

# The rain on underground porous media

## Part I. Analysis of a Richards model

by Christine Bernardi<sup>1</sup>, Adel Blouza<sup>2</sup>, and Linda El Alaoui<sup>3</sup>

**Abstract:** Richards equation models the water flow in a partially saturated underground porous medium under the surface. When it rains on the surface, boundary conditions of Signorini type must be considered on this part of the boundary. We first study this problem which results into a variational inequality. We propose a discretization by an implicit Euler's scheme in time and finite elements in space. The convergence of this discretization leads to the well-posedness of the problem.

**Résumé:** L'équation de Richards modélise l'écoulement d'eau dans un milieu poreux partiellement saturé souterrain. Lorsqu'il pleut, des conditions aux limites de type Signorini doivent être imposées sur la surface. Nous étudions tout d'abord ce problème qui se traduit par une inéquation variationnelle. Nous écrivons une discrétisation par schéma d'Euler implicite en temps et éléments finis en espace. La convergence de cette discrétisation permet de vérifier que le problème est bien posé.

---

<sup>1</sup> Laboratoire Jacques-Louis Lions, C.N.R.S. & Université Pierre et Marie Curie,  
B.C. 187, 4 place Jussieu, 75252 Paris Cedex 05, France.  
e-mail address: [bernardi@ann.jussieu.fr](mailto:bernardi@ann.jussieu.fr)

<sup>2</sup> Laboratoire de Mathématiques Raphaël Salem (U.M.R. 6085 C.N.R.S.), Université de Rouen,  
avenue de l'Université, B.P. 12, 76801 Saint-Étienne-du-Rouvray, France.  
e-mail address: [Adel.Blouza@univ-rouen.fr](mailto:Adel.Blouza@univ-rouen.fr)

<sup>3</sup> Université Paris 13, C.N.R.S., U.M.R. 7539,  
L.A.G.A., 99 avenue Jean-Baptiste Clément, 93430 Villetaneuse, France.  
e-mail address: [elalaoui@math.univ-paris13.fr](mailto:elalaoui@math.univ-paris13.fr)



## 1. Introduction.

The following equation

$$\partial_t \tilde{\Theta}(\psi) - \nabla \cdot K_w(\Theta(\psi)) \nabla(\psi + z) = 0, \quad (1.1)$$

models the flow of a wetting fluid, mainly water, in the underground surface, hence in an unsaturated medium, see L.A. Richards [14] for the introduction of this type of models. In opposite to Darcy's or Brinkman's systems (see [13] for all these models), this equation which is derived by combining Darcy's generalized equation with the mass conservation law is highly nonlinear: This follows from the fact that, due to the presence of air above the surface, the porous medium is only partially saturated with water. The unknown  $\psi$  is the difference between the pressure of water and the atmospherical pressure.

This equation is usually provided with Dirichlet or Neuman type boundary conditions. Indeed, Neumann boundary conditions on the underground part of the boundary are linked to the draining of water outside of the domain and Dirichlet boundary conditions on the surface are introduced to take into account the rain. However, when the porous media can no longer absorb the rainwater that falls, the upper surface of the domain allows to exfiltration and infiltration. In other words, the upper surface is divided into a saturated and unsaturated zone. We assume that the re-infiltration process is negligible. This leads to variational inequalities of the following type :

$$-\psi \geq 0, \quad \mathbf{v}(\psi) \cdot \mathbf{n} \geq \mathbf{v}_r \cdot \mathbf{n}, \quad \psi(\mathbf{v}(\psi) \cdot \mathbf{n} - \mathbf{v}_r \cdot \mathbf{n}) = 0, \quad (1.2)$$

where  $\mathbf{v}(\psi)$  is the flux

$$\mathbf{v}(\psi) = -K_w(\Theta(\psi)) \nabla(\psi + z). \quad (1.3)$$

Here,  $\mathbf{n}$  stands for the unit normal vector to the surface and  $\mathbf{v}_r$  for a given rain fall rate. We refer to the thesis of H. Berninger [4] for the full derivation of this model from hydrology laws and more specifically to [4, Section 1.5] for the derivation of the boundary inequalities (1.2).

It is not so easy to give a mathematical sense to the system (1.1) – (1.2). As standard, the key argument for the analysis of problem (1.1) is to use Kirchhoff's change of unknowns. Indeed, after this transformation, the new equation fits the general framework proposed in [1] (see also [6] for the analysis of a different model). Thus, the existence and uniqueness of a solution to this equation when provided with appropriate linear initial and boundary conditions can be derived from standard arguments. In order to handle the inequality in (1.2), we again use a variational formulation. We refer to [2] for a first analysis of very similar systems, see also [5]. In this paper, we prove that problem (1.1) – (1.2) is well-posed when the coefficients and the data are smooth enough but without any other restriction.

The discretization of problem (1.1) has been proposed and/or studied in many papers when provided with standard boundary conditions, see [7], [12], [15], [17] and [18] and

also [16] for a more general equation, however it seems that it has not been treated when provided with the boundary inequality (1.2). We propose here a discretization of system (1.1) – (1.2), in two steps:

- (i) We first use the Euler’s implicit scheme to build a time semi-discrete problem, where one of the nonlinear terms is treated in an explicit way for simplicity;
- (ii) We then construct a fully discrete problem that relies on the Galerkin method and finite elements in the spatial domain.

In both cases, we prove that the corresponding variational problem is well-posed.

To conclude, we prove that the solution of this discrete problem converges to a solution of the continuous one when the discretization parameters tend to zero. This makes complete our existence result, since no restrictive condition is needed here.

The outline of the paper is as follows.

- In Section 2, we present the variational formulation of the full system and investigate its wellposedness in appropriate Sobolev spaces.
- Section 3 is devoted to the description of the time semi-discrete problem and of the fully discrete problem. We check their well-posedness.
- In Section 4, we investigate the convergence of the solution of the discrete problem towards a solution of the continuous one.

## 2. The continuous problem and its well-posedness.

Let  $\Omega$  be a bounded connected open set in  $\mathbb{R}^d$ ,  $d = 2$  or  $3$ , with a Lipschitz-continuous boundary  $\partial\Omega$ , and let  $\mathbf{n}$  denote the unit outward normal vector to  $\Omega$  on  $\partial\Omega$ . We assume that  $\partial\Omega$  admits a partition without overlap into three parts  $\Gamma_B$ ,  $\Gamma_F$ , and  $\Gamma_G$  (these indices mean “bottom”, “flux” and “ground”, respectively) and that  $\Gamma_B$  has a positive measure. Let also  $T$  be a positive real number.

In order to perform the Kirchhoff’s change of unknowns in problem (1.1), we observe that, since the conductivity coefficient  $K_w$  is positive, the mapping:

$$x \mapsto \mathcal{K}(x) = \int_0^x K_w(\Theta(\xi)) d\xi,$$

is one-to-one from  $\mathbb{R}$  into itself. Thus, by setting

$$u = \mathcal{K}(\psi), \quad b(u) = \Theta \circ \mathcal{K}^{-1}(u), \quad k \circ b(u) = K_w \circ \Theta \circ \mathcal{K}^{-1}(u),$$

and thanks to an appropriate choice of the function  $\tilde{\Theta}$ , we derive the equation (more details are given in [3, Remark 2.1] for instance)

$$\alpha \partial_t u + \partial_t b(u) - \nabla \cdot \left( \nabla u + k \circ b(u) \mathbf{e}_z \right) = 0 \quad \text{in } \Omega \times ]0, T[,$$

where  $-\mathbf{e}_z$  stands for the unit vector in the direction of gravity. Moreover, the Kirchhoff’s change of unknowns has the further property of preseving the positivity:  $u$  is positive if and only if  $\psi$  is positive, negative if and only if  $\psi$  is negative, So, writing inequation (1.2) in terms of the unknown  $u$  is easy.

As a consequence, from now on, we work with the following system

$$\left\{ \begin{array}{ll} \alpha \partial_t u + \partial_t b(u) - \nabla \cdot \left( \nabla u + k \circ b(u) \mathbf{e}_z \right) = 0 & \text{in } \Omega \times ]0, T[, \\ u = u_B & \text{on } \Gamma_B \times ]0, T[, \\ - \left( \nabla u + k \circ b(u) \mathbf{e}_z \right) \cdot \mathbf{n} = f_F & \text{on } \Gamma_F \times ]0, T[, \\ \begin{array}{l} u \leq 0, \quad - \left( \nabla u + k \circ b(u) \mathbf{e}_z \right) \cdot \mathbf{n} \geq \mathbf{q}_r \cdot \mathbf{n}, \\ u \left( \nabla u + k \circ b(u) \mathbf{e}_z + \mathbf{q}_r \right) \cdot \mathbf{n} = 0 \end{array} & \text{on } \Gamma_G \times ]0, T[, \\ u|_{t=0} = u_0 & \text{in } \Omega. \end{array} \right. \quad (2.1)$$

The unknown is now the quantity  $u$ . The data are the Dirichlet boundary condition  $u_B$  on  $\Gamma_B$  and the initial condition  $u_0$  on  $\Omega$ , together with the boundary conditions  $f_F$  and  $\mathbf{q}_r$  on the normal component of the flux, with  $f_F$  corresponding to the draining of water and  $\mathbf{q}_r$  corresponding to the rain. Finally, the coefficients  $b$  and  $k$  are supposed to be known,

while  $\alpha$  is a positive constant. From now on, we assume that the function  $b$  is continuously differentiable on  $\mathbb{R}$ , with a bounded and Lipschitz-continuous first derivative  $b'$ , and that the function  $k \circ b$  is continuous, bounded, and uniformly Lipschitz-continuous on  $\mathbb{R}$ .

In what follows, we use the whole scale of Sobolev spaces  $W^{m,p}(\Omega)$ , with  $m \geq 0$  and  $1 \leq p \leq +\infty$ , equipped with the norm  $\|\cdot\|_{W^{m,p}(\Omega)}$  and seminorm  $|\cdot|_{W^{m,p}(\Omega)}$ , with the usual notation  $H^m(\Omega)$  when  $p = 2$ . As standard, the range of  $H^1(\Omega)$  by the trace operator on any part  $\Gamma$  of  $\partial\Omega$  is denoted by  $H^{\frac{1}{2}}(\Gamma)$ . For any separable Banach space  $E$  equipped with the norm  $\|\cdot\|_E$ , we denote by  $\mathcal{C}^0(0, T; E)$  the space of continuous functions from  $[0, T]$  with values in  $E$ . For each integer  $m \geq 0$ , we also introduce the space  $H^m(0, T; E)$  as the space of measurable functions on  $]0, T[$  with values in  $E$  such that the mappings:  $v \mapsto \|\partial_t^\ell v\|_E$ ,  $0 \leq \ell \leq m$ , are square-integrable on  $]0, T[$ .

To write a variational formulation for the problem, we introduce the time-dependent subset

$$\mathbb{V}(t) = \{v \in H^1(\Omega); v|_{\Gamma_B} = u_B(\cdot, t) \text{ and } v|_{\Gamma_G} \leq 0\}. \quad (2.2)$$

It is readily checked that each  $\mathbb{V}(t)$  is closed and convex, see [4, Prop. 1.5.5], when  $u_B$  belongs to  $\mathcal{C}^0(0, T; H^{\frac{1}{2}}(\Gamma_B))$ . Thus, we are led to consider the following variational problem (with obvious notation for  $L^2(0, T; \mathbb{V})$ )

Find  $u$  in  $L^2(0, T; \mathbb{V})$  such that

$$u|_{t=0} = u_0, \quad (2.3)$$

and that, for a.e.  $t$  in  $]0, T[$ ,

$$\begin{aligned} \forall v \in \mathbb{V}(t), \\ \alpha \int_{\Omega} (\partial_t u)(\mathbf{x}, t)(v - u)(\mathbf{x}, t) d\mathbf{x} + \int_{\Omega} (\partial_t b(u))(\mathbf{x}, t)(v - u)(\mathbf{x}, t) d\mathbf{x} \\ + \int_{\Omega} \left( \nabla u + k \circ b(u) \mathbf{e}_z \right)(\mathbf{x}, t) \cdot (\nabla(v - u))(\mathbf{x}, t) d\mathbf{x} \\ \geq - \int_{\Gamma_F} f_F(\boldsymbol{\tau}, t)(v - u)(\boldsymbol{\tau}, t) d\boldsymbol{\tau} - \int_{\Gamma_G} (\mathbf{q}_r \cdot \mathbf{n})(\boldsymbol{\tau}, t)(v - u)(\boldsymbol{\tau}, t) d\boldsymbol{\tau}, \end{aligned} \quad (2.4)$$

where  $\boldsymbol{\tau}$  denotes the tangential coordinates on  $\partial\Omega$ . The reason for this follows.

**Proposition 2.1.** *Problems (2.1) and (2.3) – (2.4) are equivalent, more precisely:*

(i) *Any solution of problem (2.1) in  $L^2(0, T; H^1(\Omega)) \cap H^1(0, T; L^2(\Omega))$  is a solution of (2.3) – (2.4);*

(ii) *Any solution of problem (2.3) – (2.4) is a solution of problem (2.1) in the distribution sense.*

**Proof:** We check successively the two assertions of the proposition.

1) Let  $u$  be any solution of (2.1) in  $L^2(0, T; H^1(\Omega)) \cap H^1(0, T; L^2(\Omega))$ . Obviously, it belongs

to  $L^2(0, T; \mathbb{V})$  and satisfies (2.3). Next, we observe that, for any  $v$  in  $\mathbb{V}(t)$ , the function  $v - u$  vanishes on  $\Gamma_B$ . Multiplying the first line in (2.1) by this function and integrating by parts on  $\Omega$  give

$$\begin{aligned} & \alpha \int_{\Omega} (\partial_t u)(\mathbf{x}, t)(v - u)(\mathbf{x}, t) \, d\mathbf{x} + \int_{\Omega} (\partial_t b(u))(\mathbf{x}, t)(v - u)(\mathbf{x}, t) \, d\mathbf{x} \\ & + \int_{\Omega} \left( \nabla u + k \circ b(u) \mathbf{e}_z \right)(\mathbf{x}, t) \cdot (\nabla(v - u))(\mathbf{x}, t) \, d\mathbf{x} \\ & = \int_{\Gamma_F \cup \Gamma_G} (\nabla u + k \circ b(u) \mathbf{e}_z) \cdot \mathbf{n}(\boldsymbol{\tau})(v - u)(\boldsymbol{\tau}, t) \, d\boldsymbol{\tau}. \end{aligned}$$

To conclude, we observe that, on  $\Gamma_G$ , either  $u$  is zero and  $\nabla u + k \circ b(u) \mathbf{e}_z$  is smaller than  $-\mathbf{q}_r \cdot \mathbf{n}$  or  $u$  is not zero and  $\nabla u + k \circ b(u) \mathbf{e}_z$  is equal to  $-\mathbf{q}_r \cdot \mathbf{n}$ . All this yields (2.4).

2) Conversely, let  $u$  be any solution of (2.3) – (2.4).

- By noting that, for any function  $w$  in  $\mathcal{D}(\Omega)$ ,  $(u + w)(\cdot, t)$  belongs to  $\mathbb{V}(t)$  and taking  $v$  equal to  $u \pm w$  in (2.4), we obtain the first line of (2.1) in the distribution sense.
- The second line in (2.1) follows from the definition of  $\mathbb{V}(t)$ .
- By taking  $v$  equal to  $u \pm w$  for any  $w$  in  $\mathcal{D}(\Omega \cup \Gamma_F)$ , we also derive the third line in (2.1).
- The fact that  $u$  is nonpositive on  $\Gamma_G$  comes from the definition of  $\mathbb{V}(t)$ . On the other hand, the previous equations imply that, for any  $v$  in  $\mathbb{V}(t)$ ,

$$\int_{\Gamma_G} (\nabla u + k \circ b(u) \mathbf{e}_z) \cdot \mathbf{n}(\boldsymbol{\tau})(v - u)(\boldsymbol{\tau}, t) \, d\boldsymbol{\tau} \geq - \int_{\Gamma_G} (\mathbf{q}_r \cdot \mathbf{n})(\boldsymbol{\tau}, t)(v - u)(\boldsymbol{\tau}, t) \, d\boldsymbol{\tau}.$$

Taking  $v$  equal to  $u + w$  where  $w$  vanishes on  $\Gamma_B$  and is nonpositive on  $\Gamma_G$  yields that  $-(\nabla u + k \circ b(u) \mathbf{e}_z) \cdot \mathbf{n}$  is larger than  $\mathbf{q}_r \cdot \mathbf{n}$ . Finally taking  $v$  equal to zero on  $\Gamma_G$  yields that

$$\int_{\Gamma_G} (\nabla u + k \circ b(u) \mathbf{e}_z + \mathbf{q}_r) \cdot \mathbf{n}(\boldsymbol{\tau}) u(\boldsymbol{\tau}, t) \, d\boldsymbol{\tau} \leq 0,$$

and since the two quantities  $u$  and  $(\nabla u + k \circ b(u) \mathbf{e}_z + \mathbf{q}_r)$  are nonpositive on  $\Gamma_G$ , their product is zero.

- And finally the last line of (2.1) is written in (2.3).

Proving that problem (2.3) – (2.4) is well-posed is not at all obvious. We begin with the simpler result, i.e., the uniqueness of the solution. For brevity, we set:

$$\mathbb{X} = L^2(0, T; \mathbb{V}) \cap H^1(0, T; L^2(\Omega)). \quad (2.5)$$

We also refer to [11, Chap. 1, Th. 11.7] for the definition of the space  $H_{00}^{\frac{1}{2}}(\Gamma_B)$ .

**Proposition 2.2.** *For any data  $u_B$ ,  $f_F$ ,  $\mathbf{q}_r$  and  $u_0$  satisfying*

$$\begin{aligned} u_B \in H^1(0, T; H_{00}^{\frac{1}{2}}(\Gamma_B)), \quad f_F \in L^2(0, T; L^2(\Gamma_F)), \\ \mathbf{q}_r \in L^2(0, T; L^2(\Gamma_G)^d), \quad u_0 \in H^1(\Omega), \end{aligned} \quad (2.6)$$

problem (2.3) – (2.4) has at most a solution in  $\mathbb{X}$ .

**Proof:** Let  $u_1$  and  $u_2$  be two solutions of problem (2.3) – (2.4). Thus, the function  $u = u_1 - u_2$  vanishes on  $\Gamma_B$  and at  $t = 0$ . Taking  $v$  equal to  $u_2$  in the problem satisfied by  $u_1$  and equal to  $u_1$  in the problem satisfied by  $u_2$  and subtracting the second problem from the first one, we obtain

$$\begin{aligned} \alpha \int_{\Omega} (\partial_t u)(\mathbf{x}, t) u(\mathbf{x}, t) \, d\mathbf{x} + \int_{\Omega} (\partial_t b(u_1) - \partial_t b(u_2))(\mathbf{x}, t) u(\mathbf{x}, t) \, d\mathbf{x} \\ + \int_{\Omega} (\nabla u)^2(\mathbf{x}, t) \, d\mathbf{x} + \int_{\Omega} (k \circ b(u_1) - k \circ b(u_2))(\mathbf{x}, t) \mathbf{e}_z \cdot (\nabla u)(\mathbf{x}, t) \, d\mathbf{x} \leq 0. \end{aligned} \quad (2.7)$$

We integrate this equation with respect to  $t$ , use the decomposition

$$\begin{aligned} \int_{\Omega} (\partial_t b(u_1) - \partial_t b(u_2))(\mathbf{x}, t) u(\mathbf{x}, t) \, d\mathbf{x} = \int_{\Omega} b'(u_1)(\mathbf{x}, t) (\partial_t u)(\mathbf{x}, t) u(\mathbf{x}, t) \, d\mathbf{x} \\ + \int_{\Omega} (b'(u_1) - b'(u_2))(\mathbf{x}, t) \partial_t u_2(\mathbf{x}, t) u(\mathbf{x}, t) \, d\mathbf{x}. \end{aligned}$$

Next, the nonnegativity and Lipschitz-continuity of  $b'$ , combined with the interpolation inequality (see [11, Chap. 1, Prop. 2.3]) and the Poincaré–Friedrichs inequality

$$\|u\|_{L^4(\Omega)} \leq \|u\|_{L^2(\Omega)}^{1-\frac{d}{4}} (c|u|_{H^1(\Omega)})^{\frac{d}{4}} \leq c'(1 - \frac{d}{4}) \|u\|_{L^2(\Omega)} + \frac{d}{4} |u|_{H^1(\Omega)},$$

lead to handle the second term. The Lipschitz-continuity of  $k \circ b$  together with a Young's inequality is needed to handle the fourth term. All this gives

$$\alpha \|u(\cdot, t)\|_{L^2(\Omega)}^2 + \frac{1}{2} \int_0^t |u(\cdot, s)|_{H^1(\Omega)}^2 \, ds \leq c(u_2) \int_0^t \|u(\cdot, s)\|_{L^2(\Omega)}^2 \, ds.$$

Thus, applying Grönwall's lemma yields that  $u$  is zero, whence the uniqueness result.

Proving the existence is much more complex. We begin with a basic result.

**Lemma 2.3.** *If the function  $u_B$  belongs to  $\mathcal{C}^0(0, T; H_{00}^{\frac{1}{2}}(\Gamma_B))$ , for all  $t$  in  $[0, T]$ , the convex set  $\mathbb{V}(t)$  is not empty.*

**Proof:** Denoting by  $\bar{u}_B(\cdot, t)$  the extension by zero of  $u_B(\cdot, t)$  to  $\partial\Omega$ , we observe that any lifting of  $\bar{u}_B(\cdot, t)$  in  $H^1(\Omega)$  belongs to  $\mathbb{V}(t)$ , whence the desired result.

In a first step, we consider the linear problem, for any datum  $F$  in  $L^2(0, T; L^2(\Omega))$ :



Find  $u$  in  $L^2(0, T; \mathbb{V})$  satisfying (2.3) and such that, for a.e.  $t$  in  $]0, T[$ ,

$$\begin{aligned} \forall v \in \mathbb{V}(t), \\ \alpha \int_{\Omega} (\partial_t u)(\mathbf{x}, t)(v - u)(\mathbf{x}, t) d\mathbf{x} + \int_{\Omega} (\nabla u)(\mathbf{x}, t) \cdot (\nabla(v - u))(\mathbf{x}, t) d\mathbf{x} \\ \geq - \int_{\Omega} F(\mathbf{x}, t)(v - u)(\mathbf{x}, t) d\mathbf{x} - \int_{\Gamma_F} f_F(\boldsymbol{\tau}, t)(v - u)(\boldsymbol{\tau}, t) d\boldsymbol{\tau} \\ - \int_{\Gamma_G} (\mathbf{q}_r \cdot \mathbf{n})(\boldsymbol{\tau}, t)(v - u)(\boldsymbol{\tau}, t) d\boldsymbol{\tau}. \end{aligned} \quad (2.8)$$

However a weaker formulation of this problem can be derived by integrating in time. It reads

Find  $u$  in  $L^2(0, T; \mathbb{V})$  satisfying (2.3) and such that

$$\begin{aligned} \forall v \in \mathbb{X}, \\ \alpha \int_0^T \int_{\Omega} (\partial_t u)(\mathbf{x}, t)(v - u)(\mathbf{x}, t) d\mathbf{x} dt + \int_0^T \int_{\Omega} (\nabla u)(\mathbf{x}, t) \cdot (\nabla(v - u))(\mathbf{x}, t) d\mathbf{x} dt \\ \geq - \int_0^T \int_{\Omega} F(\mathbf{x}, t)(v - u)(\mathbf{x}, t) d\mathbf{x} dt - \int_0^T \int_{\Gamma_F} f_F(\boldsymbol{\tau}, t)(v - u)(\boldsymbol{\tau}, t) d\boldsymbol{\tau} dt \\ - \int_0^T \int_{\Gamma_G} (\mathbf{q}_r \cdot \mathbf{n})(\boldsymbol{\tau}, t)(v - u)(\boldsymbol{\tau}, t) d\boldsymbol{\tau} dt. \end{aligned} \quad (2.9)$$

We recall in the next lemma the properties of this problem which are standard.

**Lemma 2.4.** *Assume that the data  $u_B$ ,  $f_F$ ,  $\mathbf{q}_r$  and  $u_0$  satisfy (2.6). Then, for any  $F$  in  $L^2(0, T; L^2(\Omega))$ , problem (2.3) – (2.9) has a unique solution  $u$  in  $L^2(0, T; \mathbb{V})$ .*

**Proof:** It follows from Lemma 2.3 and the further assumption on  $u_B$  that  $\mathbb{X}$  is a non empty closed convex set. We also consider a lifting  $\bar{u}_B$  of the extension by zero of  $u_B$  to  $\partial\Omega$  in  $H^1(0, T; H^1(\Omega))$ . Then it is readily checked that  $u - \bar{u}_B$  is the solution of a problem which satisfies all the assumptions in [10, Chap. 6, Th. 2.2], whence the existence and uniqueness result.

Any solution of (2.3) – (2.8) is a solution of (2.3) – (2.9), but the converse property is not obvious in the general case, see [10, Chap. 6]. However, in our specific case, it is readily checked by a density argument that inequation (2.9) is satisfied for any  $v$  in  $L^2(0, T; \mathbb{V})$ , so that problems (2.3) – (2.8) and (2.3) – (2.9) are fully equivalent.

To go further, we assume that the following compatibility condition holds:

$$\text{for a.e. } \mathbf{x} \in \Gamma_B, \quad u_0(\mathbf{x}) = u_B(\mathbf{x}, 0) \quad \text{and} \quad \text{for a.e. } \mathbf{x} \in \Gamma_G, \quad u_0(\mathbf{x}) \leq 0, \quad (2.10)$$

and we introduce a lifting  $u_B^*$  of an extension of  $u_B$  to  $\partial\Omega$  which belongs to  $H^1(0, T; \mathbb{V})$  and satisfies

$$\text{for a.e. } \mathbf{x} \in \Omega, \quad u_B^*(\mathbf{x}, 0) = u_0(\mathbf{x}), \quad (2.11)$$

together with the stability property

$$\|u_B^*\|_{H^1(0, T; H^1(\Omega))} \leq c \|u_B\|_{H^1(0, T; H_{00}^{\frac{1}{2}}(\Gamma_B))}. \quad (2.12)$$

Then it is readily checked that  $u$  is a solution of problem (2.3) – (2.4) if and only if the function  $u^* = u - u_B^*$  is a solution of

Find  $u^*$  in  $L^2(0, T; \mathbb{V}_0)$  with  $\partial_t u$  in  $L^2(0, T; L^2(\Omega))$  such that

$$u^*|_{t=0} = 0, \quad (2.13)$$

and that, for a.e.  $t$  in  $]0, T[$ ,

$$\begin{aligned} \forall v \in \mathbb{V}_0, \\ \alpha \int_{\Omega} (\partial_t u^*)(\mathbf{x}, t)(v - u^*)(\mathbf{x}, t) d\mathbf{x} + \int_{\Omega} (\partial_t b_*(u^*))(\mathbf{x}, t)(v - u^*)(\mathbf{x}, t) d\mathbf{x} \\ + \int_{\Omega} \left( \nabla u^* + k \circ b_*(u^*) \mathbf{e}_z \right)(\mathbf{x}, t) \cdot (\nabla(v - u^*))(\mathbf{x}, t) d\mathbf{x} \\ \geq - \int_{\Omega} F_B(\mathbf{x}, t)(v - u^*)(\mathbf{x}, t) d\mathbf{x} - \int_{\Gamma_F} f_F(\boldsymbol{\tau}, t)(v - u^*)(\boldsymbol{\tau}, t) d\boldsymbol{\tau} \\ - \int_{\Gamma_G} (\mathbf{q}_r \cdot \mathbf{n})(\boldsymbol{\tau}, t)(v - u^*)(\boldsymbol{\tau}, t) d\boldsymbol{\tau}, \end{aligned} \quad (2.14)$$

with obvious definition of the subset  $\mathbb{V}_0$ :

$$\mathbb{V}_0 = \{v \in H^1(\Omega); v|_{\Gamma_B} = 0 \text{ and } v|_{\Gamma_G} \leq 0\}, \quad (2.15)$$

and where the new application  $b_*$  is defined by  $b_*(u^*) = b(u^* + u_B^*)$ . The datum  $F_B$  is defined by, for a.e.  $t$  in  $]0, T[$ ,

$$\int_{\Omega} F_B(\mathbf{x}, t)v(\mathbf{x}) d\mathbf{x} = \alpha \int_{\Omega} (\partial_t u_B^*)(\mathbf{x}, t)v(\mathbf{x}) d\mathbf{x} + \int_{\Omega} (\nabla u_B^*)(\mathbf{x}, t) \cdot (\nabla v)(\mathbf{x}) d\mathbf{x}, \quad (2.16)$$

and clearly belongs to  $L^2(0, T; \mathbb{W}')$ , where  $\mathbb{W}$  is the smallest linear space containing  $\mathbb{V}_0$ , namely

$$\mathbb{W} = \{v \in H^1(\Omega); v|_{\Gamma_B} = 0\}. \quad (2.17)$$

It can be noted that the existence result stated in Lemma 2.4 is still valid for any  $F$  in  $L^2(0, T; \mathbb{W}')$ .

We denote by  $\mathcal{T}$  the operator which associates with any pair  $(F, D)$ , with  $F$  in  $L^2(0, T; \mathbb{W}')$  and data  $D = (0, f_F, \mathbf{q}_r, 0)$  satisfying (2.6), the solution  $u$  of problem (2.3) – (2.8). It follows from (2.13) – (2.14) that  $u^*$  satisfies

$$u^* - \mathcal{T}(F_B + F(u^*), D) = 0, \quad (2.18)$$

where the quantity  $F(u)$  is defined by duality, for a.e.  $t$  in  $[0, T]$ :

$$\langle F(u), v \rangle = \int_{\Omega} (\partial_t b(u))(\mathbf{x}, t) v(\mathbf{x}) \, d\mathbf{x} + \int_{\Omega} k \circ b(u)(\mathbf{x}, t) \mathbf{e}_z \cdot (\nabla v)(\mathbf{x}) \, d\mathbf{x}. \quad (2.19)$$

We first prove some further properties of the operator  $\mathcal{T}$ .

**Lemma 2.5.** *The operator  $\mathcal{T}$  is continuous from  $L^2(0, T; \mathbb{W}') \times L^2(0, T; L^2(\Gamma_F)) \times L^2(0, T; L^2(\Gamma_G)^d)$  into the space  $L^2(0, T; \mathbb{V}_0)$ . Moreover the following estimate holds*

$$\begin{aligned} \left( \int_0^T |\mathcal{T}(F, f_F, \mathbf{q}_r)(\cdot, t)|_{H^1(\Omega)}^2 \, dt \right)^{\frac{1}{2}} &\leq \|F\|_{L^2(0, T; \mathbb{W}')} \\ &+ c \|f_F\|_{L^2(0, T; L^2(\Gamma_F))} + c \|\mathbf{q}_r\|_{L^2(0, T; L^2(\Gamma_G)^d)}. \end{aligned} \quad (2.20)$$

**Proof:** We set:  $u = \mathcal{T}(F, f_F, \mathbf{q}_r)$  and we only prove the estimate (indeed, it is readily checked that it implies the continuity property). We take  $v$  equal to  $\frac{u}{2}$  in problem (2.8). This obviously gives

$$\begin{aligned} \frac{\alpha}{2} \int_{\Omega} (\partial_t u^2)(\mathbf{x}, t) \, d\mathbf{x} + |u(\cdot, t)|_{H^1(\Omega)}^2 \\ \leq \left( \|F(\cdot, t)\|_{\mathbb{W}'} + c \|f_F(\cdot, t)\|_{L^2(\Gamma_F)} + c \|\mathbf{q}_r(\cdot, t)\|_{L^2(\Gamma_G)^d} \right) |u(\cdot, t)|_{H^1(\Omega)}, \end{aligned}$$

where  $c$  is the norm of the trace operator. Thus, integrating with respect to  $t$  gives estimate (2.20).

**Lemma 2.6.** *The operator  $\mathcal{T}$  is continuous from  $L^2(0, T; L^2(\Omega)) \times H^1(0, T; L^2(\Gamma_F)) \times H^1(0, T; L^2(\Gamma_G)^d)$  into the space  $H^1(0, T; L^2(\Omega))$ . Moreover the following estimate holds for any positive  $\varepsilon$*

$$\begin{aligned} \alpha \|\partial_t \mathcal{T}(F, f_F, \mathbf{q}_r)\|_{L^2(0, T; L^2(\Omega))} &\leq (1 + \varepsilon) \|F\|_{L^2(0, T; L^2(\Omega))} \\ &+ c \|f_F\|_{H^1(0, T; L^2(\Gamma_F))} + c \|\mathbf{q}_r\|_{H^1(0, T; L^2(\Gamma_G)^d)}. \end{aligned} \quad (2.21)$$

**Proof:** The continuity property of  $\mathcal{T}$  is proved in [10, Chap. 6, Th. 2.1]. Next, setting:  $u = \mathcal{T}(F, f_F, \mathbf{q}_r)$ , we take  $v$  equal to  $u - \eta \partial_t u$  in (2.8) for a positive  $\eta$ . Indeed,

- since  $u$  vanishes on  $\Gamma_B$ , so does  $\partial_t u$ ;
- since  $u$  is nonpositive on  $\Gamma_G$  and  $u(\mathbf{x}, t - \eta)$  which is very near of  $u(\mathbf{x}, t) - \eta \partial_t u(\mathbf{x}, t)$  is

also nonpositive, there exists a  $\eta > 0$  such that  $u - \eta \partial_t u$  belongs to  $\mathbb{V}_0$ . This yields

$$\begin{aligned} \alpha \|\partial_t u\|_{L^2(\Omega)}^2 + \frac{1}{2} \partial_t |u|_{H^1(\Omega)}^2 &\leq \|F\|_{L^2(\Omega)} \|\partial_t u\|_{L^2(\Omega)} \\ &\quad - \int_{\Gamma_F} f_F(\boldsymbol{\tau}, t) \partial_t u(\boldsymbol{\tau}, t) d\boldsymbol{\tau} - \int_{\Gamma_G} (\mathbf{q}_r \cdot \mathbf{n})(\boldsymbol{\tau}, t) \partial_t u(\boldsymbol{\tau}, t) d\boldsymbol{\tau}. \end{aligned}$$

To bound the first term, we use Young's inequality

$$\|F\|_{L^2(\Omega)} \|\partial_t u\|_{L^2(\Omega)} \leq \frac{\alpha}{2} \|\partial_t u\|_{L^2(\Omega)}^2 + \frac{1}{2\alpha} \|F\|_{L^2(\Omega)}^2.$$

To handle the last integrals, we integrate by parts with respect to  $t$ . For instance, we have, for any  $\varepsilon > 0$ ,

$$\begin{aligned} &\int_0^t \int_{\Gamma_F} f_F(\boldsymbol{\tau}, s) \partial_t u(\boldsymbol{\tau}, s) d\boldsymbol{\tau} ds \\ &\quad = \int_{\Gamma_F} f_F(\boldsymbol{\tau}, t) u(\boldsymbol{\tau}, t) d\boldsymbol{\tau} - \int_0^t \int_{\Gamma_F} \partial_t f_F(\boldsymbol{\tau}, s) u(\boldsymbol{\tau}, s) d\boldsymbol{\tau} ds \\ &\leq \frac{1}{4} |u(\cdot, t)|_{H^1(\Omega)}^2 + c \|f_F(\cdot, t)\|_{L^2(\Gamma_F)}^2 + c \|\partial_t f_F\|_{L^2(0,t;L^2(\Gamma_F))}^2 + \varepsilon \|u\|_{L^2(0,t;H^1(\Omega))}^2. \end{aligned}$$

Thus, the desired estimate follows by combining all this and using (2.20).

We are thus in a position to prove a first existence result.

**Theorem 2.7.** *Assume that the coefficient  $\alpha$  satisfies*

$$\frac{1}{\alpha} \|b'\|_{L^\infty(\mathbb{R})} < 1. \quad (2.22)$$

For any data  $u_B$ ,  $f_F$ ,  $\mathbf{q}_r$  and  $u_0$  satisfying

$$\begin{aligned} u_B &\in H^1(0, T; H_{00}^{\frac{1}{2}}(\Gamma_B)), & f_F &\in H^1(0, T; L^2(\Gamma_F)), \\ \mathbf{q}_r &\in H^1(0, T; L^2(\Gamma_G)^d), & u_0 &\in H^1(\Omega). \end{aligned} \quad (2.23)$$

and (2.10), problem (2.3) – (2.4) has at least a solution in  $\mathbb{X}$ .

**Proof:** We proceed in several steps.

1) Let  $\mathbb{X}_0$  be the space of functions of  $\mathbb{X}$  vanishing in  $t = 0$ . We provide it with the norm:

$$\|v\|_{\mathbb{X}_0} = \|\partial_t v\|_{L^2(0,T;L^2(\Omega))}.$$

It follows from the previous Lemma 2.6 that

$$\|\mathcal{T}(F_B + F(u^*), D)\|_{\mathbb{X}_0} \leq \frac{1+\varepsilon}{\alpha} \|F(u^*)\|_{L^2(0,T;L^2(\Omega))} + c(D),$$

where the constant  $c(D)$  only depends on the data  $u_B$ ,  $f_F$  and  $\mathbf{q}_r$ . Due to the boundedness of  $b'$  and  $k \circ b$  (see (2.19) for the definition of  $F(u^*)$ ), we have

$$\|\mathcal{T}(F_B + F(u^*), D)\|_{\mathbb{X}_0} \leq \frac{1 + \varepsilon}{\alpha} \|b'\|_{L^\infty(\mathbb{R})} \|u^*\|_{\mathbb{X}_0} + c'(D).$$

Thus, due to Assumption (2.22), the application:  $u^* \mapsto \mathcal{T}(F_B + F(u^*), D)$  maps the ball in  $\mathbb{X}_0$  with radius  $R$  into itself for all  $R$  such that, for an appropriate  $\varepsilon$ ,

$$\left(1 - \frac{1 + \varepsilon}{\alpha} \|b'\|_{L^\infty(\mathbb{R})}\right) R > c'(D). \quad (2.24)$$

2) Since  $\mathbb{X}_0$  is separable, there exists an increasing sequence of finite-dimensional spaces  $\mathbb{X}_n$  which is dense in  $\mathbb{X}_0$ . If  $\Pi_n$  denotes the orthogonal projection operator (for the scalar product associated with the norm of  $\mathbb{X}_0$ ) onto  $\mathbb{X}_n$ , the mapping:  $u \mapsto \Pi_n \mathcal{T}(F_B + F(u), D)$  is continuous from  $\mathbb{X}_n$  into itself. The same arguments as previously yield that it sends the ball of  $\mathbb{X}_n$  with radius  $R$  into itself for all  $R$  satisfying (2.24). Thus, applying Brouwer's fixed point theorem (see [9, Chap. IV, Th. 1.1] for instance) yields that this mapping admits a fixed point in this same ball, namely that there exists  $u_n$  in  $\mathbb{X}_n$  satisfying the equation  $u_n = \Pi_n \mathcal{T}(F_B + F(u_n), D)$ . Moreover, it follows from Lemma 2.5 that this sequence is also bounded in  $L^2(0, T; H^1(\Omega))$ .

3) The function  $u_n$  thus satisfies

$$\begin{aligned} \forall v \in \mathbb{X}_n, \\ \alpha \int_{\Omega} (\partial_t u_n)(\mathbf{x}, t)(v - u_n)(\mathbf{x}, t) d\mathbf{x} + \int_{\Omega} (\partial_t b_*(u_n))(\mathbf{x}, t)(v - u_n)(\mathbf{x}, t) d\mathbf{x} \\ + \int_{\Omega} \left( \nabla u_n + k \circ b_*(u_n) \mathbf{e}_z \right)(\mathbf{x}, t) \cdot (\nabla(v - u_n))(\mathbf{x}, t) d\mathbf{x} \\ \geq - \int_{\Omega} F_B(\mathbf{x}, t)(v - u_n)(\mathbf{x}, t) d\mathbf{x} - \int_{\Gamma_F} f_F(\boldsymbol{\tau}, t)(v - u_n)(\boldsymbol{\tau}, t) d\boldsymbol{\tau} \\ - \int_{\Gamma_G} (\mathbf{q}_r \cdot \mathbf{n})(\boldsymbol{\tau}, t)(v - u_n)(\boldsymbol{\tau}, t) d\boldsymbol{\tau}. \end{aligned} \quad (2.25)$$

Moreover, due to the boundedness properties of the sequence  $(u_n)_n$ , there exists a subsequence still denoted by  $(u_n)_n$  for simplicity which converges to a function  $u^*$  of  $\mathbb{X}_0$  weakly in  $\mathbb{X}$  and strongly in  $L^2(0, T; L^2(\Omega))$ . Next, we observe that, for a fixed  $v$  in  $\mathbb{X}_n$ :

- The convergence of all terms in the right-hand side follows from the weak convergence in  $L^2(0, T; \mathbb{W})$ ;
- The convergence of the first term is derived by writing the expansion

$$\begin{aligned} \int_{\Omega} (\partial_t u_n)(\mathbf{x}, t)(v - u_n)(\mathbf{x}, t) d\mathbf{x} &= \int_{\Omega} (\partial_t u^*)(\mathbf{x}, t)(v - u^*)(\mathbf{x}, t) d\mathbf{x} \\ &+ \int_{\Omega} \partial_t(u_n - u^*)(\mathbf{x}, t)(v - u^*)(\mathbf{x}, t) d\mathbf{x} + \int_{\Omega} (\partial_t u_n)(\mathbf{x}, t)(u^* - u_n)(\mathbf{x}, t) d\mathbf{x}, \end{aligned}$$

and checking that the two last terms converge;

- The convergence of the term  $\int_{\Omega} (\nabla u_n)(\mathbf{x}, t) \cdot (\nabla(v - u_n))(\mathbf{x}, t) d\mathbf{x}$  is obtained by using the weak lower semi-continuity of the norm  $|u_n|_{H^1(\Omega)}$ .

Moreover, the convergence of the nonlinear terms follows from the expansions

$$\begin{aligned} \int_{\Omega} (\partial_t b_*(u_n))(\mathbf{x}, t)(v - u_n)(\mathbf{x}, t) d\mathbf{x} &= \int_{\Omega} (\partial_t b_*(u^*))(\mathbf{x}, t)(v - u^*)(\mathbf{x}, t) d\mathbf{x} \\ &+ \int_{\Omega} (\partial_t b_*(u_n) - \partial_t b_*(u^*))(\mathbf{x}, t)(v - u^*)(\mathbf{x}, t) d\mathbf{x} + \int_{\Omega} (\partial_t b_*(u_n))(\mathbf{x}, t)(u^* - u_n)(\mathbf{x}, t) d\mathbf{x}, \end{aligned}$$

and

$$\begin{aligned} \int_{\Omega} k \circ b_*(u_n)(\mathbf{x}, t) \mathbf{e}_z \cdot (\nabla(v - u_n))(\mathbf{x}, t) d\mathbf{x} &= \int_{\Omega} k \circ b_*(u^*)(\mathbf{x}, t) \mathbf{e}_z \cdot (\nabla(v - u^*))(\mathbf{x}, t) d\mathbf{x} \\ &+ \int_{\Omega} k \circ b_*(u^*)(\mathbf{x}, t) \mathbf{e}_z \cdot (\nabla(u^* - u_n))(\mathbf{x}, t) d\mathbf{x} \\ &+ \int_{\Omega} (k \circ b_*(u_n) - k \circ b_*(u^*))(\mathbf{x}, t) \mathbf{e}_z \cdot (\nabla(v - u_n))(\mathbf{x}, t) d\mathbf{x}, \end{aligned}$$

combined with the Lipschitz-continuity of  $b'$  and  $k \circ b$ . Finally, using the density of the sequence of  $\mathbb{X}_n$  in  $\mathbb{X}_0$  yields that  $u^*$  is a solution of problem (2.13) – (2.14). Thus,  $u$  is a solution of problem (2.3) – (2.4).

Condition (2.22) is rather restrictive since, in practical situations,  $\alpha$  is small. However, this condition can be relaxed when the coefficient  $b$  satisfies, for a positive constant  $b_0$ ,

$$\forall \xi \in \mathbb{R}, \quad b'(\xi) \geq b_0. \quad (2.26)$$

Indeed all the previous arguments are still valid when replacing  $\alpha$  by  $\alpha + b_0$  and the coefficient  $b(\xi)$  by  $b(\xi) - b_0 \xi$ .

**Corollary 2.8.** *Assume that the coefficient  $b$  satisfies (2.26), and that the coefficient  $\alpha$  is such that*

$$\frac{1}{\alpha + b_0} \|b' - b_0\|_{L^\infty(\mathbb{R})} < 1. \quad (2.27)$$

*For any data  $u_B$ ,  $f_F$ ,  $\mathbf{q}_r$  and  $u_0$  satisfying (2.23) and (2.10), problem (2.3) – (2.4) has at least a solution in  $\mathbb{X}$ .*

Assume that the coefficient  $b$  satisfies

$$\min_{\xi \in \mathbb{R}} b'(\xi) > 0, \quad \max_{\xi \in \mathbb{R}} b'(\xi) < 2 \min_{\xi \in \mathbb{R}} b'(\xi). \quad (2.28)$$

Under this condition, problem (2.3) – (2.4) has a solution even for  $\alpha = 0$ . We refer to [2] for another proof of this result for a similar problem.

### 3. The discrete problems.

We present successively the time semi-discrete problem constructed from the backward Euler's scheme, next a finite element discretization of this problem relying on standard, conforming, finite element spaces.

#### 3.1. The time semi-discrete problem.

Since we intend to work with non uniform time steps, we introduce a partition of the interval  $[0, T]$  into subintervals  $[t_{n-1}, t_n]$ ,  $1 \leq n \leq N$ , such that  $0 = t_0 < t_1 < \dots < t_N = T$ . We denote by  $\tau_n$  the time step  $t_n - t_{n-1}$ , by  $\tau$  the  $N$ -tuple  $(\tau_1, \dots, \tau_N)$  and by  $|\tau|$  the maximum of the  $\tau_n$ ,  $1 \leq n \leq N$ .

As already hinted in Section 1, the time discretization mainly relies on a backward Euler's scheme, where the nonlinear term  $k \circ b(u)$  is treated in an explicit way for simplicity. Thus, the semi-discrete problem reads

$$\begin{aligned} \text{Find } (u^n)_{0 \leq n \leq N} \text{ in } \prod_{n=0}^N \mathbb{V}(t_n) \text{ such that} \\ u^0 = u_0 \quad \text{in } \Omega, \end{aligned} \quad (3.1)$$

and, for  $1 \leq n \leq N$ ,

$$\begin{aligned} \forall v \in \mathbb{V}(t_n), \\ \alpha \int_{\Omega} \left( \frac{u^n - u^{n-1}}{\tau_n} \right) (\mathbf{x}) (v - u^n) (\mathbf{x}) \, d\mathbf{x} + \int_{\Omega} \left( \frac{b(u^n) - b(u^{n-1})}{\tau_n} \right) (\mathbf{x}) (v - u^n) (\mathbf{x}) \, d\mathbf{x} \\ + \int_{\Omega} (\nabla u^n + k \circ b(u^{n-1})) (\mathbf{x}) \mathbf{e}_z \cdot \nabla (v - u^n) \mathbf{x} \, d\mathbf{x} \\ \geq - \int_{\Gamma_F} f_F(\boldsymbol{\tau}, t_n) (v - u^n) (\boldsymbol{\tau}) \, d\boldsymbol{\tau} - \int_{\Gamma_G} (\mathbf{q}_r \cdot \mathbf{n})(\boldsymbol{\tau}, t_n) (v - u^n) (\boldsymbol{\tau}) \, d\boldsymbol{\tau}. \end{aligned} \quad (3.2)$$

It can be noted that this problem makes sense when both  $f_F$  and  $\mathbf{q}_r$  are continuous in time. Proving its well-posedness relies on rather different arguments as previously.

**Theorem 3.1.** *For any data  $u_B$ ,  $f_F$ ,  $\mathbf{q}_r$  and  $u_0$  satisfying*

$$\begin{aligned} u_B \in H^1(0, T; H_{00}^{\frac{1}{2}}(\Gamma_B)), \quad f_F \in \mathcal{C}^0(0, T; L^2(\Gamma_F)), \\ \mathbf{q}_r \in \mathcal{C}^0(0, T; L^2(\Gamma_G)^d), \quad u_0 \in H^1(\Omega), \end{aligned} \quad (3.3)$$

and (2.10), for any nonnegative coefficient  $\alpha$ , problem (3.1) – (3.2) has a unique solution in  $\prod_{n=0}^N \mathbb{V}(t_n)$ .

**Proof:** We proceed by induction on  $n$ . Since  $u^0$  is given by (3.1), we assume that  $u^{n-1}$  is known and we consider problem (3.2) for a fixed  $n$ , called  $(3.2)_n$ , that can equivalently be written

$$\begin{aligned} \forall v \in \mathbb{V}(t_n), \quad & \int_{\Omega} (\alpha u^n + b(u^n))(\mathbf{x})(v - u^n)(\mathbf{x}) d\mathbf{x} + \tau_n \int_{\Omega} \nabla u^n(\mathbf{x}) \cdot \nabla(v - u^n)(\mathbf{x}) d\mathbf{x} \\ & \geq \int_{\Omega} (\alpha u^{n-1} + b(u^{n-1}))(\mathbf{x})(v - u^n)(\mathbf{x}) d\mathbf{x} - \tau_n \int_{\Omega} k \circ b(u^{n-1})(\mathbf{x}) \mathbf{e}_z \cdot \nabla(v - u^n)(\mathbf{x}) d\mathbf{x} \\ & \quad - \tau_n \int_{\Gamma_F} f_F(\boldsymbol{\tau}, t_n)(v - u^n)(\boldsymbol{\tau}) d\boldsymbol{\tau} - \tau_n \int_{\Gamma_G} (\mathbf{q}_r \cdot \mathbf{n})(\boldsymbol{\tau}, t_n)(v - u^n)(\boldsymbol{\tau}) d\boldsymbol{\tau}. \end{aligned}$$

Let us now set:

$$\varphi(z) = \int_0^z (\alpha \zeta + b(\zeta)) d\zeta, \quad \Phi(v) = \int_{\Omega} \varphi(v(\mathbf{x})) d\mathbf{x}.$$

It is readily checked that, since  $b'$  is nonnegative, both  $\varphi$  and  $\Phi$  are convex and, moreover, that

$$D\Phi(u) \cdot (v - u^n) = \int_{\Omega} (\alpha u + b(u))(\mathbf{x})(v - u^n)(\mathbf{x}) d\mathbf{x}.$$

Thus, when taking

$$\begin{aligned} a(u, v) &= \int_{\Omega} \nabla u(\mathbf{x}) \cdot \nabla v(\mathbf{x}) d\mathbf{x}, \\ \ell(v) &= \int_{\Omega} (\alpha u^{n-1} + b(u^{n-1}))(\mathbf{x})v(\mathbf{x}) d\mathbf{x} - \tau_n \int_{\Omega} k \circ b(u^{n-1})(\mathbf{x}) \mathbf{e}_z \cdot \nabla v(\mathbf{x}) d\mathbf{x} \\ & \quad - \tau_n \int_{\Gamma_F} f_F(\boldsymbol{\tau}, t_n)v(\boldsymbol{\tau}) d\boldsymbol{\tau} - \tau_n \int_{\Gamma_G} (\mathbf{q}_r \cdot \mathbf{n})(\boldsymbol{\tau}, t_n)v(\boldsymbol{\tau}) d\boldsymbol{\tau}, \end{aligned}$$

problem  $(3.2)_n$  can also be written

$$\forall v \in \mathbb{V}(t_n), \quad D\Phi(u^n) \cdot (v - u^n) + a(u^n, v - u^n) - \ell(v - u^n) \geq 0.$$

We now set:  $\Psi(v) = \Phi(v) + J(v)$ , with  $J(v) = \frac{1}{2}a(v, v) - \ell(v)$ . Problem  $(3.2)_n$  can finally be written

$$\forall v \in \mathbb{V}(t_n), \quad D\Psi(u^n) \cdot (v - u^n) \geq 0,$$

or

$$\forall v \in \mathbb{V}(t_n), \quad \Psi(u^n) \leq \Psi(v).$$

So it is equivalent to the minimization of a convex functional on the convex set  $\mathbb{V}(t_n)$ , hence admits a unique solution. This concludes the proof.

It can be noted that, in contrast with the continuous problem, the existence of a solution for the semi-discrete problem (3.1) – (3.2) does not require any limitation on  $\alpha$ .



### 3.2. The fully discrete problem.

From now on, we assume that  $\Omega$  is a polygon ( $d = 2$ ) or a polyhedron ( $d = 3$ ). Let  $(\mathcal{T}_h)_h$  be a regular family of triangulations of  $\Omega$  (by triangles or tetrahedra), in the sense that, for each  $h$ :

- $\bar{\Omega}$  is the union of all elements of  $\mathcal{T}_h$ ;
- The intersection of two different elements of  $\mathcal{T}_h$ , if not empty, is a vertex or a whole edge or a whole face of both of them;
- The ratio of the diameter  $h_K$  of any element  $K$  of  $\mathcal{T}_h$  to the diameter of its inscribed circle or sphere is smaller than a constant  $\sigma$  independent of  $h$ .

As usual,  $h$  stands for the maximum of the diameters  $h_K$ ,  $K \in \mathcal{T}_h$ . We make the further and non restrictive assumption that  $\bar{\Gamma}_B$ ,  $\bar{\Gamma}_F$  and  $\bar{\Gamma}_G$  are the union of whole edges ( $d = 2$ ) or whole faces ( $d = 3$ ) of elements of  $\mathcal{T}_h$ . From now on,  $c, c', \dots$  stand for generic constants that may vary from line to line but are always independent of  $\tau$  and  $h$ .

We now introduce the finite element space

$$\bar{\mathbb{V}}_h = \{v_h \in L^2(\Omega); \forall K \in \mathcal{T}_h, v_h|_K \in \mathcal{P}_1(K)\}, \quad (3.4)$$

where  $\mathcal{P}_1(K)$  is the space of restrictions to  $K$  of affine functions on  $\mathbb{R}^d$ . Let  $\mathcal{I}_h$  denote the Lagrange interpolation operator at all the vertices of elements of  $\mathcal{T}_h$  with values in  $\bar{\mathbb{V}}_h$  and  $i_h^B$  the corresponding interpolation operator on  $\Gamma_B$ . Assuming that  $u_B$  is continuous where needed, we then define for each  $n$ ,  $0 \leq n \leq N$ , the subset of  $\bar{\mathbb{V}}_h$ :

$$\mathbb{V}_h(t_n) = \{v_h \in \bar{\mathbb{V}}_h; v_h|_{\Gamma_B} = i_h^B u_B(\cdot, t_n) \text{ and } v_h|_{\Gamma_G} \leq 0\}. \quad (3.5)$$

We are thus in a position to write the discrete problem, constructed from problem (3.1) – (3.2) by the Galerkin method,

Find  $(u_h^n)_{0 \leq n \leq N}$  in  $\prod_{n=0}^N \mathbb{V}_h(t_n)$  such that

$$u_h^0 = \mathcal{I}_h u_0 \quad \text{in } \Omega, \quad (3.6)$$

and, for  $1 \leq n \leq N$ ,

$$\begin{aligned} & \forall v_h \in \mathbb{V}_h(t_n), \\ & \alpha \int_{\Omega} \left( \frac{u_h^n - u_h^{n-1}}{\tau_n} \right) (\mathbf{x}) (v_h - u_h^n) (\mathbf{x}) d\mathbf{x} + \int_{\Omega} \left( \frac{b(u_h^n) - b(u_h^{n-1})}{\tau_n} \right) (\mathbf{x}) (v_h - u_h^n) (\mathbf{x}) d\mathbf{x} \\ & + \int_{\Omega} (\nabla u_h^n + k \circ b(u_h^{n-1})) (\mathbf{x}) \mathbf{e}_z \cdot \nabla (v_h - u_h^n) (\mathbf{x}) d\mathbf{x} \\ & \geq - \int_{\Gamma_F} f_F(\boldsymbol{\tau}, t_n) (v_h - u_h^n) (\boldsymbol{\tau}) d\boldsymbol{\tau} - \int_{\Gamma_G} (\mathbf{q}_r \cdot \mathbf{n})(\boldsymbol{\tau}, t_n) (v_h - u_h^n) (\boldsymbol{\tau}) d\boldsymbol{\tau}. \end{aligned} \quad (3.7)$$

The proof of the next theorem is exactly the same as the proof of Theorem 3.1, so we omit it.

**Theorem 3.2.** *For any data  $u_B, f_F, \mathbf{q}_r$  and  $u_0$  satisfying (3.3), (2.10) and*

$$u_B \in \mathcal{C}^0(\bar{\Gamma}_B \times [0, T]), \quad u_0 \in \mathcal{C}^0(\bar{\Omega}), \quad (3.8)$$

*for any nonnegative coefficient  $\alpha$ , problem (3.6) – (3.7) has a unique solution.*

There also, the existence result is unconditional.

#### 4. A convergence result.

The aim of this section is to prove a convergence result for the solutions  $(u_h^n)_{0 \leq n \leq N}$  of problem (3.6) – (3.7) when  $|\tau|$  and  $h$  tend to zero. In order to do that and as in Section 2, we use the lifting  $u_B^*$  of  $u_B$  which satisfies (2.11) and (2.12) and we assume moreover that it is continuous on  $\bar{\Omega} \times [0, T]$ . Indeed, if  $(u_h^n)_{0 \leq n \leq N}$  is a solution of (3.6) – (3.7), the family  $(u_h^{*n})_{0 \leq n \leq N}$ , with  $u_h^{*n} = u_h^n - \mathcal{I}_h u_B^*(t_n)$ , is a solution of

$$\begin{aligned} & \text{Find } (u_h^{*n})_{0 \leq n \leq N} \text{ in } \mathbb{V}_{h0}^{N+1} \text{ such that} \\ & u_h^{*0} = 0 \quad \text{in } \Omega, \end{aligned} \tag{4.1}$$

and, for  $1 \leq n \leq N$ ,

$$\begin{aligned} & \forall v_h \in \mathbb{V}_{h0}, \\ & \alpha \int_{\Omega} \left( \frac{u_h^{*n} - u_h^{*(n-1)}}{\tau_n} \right) (\mathbf{x}) (v_h - u_h^{*n}) (\mathbf{x}) \, d\mathbf{x} \\ & \quad + \int_{\Omega} \left( \frac{b_{*n}(u_h^{*n}) - b_{*(n-1)}(u_h^{*(n-1)})}{\tau_n} \right) (\mathbf{x}) (v_h - u_h^{*n}) (\mathbf{x}) \, d\mathbf{x} \\ & \quad + \int_{\Omega} (\nabla u_h^{*n} + k \circ b_{*(n-1)}(u_h^{*(n-1)})) (\mathbf{x}) \mathbf{e}_z \cdot \nabla (v_h - u_h^{*n}) (\mathbf{x}) \, d\mathbf{x} \\ & \geq - \int_{\Omega} F_{Bh}(\mathbf{x}, t_n) (v_h - u_h^{*n}) \, d\mathbf{x} - \int_{\Gamma_F} f_F(\boldsymbol{\tau}, t_n) (v_h - u_h^{*n}) (\boldsymbol{\tau}) \, d\boldsymbol{\tau} \\ & \quad - \int_{\Gamma_G} (\mathbf{q}_r \cdot \mathbf{n})(\boldsymbol{\tau}, t_n) (v_h - u_h^{*n}) (\boldsymbol{\tau}) \, d\boldsymbol{\tau}, \end{aligned} \tag{4.2}$$

where the convex set  $\mathbb{V}_{h0}$  and the function  $F_{Bh}$  are defined, in analogy with (2.15) and (2.16), by

$$\mathbb{V}_{h0} = \bar{\mathbb{V}}_h \cap \mathbb{V}_0, \tag{4.3}$$

and

$$\begin{aligned} & \int_{\Omega} F_{Bh}(\mathbf{x}, t) v(\mathbf{x}) \, d\mathbf{x} \\ & = \alpha \int_{\Omega} (\partial_t \mathcal{I}_h u_B^*) (\mathbf{x}, t) v(\mathbf{x}) \, d\mathbf{x} + \int_{\Omega} (\nabla \mathcal{I}_h u_B^*) (\mathbf{x}, t) \cdot (\nabla v) (\mathbf{x}) \, d\mathbf{x}, \end{aligned} \tag{4.4}$$

while each function  $b_{*n}$  is given by  $b_{*n}(\xi) = b(\xi + \mathcal{I}_h u_B^*(\cdot, t_n))$ . We now investigate the boundedness of the sequence  $(u_h^{*n})_{0 \leq n \leq N}$  in appropriate norms, we need a preliminary lemma for that.

**Lemma 4.1.** *For each part  $\Gamma$  of  $\partial\Omega$  which is the union of whole edges ( $d = 2$ ) or whole faces ( $d = 3$ ) of elements of  $\mathcal{T}_h$ , the following inequality holds for all functions  $w_h$  in  $\bar{\mathbb{V}}_h$ ,*

$$\|w_h\|_{H^{-\frac{1}{2}}(\Gamma)} \leq c \|w_h\|_{L^2(\Omega)}. \tag{4.5}$$

**Proof:** It relies on standard arguments. We have

$$\|w_h\|_{H^{-\frac{1}{2}}(\Gamma)} = \sup_{z \in H^{\frac{1}{2}}(\Gamma)} \frac{\int_{\Gamma} z(\boldsymbol{\tau}) w_h(\boldsymbol{\tau}) d\boldsymbol{\tau}}{\|z\|_{H^{\frac{1}{2}}(\Gamma)}}.$$

Let  $e$  be any edge or face of an element  $K$  of  $\mathcal{T}_h$  which is contained in  $\Gamma$ . Denoting by  $\hat{K}$  the reference triangle or tetrahedron, we have, with obvious notation for  $\hat{e}$ ,  $\hat{w}$ ,  $\hat{z}$ ,

$$\int_e z(\boldsymbol{\tau}) w_h(\boldsymbol{\tau}) d\boldsymbol{\tau} \leq c h_e^{d-1} \int_{\hat{e}} \hat{z}(\hat{\boldsymbol{\tau}}) \hat{w}_h(\hat{\boldsymbol{\tau}}) d\hat{\boldsymbol{\tau}} \leq c' h_K^{d-1} \|\hat{z}\|_{L^2(\hat{e})} \|\hat{w}_h\|_{L^2(\hat{e})}.$$

By using the equivalence of norms on  $\mathcal{P}_1(\hat{K})$  and an appropriate stable lifting  $\hat{\pi}$  which maps traces on  $\hat{e}$  into functions of  $K$  vanishing at the vertex of  $K$  which does not belong to  $\bar{\Gamma}$ , we derive

$$\int_e z(\boldsymbol{\tau}) w_h(\boldsymbol{\tau}) d\boldsymbol{\tau} \leq c' h_K^{d-1} |\hat{\pi}\hat{z}|_{H^1(\hat{K})} \|\hat{w}_h\|_{L^2(\hat{K})} \leq c' h_K^{d-1} h_K^{1-\frac{d}{2}} |\pi z|_{H^1(K)} h_K^{-\frac{d}{2}} \|w_h\|_{L^2(K)},$$

there also with an obvious definition of  $\pi$ . We conclude by summing this last inequality on the  $e$  and using a Cauchy–Schwarz inequality and the stability of  $\hat{\pi}$ :

$$\int_{\Gamma} z(\boldsymbol{\tau}) w_h(\boldsymbol{\tau}) d\boldsymbol{\tau} \leq c \|z\|_{H^{\frac{1}{2}}(\Gamma)} \|w_h\|_{L^2(\Omega)},$$

whence the desired result.

**Lemma 4.2.** For any data  $u_B$ ,  $f_F$ ,  $\mathbf{q}_r$  and  $u_0$  satisfying

$$\begin{aligned} u_B &\in H^1(0, T; H_{00}^{\frac{1}{2}}(\Gamma_B)), & f_F &\in \mathcal{C}^0(0, T; H^{\frac{1}{2}}(\Gamma_F)), \\ \mathbf{q}_r &\in \mathcal{C}^0(0, T; H^{\frac{1}{2}}(\Gamma_G)^d), & u_0 &\in H^1(\Omega), \end{aligned} \quad (4.6)$$

and (2.10), the sequence  $(u_h^{*n})_{0 \leq n \leq N}$  satisfies the following bound, for  $1 \leq n \leq N$ ,

$$\begin{aligned} \alpha \sum_{m=1}^n \tau_m \left\| \frac{u_h^{*m} - u_h^{*m-1}}{\tau_m} \right\|_{L^2(\Omega)}^2 + |u_h^{*n}|_{H^1(\Omega)}^2 \\ \leq c \left( 1 + \|\mathcal{I}_h u_B^*\|_{H^1(0, T; H^1(\Omega))}^2 + \|f_F\|_{\mathcal{C}^0(0, T; H^{\frac{1}{2}}(\Gamma_F))}^2 + \|\mathbf{q}_r\|_{\mathcal{C}^0(0, T; H^{\frac{1}{2}}(\Gamma_G)^d)}^2 \right). \end{aligned} \quad (4.7)$$

**Proof:** Taking  $v$  equal to  $u_h^{*n-1}$  in (4.2) yields

$$\begin{aligned} \alpha \tau_n \left\| \frac{u_h^{*n} - u_h^{*n-1}}{\tau_n} \right\|_{L^2(\Omega)}^2 + \int_{\Omega} \nabla u_h^{*n}(\mathbf{x}) \cdot \nabla (u_h^{*n} - u_h^{*n-1})(\mathbf{x}) d\mathbf{x} \\ \leq - \int_{\Omega} \left( \frac{b_{*n}(u_h^{*n}) - b_{*n-1}(u_h^{*n-1})}{\tau_n} \right)(\mathbf{x}) (u_h^{*n} - u_h^{*n-1})(\mathbf{x}) d\mathbf{x} \\ - \int_{\Omega} k \circ b_{*n-1}(u_h^{*n-1})(\mathbf{x}) \mathbf{e}_z \cdot \nabla (u_h^{*n} - u_h^{*n-1})(\mathbf{x}) d\mathbf{x} + \langle \mathcal{G}, u_h^{*n} - u_h^{*n-1} \rangle, \end{aligned}$$

where the data depending quantity  $\mathcal{G}$  is defined by

$$\langle \mathcal{G}, v \rangle = - \int_{\Omega} F_{Bh}(\mathbf{x}, t_n) v(\mathbf{x}) d\mathbf{x} - \int_{\Gamma_F} f_F(\boldsymbol{\tau}, t_n) v(\boldsymbol{\tau}) d\boldsymbol{\tau} - \int_{\Gamma_G} (\mathbf{q}_r \cdot \mathbf{n})(\boldsymbol{\tau}, t_n) v(\boldsymbol{\tau}) d\boldsymbol{\tau}.$$

To handle the second term, we use the identity

$$\int_{\Omega} \nabla u_h^{*n} \cdot \nabla (u_h^{*n} - u_h^{*n-1})(\mathbf{x}) d\mathbf{x} = \frac{1}{2} \left( |u_h^{*n}|_{H^1(\Omega)}^2 + |u_h^{*n} - u_h^{*n-1}|_{H^1(\Omega)}^2 - |u_h^{*n-1}|_{H^1(\Omega)}^2 \right).$$

To handle the third term, we write the expansion

$$\begin{aligned} & \int_{\Omega} \left( \frac{b_{*n}(u_h^{*n}) - b_{*n-1}(u_h^{*n-1})}{\tau_n} \right)(\mathbf{x}) (u_h^{*n} - u_h^{*n-1})(\mathbf{x}) d\mathbf{x} \\ &= \int_{\Omega} \left( \frac{b(u_h^{*n} + \mathcal{I}_h u_B^*(t_n)) - b(u_h^{*n-1} + \mathcal{I}_h u_B^*(t_n))}{\tau_n} \right)(\mathbf{x}) (u_h^{*n} - u_h^{*n-1})(\mathbf{x}) d\mathbf{x} \\ & \quad + \int_{\Omega} \left( \frac{b(u_h^{*n-1} + \mathcal{I}_h u_B^*(t_n)) - b(u_h^{*n-1} + \mathcal{I}_h u_B^*(t_{n-1}))}{\tau_n} \right)(\mathbf{x}) (u_h^{*n} - u_h^{*n-1})(\mathbf{x}) d\mathbf{x}, \end{aligned}$$

and, by using the nonnegativity of  $b'$  together with the Lipschitz-continuity of  $b$ , we derive

$$\begin{aligned} & \int_{\Omega} \left( \frac{b_{*n}(u_h^{*n}) - b_{*n-1}(u_h^{*n-1})}{\tau_n} \right)(\mathbf{x}) (u_h^{*n} - u_h^{*n-1})(\mathbf{x}) d\mathbf{x} \\ & \leq \frac{\alpha}{4} \tau_n \left\| \frac{u_h^{*n} - u_h^{*n-1}}{\tau_n} \right\|_{L^2(\Omega)}^2 + \frac{1}{\alpha} \tau_n \left\| \frac{\mathcal{I}_h u_B^*(t_n) - \mathcal{I}_h u_B^*(t_{n-1})}{\tau_n} \right\|_{L^2(\Omega)}^2. \end{aligned}$$

Finally, evaluating the last term is an easy consequence of Lemma 4.1:

$$\begin{aligned} \langle \mathcal{G}, u_h^{*n} - u_h^{*n-1} \rangle & \leq \frac{\alpha}{4} \tau_n \left\| \frac{u_h^{*n} - u_h^{*n-1}}{\tau_n} \right\|_{L^2(\Omega)}^2 \\ & \quad + c\tau_n \left( \|F_{Bh}(\cdot, t_n)\|_{L^2(\Omega)}^2 + \|f_F(\cdot, t_n)\|_{H^{\frac{1}{2}}(\Gamma_F)}^2 + \|\mathbf{q}_r(\cdot, t_n)\|_{H^{\frac{1}{2}}(\Gamma_G)^d}^2 \right). \end{aligned}$$

By combining all this, we obtain

$$\begin{aligned} & \frac{\alpha}{2} \tau_n \left\| \frac{u_h^{*n} - u_h^{*n-1}}{\tau_n} \right\|_{L^2(\Omega)}^2 + \frac{1}{2} |u_h^{*n}|_{H^1(\Omega)}^2 \leq \frac{1}{2} |u_h^{*n-1}|_{H^1(\Omega)}^2 \\ & \quad + c' \tau_n \left( \|F_{Bh}(\cdot, t_n)\|_{L^2(\Omega)}^2 + \|f_F(\cdot, t_n)\|_{H^{\frac{1}{2}}(\Gamma_F)}^2 + \|\mathbf{q}_r(\cdot, t_n)\|_{H^{\frac{1}{2}}(\Gamma_G)^d}^2 \right) \\ & \quad - \int_{\Omega} k \circ b_{*n-1}(u_h^{*n-1})(\mathbf{x}) \mathbf{e}_z \cdot \nabla (u_h^{*n} - u_h^{*n-1})(\mathbf{x}) d\mathbf{x}. \end{aligned}$$

We sum this inequality on  $n$ . To handle the last term, we observe that

$$\begin{aligned} & - \sum_{m=1}^n \int_{\Omega} k \circ b_{*m-1}(u_h^{*m-1})(\mathbf{x}) \mathbf{e}_z \cdot \nabla (u_h^{*m} - u_h^{*m-1})(\mathbf{x}) d\mathbf{x} \\ &= - \int_{\Omega} k \circ b_{*n-1}(u_h^{*n-1})(\mathbf{x}) \mathbf{e}_z \cdot \nabla u_h^{*n}(\mathbf{x}) d\mathbf{x} \\ & \quad + \sum_{m=1}^{n-1} \int_{\Omega} (k \circ b_{*m}(u_h^{*m}) - k \circ b_{*m-1}(u_h^{*m-1}))(\mathbf{x}) \mathbf{e}_z \cdot \nabla u_h^{*m}(\mathbf{x}) d\mathbf{x}, \end{aligned}$$

whence, thanks to the boundedness of  $k$  and Lipschitz continuity of  $k \circ b$ ,

$$\begin{aligned}
& - \sum_{m=1}^n \int_{\Omega} k \circ b_{*m-1}(u_h^{m-1*})(\mathbf{x}) \mathbf{e}_z \cdot \nabla(u_h^{*m} - u_h^{*m-1})(\mathbf{x}) \, d\mathbf{x} \\
& = c + \frac{1}{4} |u_h^{*n}|_{H^1(\Omega)}^2 + \frac{\alpha}{4} \sum_{m=1}^{n-1} \tau_m \left\| \frac{u_h^{*m} - u_h^{*m-1}}{\tau_m} \right\|_{L^2(\Omega)}^2 \\
& \quad + c' \sum_{m=1}^{n-1} \tau_m \left\| \frac{\mathcal{I}_h u_B^*(t_m) - \mathcal{I}_h u_B^*(t_{m-1})}{\tau_m} \right\|_{L^2(\Omega)}^2 + c'' \sum_{m=1}^{n-1} \tau_m |u^{*m}|_{H^1(\Omega)}^2.
\end{aligned}$$

We conclude by using the discrete Grönwall's lemma, see [8, Chap. V, Lemma 2.4] for instance.

Let us now introduce the function  $u_{h\tau}^*$  which is affine on each interval  $[t_{n-1}, t_n]$ ,  $1 \leq n \leq N$ , and equal to  $u_h^{*n}$  at time  $t_n$ ,  $0 \leq n \leq N$ . When the data  $u_B$ ,  $f_F$ ,  $\mathbf{q}_r$  and  $u_0$  satisfy

$$\begin{aligned}
u_B & \in H^1(0, T; H^s(\Gamma_B)), & f_F & \in \mathcal{C}^0(0, T; H^{\frac{1}{2}}(\Gamma_F)), \\
\mathbf{q}_r & \in \mathcal{C}^0(0, T; H^{\frac{1}{2}}(\Gamma_G)^d), & u_0 & \in H^{s+\frac{1}{2}}(\Omega),
\end{aligned} \tag{4.8}$$

for some  $s > \frac{d-1}{2}$  (in order to ensure the stability the operator  $\mathcal{I}_h$ ), it follows from Lemma 4.2 that this function belongs to the set  $\mathbb{X}_0 = L^2(0, T; \mathbb{V}_0) \cap H^1(0, T; L^2(\Omega))$ , see (2.5) and (2.14); more precisely it satisfies

$$\|u_{h\tau}^*\|_{L^2(0, T; H^1(\Omega)) \cap H^1(0, T; L^2(\Omega))} \leq c(u_B, f_F, \mathbf{q}_r), \tag{4.9}$$

where the constant  $c(u_B, f_F, \mathbf{q}_r)$  only depends on the data. Thus, we are in a position to derive the next result.

**Theorem 4.3.** *For any data  $u_B$ ,  $f_F$ ,  $\mathbf{q}_r$  and  $u_0$  satisfying (4.8) and (2.10), and for any positive coefficient  $\alpha$ , problem (2.3) – (2.4) has at least a solution in  $\mathbb{X}$ .*

**Proof:** Thanks to (4.9), the family of functions  $u_{h\tau}^*$  is bounded in  $\mathbb{X}_0$  independently of  $h$  and  $\tau$ . Thus, there exist a sequence  $(\mathcal{T}_{hk})_k$  of triangulations  $\mathcal{T}_h$  and a sequence  $(\tau_k)_k$  of parameters  $\tau$  such that, with obvious notation, the sequence  $(u_k^*)_k$  converges to a function  $u^*$  of  $\mathbb{X}_0$  weakly in  $L^2(0, T; H^1(\Omega) \cap H^1(0, T; L^2(\Omega)))$  and strongly in  $L^2(0, T; L^2(\Omega))$ . We now intend to prove that  $u^*$  is a solution of problem (2.13) – (2.14). Since it obviously satisfies (2.13), we now investigate the convergence of all terms in (4.2). For clarity, we keep the notation  $u_h^{*n}$  for  $u_k^*(t_n)$ .

1) The convergence of the first term follows from the expansion

$$\begin{aligned}
& \alpha \int_{\Omega} \left( \frac{u_h^{*n} - u_h^{*n-1}}{\tau_n} \right) (\mathbf{x}) (v_h - u_h^{*n})(\mathbf{x}) \, d\mathbf{x} = \alpha \int_{\Omega} (\partial_t u^*)(\mathbf{x}, t_n) (v_h - u^*)(\mathbf{x}, t_n) \, d\mathbf{x} \\
& \quad + \alpha \int_{\Omega} (\partial_t (u_k^* - u^*))(\mathbf{x}, t_n) (v_h - u^*)(\mathbf{x}, t_n) \, d\mathbf{x} + \alpha \int_{\Omega} (\partial_t u_k^*)(\mathbf{x}, t_n) (u^* - u_h^{*n})(\mathbf{x}, t_n) \, d\mathbf{x}.
\end{aligned}$$

2) To prove the convergence of the term

$$\int_{\Omega} \left( \frac{b_{*n}(u_h^{*n}) - b_{*n-1}(u_h^{*(n-1)})}{\tau_n} \right)(\mathbf{x})(v_h - u_h^{*n})(\mathbf{x}) d\mathbf{x},$$

we use a rather complex expansion that we skip for brevity, combined with the dominated convergence theorem of Lebesgue: Indeed, since  $(u_k^*)_k$  converges to a function  $u^*$  in  $L^2(0, T; L^2(\Omega))$ , it converges almost everywhere in  $\Omega \times ]0, T[$ , so that  $(b'(u_k^*))_k$  also converges a.e. to  $b'(u^*)$ ; thus, since  $b'$  is bounded,  $(b'(u_k^*))_k$  also converges to  $b'(u^*)$  in  $L^2(0, T; L^2(\Omega))$ .

3) The convergence of the term  $\int_{\Omega} \nabla u_h^{*n}(\mathbf{x}, t_n) \mathbf{e}_z \cdot \nabla(v_h - u_h^{*n})(\mathbf{x}, t_n) d\mathbf{x}$  is a consequence of the weak lower semi-continuity of the norm.

4) The convergence of the term  $\int_{\Omega} k \circ b_{*n-1}(u_h^{*(n-1)}) (\mathbf{x}) \mathbf{e}_z \cdot \nabla(v_h - u_h^{*n})(\mathbf{x}) d\mathbf{x}$  is easily derived from the expansion

$$\begin{aligned} & \int_{\Omega} k \circ b_{*n-1}(u_h^{*(n-1)})(\mathbf{x}) \mathbf{e}_z \cdot \nabla(v_h - u_h^{*n})(\mathbf{x}) d\mathbf{x} \\ &= \int_{\Omega} k \circ b_*(u^*)(\mathbf{x}, t_n) \mathbf{e}_z \cdot \nabla(v_h - u^*)(\mathbf{x}, t_n) d\mathbf{x} \\ & \quad + \int_{\Omega} (k \circ b_{*n-1} - k \circ b_*)(u^*)(\mathbf{x}, t_n) \mathbf{e}_z \cdot \nabla(v_h - u^*)(\mathbf{x}, t_n) d\mathbf{x} \\ & \quad + \int_{\Omega} k \circ b_{*n-1}(u^*)(\mathbf{x}, t_n) \mathbf{e}_z \cdot \nabla(u^* - u_h^{*n})(\mathbf{x}) d\mathbf{x} \\ & \quad + \int_{\Omega} (k \circ b_{*n-1}(u_h^{*(n-1)}) - k \circ b_{*n-1}(u^*))(\mathbf{x}) \mathbf{e}_z \cdot \nabla(v_h - u_h^{*n})(\mathbf{x}, t_n) d\mathbf{x}, \end{aligned}$$

and there also from the dominated convergence theorem of Lebesgue.

5) The convergence of all terms in the right-hand side of (4.2) is obviously derived from the weak convergence of the sequence  $(u_k^*)_k$ .

Finally, using the density of the union of the  $\mathbb{V}_{h0}$  in  $\mathbb{V}_0$ , we derive that  $u^*$  is a solution of problem (2.13) – (2.14). Thus, the function  $u = u^* + u_B^*$  is a solution of problem (2.3) – (2.4).

Even if this requires a slightly different regularity of the data, Theorem 4.3 combined with Proposition 2.2 yields that, for any positive coefficient  $\alpha$ , problem (2.3) – (2.4) is well-posed in  $\mathbb{X}$ . This of course is a great improvement of the results in Section 2 and leads to think that the discretization proposed in Section 3 is rather efficient. We shall check this in the second part of this work.

## References

- [1] H.W. Alt, S. Luckhaus — Quasilinear elliptic-parabolic differential equations, *Math. Z.* **183** (1983), 311–341.
- [2] H.W. Alt, S. Luckhaus, A. Visintin — On nonstationary flow through porous media, *Ann. Mat. Pura Appl.* **136** (1984), 303–316.
- [3] C. Bernardi, L. El Alaoui, Z. Mghazli — A posteriori analysis of a space and time discretization of a nonlinear model for the flow in variably saturated porous media, submitted.
- [4] H. Berninger — Domain decomposition methods for elliptic problems with jumping nonlinearities and application to the Richards equation, Ph.D. Thesis, Freie Universität Berlin (2007).
- [5] F. Brezzi, W.W. Hager, P.-A. Raviart — Error estimates for the finite element solution of variational inequalities. II. Mixed methods, *Numer. Math.* **31** (1978/79), 1–16.
- [6] P. Fabrié, T. Gallouët — Modelling wells in porous media flows, *Math. Models Methods Appl. Sci.* **10** (2000), 673–709.
- [7] M. Gabbouhy — Analyse mathématique et simulation numérique des phénomènes d’écoulement et de transport en milieux poreux non saturés. Application à la région du Gharb, Ph.D. Thesis, Université Ibn Tofail, Kénitra, Maroc (2000).
- [8] V. Girault, P.-A. Raviart — *Finite Element Approximation of the Navier–Stokes Equations*, Lecture Notes in Mathematics **749**, Springer–Verlag (1979).
- [9] V. Girault, P.-A. Raviart — *Finite Element Methods for Navier–Stokes Equations, Theory and Algorithms*, Springer–Verlag (1986).
- [10] R. Glowinski, J.L. Lions, R. Trémolières — *Analyse numérique des inéquations variationnelles. 2. Applications aux phénomènes stationnaires et d’évolution*, Collection “Méthodes Mathématiques de l’Informatique” 5, Dunod (1976).
- [11] J.-L. Lions, E. Magenes — *Problèmes aux limites non homogènes et applications*, Vol. I, Dunod, Paris (1968).
- [12] F. Radu, I.S. Pop, P. Knabner — Order of convergence estimates for an Euler implicit, mixed finite element discretization of Richards’ equation, *SIAM J. Numer. Anal.* **42** (2004), 1452–1478.
- [13] K.R. Rajagopal — On a hierarchy of approximate models for flows of incompressible fluids through porous solid, *Math. Models Methods Appl. Sci.* **17** (2007), 215–252.
- [14] L.A. Richards — Capillary conduction of liquids through porous mediums. *Physics* **1** (1931), 318–333.
- [15] E. Schneid, P. Knabner, F. Radu — A priori error estimates for a mixed finite element discretization of the Richards’ equation, *Numer. Math.* **98** (2004), 353–370.
- [16] P. Sochala, A. Ern — Numerical methods for subsurface flows and coupling with runoff, to appear.
- [17] P. Sochala, A. Ern, S. Piperno — Mass conservative BDF-discontinuous Galerkin/explicit finite volume schemes for coupling subsurface and overland flows, *Comput. Methods Appl. Mech. Engrg.* **198** (2009), 2122–2136.
- [18] C.S. Woodward, C.N. Dawson — Analysis of expanded mixed finite element methods for a nonlinear parabolic equation modeling flow into variably saturated porous media, *SIAM J. Numer. Anal.* **37** (2000), 701–724.