Spectral discretization of Darcy’s equations coupled with the heat equation
Christine Bernardi, Sarra Maarouf, Driss Yakoubi

To cite this version:
Christine Bernardi, Sarra Maarouf, Driss Yakoubi. Spectral discretization of Darcy’s equations coupled with the heat equation. 2014. hal-01085011

HAL Id: hal-01085011
https://hal.sorbonne-universite.fr/hal-01085011
Preprint submitted on 20 Nov 2014

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L’archive ouverte pluridisciplinaire HAL, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d’enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.
Spectral discretization of Darcy’s equations coupled with the heat equation

Christine Bernardi∗, Sarra Maarouf∗, Driss Yakoubi∗†

Abstract: In this paper we consider the heat equation coupled with Darcy’s law with a nonlinear source term describing heat production due to an exothermic chemical reaction. Existence and uniqueness of a solution are established. Next, a spectral discretization of the problem is presented and thoroughly analysed. Finally, we present some numerical experiments which confirm the interest of the discretization.

Résumé: Dans cet article, nous considérons l’équation de la chaleur couplée avec la loi de Darcy avec un terme source non linéaire décrivant la production de chaleur due à une réaction chimique exothermique. L’existence et l’unicité d’une solution sont établies. Ensuite, une discrétisation par méthode spectrale du problème est décrite et analysée en détail. Enfin, nous présentons quelques expériences numériques qui confirment l’intérêt de la discrétisation.

†Département de Mathématiques et Statistiques, Université Laval, Québec, Canada.
e-mails: bernardi@ann.jussieu.fr, maarouf@ann.jussieu.fr, yakoubi@giref.ulaval.ca
1 Introduction

Many works study heat convection in a gaseous or liquid medium with its motion described by the Navier-Stokes equations coupled with the heat equation (see for example [8, 15, 14]) under the Boussinesq approximation, proposed in 1903 by J. Boussinesq [9]. In this work we study distribution of the temperature in a fluid running through a porous medium modeled by a convection-diffusion equation coupled with Darcy’s law, which leads to the following system

\[
\begin{align*}
\alpha u + \nabla p &= F(T) \quad \text{in} \quad \Omega, \\
\nabla \cdot u &= 0 \quad \text{in} \quad \Omega, \\
- \lambda \Delta T + (u \cdot \nabla)T &= h \quad \text{in} \quad \Omega,
\end{align*}
\]  

(1.1)

where \( \Omega \) is a bounded open set in \( \mathbb{R}^d \), \( d = 2 \) or 3, with a Lipschitz-continuous boundary \( \partial \Omega \) divided in two parts \( \Gamma_* \) and \( \Gamma^* = \partial \Omega \setminus \Gamma_* \). The unknowns are the velocity \( u \), the pressure \( p \) and the temperature \( T \) of the fluid. The function \( h \) represents an external heat source. The parameters \( \alpha \) and \( \lambda \) are positive constants which correspond to the permeability of the medium and diffusion coefficient respectively. This system of equations is supplemented by the boundary conditions

\[
\begin{align*}
u \cdot n &= 0 \quad \text{on} \quad \partial \Omega, \\
T &= T_* \quad \text{on} \quad \Gamma_*, \\
\frac{\partial T}{\partial n} &= \theta^* \quad \text{on} \quad \Gamma^*.
\end{align*}
\]

(1.2)

The first and second equations of motion of (1.1) are written under the Boussinesq approximation and quasi-stationary approximation. The Boussinesq approximation signifies that the density of the fluid is everywhere constant except for the buoyancy term, this appears the right hand term in the first equation of (1.1). Quasi-stationary approximation in the Darcy law, though often used for fluids in porous medium, should be verified for the problem of heat explosion (for instance see [3]).

We first write an equivalent variational formulation of system (1.1)-(1.2) and we prove the existence of the solution. In fact, we establish two results: In the first one we assume some conditions on the data; the second one is less restrictive than the first one, we just make some assumptions on the geometry of the domain.

Next, we are interested in approximating this problem. We propose a discretization by spectral methods when the domain is a square or a cube. We propose a discrete problem and perform its numerical analysis. The main tool here is the theorem of discrete implicit functions due to F. Brezzi, J. Rappaz and P.-A. Raviart [11]. We prove a priori error estimates for both the velocity, pressure and temperature.

Finally, we propose an iterative algorithm for solving the nonlinear problem and prove its convergence under some conditions. We present some numerical experiments that confirm the interest of the model and its discretization.

The paper is organized as follows:

- Section 2 presents the problem setting and its analysis.
• Section 3 is devoted to the description of the discrete problem using a spectral method.
• In Section 4, we perform the a priori analysis of the discretization and prove optimal error estimates.
• Finally, we describe some numerical tests in Section 5. These preliminary tests were realized with FreeFem3d and are in agreement with our theoretical results.

2 Analysis and well-posedness

2.1 The variational formulation

From now on, we assume that

(i) the intersection $\Gamma_\ast \cap \bar{\Gamma}_\sharp$ is a Lipschitz-continuous submanifold of $\partial \Omega$;
(ii) $\Gamma_\ast$ has a positive $(d - 1)$-measure in $\partial \Omega$;
(iii) the function $F$ is continuously differentiable on $\mathbb{R}$ with bounded derivatives and there exists a real number $T_0$ where $F$ vanishes.

We set
\[
\gamma = \sup_{\zeta \in \mathbb{R}} |F'(\zeta)|.
\]

We perform the change of variable $\theta = T - T_0$ and we set $f(\theta) = \frac{1}{\gamma} F(T)$, so that the function $f$ vanishes at 0 and is continuously differentiable on $\mathbb{R}$ with the norm of its derivative $\leq 1$. With this change, the equations (1.1) are written as
\[
\begin{align*}
\alpha \, u + \nabla p &= \gamma f(\theta) & \text{in } \Omega, \\
\nabla \cdot u &= 0 & \text{in } \Omega, \\
- \lambda \Delta \theta + (u \cdot \nabla) \theta &= h & \text{in } \Omega.
\end{align*}
\]

So if we look at the boundary conditions (1.2), it is seen that only the boundary condition on the temperature is changed by the change of variable: The equations (1.2) is now replaced by
\[
\begin{align*}
\theta &= \theta_\ast = T_\ast - T_0 & \text{on } \Gamma_\ast, \\
\frac{\partial \theta}{\partial n} &= \theta_\sharp & \text{on } \Gamma_\sharp.
\end{align*}
\]

In what follows, the scalar product defined on $L^2(\Omega)$ or $L^2(\Omega)^d$ is denoted by $(\cdot, \cdot)$. As usual, $H^s(\Omega), s \in \mathbb{R}$, denotes the real Sobolev space equipped with the norm $\| \cdot \|_{H^s(\Omega)}$ and semi-norm $| \cdot |_{H^s(\Omega)}$ (see for instance [2, Chap. III and VII]).
We introduce the dual space \((H^{1/2}_{00}(\Gamma_z))^\prime\) of \(H^{1/2}_{00}(\Gamma_z)\) (see \cite[Chap.1, §11]{17} for the definition of this space), and denote by \(\langle \cdot, \cdot \rangle_{\Gamma}\) the duality pairing between \((H^{1/2}_{00}(\Gamma_z))^\prime\) and \(H^{1/2}_{00}(\Gamma_z)\). Thus, we consider the space
\[
H^1_0(\Omega) = \{ \varphi \in H^1(\Omega), \quad \varphi = 0 \text{ on } \Gamma_+ \}.
\]
Assumption (i) implies that \(D(\Omega \cup \Gamma_z)\) dense in \(H^1_0(\Omega)\) (sufficient conditions for this are given in \cite{5} for instance).

To write Darcy’s equation (2.1) in variational form, we introduce the pressure space defined as follows:
\[
H^1_0(\Omega) = \left\{ q \in H^1(\Omega) ; \int_\Omega q \, dx = 0 \right\}.
\]
(2.3)

We denote by \(c_p\) the constant of the Poincaré-Friedrichs inequality. We recall from Bernardi et al. \cite[Chap. XIII]{7} and Achdou, Bernardi and Coquel \cite{1}, that Darcy’s equations admit several variational formulations. We have chosen here the formulation which enables us to treat the boundary condition on \(u\) as a natural one. In the following, we assume that
\[
h \in L^2(\Omega), \quad \theta_\ast \in H^{1/2}(\Gamma_+) \quad \text{and} \quad \theta_\sharp \in (H^{1/2}_{00}(\Gamma_z))^\prime.
\]
(2.4)

The equivalent variational formulation reads

Find \(u, p, \theta \in L^3(\Omega)^d \times H^1_0(\Omega) \times H^1(\Omega)\) such that
\[
\theta = \theta_\ast \text{ on } \Gamma_+,
\]
and that,
\[
\forall v \in L^2(\Omega)^d, \quad \alpha \int_\Omega u \cdot v \, dx + \int_\Omega v \cdot \nabla p \, dx = \gamma \int_\Omega f(\theta) \cdot v \, dx,
\]
(2.5)
\[
\forall q \in H^1_0(\Omega), \quad \int_\Omega u \cdot \nabla q \, dx = 0,
\]
(2.6)
\[
\forall \varphi \in H^1_0(\Omega), \quad \lambda \int_\Omega \nabla \theta \cdot \nabla \varphi \, dx + \int_\Omega (u \cdot \nabla) \theta \varphi \, dx = \int_\Omega h \varphi \, dx + \lambda \vartheta_\sharp, \varphi >_{\Gamma_z}.
\]
(2.7)

It is readily checked from the density of \(D(\Omega \cup \Gamma_z)\) in \(H^1_0(\Omega)\) that this problem is equivalent to (2.1)-(2.2) (in the distribution sens).

One of the key points for the study of this problem is the inf-sup condition, its proof consists in taking \(v\) equal to \(\nabla q\) (see \cite[Chap I, Cor 2.4]{16}).

**Proposition 2.1** The following inf-sup condition is satisfied
\[
\forall q \in H^1(\Omega), \quad \sup_{v \in L^2(\Omega)^d} \frac{\int_\Omega v \cdot \nabla q \, dx}{\|v\|_{L^2(\Omega)^d}} \geq \|\nabla q\|_{L^2(\Omega)^d}.
\]
(2.8)

### 2.2 A priori estimate

We consider the space
\[
V(\Omega) = \left\{ v \in L^2(\Omega)^d, \quad \forall q \in H^1_0(\Omega), \quad \int_\Omega v \cdot \nabla q \, dx = 0 \right\}.
\]
which is characterized by
\[ V(\Omega) = \{ v \in L^2(\Omega)^d, \quad \nabla \cdot v = 0 \text{ in } \Omega \quad \text{and} \quad v \cdot n = 0 \text{ in } \partial \Omega \}. \]

Thanks to the inf-sup condition (2.8), problem (2.5)-(2.6) is equivalent to: Find \( u \) in \( V(\Omega) \) solution of
\[ \forall v \in V(\Omega), \quad \alpha \int_{\Omega} u \cdot v \, dx = \gamma \int_{\Omega} f(\theta) \cdot v \, dx, \tag{2.9} \]

We see that the couple \((u, \theta)\) in \((V(\Omega) \cap L^2(\Omega)^d) \times H^1(\Omega)\) is solution of (2.9)-(2.7).

**Proposition 2.2** For any data \( h, \theta_* \) and \( \theta_\sharp \) satisfying (2.4), there exist two positive constants \( c_1, c_2 \) such that any solution \((u, \theta)\) in \((V(\Omega) \cap L^2(\Omega)^d) \times H^1(\Omega)\) of Problem (2.9)-(2.7) satisfies
\[
\begin{align*}
||u||_{L^2(\Omega)^d} & \leq c_1 \left(||h||_{L^2(\Omega)} + ||\theta_*||_{H^{\frac{1}{2}}(\Gamma_\alpha)} + ||\theta_\sharp||_{H^\frac{1}{2}(\Gamma_\alpha)}\right), \tag{2.10} \\
||\theta||_{H^1(\Omega)} & \leq c_2 \left(||h||_{L^2(\Omega)} + ||\theta_*||_{H^{\frac{1}{2}}(\Gamma_\alpha)} + ||\theta_\sharp||_{H^\frac{1}{2}(\Gamma_\alpha)}\right). \tag{2.11}
\end{align*}
\]

**Proof.** 1) We take \( v \) equal to \( u \) in (2.9), we obtain
\[ \alpha ||u||_{L^2(\Omega)^d}^2 \leq \gamma ||f(\theta)||_{L^2(\Omega)^d} ||u||_{L^2(\Omega)^d}, \]

The properties of the mapping \( f \) and the mean value theorem imply that for any real number \( \zeta \),
\[ |f(\zeta)| = |f(\zeta) - f(0)| \leq |\zeta|. \]

This gives
\[ \alpha ||u||_{L^2(\Omega)^d} \leq \gamma ||\theta||_{L^2(\Omega)}. \tag{2.12} \]

2) We refer to Hopf lemma, see [16, Chap. IV, Lemma 2.3], for the following result: For any \( \varepsilon > 0 \), there exists a lifting \( \tilde{\theta}_* \) of \( \theta_* \) with \( \tilde{\theta}_\sharp \) which satisfies
\[
\begin{align*}
||\tilde{\theta}_*||_{H^1(\Omega)} & \leq c ||\theta_*||_{H^{\frac{1}{2}}(\Gamma_\alpha)} \quad \text{and} \quad ||\tilde{\theta}_\sharp||_{L^3(\Omega)} \leq \varepsilon ||\theta_*||_{H^{\frac{1}{2}}(\Gamma_\alpha)}. \tag{2.13}
\end{align*}
\]

We set \( \tilde{\theta} = \theta - \tilde{\theta}_\sharp \) and take \( \varphi \) equal to \( \tilde{\theta} \) in (2.7), we see
\[ \lambda ||\nabla \tilde{\theta}||_{L^2(\Omega)^d} \leq ||h||_{L^2(\Omega)} \tilde{\theta} ||||L^2(\Omega) + \lambda ||\theta_*||_0^{\frac{1}{2}}(\Gamma_\alpha)'||\tilde{\theta}||_{H^1(\Omega)} \]
\[ + \lambda ||\nabla \tilde{\theta}||_{L^2(\Omega)^d} ||\tilde{\theta}||_{L^2(\Omega)^d} + ||u||_{L^2(\Omega)^d} ||\nabla \tilde{\theta}||_{L^2(\Omega)^d} ||\tilde{\theta}||_{L^2(\Omega)^d}. \]

Inserting (2.13) yields
\[ \lambda ||\nabla \tilde{\theta}||_{L^2(\Omega)^d} \leq ||h||_{L^2(\Omega)} \tilde{\theta} ||||L^2(\Omega) + \lambda ||\theta_*||_0^{\frac{1}{2}}(\Gamma_\alpha)'||\tilde{\theta}||_{H^1(\Omega)} \]
\[ + \lambda ||\nabla \tilde{\theta}||_{L^2(\Omega)^d} ||\tilde{\theta}||_{L^2(\Omega)^d} + \varepsilon ||u||_{L^2(\Omega)^d} ||\nabla \tilde{\theta}||_{L^2(\Omega)^d} ||\tilde{\theta}||_{L^2(\Omega)^d}. \]

We choose \( \varepsilon = \frac{1}{||u||_{L^3(\Omega)^d}} \) and, using the Poincaré–Friedrichs inequality, we obtain
\[ \lambda ||\nabla \tilde{\theta}||_{L^2(\Omega)^d} \leq c_\theta ||h||_{L^2(\Omega)} + \lambda (1 + c_\theta^2)^{\frac{1}{2}} ||\theta_*||_0^{\frac{1}{2}}(\Gamma_\alpha)' + \lambda ||\nabla \tilde{\theta}||_{L^2(\Omega)^d} + ||\tilde{\theta}||_{H^\frac{1}{2}(\Gamma_\alpha)}, \]

whence (2.11). By combining (2.11) and (2.12) and using Poincaré-Friedrichs inequality on \( \tilde{\theta} \), we obtain (2.10).
2.3 Well-posedness result

We refer to [7, Chap. XIII, Th. 1.9] for the next result.

**Proposition 2.3** For any $\theta$ in $L^2(\Omega)$, there exists a unique $(u, p)$ in $L^2(\Omega)^d \times H^1(\Omega)$ solution of problem (2.5) – (2.6).

The regularity properties of the solution $(u, p)$ with respect to the space variable are easy to prove.

**Theorem 2.4** Let

$$s = \begin{cases} 
1 & \text{if } \Omega \text{ is convex}, \\
1 \frac{1}{2} & \text{otherwise}.
\end{cases}$$

We assume that $\theta$ belongs to $H^1(\Omega)$. Then, the solution $(u, p)$ of problem (2.5) – (2.6) belongs to the space $H^s(\Omega)^d \times H^{s+1}(\Omega)$. Moreover, there exists a positive constant $c_R$ such that,

$$\|u\|_{H^s(\Omega)^d} \leq c_R \|\theta\|_{H^1(\Omega)}. \quad (2.14)$$

**Proof.** Let the space

$$X_T(\Omega) = \left\{ v \in L^2(\Omega)^d, \quad \nabla \cdot v \in L^2(\Omega), \quad \text{curl} \ v \in L^2(\Omega)^{d(d-1)/2} \quad \text{and} \quad v \cdot n = 0 \text{ sur } \partial \Omega \right\}.$$ 

We recall that this space is embedded in $H^s(\Omega)^d$ such that $s = \frac{1}{2}$ (see Costabel [12]) and $s = 1$ if $\Omega$ is convex (see Amrouche et al. [4, Thm 2.17]). By applying the curl operator on the first equation of (2.1), we obtain

$$\text{curl} \ u = \gamma \text{ curl} \ f(\theta) \quad \text{in} \quad \Omega.$$ 

Since $\text{curl} \ f(\theta)$ belongs to $L^2(\Omega)^{d(d-1)/2}$, $u$ belongs to $X_T(\Omega)$, whence the desired result.

**Remark 2.5** Since $H^s(\Omega)^d$, $s \geq \frac{1}{2}$, is imbedded in $L^q(\Omega)^d$ such that $\frac{1}{q} = \frac{d - 2s}{2d}$, we deduce that $u$ belongs to $L^3(\Omega)^d$.

The next result is a direct consequence of Lax-Milgram theorem [10, Chap. VIII]

**Proposition 2.6** We assume that data $h, \theta_\star$ and $\theta_\sharp$ satisfy (2.4). For any divergence-free function $u$ in $L^3(\Omega)^d$, Problem (2.7) admits a unique solution $\theta$ in $H^1(\Omega)$.

We are in a position to state the main result.

**Theorem 2.7** We assume that $h, \theta_\star$ and $\theta_\sharp$ satisfy (2.4) and

$$\frac{C_2 c_R}{\lambda} \left(1 + c_p^2 \right) \left(\|h\|_{L^2(\Omega)} + \|\theta_\star\|_{H^1_0(\Gamma_\star)} + \|\theta_\sharp\|_{H^1/2(\Gamma_\sharp)} \right)^2 < 1. \quad (2.15)$$

Then, Problem (2.5)-(2.6)-(2.7) admits at least a solution $(u, p, \theta)$ in $L^3(\Omega)^d \times H^1_0(\Omega) \times H^1(\Omega)$. 


Proof. Let the mapping \( \mathcal{F}_1 : \theta \mapsto u \), such that \( u \) is the part of \((u, p)\) solution of (2.5)-(2.6) which we have exhibited in Proposition 2.3. This mapping is continuous from \( H^1(\Omega) \) into \( L^3(\Omega)^d \) from Theorem 2.4. Thus, we introduce the mapping \( \mathcal{F}_2 \) from \( L^3(\Omega)^d \) into \( H^1(\Omega) \), which associates with any \( u \) in \( L^3(\Omega)^d \) the solution \( \theta \) of (2.7). We set \( \mathcal{F} = \mathcal{F}_2 \circ \mathcal{F}_1 \). Let \( \theta_1, \theta_2, \tilde{\theta}_1 \) and \( \tilde{\theta}_2 \) in \( H^1(\Omega) \) and \( u_1, u_2 \) in \( L^3(\Omega)^d \) such that

\[
\mathcal{F}_1(\theta_1) = u_1, \quad \mathcal{F}_1(\theta_2) = u_2,
\]

\[
\mathcal{F}_2(u_1) = \tilde{\theta}_1, \quad \mathcal{F}_2(u_2) = \tilde{\theta}_2.
\]

According to (2.14), we have

\[
\| u_1 - u_2 \|_{L^3(\Omega)^d} \leq c_R \| \theta_1 - \theta_2 \|_{H^1(\Omega)}.
\]

On the other hand,

\[
\lambda \int \nabla (\tilde{\theta}_1 - \tilde{\theta}_2) \nabla \varphi \, dx + \int (u_1 \cdot \nabla) \tilde{\theta}_1 \varphi \, dx - \int (u_2 \cdot \nabla) \tilde{\theta}_2 \varphi \, dx = 0
\]

Taking \( \varphi = \tilde{\theta}_1 - \tilde{\theta}_2 \),

\[
\lambda \| \nabla (\tilde{\theta}_1 - \tilde{\theta}_2) \|_{L^2(\Omega)^d}^2 \leq \| u_1 - u_2 \|_{L^3(\Omega)^d} \| \tilde{\theta}_1 \|_{H^1(\Omega)} \| \nabla (\tilde{\theta}_1 - \tilde{\theta}_2) \|_{L^2(\Omega)}.
\]

As \( \tilde{\theta}_1 - \tilde{\theta}_2 \) is equal to 0 on \( \Gamma_* \), we deduce

\[
\lambda \| \tilde{\theta}_1 - \tilde{\theta}_2 \|_{H^1(\Omega)} \leq (1 + c_R^2) \frac{1}{2} \| \tilde{\theta}_2 \|_{H^1(\Omega)} \| u_1 - u_2 \|_{L^3(\Omega)^d}.
\]

Owing to (2.11) and (2.15), we use the fixed point of Banach theorem to conclude.

Condition (2.15) is a little restrictive. So we now prove another result of existence of the solution which only requires some hypothesis on the geometry of the domain.

**Theorem 2.8** We assume that \( \Omega \) is a bounded open set with a Lipschitz boundary of \( \mathbb{R}^2 \), or a convex set or a polyhedron of \( \mathbb{R}^3 \). The problem (2.5)-(2.6)-(2.7) admits a solution \((u, p, \theta)\) in \( L^3(\Omega)^d \times H^1_0(\Omega) \times H^1(\Omega) \).

**Proof.** The existence of \((u, p)\) is exhibited in Proposition 2.3. We introduce the mapping \( \mathcal{F} \) from \( H^1(\Omega) \) into \( L^3(\Omega)^d \) which associates with any \( \theta \) in \( H^1(\Omega) \) the solution \( u \) in \( L^3(\Omega)^d \) of (2.5). Taking \( \tilde{\theta} = \theta - \hat{\theta}_* \), we define the mapping \( \phi(\tilde{\theta}) \) by

\[
\forall \varphi \in H^1_0(\Omega), \quad < \phi(\tilde{\theta}), \varphi > = \lambda \int \nabla \tilde{\theta} \cdot \nabla \varphi \, dx + \int \mathcal{F}(\theta + \hat{\theta}_*) \cdot \nabla \varphi \, dx - \int h \varphi \, dx
\]

\[
- \lambda < \theta_*, \varphi >_{\Gamma_*} + \lambda \int \nabla \hat{\theta}_* \cdot \nabla \varphi \, dx + \int (\mathcal{F}(\tilde{\theta} + \hat{\theta}_*) \cdot \nabla \hat{\theta}_* \varphi \, dx.
\]

This mapping is continuous in \( H^1_0(\Omega) \). By taking \( \varphi = \tilde{\theta} \) we have

\[
< \phi(\tilde{\theta}), \tilde{\theta} > \geq \frac{\lambda}{(1 + c_R^2)} \| \tilde{\theta} \|_{H^1(\Omega)}^2 - (\| h \|_{L^2(\Omega)} + \lambda \| \theta_* \|_{H^1(\Omega)} + \lambda \| \hat{\theta}_* \|_{H^1(\Omega)}) \| \tilde{\theta} \|_{H^1(\Omega)}
\]

\[
- \varepsilon c_R \| \tilde{\theta} \|_{H^1(\Omega)} + c \| \theta_* \|_{H^1(\Omega)} \| \tilde{\theta} \|_{H^1(\Omega)} + \| \tilde{\theta} \|_{H^1(\Omega)} \| \theta_* \|_{H^1(\Omega)} + \| \hat{\theta}_* \|_{H^1(\Omega)}.
\]
Then, we choose \( \varepsilon = \frac{\lambda}{2c(1+c_p^2)} \), this yields that \( <\phi(\tilde{\theta}), \tilde{\theta}> \) is nonnegative on the sphere of \( H^1_0(\Omega) \) with radius
\[
\mu = \frac{2(1+2c_p^2)}{\lambda} \left( \| \theta \|_{L^2(\Omega)} + \lambda \| \theta \|_{H^1_0(\Gamma_2)} \right) + \lambda \left( 1 + \frac{c_p^2}{2} \right) \| \theta \|_{H^1(\Gamma_2)}.
\]

2) As \( D(\Omega \cup \Gamma_2) \) is dense in \( H^1_0(\Omega) \), there exists an increasing sequence \( (\tilde{\phi}_n) \) of finite dimensional subspaces of \( H^1_0(\Omega) \) such that \( \cup_n \tilde{\phi}_n \) is dense in \( H^1_0(\Omega) \). Moreover, the properties of the mapping \( \phi \) established above still hold with \( H^1_0(\Omega) \) replaced by \( \tilde{\phi}_n \). Thus applying Brouwer’s fixed point theorem (see [16, Chap. IV, Cor. 1.1] for instance) yields that, for each \( n \), there exists \( \tilde{\phi}_n \) satisfying
\[
\forall \tilde{\phi}_n \in H^1_0(\Omega), \quad <\phi(\tilde{\phi}_n), \tilde{\phi}_n> = 0 \quad \text{and} \quad \| \tilde{\phi}_n \|_{H^1(\Omega)} \leq \mu.
\]

3) Since \( (\tilde{\phi}_n) \) is bounded in \( H^1_0(\Omega) \), there exists a subsequence still denoted \( (\tilde{\phi}_n) \) for simplify, which converges to \( \tilde{\theta} \) weakly in \( H^1(\Omega) \). For \( m \leq n, \theta_n \) satisfies
\[
\forall \varphi_m \in \tilde{\phi}_m, \quad <\phi(\tilde{\phi}_n), \varphi_m > = 0.
\]

Passing to the limit on \( n \) is obvious for the linear terms. Recalling that \( \mathcal{F} \) is continuous from \( H^1(\Omega) \) into \( L^\infty(\Omega)^d \), it is readily checked that \( H^1(\Omega)^d, s = 1 \) if \( \Omega \) is convex and \( s = \frac{1}{2} + \varepsilon, \varepsilon > 0 \) if \( \Omega \) is a polyhedron (see [4, Prop. 3.7]), is compactly imbedded in \( L^3(\Omega)^d \). Thus, there exists a subsequence still denote \( (\tilde{\theta}_n) \) which converges to \( \tilde{\theta} \) weakly in \( H^1(\Omega) \), so that the sequence \( (\mathcal{F}(\tilde{\theta}_n)) \) converge strongly to \( \mathcal{F}(\tilde{\theta}) \) in \( L^3(\Omega)^d \). We obtain
\[
\int_\Omega (\mathcal{F}(\tilde{\theta}_n) \cdot \nabla) \theta_n \varphi_m \, dx - \int_\Omega (\mathcal{F}(\theta) \cdot \nabla) \varphi_m \, dx \\
\leq \| \mathcal{F}(\theta) - \mathcal{F}(\tilde{\theta}) \|_{L^3(\Omega)^d} \| \theta_n \|_{H^1(\Omega)} \| \varphi \|_{H^1(\Omega)} + \int_\Omega (\mathcal{F}(\theta) \cdot \nabla)(\theta_n - \theta) \varphi_m \, dx.
\]

So, \( (\mathcal{F}(\tilde{\theta}_n) \nabla \tilde{\theta}_n) \) converges weakly to \( \mathcal{F}(\tilde{\theta}) \nabla \tilde{\theta} \). Finally, thanks to the density argument \( \theta = \tilde{\theta} + \tilde{\theta} \) is a solution of (2.7).

**Remark 2.9** If the data \( h, \theta_*, \) and \( \theta_\ast \) satisfy (2.4) and (2.15), then the solution exhibited in Theorem 2.8 is unique.

### 3 The discrete problem

From now on, we assume that the domain \( \Omega \) is the square or the cube \( [-1,1]^d, d = 2 \) or 3, \( \Gamma_* \) and \( \Gamma_2 \) are union of whole sides in dimension \( d = 2 \) or whole faces in dimension \( d = 3 \) of \( \partial \Omega \) and that all data \( h, \theta_2 \) and \( \theta_* \) are continuous on \( \overline{\Omega}, \overline{\Gamma}_2 \) and \( \overline{\Gamma}_* \), respectively. We first describe the discrete problem.

For each nonnegative integer \( n \), we introduce the space \( P_n(\Omega) \) of restrictions to \( \Omega \) of polynomials with \( d \) variables and degree with respect to each variable \( \leq n \).

Let \( N \) be a fixed positive integer, we introduce the discrete spaces
\[
X_N = P_N(\Omega)^d, \quad Y_N = P_N(\Omega), \quad Y_N^* = P_N(\Omega) \cap H^1_0(\Omega), \quad Y_N^* = P_N(\Omega) \cap H^1_0(\Omega).
\]
We recall that there exist a unique set of $N + 1$ nodes $\xi_j$, $0 \leq j \leq N$, with $\xi_0 = -1$ and $\xi_N = 1$, and a unique set of $N + 1$ weights $\rho_j$, $0 \leq j \leq N$, such that the following Gauss-Lobatto formula holds
\[
\forall \phi \in \mathbb{P}_{2N-1}(-1,1), \quad \int_{-1}^{1} \phi(\zeta) \, d\zeta = \sum_{i=0}^{N} \phi(\xi_i) \rho_i.
\]

(3.2)

We also recall [7, Chap IV, Cor. 1.10], the following property, which is useful in what follows
\[
\forall \phi_N \in \mathbb{P}_N(-1,1), \quad \|\phi_N\|^2_{L^2(-1,1)} \leq \sum_{i=0}^{N} \phi_N^2(\xi_i) \rho_i \leq 3\|\phi_N\|^2_{L^2(-1,1)}.
\]

(3.3)

We introduce the grid
\[
\Xi = \begin{cases} 
\{ (\xi_i, \xi_j); \ 0 \leq i, j \leq N \} & \text{in dimension } d = 2, \\
\{ (\xi_i, \xi_j, \xi_k); \ 0 \leq i, j, k \leq N \} & \text{in dimension } d = 3.
\end{cases}
\]

(3.4)

We denote by $I_N$ the Lagrange interpolation operator at the nodes of the grid $\Xi$ with values in $\mathbb{P}_N(\Omega)$, and by $i_N^{\star}$ the Lagrange interpolation operator at the nodes of $\Xi \cap \Gamma_*$ with values in the space of traces of functions in $\mathbb{P}_N(\Omega)$ on $\Gamma_*$. Finally, we introduce the discrete product which is a scalar product on $\mathbb{P}_N(\Omega)$ from (3.3), defined for all continuous functions $u$ and $v$ on $\overline{\Omega}$ by
\[
(u, v)_N = \begin{cases} 
\sum_{i=0}^{N} \sum_{j=0}^{N} u(\xi_i, \xi_j) v(\xi_i, \xi_j) \rho_i \rho_j & \text{if } d = 2 \\
\sum_{i=0}^{N} \sum_{j=0}^{N} \sum_{k=0}^{N} u(\xi_i, \xi_j, \xi_k) v(\xi_i, \xi_j, \xi_k) \rho_i \rho_j \rho_k & \text{if } d = 3.
\end{cases}
\]

(3.5)

On each edge or face $\Gamma_\ell$ of $\Omega$, we define a discrete product: For instance, if $\Gamma_\ell$ is the edge $\{ -1 \} \times [-1,1]^{d-1}$
\[
(u, v)_{N, \ell} = \begin{cases} 
\sum_{i=0}^{N} u(\xi_0, \xi_j) v(\xi_0, \xi_j) \rho_j & \text{if } d = 2, \\
\sum_{j=0}^{N} \sum_{k=0}^{N} u(\xi_0, \xi_j, \xi_k) v(\xi_0, \xi_j, \xi_k) \rho_j \rho_k & \text{if } d = 3.
\end{cases}
\]

(3.6)

A global product on $\Gamma_\ell$ is then defined by
\[
(u, v)_{\ell,M} = \sum_{\ell \in \mathcal{L}} (u, v)_{\ell,M}.
\]

where $\mathcal{L}$ stands for the set of indices $\ell$ such that $\Gamma_\ell$ is contained in $\Gamma_*$. The discrete problem is built from problem (2.5)-(2.6)-(2.7) by the Galerkin method with numerical integration. It is written as

\[
\text{Find } (u_N, p_N, \theta_N) \in X_N \times Y_N^0 \times Y_N, \text{such that}
\]
\[
\theta_N = i_N^{\star} \theta_0 \quad \text{on} \quad \Gamma_*, \\
\forall v_N \in X_N, \quad a(u_N, v_N)_N + (v_N, \nabla p_N)_N = \gamma(f(\theta_N), v_N)_N, \\
\forall q_N \in Y_N^0, \quad (u_N, \nabla q_N)_N = 0, \\
\forall \varphi_N \in Y_N^0, \quad \lambda(\nabla \theta_N, \nabla \varphi_N)_N + ((u_N \cdot \nabla) \theta_N, \varphi_N)_N
\]
\[
= (h, \varphi_N)_N + \lambda(\theta_N, \varphi_N)_N^{\star}.
\]

(3.7)

The well-posedness of this problem will be established later on with the a priori error estimates.
4 A priori error estimate

To apply the theorem due to Brezzi, Rappaz and Raviart [11], we need to write another formulation of both the continuous and discrete problems.

4.1 Another formulation

We define the linear operator $T$, which associates with any data $f$ in $L^2(\Omega)^d$, the solution $\tilde{U} = (u, p)$ of the following problem

$$\forall v \in L^2(\Omega)^d, \quad \alpha \int_{\Omega} u \cdot v \, dx + \int_{\Omega} v \cdot \nabla p \, dx = \gamma \int_{\Omega} f \cdot v \, dx, \quad (4.1)$$

$$\forall q \in H^1_0(\Omega), \quad \int_{\Omega} u \cdot \nabla q \, dx = 0.$$  

It is readily checked that problem (4.1) has a unique solution, see [7, Chap. XIII, Th. 1.9]. We introduce the linear operator $L$ which associates with any data $(h, \theta^\#, \theta^\star)$ in $L^2(\Omega) \times (H^1_0(\Gamma^\#))' \times H^1_0(\Gamma^\star)$ the solution $\theta$ in $H^1(\Omega)$ of problem

$$\theta = \theta^\star \quad \text{on} \quad \Gamma^\star,$$

$$\forall \varphi \in H^1_0(\Omega), \quad \lambda \int_{\Omega} \nabla \theta \cdot \nabla \varphi \, dx = \int_{\Omega} h \varphi \, dx + \lambda <\theta^\#, \varphi>_{\Gamma^\#}. \quad (4.2)$$

Thus, it is readily checked, when setting $U = (\tilde{U}, \theta)$ and assuming that $u$ belongs to $L^3(\Omega)^d$, problem (2.5)-(2.6)-(2.7) can be written equivalently as

$$\mathcal{F}(U) = U - \begin{pmatrix} T & 0 \\ 0 & \mathcal{L} \end{pmatrix} \begin{pmatrix} G^1(U) \\ G^2(U) \end{pmatrix} = 0 \quad (4.3)$$

with $G^1(U) = f(\theta)$ and $G^2(U) = (h - (u \cdot \nabla)\theta, \theta^\#, \theta^\star)$.

We now define the discrete operators $T_N$ and $L_N$ as follows: For any datum $f$ in $L^2(\Omega)^d$, we set

$$T_N(f) = \tilde{U}_N = (u_N, p_N),$$

where $(u_N, p_N)$ is the solution of the following problem

$$\begin{cases} \forall v_N \in X_N, \quad \alpha (u_N, v_N)_N + (v_N, \nabla p_N)_N = \int_{\Omega} f \cdot v_N \, dx, \\
\forall q_N \in Y_N, \quad (u_N, \nabla q_N)_N = 0. \end{cases}$$

Next, we introduce the discrete operator $L_N$ which associates with any data $h$ in $L^2(\Omega)$, $\theta^\#$ in $H^1_0(\Gamma^\#)$ and $\theta^\star$ continuous on $\Gamma^\star$, the solution $\theta_N$ of problem

$$\begin{cases} \theta_N = i_N \theta^\star \quad \text{on} \quad \Gamma^\star, \\
\forall \varphi_N \in Y_N, \quad \lambda (\nabla \theta_N, \nabla \varphi_N)_N = \int_{\Omega} h \varphi_N \, dx + \lambda <\theta_N, \varphi_N>_{\Gamma^\#}. \end{cases} \quad (4.4)$$
So, taking $U_N = (\hat{U}_N, \theta_N)$, problem (3.7) can be equivalently writing as

$$
F_N(U_N) = U_N - \left( \begin{array}{cc} T_N & 0 \\ 0 & L_N \end{array} \right) \left( \begin{array}{c} G^1_N(U_N) \\ G^2_N(U_N) \end{array} \right),
$$

(4.5)

with

$$
G^1_N(U_N) = (\hat{G}^1_N(U_N), \tilde{\theta}_1, \theta_*),
$$

where the components $G^1_N(U_N)$, $G^2_N(U_N)$ and $\tilde{\theta}_2$ are defined by

$$
\forall V_N \in X_N \times Y_N, \quad < G^1_N(U_N), V_N > = (f(\theta_N), v_N)_N,
$$

$$
\forall \varphi_N \in Y_N, \quad < G^2_N(U_N), \varphi_N > = (h, \varphi_N)_N - ((u_N \cdot \nabla)\theta_N, \varphi_N)_N,
$$

and

$$
< \tilde{\theta}_2, \varphi_N > = \lambda(\theta_1, \varphi_N)^{r_2}_N.
$$

We recall the stability properties and error estimates of the linear operators from [7, Chap. V & XIII].

**Proposition 4.1** There exists a nonnegative constant $c$ such that

(i) For any datum $f$ in $L^2(\Omega)^d$,

$$
\|T_N(f)\|_{L^2(\Omega)^d \times H^1(\Omega)} \leq c\|f\|_{L^2(\Omega)^d}.
$$

(4.6)

(ii) For any datum $h$ in $L^2(\Omega)$,

$$
\|L_N(h, 0, 0)\|_{H^1(\Omega)} \leq c\|h\|_{L^2(\Omega)}.
$$

(4.7)

**Proposition 4.2** For any datum $f$ continuous on $\overline{\Omega}$, we assume that the solution $T(f)$ belongs to $H^s(\Omega)^d \times H^{s+1}(\Omega)$, $s \geq 0$. Thus we have

$$
\|(T - T_N)(f)\|_{L^2(\Omega)^d \times H^s(\Omega)} \leq cN^{-s}\|T(f)\|_{H^s(\Omega)^d \times H^{s+1}(\Omega)}.
$$

(4.8)

For any data $h$, $\theta_1$ and $\theta_*$ continuous on $\overline{\Omega}$ and $\overline{\Gamma}_2$ and $\overline{\Gamma}_*$ resp., we assume that $L(h, \theta_1, \theta_*)$ belongs to $H^{s+1}(\Omega)$, $s \geq 0$. Thus we have

$$
\|(L - L_N)(h, \theta_1, \theta_*)\|_{H^s(\Omega)} \leq cN^{-s}\|L(h, \theta_1, \theta_*)\|_{H^{s+1}(\Omega)}.
$$

(4.9)

### 4.2 Preliminary lemmas

To apply the theorem due to Brezzi, Rappaz and Raviart [11], we need some further lemmas. Let us choose an approximation $U_N^\ell = (U_N^\ell, \theta_N^\ell)$ with $U_N^\ell = (u_N^\ell, p_N^\ell)$ of $(u, p, \theta)$ in $X_N \times Y_N \times Y_N$ which satisfies for real numbers $\ell$, $s$ such that $0 \leq \ell \leq s$

$$
\|u - u_N\|_{H^\ell(\Omega)^d} \leq cN^{\ell-s}\|u\|_{H^s(\Omega)^d},
$$

(4.10)

$$
\|p - q_N\|_{H^{\ell+1}(\Omega)} \leq cN^{\ell-s}\|p\|_{H^{s+1}(\Omega)},
$$

(4.11)

$$
\|\theta - \theta_N\|_{H^{\ell+1}(\Omega)} \leq cN^{\ell-s}\|\theta\|_{H^{s+1}(\Omega)}.
$$

(4.12)

The existence of such an approximation is stated in [6, Thm. 7.4] (see also [7, Chap. III, Th. 2.4 & Chap. VI, Th. 2.5]).
From now on, we denote by
\[ \mathcal{W} = \mathcal{L}^3(\Omega)^d \times H^1_0(\Omega) \times H^1(\Omega), \]
and
\[ \mathcal{W}_N = \mathcal{X}_N \times \mathcal{Y}_N \times \mathcal{Y}_N. \]
We also denote by \( D \) the differential operator with respect to \((u, p, \theta)\).

**Assumption 4.3** The solution \((u, p, \theta)\) of problem (2.5)-(2.6)-(2.7)
(i) belongs to \( H^s(\Omega)^d \times H^{s+1}(\Omega) \times H^{s+1}(\Omega), s > \frac{d}{2} \),
(ii) is such that \( D \mathcal{F}(U) \) is an isomorphism of the space \( \mathcal{W} \).

In a first step, we must prove the analogue of part (ii) of Assumption 4.3 for the discrete operator. Let us denote by \( \mathcal{E}(\mathcal{W}) \) the space of endomorphisms of \( \mathcal{W} \).

**Lemma 4.4** Assume that the mapping \( f(\cdot) \) is of class \( C^2 \) with bounded derivatives. If Assumption 4.3 holds, there exists a positive real number \( N_0 \) such that for all \( N \geq N_0 \), the operator \( D \mathcal{F}_N(U^N, \theta^N) \) is an isomorphism of \( \mathcal{W}_N \) and the norm of its inverse is bounded independently of \( N \).

**Proof.** We have
\[ D \mathcal{F}_N(U^N) = D \mathcal{F}(U) + \begin{pmatrix} \mathcal{T} - \mathcal{T}_N & 0 \\ 0 & \mathcal{L} - \mathcal{L}_N \end{pmatrix} \begin{pmatrix} D \mathcal{G}^1(U) \\ D \mathcal{G}^2(U) \end{pmatrix} 
+ \begin{pmatrix} \mathcal{T}_N & 0 \\ 0 & \mathcal{L}_N \end{pmatrix} \begin{pmatrix} D \mathcal{G}^1(U) - D \mathcal{G}^1(U^N) \\ D \mathcal{G}^2(U) - D \mathcal{G}^2(U^N) \end{pmatrix} 
+ \begin{pmatrix} \mathcal{T}_N & 0 \\ 0 & \mathcal{L}_N \end{pmatrix} \begin{pmatrix} D \mathcal{G}^1(U^N) - D \mathcal{G}^1(U^{N}) \\ D \mathcal{G}^2(U^N) - D \mathcal{G}^2(U^{N}) \end{pmatrix}. \] (4.13)

It suffices to prove that the last three terms in the right side tend to 0 when \( N \to 0 \). So, let \( W_N = (w_N, \rho_N, \zeta_N) \) be in the unit sphere of \( \mathcal{W}_N \). We proceed in three steps.
1) We have
\[ \begin{pmatrix} D \mathcal{G}^1(U) \cdot W_N \\ D \mathcal{G}^2(U) \cdot W_N \end{pmatrix} = \begin{pmatrix} f'(\theta)\zeta_N \\ - (u \cdot \nabla)\zeta_N - (w \cdot \nabla)\theta, 0, 0 \end{pmatrix} \]
We deduce from (4.6) and (4.8) that, if \( \mathcal{K} \) is compact of \( L^2(\Omega)^d \), we have the convergence
\[ \lim_{N \to \infty} \sup_{f \in \mathcal{K}} \| (\mathcal{T} - \mathcal{T}_N)f \|_{L^2(\Omega)^d \times H^1(\Omega)} = 0. \]

So, the image of the unit ball of \( \mathcal{W}_N \) by the fonction
\[ (w_N, \rho_N, \zeta_N) \mapsto f'(\theta)\zeta_N, \]
is compact of \( L^2(\Omega)^d \). On the other hand, \((u \cdot \nabla)\zeta_N + (w_N \cdot \nabla)\theta\) belongs to a compact subsequence of \( L^2(\Omega)^d \). This yields
\[ \lim_{N \to \infty} \left\| \begin{pmatrix} \mathcal{T} - \mathcal{T}_N & 0 \\ 0 & \mathcal{L} - \mathcal{L}_N \end{pmatrix} \begin{pmatrix} D \mathcal{G}^1(U) \\ D \mathcal{G}^2(U) \end{pmatrix} \right\|_{\mathcal{E}(\mathcal{W})} = 0. \]
2) For the next term, we prove the convergence of

\[
(f'(\theta) - f'(\theta_N^*) )\cdot \zeta_N \quad \text{and} \quad (u - u_N^*) \cdot \nabla)\cdot \zeta_N - (w_N \cdot \nabla)(\theta - \theta_N^*).
\]

Let \( V' \) be in the unit sphere of \( W_N \), to evaluate the first one we use the continuous imbedding of \( H^1(\Omega) \) in \( L^4(\Omega) \), we obtain

\[
< (D\theta^1(U) - D\theta^1(U_N^*)) \cdot W_N, V_N > \leq cN^{-s}\|\theta\|_{H^{s+1}(\Omega)}\|v_N\|_{L^2(\Omega)^d},
\]

also for the second one,

\[
< (D\theta^2(U) - D\theta^2(U_N^*)) \cdot W_N, V_N > \leq \|u - u_N^*\|_{L^2(\Omega)^d}\|\nabla \zeta_N\|_{L^2(\Omega)^d}\|\varphi_N\|_{L^6(\Omega)} + \|w_N\|_{L^2(\Omega)^d}\|\nabla(\theta - \theta_N^*)\|_{L^2(\Omega)^d}\|\varphi_N\|_{L^6(\Omega)}.
\]

We choose \( \ell \) equal to \( \frac{d}{6} \) in (4.10) for \( H^4(\Omega) \) to be imbedded in \( L^3(\Omega) \). Thus

\[
\lim_{N \to \infty} \left\| \begin{pmatrix} T_N & 0 \\ 0 & \mathcal{L}_N \end{pmatrix} \begin{pmatrix} D\theta^1(U) - D\theta^1(U_N^*) \\ D\theta^2(U) - D\theta^2(U_N^*) \end{pmatrix} \right\|_{\mathcal{L}(W)} = 0.
\]

3) Evaluating the last term, we have

\[
< (D\theta^1(U_N^*) - D\theta^1(U_N^*)) \cdot W_N, V_N > = \int_{\Omega} f'(\theta_N^*) \cdot v_N \, dx - (f'(\theta_N^*) \cdot \zeta_N, v_N) = \int_{\Omega} f'(\theta_N^*) \cdot v_N \, dx - (\mathcal{T}_N f'(\theta_N^*)) \cdot v_N).
\]

We take \( N^* \) equal to the integer part of \( \frac{N - 1}{2} \), we introduce an approximation \( f_{N^*} \) of \( f'(\theta_N^*) \) in \( X_{N^*} \) and \( \zeta_{N^*} \) of \( \zeta_N \) in \( Y_{N^*} \). Thanks to the exactitude of the quadrature formula (3.2)

\[
< (D\theta^1(U_N^*) - D\theta^1(U_N^*)) \cdot W_N, V_N > \leq (1 + 3^d)\|f'(\theta_N^*) - f_{N^*}\|_{L^p(\Omega)^d}\|\zeta_{N^*}\|_{L^p(\Omega)^d}\|v_N\|_{L^2(\Omega)^d} + (1 + 3^d)\|\zeta_N - \zeta_{N^*}\|_{L^p(\Omega)^d}\|f_{N^*}\|_{L^p(\Omega)^d}\|v_N\|_{L^2(\Omega)^d} + 3^d\|f'(\theta_N^*) - \mathcal{T}_N f'(\theta_N^*)\|_{L^p(\Omega)^d}\|v_N\|_{L^2(\Omega)^d}.
\]

On the other hand, we introduce \( u_{N^*}, w_{N^*} \) approximations of \( u_N^* \) and \( w_N^* \) in \( X_{N^*} \) and \( \theta_{N^*} \) of \( \theta_N^* \) in \( Y_{N^*} \), we see that

\[
< (D\theta^2(U_N^*) - D\theta^2(U_N^*)) \cdot W_N, V_N > \leq (1 + 3^d)\|u_{N^*} - u_{N^*}\|_{L^2(\Omega)^d}\|\nabla \zeta_{N^*}\|_{L^2(\Omega)^d} + \|u_{N^*}\|_{L^3(\Omega)^d}\|\nabla(\zeta_{N^*} - \zeta_N)\|_{L^2(\Omega)^d} + \|w_{N^*} - w_{N^*}\|_{L^3(\Omega)^d}\|\nabla \theta_{N^*}\|_{L^2(\Omega)^d} + \|w_{N^*}\|_{L^3(\Omega)^d}\|\nabla(\theta_{N^*} - \theta_{N^*})\|_{L^2(\Omega)^d}\|\varphi_N\|_{L^6(\Omega)}.
\]

The stability properties of the approximations [7, Chap. III] yield the convergence of the last term

\[
\lim_{N \to \infty} \left\| \begin{pmatrix} T_N & 0 \\ 0 & \mathcal{L}_N \end{pmatrix} \begin{pmatrix} D\theta^1(U_N^*) - D\theta^1(U_N^*) \\ D\theta^2(U_N^*) - D\theta^2(U_N^*) \end{pmatrix} \right\|_{\mathcal{L}(W)} = 0.
\]

Combining all this yields the desired result.

**Lemma 4.5** If Assumption 4.3 holds and the mapping \( f(\cdot) \) is of class \( C^2 \) with bounded derivatives, there exist a neighborhood of \( U_N \) in \( W_N \) and a positive constant \( c \), such that the following property holds for any \( Z_N \) in this neighborhood

\[
\|D\mathcal{F}_N(V_N) - D\mathcal{F}_N(Z_N)\|_{\mathcal{L}(W)} \leq c\|V_N - Z_N\|_W. \quad (4.14)
\]
Proof. We set $Z_N = (z_N, \sigma_N, \zeta_N)$, we have
\[
DF_N(U_N^0) - DF_N(Z_N) = (U_N^0 - Z_N) - \left( \begin{array}{cc} \mathcal{T}_N & 0 \\ 0 & \mathcal{L}_N \end{array} \right) \left( \begin{array}{c} DG^1_N(U_N^0) - DG^1_N(Z_N) \\ DG^2_N(U_N^0) - DG^2_N(Z_N) \end{array} \right).
\]
with
\[
< (DG^1_N(U_N^0) - DG^1_N(Z_N)) \cdot W_N, V_N > \leq 3^d \| \theta^*_N - \sigma_N \|_{H^1(\Omega)} \| \zeta_N \|_{H^1(\Omega)} \| u_N \|_{L^2(\Omega)^d}.
\]
and
\[
< (DG^2_N(U_N^0) - DG^2_N(Z_N)) \cdot W_N, V_N > \leq 3^d (\| u_N \|_{L^2(\Omega)^d} \| \nabla (\theta^*_N - \sigma_N) \|_{L^2(\Omega)^d} + \| u_N^0 - z_N \|_{L^2(\Omega)^d} \| \nabla \zeta_N \|_{L^2(\Omega)^d} + \| u \|_{H^1(\Omega)}).
\]
Combining all this with the stability properties of $\mathcal{T}_N$ and $\mathcal{L}_N$, yields the desired result.

Lemma 4.6 If the assumptions of the previous lemma hold, and if the data $(h, \theta, \theta_\sigma)$ belong to $H^0(\Omega) \times H^0(\Gamma_1) \times H^{\alpha+\frac{1}{2}}(\Gamma_\sigma)$, $\sigma > \frac{d}{2}$, then the following estimate is satisfied
\[
\| F_N(U_N^0) \|_W \leq c N^{\frac{d}{2} - s} \| u \|_{H^*(\Omega)^d} + c N^{s} (\| p \|_{H^{s+1}(\Omega)} + \| \theta \|_{H^{s+1}(\Omega)}) + c N^{-s} (\| h \|_{H^s(\Omega)} + \| \theta_\sigma \|_{H^s(\Gamma_1)}). \tag{4.15}
\]

Proof. Thanks to (4.3), we have
\[
F_N(U_N^0) = -(U - U_N^0) + \left( \begin{array}{cc} \mathcal{T} - \mathcal{T}_N & 0 \\ 0 & \mathcal{L} - \mathcal{L}_N \end{array} \right) \left( \begin{array}{c} G_1(U) \\ G_2(U) \end{array} \right)
+ \left( \begin{array}{cc} \mathcal{T}_N & 0 \\ 0 & \mathcal{L}_N \end{array} \right) \left( \begin{array}{c} G^1(U) - G^1(U_N^0) \\ G^2(U) - G^2(U_N^0) \end{array} \right)
+ \left( \begin{array}{cc} \mathcal{T}_N & 0 \\ 0 & \mathcal{L}_N \end{array} \right) \left( \begin{array}{c} G^1(U_N^0) - G^1(U_N^0) \\ G^2(U_N^0) - G^2(U_N^0) \end{array} \right).
\]

To bound the first term, we use imbedding of $H^\ell(\Omega)$ in $L^3(\Omega)$ when $\ell = \frac{d}{6}$ combining with (4.10), (4.11) and (4.12), we obtain
\[
\| U - U_N^0 \|_W \leq c N^{\frac{d}{6} - s} \| u \|_{H^*(\Omega)^d} + N^{-s} (\| p \|_{H^{s+1}(\Omega)} + \| \theta \|_{H^{s+1}(\Omega)}).
\]
The second term is bounded in (4.8) and (4.9). To evaluate the third term we use the triangle inequality and estimate (4.10) and (4.12). Finally, to prove the estimate of the fourth term we use the same arguments in the proof of Lemma 4.4.

4.3 Final result

Owing to Lemmas 4.4 to 4.6, all the assumptions needed to apply the theorem of Brezzi, Rappaz and Raviart [11, Thm. 7] (see also [16, Chap.IV, Thm. 3.1]) are satisfied.
**Theorem 4.7** Let \((u, p, \theta)\) be a solution of problem (2.5)-(2.6)-(2.7) satisfying Assumption 4.3. If the function \(f\) is of class \(C^2\) with bounded derivatives and if the data \((h, \theta_\star, \theta_\sharp)\) belong to \(H^\sigma(\Omega) \times H^\sigma(\Gamma_\star) \times H^{\sigma + \frac{1}{2}}(\Gamma_\sharp)\) for any real number \(\sigma > \frac{1}{2}\), there exists an integer \(N_\star\) and a constant \(c > 0\) such that for all \(N \geq N_\star\), Problem (3.7) has a unique solution \((u_N, p_N, \theta_N)\) in a neighborhood of \((u, p, \theta)\). Moreover, this solution satisfies

\[
\|u - u_N\|_{L^3(\Omega)} + \|p - p_N\|_{H^1(\Omega)} + \|\theta - \theta_N\|_{H^1(\Omega)} \\
\leq c N_\star^{-\sigma} \|u\|_{H^s(\Omega)}^s \\
+ c N_\star^{-\sigma} (\|p\|_{H^{s+1}(\Omega)} + \|\theta\|_{H^{s+1}(\Omega)}) \\
+ c N_\star^{-\sigma} (\|h\|_{H^s(\Omega)} + \|\theta_\star\|_{H^{s+\frac{1}{2}}(\Gamma_\star)} + \|\theta_\sharp\|_{H^s(\Gamma_\sharp)}). \tag{4.16}
\]

Estimate (4.16) is fully optimal and the assumptions which are required on the solution \((u, p, \theta)\) seem likely and are not at all restrictive.

5 Numerical experiments

We present in this section some numerical tests in order to validate and confirm the interest of our spectral approximation of the coupled problem (1.1), from both numerical and physical points of view. Firstly, in order to check rate convergence of the proposed spectral approximation, we solve this nonlinear coupled problem numerically in two dimensions. Secondly, we treat a real case, so we propose Horton-Rogers-Lapwood Problem to simulate. All the computations have been performed on the code FreeFEM3D, spectral version, developed during the thesis of D. Yakoubi [19] in collaboration with S. Del Pino [13]. Note that our results are obtained using a \(P_N \times P_N \times P_N\) space discretization of \((u, p, \theta)\).

5.1 Iterative scheme

To solve the nonlinear coupled system, we adopt the following iterative procedure based on a decoupled computation of the Darcy and heat equations that we present in the continuous case for simplicity: We choose \(T^0\) as a solution of problem

\[
\begin{aligned}
-\lambda T^0 &= h^0 \quad \text{in} \quad \Omega, \\
T^0 &= T_\star \quad \text{on} \quad \Gamma_\star, \\
\frac{\partial T}{\partial n} &= \theta_\sharp \quad \text{on} \quad \Gamma_\sharp.
\end{aligned}
\]

The iterative scheme reads

Find \((u^n, p^n, T^n) \in L^3(\Omega)^d \times H^1(\Omega) \times H^1(\Omega)\) with \(T - T_\star \in H^1(\Omega)\) such that

\[
\forall v \in L^2(\Omega)^d, \quad \alpha \int_\Omega u^n \cdot v \, dx + \int_\Omega v \cdot \nabla p^n \, dx = \int_\Omega F(T^{n-1}) \cdot v \, dx; \tag{5.1}
\]

\[
\forall q \in H^1(\Omega), \quad \int_\Omega u^n \cdot \nabla q \, dx = 0; \tag{5.2}
\]

\[
\forall \varphi \in H^1(\Omega), \quad \lambda \int_\Omega \nabla T^n \cdot \nabla \varphi \, dx + \int_\Omega (u^n \cdot \nabla)T^n \varphi \, dx = \int_\Omega h \varphi \, dx. \tag{5.3}
\]
Theorem 5.1  Let \( h \in L^2(\Omega), \theta_x \in (H_0^1(\Gamma))', \) and \( T_* \in H^{\frac{1}{2}}(\Gamma_*) \) such that
\[
\frac{2c_R}{\lambda^2(1 + c_2^2)^{\frac{1}{2}}} (c_p||h||_{L^2(\Omega)} + 2\lambda||T_*||_{H^{\frac{1}{2}}(\Gamma_*)} + \lambda||\theta_x||_{L^2(\Gamma_*)}) < 1.
\]
Then, the sequences \((u^n)\), \((p^n)\), and \((T^n)\) converge. Moreover, the velocity sequence \((u^n)\) and the temperature sequence \((T^n)\) satisfy
\[
\|u^{n+1} - u^n\|_{L^2(\Omega)} \leq \frac{1}{\alpha} \|T^n - T^{n-1}\|_{L^2(\Omega)},
\]
\[
\|T^{n+1} - T^n\|_{H^1(\Omega)} \leq \frac{2c_R}{\lambda^2(1 + c_2^2)^{\frac{1}{2}}} (c_p||h||_{L^2(\Omega)} + 2\lambda||T_*||_{H^{\frac{1}{2}}(\Gamma_*)} + \lambda||\theta_x||_{L^2(\Gamma_*)}) \|T^n - T^{n-1}\|_{H^1(\Omega)}.
\]

Proof. We choose \( v = u^{n+1} - u^n \) and inject it in (5.1) in steps \( n \) and \( n + 1 \). The difference between the two equations gives
\[
\alpha\|u^{n+1} - u^n\|^2_{L^2(\Omega)^d} = \int_\Omega \left(F(T^n) - F(T^{n-1})\right)(u^{n+1} - u^n) \, dx
\leq ||F(T^n) - F(T^{n-1})||_{L^2(\Omega)^d} ||u^{n+1} - u^n||_{L^2(\Omega)^d}.
\]
By the same argument, we choose \( \varphi = T^{n+1} - T^n \), we obtain
\[
\lambda||\nabla(T^{n+1} - T^n)||_{L^2(\Omega)^d} \leq ||u^{n+1} - u^n||_{L^3(\Omega)^d} ||T^{n+1}||_{H^1(\Omega)}.
\]
Since \((T^n)\) is bounded, see (2.11), we obtain
\[
\lambda||\nabla(T^{n+1} - T^n)||_{L^2(\Omega)^d} \leq c_2 \left(||h||_{L^2(\Omega)} + ||\theta_x||_{H^{\frac{1}{2}}(\Gamma_*)} + ||\theta_x||_{L^2(\Gamma_*)}\right) ||u^{n+1} - u^n||_{L^3(\Omega)^d}.
\]
Thanks to (2.14)
\[
||u^{n+1} - u^n||_{L^3(\Omega)} \leq c_R ||T^n - T^{n-1}||_{H^1(\Omega)}.
\]
Combining the last two inequalities gives the desired result.

5.2 Accuracy test

In this test, we are interested in the convergence of the solution with respect to polynomial degree \( N \) for the triplet \((u, p, T)\) based on an analytic solution in dimension 2 defined in a square \( \Omega = [-1, 1]^2 \) by
\[
\begin{align*}
u_1(x, y) &= -\sin(\pi x)\cos(\pi y), & u_2(x, y) &= \cos(\pi x)\sin(\pi y) \\
p(x, y) &= -\frac{1}{\pi}\sin(\pi x)\cos(\pi y), & T(x, y) &= 2\cos(\pi x)\sin(\pi y).
\end{align*}
\]
We choose to work with a non-constant permeability coefficient \( \alpha \) and strongly depending on temperature as follow \( \alpha(x, y) = \frac{1}{T^2(x, y) + 1} \), however the thermal conductivity \( \lambda \) is taken equal to 1. Numerically, the function \( \alpha \) was replaced by it Lagrange interpolates and the integral \( \int_\Omega u^n \cdot v \, dx \) was replaced by \( \int_\Omega (I_N \alpha) u^n \cdot v \, dx \). The corresponding source terms are:
\[
F_1(T) = \frac{-1}{T^2 + 1} \sin(\pi x)\cos(\pi y) - \cos(\pi x)\cos(\pi y)
\]
\[
F_2(T) = \frac{1}{T^2 + 1} \cos(\pi x)\sin(\pi y) + \sin(\pi x)\sin(\pi y)
\]
\[
h = 4\pi^2\cos(\pi x)\sin(\pi y) + 2\pi\cos(\pi y)\sin(\pi y).
\]
We plot in Figure 1 the $L^2$-norm error of velocity, pressure and temperature in the left panel and $H^1$-error of pressure and temperature in the right panel between the numerical solution and the exact solution with a successive polynomial degree from $N = 5$ to $N = 25$. These results confirm the expected rates of convergence. Hence the spectral convergence for all unknowns is obtained which is consistent with the error estimate (4.16). We observe that beyond degree 20, the leading error is due to the accuracy of the machine, so that the curve stops decreasing.

![Figure 1: Convergence rate with respect to polynomial degree](image1)

In Figure 2, we present the exact solutions in the left panel and the discrete solutions in the right one issued from (5.6), where the discrete one are computed with $N = 17$. As standard in spectral methods, the results in the two parts of the figure can not be distinguished.

![Figure 2](image2)
Figure 2: Exact Solution (left) versus Spectral Solution for $N = 17$ (right)
5.3 The Horton-Rogers-Lapwood Problem

This section deals with the simulations of the so-called Horton-Rogers-Lapwood Problem (HRL), see for instance [18, Chap. 6], when the fluid-saturated porous medium has a horizontal layer geometry, and is heated uniformly from the bottom of the layer, as explain in Figure 3. Consequently, there may exist a temperature difference between the top and bottom boundaries of the layer. Since the positive direction of the temperature gradient due to this temperature difference is opposite to that of the gravity acceleration, there is no natural convection for a small temperature gradient in the porous medium. In this case, heat energy is solely transferred from the high-temperature region (the bottom of the horizontal layer) to the low-temperature region (the top of the horizontal layer) by thermal conduction. However, if the temperature difference is large enough, it may trigger natural convection in the fluid-saturated porous medium.

The appropriate equations are in the simple configuration:

\[
\begin{align*}
\frac{\partial t}{\partial t}u + \frac{\mu}{K} u + \nabla p &= \rho_0 (1 - \beta (T - T_0)) g \\
\nabla \cdot u &= 0 \\
\frac{\partial t}{\partial t}T + u \cdot \nabla T - \Delta T &= 0,
\end{align*}
\]

where \(\rho_0\) and \(\mu\) denote the density and viscosity of the fluid, \(K\) the permeability, \(\beta\) the thermal expansion coefficient, \(\rho_0\) the density and \(g = (0, g)\) the gravitational acceleration. It is clear that System (5.7) have a following steady state solution, which satisfies the boundary conditions: \(T|_{y=0} = T_0 + \delta T\) and \(T|_{y=H} = T_0\):

\[
T = T_0 + \delta T \left(1 - \frac{y}{H}\right) \\
\quad u = (0, 0), \quad p = P_0 - \rho_0 g \left(y + \frac{1}{2} \beta \delta T \left(\frac{y^2}{H} - 2y\right)\right)
\]

This solution describes the conduction state, one in which the heat transfer is solely by thermal conduction, i.e when \(\delta T\) is small enough.

The proposed iterative procedure (5.1), (5.3) associated with the spectral method for solving the HRL problem in a fluid-saturated porous medium is validated by two realizations. A rectangular domain \(\Omega = [0, 2] \times [0, 1]\) is considered in the calculation. In the first test, the physical parameters are taken as follows: \(\delta T = 1\), \(\beta = 0.1\), \(g = 10\), \(\mu = K = 1\). As shown in Figure 4, the convergence of iterative scheme (5.1), (5.3) is confirmed, clearly the error \(L^2\)-error(u)=\(\|u^{n+1} - u^n\|\) is less than \(L^2\)-error(T)=\(\|T^n - T^{n-1}\|\) which is compatible with estimate (5.4).

In the second test, we increase the temperature difference \(\delta T = 100\) between the top and the bottom of domain. Figure 5 shows that when \(\delta T\) is higher, instability appears as convection in the form of a cellular motion.
6 Conclusion

In this paper, we dealt with the spectral approximation of the steady model for the coupling of the heat equation with the Darcy equations. We prove that both continuous and discrete problems admit at least a solution in suitable spaces. The a priori estimates are then proved in appropriate norms. The coupling problem is nonlinear, we introduce an iterative scheme based on the fixed point algorithm whose convergence is established. Numerical tests confirm these theoretical findings.

Furthermore, we simulate the well known problem so-called Horton-Rogers-Lapwood, where the obtained results are in total concordance with the well known results, i.e. the natural convection is observed when the difference of the temperature $\delta T$ is higher.

References


