } \mathrm{L}^\infty((0,T),\mathrm{L}^2(\Omega)) \text{ weak } *, \\ \partial_t U_h \rightharpoonup \partial_t U & \text{in } \mathrm{L}^2((0,T),V^*) \text{ weak}. \end{cases}$$

Thanks to Lemma 4.1 we deduce:

$$\begin{cases} \mathbb{P}^0 U^n_{i,j} \to U & \text{in } \mathrm{L}^2((0,T),\mathrm{L}^2(\Omega)) \text{ strong}, \\ \mathbb{P}^0\left(D^{(1)} U^n_{i,j}\right) \rightharpoonup \partial_t U & \text{in } \mathrm{L}^2((0,T),V^*) \text{ weak}, \\ \mathbb{P}^0\left(D_{0,1}\, U^n_{i,j}\right) \rightharpoonup \partial_x U \text{ and } \mathbb{P}^0\left(D_{1,0}\, U^n_{i,j}\right) \rightharpoonup \partial_z U & \text{in } \mathrm{L}^2((0,T),\mathrm{L}^2(\Omega)) \text{ weak}. \end{cases} \tag{4.8}$$

Let us now rewrite (3.26) by transforming the discrete sums on $\tilde{\Omega}_h$ into integrals thanks to Lemma (4.1) *(a)*. Let $\varphi \in \mathrm{C}^\infty(\overline{\Omega})$ which vanishes on $\Gamma_2 \cup \Gamma_4$. The sequence $\varphi_{i,j} = \varphi(X_{i,j})$ is in $V_h$ and we have

$$\begin{aligned} \int_\Omega \left(\mathbb{P}^0\left(D^{(1)} U^n_{i,j}\right)\right)\left(\mathbb{P}^0\, \varphi_{i,j}\right) = & -\int_\Omega \left(\mathbb{P}^0\left(D_{0,1}\, U^n_{i,j}\right)\right)\left(\mathbb{P}^0\left(D_{0,1}\, \varphi_{i,j}\right)\right) \\ & -\int_\Omega \left(\mathbb{P}^0\left(D_{1,0}\, U^n_{i,j}\right)\right)\left(\mathbb{P}^0\left(D_{1,0}\, \varphi_{i,j}\right)\right) \\ & +\frac{K}{2}\int_\Omega \left(\mathbb{P}^0\left(U^n_{i,j}+U^n_{i+1,j}\right)\right)\left(\mathbb{P}^0\left(D_{1,0}\, U^n_{i,j}\right)\right)\left(\mathbb{P}^0\left(D_{0,1}\left(\varphi_{i+1,j}\,\theta_{i+1,j}\right)\right)\right) \\ & -\frac{K}{2}\int_\Omega \left(\mathbb{P}^0\left(U^n_{i,j}+U^n_{i,j+1}\right)\right)\left(\mathbb{P}^0\left(D_{0,1}\, U^n_{i,j}\right)\right)\left(\mathbb{P}^0\left(D_{1,0}\left(\varphi_{i,j+1}\,\theta_{i,j+1}\right)\right)\right) \\ & +\widetilde{\mathcal{R}}_n + \widetilde{\mathcal{S}}_n, \end{aligned} \tag{4.9}$$

with

$$\widetilde{\mathcal{R}}_n = -\Delta z \Delta x \sum_{\Gamma_{3h}} \left(D^{(1)} U^n_{i,j}\right)\varphi_{i,j}, \quad \widetilde{\mathcal{S}}_n = -\Delta z \Delta x \sum_{i=0}^{N_z-1}\left(D_{1,0}\, U^n_{i,N_x}\right)\left(D_{1,0}\,\varphi_{i,N_x}\right).$$

By (4.8) and some standard results of approximation (the functions $\varphi$ and $\theta$ are smooth) we can pass to the limit in the integrals of (4.9). Note that the strong convergence in $\mathrm{L}^2$ is required to treat the two nonlinear terms. It remains to show that the remainders $\widetilde{\mathcal{R}}_n$ and $\widetilde{\mathcal{S}}_n$ converge to 0.

Consider first $\widetilde{\mathcal{R}}_n$ which can be interpolated in time by $\mathbb{P}^0_t$. Let $\zeta(t) \in \mathrm{C}^\infty_c(]0,T[)$. We have

$$\begin{aligned}
\int_0^T \left(\mathbb{P}^0_t \widetilde{\mathcal{R}}_n\right)(t)\,\zeta(t)dt &= -\Delta z\Delta x \sum_{n=0}^{N-1}\left[\sum_{\Gamma_{3h}}\left(\frac{U^{n+1}_{i,j}-U^n_{i,j}}{\Delta t}\,\varphi_{i,j}\right)\int_{t^n}^{t^{n+1}}\zeta(t)dt\right] \\
&= \Delta x \sum_{n=0}^{N-1}\left[\Delta z\sum_{\Gamma_{3h}}\left(U^n_{i,j}\,\varphi_{i,j}\right)\int_{t^{n-1}}^{t^n}\frac{\zeta(t+\Delta t)-\zeta(t)}{\Delta t}\,dt\right].
\end{aligned}$$

Note that for $\Delta t$ small enough, the discrete integration by parts in time produces no boundary term since $\zeta$ vanishes on $[0,\Delta t]$ and on $[T-\Delta t, T]$. Hence from $0 \le U^n_{i,j} \le 1$ we deduce

$$\left|\int_0^T \left(\mathbb{P}^0_t \widetilde{\mathcal{R}}_n\right)\,\zeta\right| \le \Delta x\, T\, \|\varphi\|_{\mathrm{L}^\infty}\|\zeta'\|_{\mathrm{L}^\infty} \le C\Delta x,$$

thus $\widetilde{\mathcal{R}}_n$ converges to 0 in $\mathcal{D}'(0,t^N)$. The second term, $\widetilde{\mathcal{S}}_n$, converges to 0 uniformly with respect to $n$. Indeed, an integration by parts gives

$$\begin{aligned}
\widetilde{\mathcal{S}}_n &= \Delta z\Delta x\sum_{i=1}^{N_z-1} U^n_{i,N_x}\,\left(D_{-1,0}\,D_{1,0}\,\varphi_{i,N_x}\right) + \Delta x\, D_{1,0}\,\varphi_{0,N_x} \\
&\le \Delta x\,\left(\|\partial_{zz}\varphi\|_{\mathrm{L}^\infty} + \|\partial_z\varphi\|_{\mathrm{L}^\infty}\right).
\end{aligned}$$

Finally letting $h \to 0$ in (4.9) gives

$$\frac{d}{dt}\int_\Omega U\,\varphi = -\int_\Omega \nabla U.\nabla\varphi + K\int_\Omega U(\partial_z U)\,\partial_x(\varphi\,\theta) - K\int_\Omega U(\partial_x U)\,(\partial_z\varphi)\theta \qquad (4.10)$$

in the $\mathcal{D}'(0,t^N)$ sense. By remarking that $\partial_x\theta_x = -1/l$, we conclude that $U$ verifies (1.9). The continuity of $U$ with respect to time in (4.6) is deduced from an Aubin-Lions theorem and the estimates (4.7) can be obtained easily by passing to the limit in (3.2) and (3.7). Actually, we have not proved yet that all the sequence converges, but only a subsequence of $U_h$. This result will be deduced from the uniqueness result of Theorem 4.3. □

## 4.3 Uniqueness of the variational solution

**Theorem 4.3** *There exists a unique weak solution of* (1.2)–(1.7)

$$U \in \mathrm{L}^2((0,T),\mathrm{H}^1(\Omega)) \cap \mathrm{L}^\infty((0,T)\times\Omega)$$

*in the class of the solutions that verify*

$$(\partial_z U)^+ \in \mathrm{L}^\infty((0,T)\times\Omega). \qquad (4.11)$$

**Proof.** A weak solution of (1.2)–(1.7) verifies (1.8) for every function $\varphi \in \mathrm{H}^1(\Omega)$:

$$\int_\Omega \partial_t U \,\varphi + \int_\Omega \nabla U.\nabla\varphi + \frac{K}{2}\left\langle \partial_z(U^2),\, \varphi\right\rangle_{\mathrm{H}^{1/2}(\Gamma_1),\mathrm{H}^{1/2}(\Gamma_1)} = 0,$$

where $\partial_t U \in \mathrm{L}^2((0,T),V^*)$ and $\partial_z(U^2) \in \mathrm{L}^2((0,T),\mathrm{H}^{-1/2}(\Gamma_1))$.

Let $U_1$ and $U_2$ be two weak solution of this problem which verify (4.11). We set $w = U_1 - U_2$. The idea of the proof is to take as a test-function a regularization of $sign(w)$. We define the piecewise $\mathrm{C}^1$ function:

$$S_\delta(w) = \begin{cases} -w - \dfrac{\delta}{2} & \text{for } w \le -\delta \\ \dfrac{w^2}{2\delta} & \text{for } -\delta \le w \le \delta \\ w - \dfrac{\delta}{2} & \text{for } w \ge \delta \end{cases}$$

then we take $\varphi = S'_\delta(w)$ and remark that this function belongs to $\mathrm{L}^2((0,T),V)$. After some calculations we get

$$\int_\Omega S_\delta(w) + \int_{t=0}^{t=T}\int_\Omega S''_\delta(w)\,|\nabla w|^2 = -\frac{K}{2}\int_{t=0}^{t=T}\mathcal{B}(t)dt, \tag{4.12}$$

with

$$\mathcal{B} = \left\langle \partial_z(U_1+U_2)\,,\, wR'_\delta(w) - R_\delta(w)\right\rangle_{\mathrm{H}^{-1/2},\mathrm{H}^{1/2}}.$$

The condition (4.11) will enable us to estimate $\mathcal{B}$. We remark that

$$\forall v \in \mathbb{R}, \quad 0 \le vR'_\delta(v) - R_\delta(v) \le \frac{\delta}{2},$$

thus we have, almost everywhere on $(0,T)$,

$$\mathcal{B} \le \int_{\Gamma_1}\left(\|(\partial_z U_1)^+\|_{L^\infty} + \|(\partial_z U_2)^+\|_{L^\infty}\right)\frac{\delta}{2} \le C\,\delta.$$

Therefore, the function $S''_\delta$ being nonnegative, (4.12) gives $\int_\Omega S_\delta(w) \le C\,\delta$. To conclude we let $\delta \to 0$ in this inequality thanks to the Lebesgue dominated convergence theorem (we have $0 \le S_\delta(w) \le |w|$ a.e. in $\Omega$). $\square$

# 5 Numerical results

## 5.1 Fast propagation at the boundary

The main effect that we wanted to observe was the fast diffusion of the quantity $U$ (the magnetic field in the application modelled here) near the boundary. In order to show this phenomenon, we provide here several numerical results obtained for different values of the parameter $K$.

The functions represented below are the following ones:

- Figure 3: the initial data $U^0(z,x)$.
- Figure 4: the solution $U(t^N,z,x)$ computed for $K=0$ at $t^N = 200\Delta t$. This is the solution of the one-dimensional heat equation (the function is invariant along the direction $x$) which can be used as the benchmark to be compared with the solutions for $K>0$.
- Figure 5: the solution $U(t^N,z,x)$ computed for $K=3$ at $t^N = 200\Delta t$.
- Figure 6: the solution $U(t^N,z,x)$ computed for $K=10$ at $t^N = 200\Delta t$.

The common parameters of these computations are

| $L$ | $l$ | $N_z$ | $N_x$ | $\Delta t$ | $N$ |
|---|---|---|---|---|---|
| 2 | 1 | 50 | 25 | $2\times 10^{-5}$ | 200 |

We recall the $\mathrm{L}^\infty$ stability criterium (3.1) and we set

$$\nu = \left(2+\max\left(0,\ K\frac{\Delta x}{\Delta z}-1\right)\right)\frac{\Delta t}{\Delta x^2}+2\frac{\Delta t}{\Delta z^2}.$$

The stability condition is fulfilled in the three cases $K=0$, 3 and 10, since we have

| $K$ | 0 | 3 | 10 |
|---|---|---|---|
| $\nu$ | 0.05 | 0.08 | 0.16 |

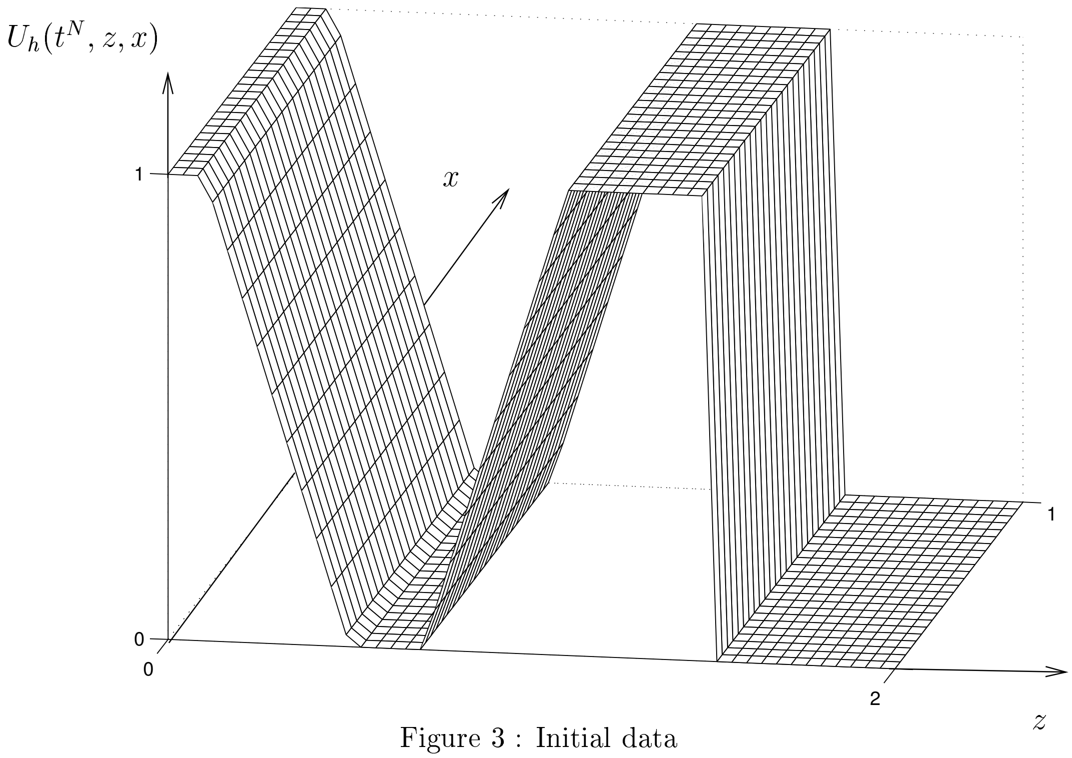


Figure 3 : Initial data

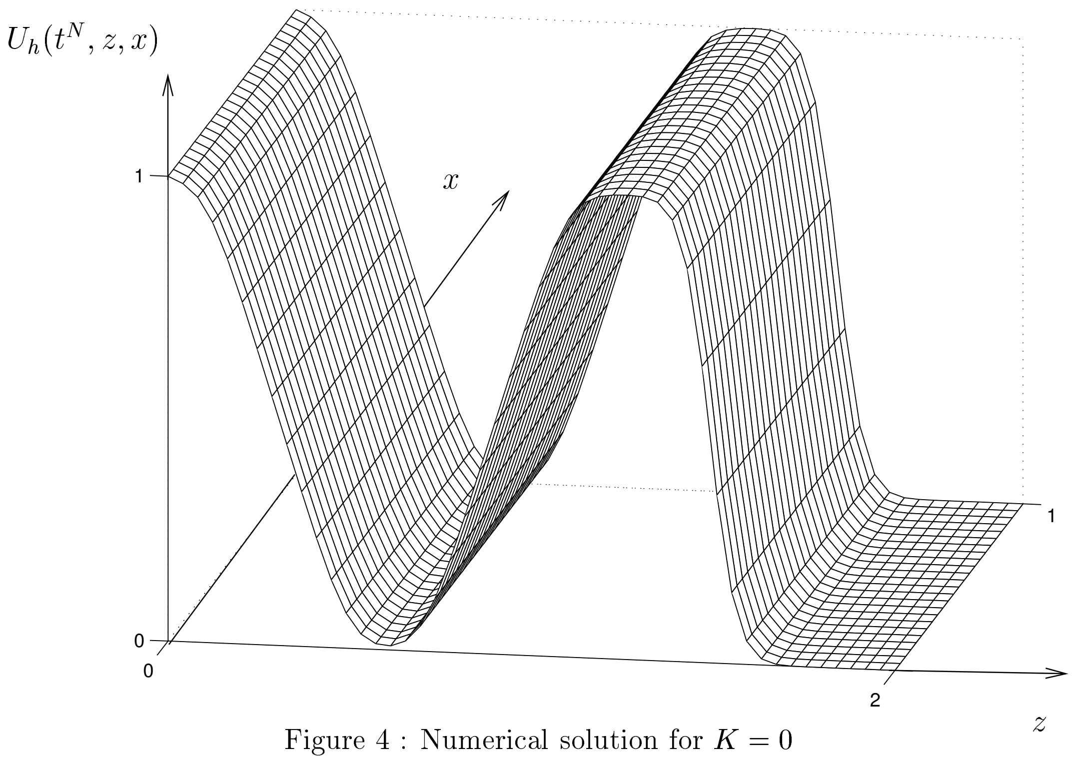


Figure 4 : Numerical solution for $K = 0$

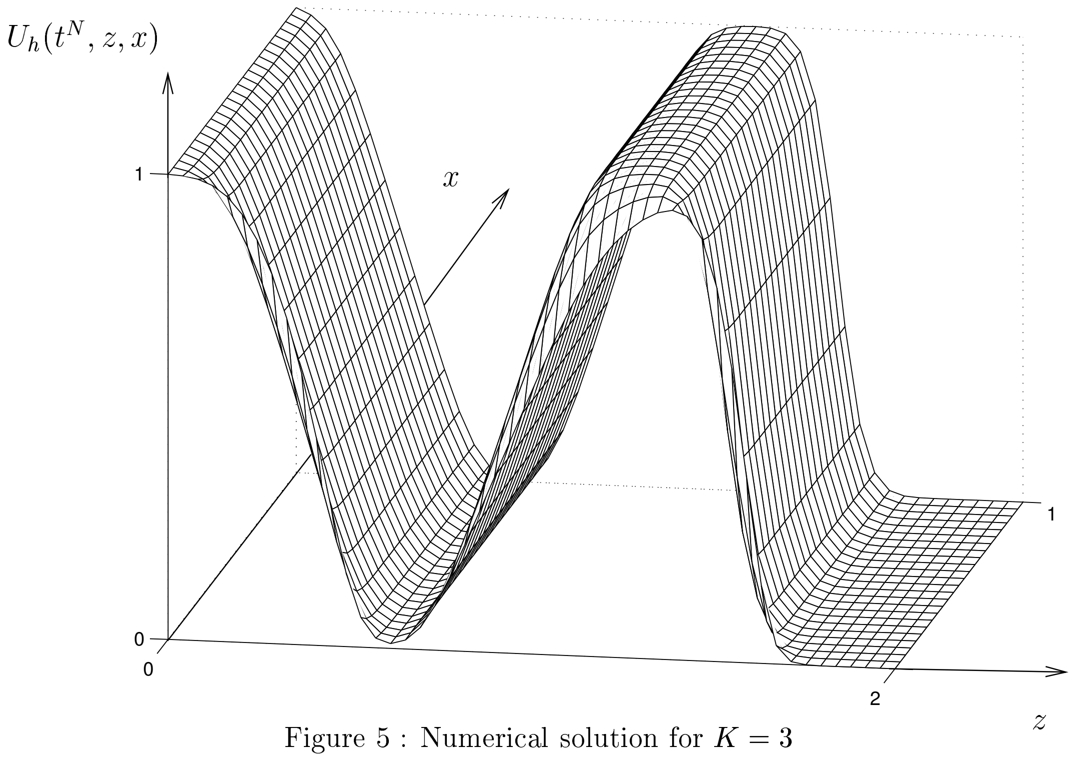


Figure 5 : Numerical solution for $K = 3$

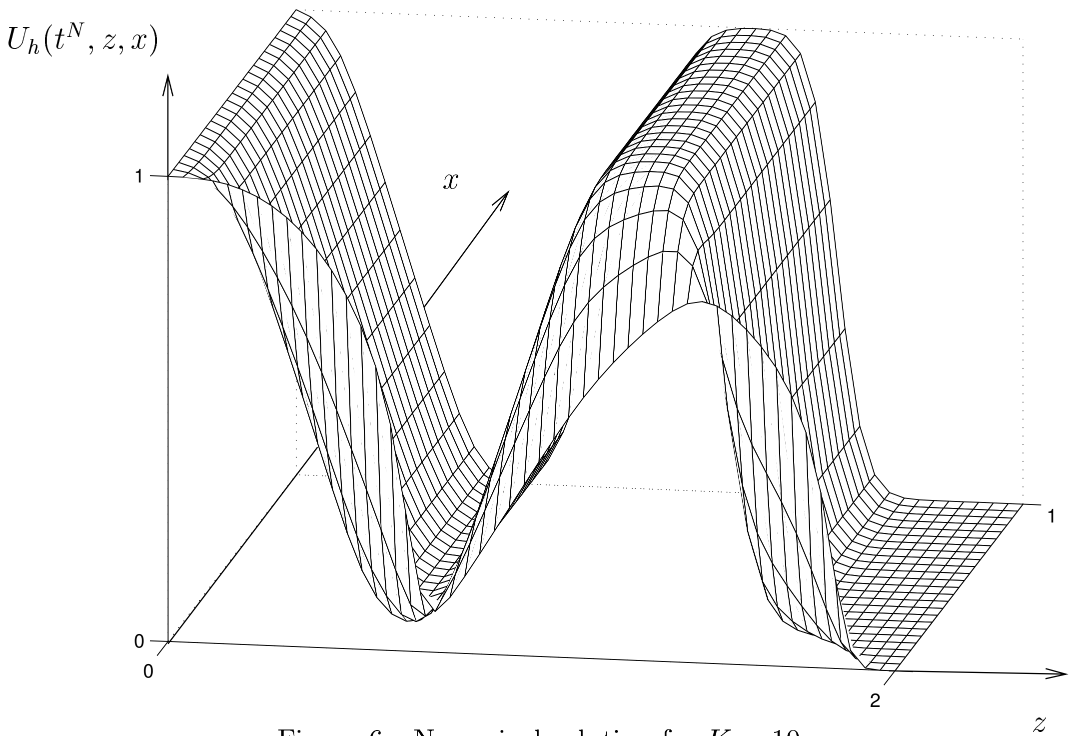


Figure 6 : Numerical solution for $K = 10$

From these numerical results, one can make the following two observations.

- When $K > 0$ a propagation from the left to the right can be observed near the boundary $x = 0$: this is the fast propagation of the field.

- The larger is $K$, the faster is this diffusion near the boundary.

## 5.2 Numerical stability

In this section we show –numerically– that the theorical $\mathrm{L}^\infty$ stability condition (3.15) is optimal. The gridsteps $\Delta z$ and $\Delta x$ and an initial data are fixed here. The parameters which vary are $K$ and $\Delta t$. Let

$$\mathrm{stab}(K) = \frac{1}{\left(2 + \max\left(0,\, K\frac{\Delta x}{\Delta z} - 1\right)\right)\frac{1}{\Delta x^2} + \frac{2}{\Delta z^2}}.$$

The theorical $\mathrm{L}^\infty$ stability condition (3.1) writes

$$\Delta t \leq \mathrm{stab}(K).$$

The following figure represents the stability diagram, where we plotted the theorical stability condition $\mathrm{stab}(K)$. We applied a grid on the domain in $(K, \Delta t)$ and used our numerical scheme for the points of this grid. On Figure 7, a cross represents a pair $(K, \Delta t)$ which lead to a numerical instability. For the sake of clarity of the figure, we did not represent the pairs that lead to a stable calculation.

Numerically, one can distinguish a *stable domain* and an *unstable domain*, separated by a curve. This curve –the bottom line of the unstable domain represented on Figure 7– is the numerical stability condition. We refined the mesh in $K$ and $\Delta t$ near this curve. As a conclusion, one can note a very good agreement between the theorical stability condition and the numerical results.

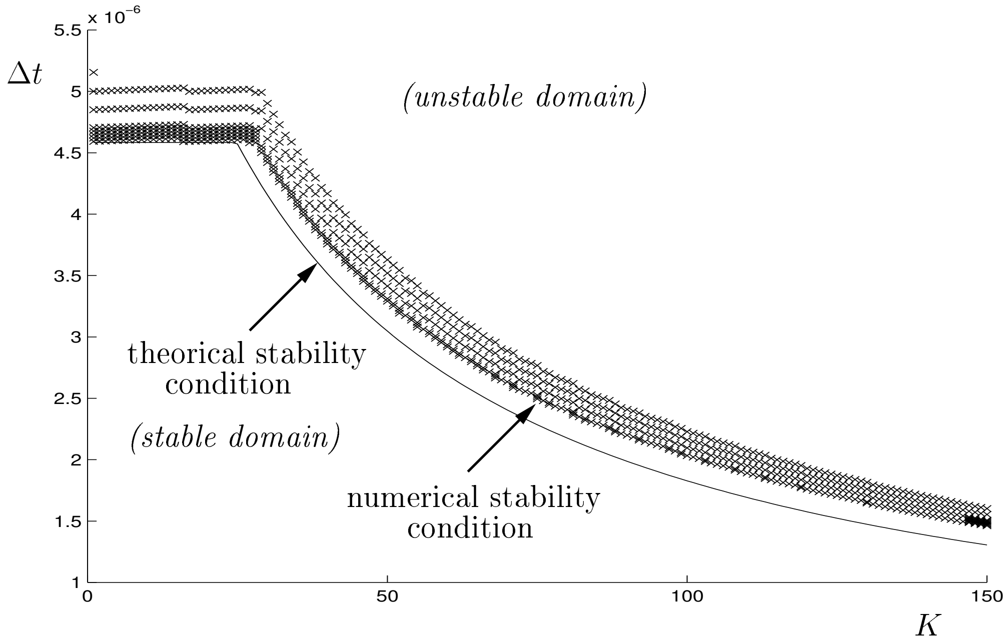


Figure 7 : Stability of the scheme

# References


[1] L. Caffarelli, J.-M. Roquejoffre, *A nonlinear oblique derivative boundary value problem for the heat equation: analogy with the porous medium equation*, Ann. Inst. H. Poincaré Anal. Non Linéaire **19** (2002), no. 1, 41–80.

[2] G. Dong, *Initial and nonlinear oblique boundary value problems for fully nonlinear parabolic equations*, J. Partial Differential Equations Ser. A **1** (1988), no. 2, 12-42.

[3] E. Godlewski, P.-A. Raviart, *Hyperbolic systems of conservation laws*, Mathématiques & Applications, Ellipse, Paris (1991).

[4] B. Larrouturou, *Modélisation mathématique et numérique pour les sciences de l'ingénieur*, Éditions de l'École polytechnique (to appear).

[5] R. J. LeVeque, *Numerical Methods for Conservation Laws*, Lectures in Mathematics, Birkhäuser Verlag (1990).

[6] J.-L. Lions, *Quelques méthodes de résolution des problèmes aux limites non linéaires*, Études Mathématiques, Dunod, Gauthier-Villars (1969).

[7] J.-L. Lions, E. Magenes, *Problèmes aux limites non homogènes et applications*, vol. 1, Travaux et recherches Mathématiques, Dunod (1968).

[8] F. Méhats, *Étude de problèmes aux limites en physique du transport des particules chargées*, thèse de doctorat (1997).

[9] F. Méhats, J.-M. Roquejoffre, *A nonlinear oblique derivative boundary value problem for the heat equation, part 1 and part 2*, Ann. Inst. H. Poincaré Anal. Non Linéaire **16** (1999) 221-253 and 691-724.

[10] A. I. Nazarov, N. N. Ural'tseva, *A Problem with an Oblique Derivative for a Quasilinear Parabolic Equation*, J. Math. Sci. **77** (1995), no. 3, 3212–3220.