An Optimal Control Problem for a Generalized Boussinesq Model: The Time Dependent Case

Jose Luiz Boldrini, Enrique Fernández-Cara, and Marko Antonio Rojas-Medar

Department of Mathematics,
IMECC-UNICAMP, CP 6065, 130813-859,
Campinas-SP — Brazil
boldrini@ime.unicamp.br

Department of Differential Equations
and Numerical Analysis,
University of Sevilla, Aptdo. 1160,
41080 Sevilla — Spain
cara@us.es

Department of Applied Mathematics,
IMECC-UNICAMP, CP 6065, 13083-859,
Campinas-SP — Brazil
marko@ime.unicamp.br

Received: January 9, 2006
Accepted: January 15, 2007

ABSTRACT

We consider an optimal control problem governed by a system of nonlinear partial differential equations modelling viscous incompressible flows submitted to variations of temperature. We use a generalized Boussinesq approximation. We obtain the existence of the optimal control as well as first order optimality conditions of Pontryagin type by using the Dubovitskii-Milyutin formalism.

Key words: optimal control, Boussinesq model, Dubovitskii-Milyutin formalism.

2000 Mathematics Subject Classification: 49J20, 76D03.

Partially supported by MEC (Spain, Grants BFM2003-06446 and HBP2005-0014), CNPq (Brazil, Grant 301354/03-0) and FAPESP (Brazil; Grant FAPESP 2005/50705-4).

Introduction

In this work we consider an optimal control problem for the equations satisfied by the velocity field $\mathbf{u}$, the pressure p and the temperature θ of a viscous incompressible fluid subject to heat effects. We use a generalized Boussinesq approximation, where the viscosity is assumed to be temperature dependent.

Roughly speaking, we intend to determine the least amount of heat to be applied in order that the velocity and the temperature be as good as possible in certain subdomains.

To be more precise, let the flow domain be a bounded open set $\Omega \subset \mathbb{R}^N$, with $N = 2$ or 3 and let $0 < T < +\infty$ be the final time of observation. We will set $Q = \Omega \times (0, T)$ and $\Sigma = \partial\Omega \times (0, T)$.

Let us introduce the initial data $\mathbf{u}_0$ and θ_0. Also, let us introduce two fixed nonempty open sets

$$\omega_{\mathbf{u}},\, \omega_\theta \subset \Omega$$

and let us assume that a velocity field $\mathbf{u}_d$ defined on $\omega_{\mathbf{u}} \times (0, T)$, a temperature θ_d on $\omega_\theta \times (0, T)$ and two external sources $\mathbf{h}$ and f are given. Our problem is to find a suitable heat source v (the control) in the set of admissible controls $\mathcal{U}$ such that the associated fluid velocity and temperature, i.e. the solution together with p of the system

$$\begin{cases} \mathbf{u}_t - \operatorname{div}(\nu(\theta)\nabla\mathbf{u}) + (\mathbf{u}\cdot\nabla)\mathbf{u} + \nabla p = \theta\mathbf{g} + \mathbf{h}, & (x,t) \in Q, \\ \operatorname{div}\mathbf{u} = 0, & (x,t) \in Q, \\ \theta_t - k\,\Delta\theta + \mathbf{u}\cdot\nabla\theta = f + v, & (x,t) \in Q, \\ \mathbf{u} = 0, \quad \theta = 0, & (x,t) \in \Sigma, \\ \mathbf{u}(0,x) = \mathbf{u}_0(x), \quad \theta(0,x) = \theta_0(x), & x \in \Omega, \end{cases} \tag{1}$$

minimizes the functional

$$J(\mathbf{u},\theta,v) = \frac{\alpha_1}{2}\iint_{\omega_{\mathbf{u}}\times(0,T)} |\mathbf{u}-\mathbf{u}_d|^2 + \frac{\alpha_2}{2}\iint_{\omega_\theta\times(0,T)} |\theta-\theta_d|^2 + \frac{\mu}{2}\iint_Q |v|^2 \tag{2}$$

subject to $v \in \mathcal{U}$, where $\alpha_1 \geq 0$, $\alpha_2 \geq 0$ and $\mu > 0$ are given constants.

The first goal of this paper is to show that this problem admits at least one optimal solution. The second one is to characterize such solution in terms of first order optimality conditions, that is to say, to deduce a system of equations that the optimal solution and the associated adjoint state must satisfy. In the process, we will also find the corresponding Pontryagin minimum principle for the problem.

The proof of the existence of an optimal solution is (more or less) standard. It relies on suitable existence results and *a priori* estimates for (1).

In what concerns the optimality conditions, the situation is more complex. The techniques usually employed for distributed control problems (see for instance [1, 3, 4, 14, 15]) are more difficult to apply in this case because the viscosity ν depends on θ. Although it is not strictly necessary, in this paper we will use an alternative method: the so called formalism of Dubovitskii and Milyutin. This approach was introduced in the context of mathematical programming. Later, it was shown to be very useful for optimal control problems for ordinary differential equations. A good presentation of its applications to these areas can be found in Girsanov [10]; see also Flett [9]. Recently, these techniques have been applied in a very promising way to some distributed control problems.

Some basic ideas that can be used to explain the formalism are the following. At a local minimizer, the cone of *descent directions* associated to J must be disjoint of the intersection of the cones of *feasible directions* and *tangent directions*, respectively determined by $\mathcal{U}$ and (1). Consequently, from Hahn-Banach's theorem and some additional arguments, it follows that there must exist elements in the associated dual cones, not all them zero, that add up to zero. This algebraic condition is just the Euler-Lagrange system of the extremal problem at hand. When it is possible to identify the previous primal and dual cones, this system provides the first order optimality conditions in a systematic way. In the case of an optimal control problem, it also leads to the corresponding Pontryagin minimum (or maximum) principle.

Thus, a major task in our problem is the identification of the cones mentioned above in terms of the involved partial differential equations. Since the difficulty level of this task is related to the highly nonlinear behavior of (1), we will now make some comments on the physical meaning of the variables and constants and we will also recall some mathematical results that are known for these equations.

A derivation of the equations (1) can be found for instance in Drazin and Reid [8]. The physical variables are the following: $\mathbf{u}(x,t) \in \mathbb{R}^N$ denotes the velocity of the fluid at point $x \in \Omega$ and time $t \in [0,T]$; $p(x,t) \in \mathbb{R}$ is the hydrostatic pressure; $\theta(x,t) \in \mathbb{R}$ is the temperature; $\mathbf{g}(x,t)$ is the external force per unit-mass; $\nu(\cdot) > 0$ and $k > 0$ are respectively the kinematic viscosity and thermal conductivity; $\mathbf{h}$ and f are given external sources.

In this paper, the expressions ∇, Δ, and div respectively denote the gradient, Laplace, and divergence operator; the i-th component in cartesian coordinates of $(\mathbf{u}\cdot\nabla)\mathbf{u}$ is given by $((\mathbf{u}\cdot\nabla)\mathbf{u})_i = \sum_{j=1}^N u_j \partial_{x_j} u_i$; also, $\mathbf{u}\cdot\nabla\theta = \sum_{j=1}^n u_j \partial_{x_j}\theta$.

For simplicity, in this work we consider homogeneous Dirichlet boundary conditions; the more general case of nonzero Dirichlet conditions on Σ can be reduced to this one by assuming suitable smoothness on the boundary data (in connection to this, see for instance Lorca and Boldrini [18]). Such a reduction leads only to changes in the right-hand sides of (1) (where some linear and nonlinear terms have to be added), which have no influence on the proofs of our results in an essential way.

Recall that the classical Boussinesq equations correspond to the important special case where ν and k are positive constants; see for instance Morimoto [19], Ōeda [20], Hishida [11], and Shinbrot and Kotorynski [22]. For some results concerning the optimal control of the classical evolution Boussinesq equations, see for instance Lee and Shin [12] and Li and Wang [13]. For certain fluids, however, we cannot neglect the variation of the viscosity with temperature, this being important in the determination of the details of the flow. In particular, it is believed that the temperature dependence of the viscosity is responsible for the fact that the direction of the flow in the middle of a convection cell is usually different for gases and liquids; see Lorca and Boldrini [16] and the references therein. Thus, it is also important to know well the properties of the equations (1).

From the mathematical point of view, (1) has been less studied and a rigorous analysis is more difficult than in the case of the classical Boussinesq equations. In the more general case where both the viscosity and thermal conductivity are temperature dependent, the spectral Galerkin method was used by Lorca and Boldrini [16] to prove the existence of stationary solutions; the same authors found local in time strong solutions in [17]. The global existence and regularity of the solutions is considered in [18]; another global existence result, under somewhat different conditions, is obtained in [6]. Other existence results, under different situations and conditions, are for instance given in Shilkin [21], Zabrodzki [23], and Díaz and Galiano [7].

To end this introductory section, let us mention that, in the last section of this paper, we will present some possible extensions and variants of our optimal control problem. In particular, we will consider the interesting case in which the controls are localized in space. We will also make some comments on the difficulties of using weak solutions, the possibilities of using large controls and the uniqueness of the control-to-state mapping.

1. Preliminaries and hypotheses

We begin by fixing the notation and recalling certain definitions and results that will be used later on. In what follows, the functions are either $\mathbb{R}$ or $\mathbb{R}^N$ valued ($N = 2$ or 3); $H^m(\Omega) = W^{m,2}(\Omega)$ and $W^{k,p}(\Omega)$ are the usual Sobolev spaces (see Adams [2] for their properties); $H_0^1(\Omega)$ is the closure of $C_0^\infty(\Omega)$ in the H^1-norm; when ℓ is a nonnegative integer, $C_b^\ell(\mathcal{O})$ denotes the set of real functions with have bounded derivatives up to the ℓ-th order.

Let B be an arbitrary Banach space; then the norm in B will be denoted by $\|\cdot\|_B$; the topological dual of B will be denoted by B'; if $K \subset B$ is a cone, the *dual cone* of K is defined by

$$K^* = \{ f \in B' \, : \, f(v) \geq 0 \quad \forall v \in K \}.$$

Let A be a subset of the Banach space B and assume that $v_0 \in A$. It will be said that a nonzero $f \in B'$ is a *support functional* for A at v_0 if $f(v) \geq f(v_0)$ for any $v \in A$.

As usual, we will denote by $L^q(0,T;B)$ the Banach space of the B-valued (classes of) functions $f : [0,T] \mapsto B$ that are L^q-integrable in the sense of Bochner, with the standard norm $\|\cdot\|_{L^q(0,T;B)}$.

The next result, whose proof can be found for instance in [5], will be useful to apply the Dubovitskii-Milyutin formalism:

Proposition 1.1. *Let X and Z be Banach spaces and let U be a neighborhood of $u_0 \in X$. Let $F : U \to Z$ be a mapping and assume that $F(u_0) = 0$, F is strictly differentiable at u_0 and $R(F'(u_0)) = Z$, that is, $F'(u_0)$ is an epimorphism. Then the tangent space to the set $M = \{ v \in U : F(v) = 0 \}$ at u_0 is given by*

$$TC(M; x_0) = N(F'(u_0)) = \{ h \in X \ : \ F'(u_0)h = 0 \}.$$

Next, we introduce some functional spaces which are useful to model the flow of incompressible fluids. First, we set

$$\mathcal{V} = \{ \mathbf{v} \in C_0^\infty(\Omega)^N \ : \ \operatorname{div} \mathbf{v} = 0 \quad \text{in} \quad \Omega \};$$

we will denote by V be the closure of $\mathcal{V}$ in $H_0^1(\Omega)$; on the other hand, the closure of $\mathcal{V}$ in $L^2(\Omega)$ will be denoted by H; P will stand for the usual orthogonal projection from $L^2(\Omega)$ onto H.

The following result, proved by Lorca and Boldrini ([18, lemma 3.4]), provides a suitable estimate for the pressure associated to a Helmholtz decomposition. It will be useful to obtain high order estimates for the fluid velocity:

Proposition 1.2. *Let $\mathbf{v} \in H^2(\Omega) \cap V$ and consider the Helmholtz decomposition of $-\Delta \mathbf{v}$, given by $-\Delta \mathbf{v} = A\mathbf{v} + \nabla q$, where $q \in H^1(\Omega)$ is taken such that $\int_\Omega q \, dx = 0$ and A is the Stokes operator. Then, for every $\epsilon > 0$, there exists a positive constant C_ϵ independent of $\mathbf{v}$ such that*

$$\|q\|_{L^2(\Omega)} \leq C_\epsilon \|\nabla \mathbf{v}\|_{L^2(\Omega)} + \epsilon \|A\mathbf{v}\|_{L^2(\Omega)}.$$

In the sequel, we will denote by $D(A) = H^2(\Omega) \cap V$ the domain of the Stokes operator A. This is a Hilbert space for the usual norm in $H^2(\Omega)$. Recall that, for any $v \in D(A)$, Av is characterized by

$$(Av, w) = \int_\Omega (-\Delta v) \cdot w \quad \forall w \in H, \quad Av \in H.$$

We will also denote by $D(-\Delta) = H^2(\Omega) \cap H_0^1(\Omega)$ the domain of the usual Laplace-Dirichlet operator in Ω.

Now, consider the generalized Boussinesq system (1). The following existence theorem holds:

Proposition 1.3. *Let Ω be a bounded domain in $\mathbb{R}^N$ with a C^3 boundary. Suppose that*

(i) *k is a positive constant,*

(ii) *$\nu \in C^{0,1}(\mathbb{R})$ and $\nu(\sigma) \geq \nu_0 > 0$ for all $\sigma \in \mathbb{R}$,*

(iii) *$\mathbf{g} \in L^\infty(0, T; L^3(\Omega))$,*

(iv) $\mathbf{h} \in L^2(0,T;L^2(\Omega))$,

(v) $f \in L^\infty(0,T;L^2(\Omega))$ *and* $\nabla f \in L^2(0,T;L^2(\Omega))$,

(vi) $\mathbf{u}_0 \in V$,

(vii) $\theta_0 \in D(-\Delta)$.

Then there exists a positive number $T^* \le T$ *such that* (1) *possesses a unique solution* (u,p,θ) *satisfying*

$$\mathbf{u} \in L^\infty(0,T^*;V) \cap L^2(0,T^*;H^2(\Omega)), \quad \mathbf{u}_t \in L^2(0,T^*;L^2(\Omega)),$$
$$\theta \in L^\infty(0,T^*;H^2(\Omega)) \cap L^2(0,T;H^3(\Omega)), \quad \theta_t \in L^\infty(0,T^*;L^2(\Omega)),$$
$$\mathbf{u}(t) \to u_0 \text{ strongly in } V \text{ and } \theta(t) \to \theta_0 \text{ weakly in } H^2(\Omega) \text{ as } t \to 0^+.$$

Furthermore, there exists a constant C*, only depending on* T^**,* k*,* ν_0*,* $\|\nu\|_{C^2_b(\mathbb{R})}$*,* $\|\mathbf{g}\|_{L^\infty(0,T;L^3(\Omega))}$*,* $\|\mathbf{h}\|_{L^2(0,T;L^2(\Omega))}$*,* $\|f\|_{L^\infty(0,T;L^2(\Omega))}$*,* $\|\nabla f\|_{L^2(0,T;L^2(\Omega))}$*,* $\|\mathbf{u}_0\|_V$*, and* $\|\varphi_0\|_{H^2(\Omega)}$*, such that*

$$\|\mathbf{u}\|_{L^\infty(0,T^*;V)} + \|\mathbf{u}\|_{L^2(0,T^*;H^2(\Omega))} + \|\mathbf{u}_t\|_{L^2(0,T^*;L^2(\Omega))} \le C,$$
$$\|\theta\|_{L^\infty(0,T^*;H^2(\Omega))} + \|\theta\|_{L^2(0,T^*;H^3(\Omega))} + \|\theta_t\|_{L^\infty(0,T^*;L^2(\Omega))} \le C.$$

Finally, there exists $\delta > 0$ *such that, when* $\|\mathbf{h}\|_{L^2(0,T;L^2(\Omega))}$*,* $\|f+v\|_{L^\infty(0,T;L^2(\Omega))}$*,* $\|\nabla f+\nabla v\|_{L^2(0,T;L^2(\Omega))}$*,* $\|\mathbf{u}_0\|_V$*, and* $\|\varphi_0\|_{H^2(\Omega)}$ *are less than or equal to* δ*, the solution* $(\mathbf{u},p,\theta)$ *exists globally in time, that is, we can take* $T^* = T$.

This result can be proved arguing as in [6] (see also [17,18]). It suffices to establish suitable uniform estimates for spectral Galerkin approximations of $\mathbf{u}$ and θ. Obviously, the previous conditions are used to prove such estimates; the requirements on ν, for instance, are used to control the extra terms that appear due to the nonlinearity in the higher order term of the first equation; also, the fact that $\partial\Omega$ is of class C^3 is used to guarantee that the H^1_0-norm of the $\Delta\theta$ controls the norm of θ in $H^3 \cap H^1_0$. (Actually, this is used for the semi-Galerkin approximations of θ.) The same sort of conditions will be necessary for the estimates in what follows.

To end this section, we summarize the assumptions on the data we have imposed in Proposition 1.3. They will be assumed throughout this paper (with the exception of section 4).

(H$_1$) Ω is a bounded domain of class C^3 in $\mathbb{R}^N$ ($N = 2$ or $N = 3$);

(H$_2$) k is a positive constant;

(H$_3$) $\nu \in C^{0,1}(\mathbb{R})$ and $\nu(\sigma) \ge \nu_0 > 0$ for all $\sigma \in \mathbb{R}$;

(H$_4$) $\mathbf{g} \in L^\infty(0,T;L^3(\Omega))$;

($\mathbf{H_5}$) $\delta > 0$ is the constant appearing in proposition 1.3;

($\mathbf{H_6}$) $\mathbf{u}_0 \in V$ and $\|\mathbf{u}_0\|_V \leq \delta$;

($\mathbf{H_7}$) $\theta_0 \in H_0^1(\Omega) \cap H^2(\Omega)$ and $\|\theta_0\|_{H^2(\Omega)} \leq \delta$;

($\mathbf{H_8}$) $\mathbf{h} \in L^2(0,T;L^2(\Omega))$ and $\|\mathbf{h}\|_{L^2(0,T;L^2(\Omega))} \leq \delta$;

($\mathbf{H_9}$) $f \in L^\infty(0,T;L^2(\Omega))$, $\nabla f \in L^2(0,T;L^2(\Omega))$, and

$$\|f\|_{L^\infty(0,T;L^2(\Omega))} + \|\nabla f\|_{L^2(0,T;L^2(\Omega))} \leq \delta/2;$$

($\mathbf{H_{10}}$) $\mathbf{u}_d \in L^2(\omega_{\mathbf{u}} \times (0,T))^N$ and $\theta_d \in L^2(\omega_\theta \times (0,T))$.

2. Setting of the problem and existence of optimal solutions

In this section we will define in precise mathematical terms the optimal control problem associated to (1), (2). First, we introduce the following functional spaces:

$$W_{\mathbf{u}} = \{\, \mathbf{w} \in L^2(0,T;D(A)) \,:\, \mathbf{w}_t \in L^2(0,T;L^2(\Omega)) \,\},$$
$$W_\theta = \{\, \phi \in L^2(0,T;H^3(\Omega)) \,:\, \phi_t \in L^2(0,T;H^1(\Omega)), \quad \phi = 0 \text{ on } \Sigma \,\},$$
$$W_c = L^2(0,T;H^1(\Omega)) \cap L^\infty(0,T;L^2(\Omega)).$$

Then $W_{\mathbf{u}}$ is a Hilbert space for the norm

$$\|\mathbf{w}\|_{W_{\mathbf{u}}} = \big(\|\mathbf{w}\|^2_{L^2(0,T;D(A))} + \|\mathbf{w}_t\|^2_{L^2(0,T;L^2(\Omega))}\big)^{1/2}.$$

On the other hand, W_θ is a Hilbert space for the norm

$$\|\phi\|_{W_\theta} = \big(\|\phi\|^2_{L^2(0,T;H^3(\Omega))} + \|\nabla\phi_t\|^2_{L^2(0,T;L^2(\Omega))}\big)^{1/2}$$

and W_c is a Banach space for the norm

$$\|f\|_{W_c} = \|\nabla f\|_{L^2(0,T;L^2(\Omega))} + \|f\|_{L^\infty(0,T;L^2(\Omega))}.$$

In the sequel, we will denote by W the product space

$$W = W_{\mathbf{u}} \times W_\theta \times W_c.$$

Let us also introduce $W_{ic}^{\mathbf{u}} = V$ and $W_{ic}^\theta = H^2(\Omega) \cap H_0^1(\Omega)$ with their natural norms and let us set $W_{ic} = W_{ic}^{\mathbf{u}} \times W_{ic}^\theta$.

Next, we define the set of admissible controls. We set

$$\mathcal{U} = \{\, v \in W_c \,:\, \|v\|_{W_c} \leq \delta/2 \,\}, \tag{3}$$

where δ is furnished by proposition 1.3.

Under hypotheses $(\mathbf{H_1})$–$(\mathbf{H_{10}})$, if the control v belongs to $\mathcal{U}$, (1) possesses exactly one strong solution (since $\|\mathbf{h}\|_{L^\infty(0,T;L^2(\Omega))} \leq \delta$ and $\|f + v\|_{W_c} \leq \delta$).

Let us denote by $\widetilde{W}$ the product space

$$\widetilde{W} = L^2(0,T;H) \times W_c \times W_{ic}$$

and let the mapping $M : W \mapsto \widetilde{W}$ be given by

$$M(\mathbf{w}, \phi, v) = (\psi_1, \psi_2, \psi_3, \psi_4),$$

where $(\psi_1, \psi_2, \psi_3, \psi_4)$ is defined as follows:

$$\begin{cases} \partial_t \mathbf{w} - P(\mathrm{div}(\nu(\phi)\nabla\mathbf{w}) + (\mathbf{w}\cdot\nabla)\mathbf{w} - \alpha\phi\mathbf{g} - \mathbf{h}) = \psi_1, \\ \partial_t\phi - k\Delta\phi + \mathbf{w}\cdot\nabla\phi - f - v = \psi_2, \\ \mathbf{w}|_{t=0} - \mathbf{u}_0 = \psi_3, \\ \phi|_{t=0} - \theta_0 = \psi_4. \end{cases}$$

Notice that M is well defined. The optimal control problem we want to solve is the following: find $(\mathbf{u}, \theta, v) \in \mathcal{Q}$ such that

$$J(\mathbf{u}, \theta, v) = \min_{(\mathbf{w},\psi,\bar{v})\in\mathcal{Q}} J(\mathbf{w}, \psi, \bar{v}), \tag{4}$$

where J is given by (2) and $\mathcal{Q}$ is the non-empty set

$$\mathcal{Q} = \{\, (\mathbf{w}, \psi, \bar{v}) \in W \,:\, \bar{v} \in \mathcal{U},\ M(\mathbf{w}, \psi, \bar{v}) = 0 \,\}.$$

We have:

Theorem 2.1. *Under hypotheses* $(\mathbf{H_1})$–$(\mathbf{H_{10}})$*, problem* (4) *possesses at least one optimal solution.*

Proof. The proof is standard, so we just sketch it. Since $\mathcal{Q}$ is non-empty and $J \geq 0$, we can find a minimizing sequence $\{(\mathbf{u}_n, \theta_n, v_n)\}$ in $\mathcal{Q}$, with

$$\lim_{n\to\infty} J(\mathbf{u}_n, \theta_n, v_n) = \inf\{\, J(\mathbf{w}, \psi, \bar{v}) \,:\, (\mathbf{w}, \psi, \bar{v}) \in \mathcal{Q} \,\}.$$

Since $(\mathbf{u}_n, \theta_n, v_n) \in \mathcal{Q}$, we have $\|f + v_n\|_{W_c} \leq \delta$ for all n. Thus, from proposition 1.3, we see that $\|\mathbf{u}_n\|_{W_\mathbf{u}}$ and $\|\theta_n\|_{W_\theta}$ are also uniformly bounded with respect to n. From this fact and Aubin-Lions' lemma, we conclude that there exist $(\mathbf{u}, \theta, v) \in W_\mathbf{u} \times W_\theta \times \mathcal{U}$ and a subsequence $\{(\mathbf{u}_{n_k}, \theta_{n_k}, v_{n_k})\}$ converging to $(\mathbf{u}, \theta, v)$ in an appropriate sense. This and the convexity and continuity of J, see (2), are enough to conclude that

$$\liminf_{k\to\infty} J(\mathbf{u}_{n_k}, \theta_{n_k}, v_{n_k}) \geq J(\mathbf{u}, \theta, v).$$

Since $M(\mathbf{u}_{n_k}, \theta_{n_k}, v_{n_k}) = 0$ for all k, we can pass to the limit and obtain that $M(\mathbf{u}, \theta, v) = 0$. Thus, $(\mathbf{u}, \theta, v) \in \mathcal{Q}$ and $J(\mathbf{u}, \theta, v) = \inf\{\, J(\mathbf{w}, \psi, \bar{v}) : (\mathbf{w}, \psi, \bar{v}) \in \mathcal{Q} \,\}$.

This ends the proof. □

3. First order optimality conditions and minimum principle

Our main contribution in this paper is the following optimality characterization:

Theorem 3.1. *Assume* $(\mathbf{H_1})$–$(\mathbf{H_{10}})$. *Let* $(\mathbf{u}, \theta, v) \in \mathcal{Q}$ *be an optimal solution of* (4). *Then one has*

$$\begin{cases} \mathbf{u}_t - \ P(\operatorname{div}(\nu(\theta)\nabla\mathbf{u}) + (\mathbf{u}\cdot\nabla)\mathbf{u} - \alpha\theta\mathbf{g}) = \mathbf{h}, \\ \theta_t - \ k\,\Delta\theta + \mathbf{u}\cdot\nabla\theta = f + v, \\ \mathbf{u} = 0, \quad \theta = 0 \quad on \quad \Sigma, \\ \mathbf{u}|_{t=0} = \mathbf{u}_0, \quad \theta|_{t=0} = \theta_0 \end{cases}$$

and there are variables $\mathbf{q} \in L^2(0,T;H)$ *and* $\zeta \in L^2(0,T;L^2(\Omega))$ *that solve in the transposition sense the following adjoint equations:*

$$\begin{cases} -\mathbf{q}_t - P(\operatorname{div}(\nu(\theta)\nabla\mathbf{q}) - (\mathbf{u}\cdot\nabla)\mathbf{q} - \nabla\mathbf{q}^t\cdot\mathbf{u} - \theta\nabla\zeta) \\ \qquad\qquad + P(\alpha_1(\mathbf{u}-\mathbf{u}_d)\chi_{\omega_\mathbf{u}}) = 0, \quad i = 1,\dots,N, \\ -\zeta_t - k\Delta\zeta - \mathbf{u}\cdot\nabla\zeta + \nu'(\theta)\nabla\mathbf{u} : \nabla q + \alpha_2(\theta-\theta_d)\chi_{\omega_\theta} = 0, \\ \mathbf{q} = 0, \quad \zeta = 0 \quad on \quad \Sigma, \\ \mathbf{q}|_{t=T} = 0, \quad \zeta|_{t=T} = 0. \end{cases} \tag{5}$$

Furthermore, the following minimum principle is satisfied:

$$\iint_Q (-\zeta + \mu v)(\bar{v} - v) \le 0 \quad \forall \bar{v} \in \mathcal{U}, \quad v \in \mathcal{U}. \tag{6}$$

In this theorem and in the sequel, for any set G, χ_G denotes the associated characteristic function.

Remark 3.2. It will be seen below that the couple $(\mathbf{q}, \zeta)$ is in fact more regular than stated in this result and solves (5) in a stronger sense; see lemma 3.8.

The proof of this theorem will be obtained by applying the Dubovitskii-Milyutin formalism. To this end, we will need some auxiliary results that will be given now.

Lemma 3.3. *The mapping* M *is* C^1 *in a neighborhood of any point* $(\mathbf{u}, \theta, v) \in W$. *Moreover, its Fréchet derivative* $DM = (DM^{(1)}, DM^{(2)}, DM^{(3)}, DM^{(4)})$ *is given by:*

$$\begin{aligned} DM^{(1)}(\mathbf{u},\theta,v)(\mathbf{w},\phi,\bar{v}) = \mathbf{w}_t &- P\operatorname{div}(\nu(\theta)\nabla\mathbf{w} + \nu'(\theta)\phi\nabla\mathbf{u}) \\ &+ P((\mathbf{u}\cdot\nabla)\mathbf{w} + (\mathbf{w}\cdot\nabla)\mathbf{u} - \alpha\phi\mathbf{g}), \end{aligned} \tag{7}$$

$$DM^{(2)}(\mathbf{u},\theta,v)(\mathbf{w},\phi,\bar{v}) = \phi_t - k\Delta\phi + \mathbf{u}\cdot\nabla\phi + \mathbf{w}\cdot\nabla\theta - \bar{v}, \tag{8}$$

$$DM^{(3)}(\mathbf{u},\theta,v)(\mathbf{w},\phi,\bar{v}) = \mathbf{w}|_{t=0} \tag{9}$$

and

$$DM^{(4)}(\mathbf{u},\theta,v)(\mathbf{w},\phi,\bar{v}) = \phi|_{t=0} \tag{10}$$

for any $(\mathbf{u},\theta,v), (\mathbf{w},\phi,\bar{v}) \in W$.

Proof. First of all, we observe that $DM^{(j)}(\mathbf{u}, \theta, v)$is a bounded linear operator in W for $j = 1, \dots, 4$. Indeed, we have for instance

$$\begin{aligned}
&\|\phi_t - k\Delta\phi + \mathbf{u}\cdot\nabla\phi + \mathbf{w}\cdot\nabla\theta - \bar{v}\|_{L^2(\Omega)} \\
&\qquad \le \|\phi_t\|_{L^2(\Omega)} + k\|\Delta\phi\|_{L^2(\Omega)} + \|\mathbf{u}\|_{L^4(\Omega)}\|\nabla\phi\|_{L^4(\Omega)} \\
&\qquad\quad + \|\mathbf{w}\|_{L^4(\Omega)}\|\nabla\theta\|_{L^4(\Omega)} + \|\bar{v}\|_{L^2(\Omega)} \\
&\qquad \le \|\phi_t\|_{L^2(\Omega)} + k\|\Delta\phi\|_{L^2(\Omega)} + C\|\mathbf{u}\|_V\|\phi\|_{H^2(\Omega)} \\
&\qquad\quad + C\|\mathbf{w}\|_V\|\theta\|_{H^2(\Omega)} + \|\bar{v}\|_{L^2(\Omega)},
\end{aligned}$$

whence

$$\begin{aligned}
&\|DM^{(2)}(\mathbf{u}, \theta, v)(\mathbf{w}, \phi, \bar{v})\|_{L^\infty(0,T;L^2(\Omega))} \\
&\qquad = \|\phi_t - k\Delta\phi + \mathbf{u}\cdot\nabla\phi + \mathbf{w}\cdot\nabla\theta - \bar{v}\|_{L^\infty(0,T;L^2(\Omega))} \\
&\qquad \le \|\phi_t\|_{L^\infty(0,T;L^2(\Omega))} + k\|\Delta\phi\|_{L^\infty(0,T;L^2(\Omega))} \\
&\qquad\quad + C(\|\mathbf{u}\|_{L^\infty(0,T;V)}\|\phi\|_{L^\infty(0,T;H^2(\Omega))} \\
&\qquad\quad + \|\mathbf{w}\|_{L^\infty(0,T;V)}\|\theta\|_{L^\infty(0,T;H^2(\Omega))} + \|\bar{v}\|_{L^\infty(0,T;L^2(\Omega))}) \\
&\qquad \le C(1 + k + \|\mathbf{u}\|_{W_\mathbf{u}} + \|\theta\|_{W_\theta})(\|\mathbf{w}\|_{W_\mathbf{u}} + \|\phi\|_{W_\theta} + \|\bar{v}\|_{W_c}).
\end{aligned}$$

Similarly, we have

$$\begin{aligned}
&\|\nabla(\phi_t - k\Delta\phi + \mathbf{u}\cdot\nabla\phi + \mathbf{w}\cdot\nabla\theta - \bar{v})\|_{L^2(\Omega)} \\
&\qquad \le \|\nabla\phi_t\|_{L^2(\Omega)} + k\|\nabla\Delta\phi\|_{L^2(\Omega)} \\
&\qquad\quad + \|\nabla\mathbf{u}\|_{L^4(\Omega)}\|\nabla\phi\|_{L^4(\Omega)} + \|\mathbf{u}\|_{L^4(\Omega)}\|\nabla^2\phi\|_{L^4(\Omega)} \\
&\qquad\quad + \|\nabla\mathbf{w}\|_{L^4(\Omega)}\|\nabla\theta\|_{L^4(\Omega)} + \|\mathbf{w}\|_{L^4(\Omega)}\|\nabla^2\theta\|_{L^4(\Omega)} + \|\nabla\bar{v}\|_{L^2(\Omega)} \\
&\qquad \le \|\nabla\phi_t\|_{L^2(\Omega)} + k\|\nabla\Delta\phi\|_{L^2(\Omega)} \\
&\qquad\quad + C(\|\mathbf{u}\|_{H^2(\Omega)}\|\phi\|_{H^2(\Omega)} + \|\mathbf{u}\|_V\|\phi\|_{H^3(\Omega)} \\
&\qquad\quad + \|\mathbf{w}\|_{H^2(\Omega)}\|\theta\|_{H^2(\Omega)} + \|\mathbf{w}\|_V\|\theta\|_{H^3(\Omega)} + \|\nabla\bar{v}\|_{L^2(\Omega)}),
\end{aligned}$$

whence

$$\begin{aligned}
&\|\nabla DM^{(2)}(\mathbf{u}, \theta, v)(\mathbf{w}, \phi, \bar{v})\|_{L^2(0,T;L^2(\Omega))} \\
&\qquad = \|\nabla(\phi_t - k\Delta\phi + \mathbf{u}\cdot\nabla\phi + \mathbf{w}\cdot\nabla\theta - \bar{v})\|_{L^2(0,T;L^2(\Omega))} \\
&\qquad \le C(\|\nabla\phi_t\|_{L^2(0,T;L^2(\Omega))} + k\|\phi\|_{L^2(0,T;H^3(\Omega))} \\
&\qquad\quad + \|\mathbf{u}\|_{L^2(0,T;H^2(\Omega))}\|\phi\|_{L^\infty(0,T;H^2(\Omega))} + \|\mathbf{u}\|_{L^\infty(0,T;V)}\|\phi\|_{L^2(0,T;H^3(\Omega))} \\
&\qquad\quad + \|\mathbf{w}\|_{L^2(0,T;H^2(\Omega))}\|\theta\|_{L^\infty(0,T;H^2(\Omega))} + \|\mathbf{w}\|_{L^\infty(0,T;V)}\|\theta\|_{L^2(0,T;H^3(\Omega))} \\
&\qquad\quad + \|\nabla\bar{v}\|_{L^2(\Omega)}) \\
&\qquad \le C(1 + k + \|\mathbf{u}\|_{W_\mathbf{u}} + \|\theta\|_{W_\theta})(\|\mathbf{u}\|_{W_\mathbf{u}} + \|\phi\|_{W_\theta} + \|\bar{v}\|_{W_c}).
\end{aligned}$$

Therefore, $DM^{(2)} : W \mapsto W_c$ is a linear bounded operator. Similar estimates can be obtained for the other $DM^{(j)}$.

Now, we have

$$M^{(1)}(\mathbf{u}+\mathbf{w},\theta+\phi,v+\bar{v}) - M^{(1)}(\mathbf{u},\theta,v) - DM^{(1)}(\mathbf{u},\theta,v)(\mathbf{w},\phi,\bar{v}) \\ = I_1 + I_2 + I_3,$$

where

$$\begin{aligned} I_1 &= -\operatorname{div}((\nu(\theta+\phi) - \nu(\theta) - \nu'(\theta)\,\phi)\nabla\mathbf{u}), \\ I_2 &= -\operatorname{div}((\nu(\theta+\phi) - \nu(\theta))\nabla\mathbf{w}), \\ I_3 &= P((\mathbf{w}\cdot\nabla)\mathbf{w}). \end{aligned}$$

We can estimate the previous terms as follows:

$$\begin{aligned} &\|I_1\|^2_{L^2(0,T;L^2(\Omega))} \\ &\qquad \le 2\iint_Q (|\nu(\theta+\phi) - \nu(\theta) - \nu'(\theta)\,\phi|^2|\Delta\mathbf{u}|^2 \\ &\qquad\quad + |\nabla(\nu(\theta+\phi) - \nu(\theta) - \nu'(\theta)\,\phi)|^2|\nabla\mathbf{u}|^2) \\ &\qquad \le C\iint_Q \big(|\phi|^2|\Delta\mathbf{u}|^2 + (|\phi|^2|\nabla\phi|^2 + |\phi|^2|\nabla\theta|^2 + |\theta|^2|\nabla\phi|^2)|\nabla\mathbf{u}|^2\big) \\ &\qquad \le C\|\phi\|^2_{L^\infty(0,T;L^\infty(\Omega))}\big(1 + \|\phi\|^2_{L^\infty(0,T;W^{1,\infty}(\Omega))} \\ &\qquad\quad + \|\theta\|^2_{L^\infty(0,T;W^{1,\infty}(\Omega))} + \|\phi\|^2_{L^4(0,T;W^{1,\infty}(\Omega))}\|\theta\|^2_{L^4(0,T;W^{1,\infty}(\Omega))}\big)\|\mathbf{u}\|^2_{W_{\mathbf{u}}} \\ &\qquad \le C\|\mathbf{u}\|^2_{W_{\mathbf{u}}}\left(1 + \|\phi\|^2_{W_\theta} + \|\theta\|^2_{W_\theta}\right)\|\phi\|^2_{W_\theta}, \end{aligned}$$

$$\begin{aligned} &\|I_2\|^2_{L^2(0,T;L^2(\Omega))} \\ &\quad \le 2\iint_Q (\|\nu(\theta+\phi) - \nu(\theta)\|^2\|\Delta\mathbf{w}\|^2 + \|\nabla(\nu(\theta+\phi) - \nu(\theta))\|^2\|\nabla\mathbf{w}\|^2) \\ &\quad \le C\iint_Q \big(\|\phi\|^2\|\nabla\mathbf{w}\|^2 + (\|\phi\|^2\|\nabla\phi\|^2 + \|\nabla\theta\|^2)\|\nabla\mathbf{w}\|^2\big) \\ &\quad \le C\big(\|\phi\|^2_{L^2(0,T;L^\infty(\Omega))} + \|\phi\|^4_{L^\infty(0,T;W^{1,\infty}(\Omega))} + \|\theta\|^2_{L^\infty(0,T;W^{1,\infty}(\Omega))}\big)\|\mathbf{w}\|^2_{W_{\mathbf{u}}} \\ &\quad \le C\left(\|\theta\|^2_{W_\theta} + \|\phi\|^2_{W_\theta}(1 + \|\phi\|^2_{W_\theta})\right)\|\mathbf{w}\|^2_{W_{\mathbf{u}}} \end{aligned}$$

and

$$\begin{aligned} \|I_3\|^2_{L^2(0,T;L^2(\Omega))} &\le \iint_Q |\mathbf{w}|^2|\nabla\mathbf{w}|^2 \\ &\le \|\mathbf{w}\|^2_{L^\infty(0,T;L^\infty(\Omega)}\|\mathbf{w}\|^2_{L^2(0,T;V)} \\ &\le C\|\mathbf{w}\|^4_{W_{\mathbf{u}}}. \end{aligned}$$

From these estimates, we conclude that $M^{(1)}$ is Fréchet-differentiable and its derivative is $DM^{(1)}(\mathbf{u},\theta,v)(\mathbf{w},\phi,\bar{v})$.

Next, we consider $M^{(2)}$. We have the following:

$$M^{(2)}(\mathbf{u}+\mathbf{w},\theta+\phi,v+\bar{v}) - M^{(2)}(\mathbf{u},\theta,v) - DM^{(2)}(\mathbf{u},\theta,v)(\mathbf{w},\phi,\bar{v}) = \mathbf{w}\cdot\nabla\phi.$$

But

$$\|\mathbf{w}\cdot\nabla\phi\|_{L^\infty(0,T;L^2(\Omega))} \le C\|\mathbf{w}\|_{L^\infty(0,T;V)}\|\phi\|_{L^\infty(0,T;H^2(\Omega))} \le C\|\mathbf{w}\|_{W_\mathbf{u}}\|\phi\|_{W_\theta}$$

and

$$\begin{aligned}
&\|\nabla(\mathbf{w}\cdot\nabla\phi)\|^2_{L^2(\Omega)} \\
&\qquad \le \iint_Q (|\nabla(\mathbf{w})|^2|\nabla\phi|^2 + |\mathbf{w}|^2|D^2\phi|^2) \\
&\qquad \le C\big(\|\phi\|^2_{L^2(0,T;W^{1,\infty}(\Omega))}|\mathbf{w}|^2_{L^\infty(0,T;V} + \|\phi\|^2_{L^2(0,T;H^2(\Omega))}|\mathbf{w}|^2_{L^\infty(0,T;L^\infty(\Omega))}\big) \\
&\qquad \le C\|\mathbf{w}\|^2_{W_\mathbf{u}}\|\phi\|^2_{W_\theta}.
\end{aligned}$$

Thus, $M^{(2)}$ is Fréchet-differentiable and its derivative is $DM^{(2)}(\mathbf{u},\theta,v)(\mathbf{w},\phi,\bar{v})$.

On the other hand, that $M^{(3)}$ and $M^{(4)}$ are Fréchet-differentiable and their derivatives are given by $DM^{(3)}(\mathbf{u},\theta,v)$ and $DM^{(4)}(\mathbf{u},\theta,v)$ are immediate consequences of linearity and continuity.

We conclude that M is Fréchet-differentiable and its derivative is given by the linear mapping $DM(\mathbf{u},\theta,v)$.

Now, we prove that the mapping $(\mathbf{u},\theta,v)\mapsto DM(\mathbf{u},\theta,v)$ is continuous.

Let $(\mathbf{u},\theta,v)$, $(\mathbf{u}_1,\theta_1,v_1)$, and $(\mathbf{w},\phi,\bar{v})\in W$ be given. Observe that

$$\begin{aligned}
&DM^{(1)}(\mathbf{u}_1,\theta_1,v_1)(\mathbf{w},\phi,\bar{v}) - DM^{(1)}(\mathbf{u},\theta,v)(\mathbf{w},\phi,\bar{v}) \\
&\quad = -P\operatorname{div}\big((\nu(\theta_1)-\nu(\theta))\nabla\mathbf{w} + (\nu'(\theta_1)-\nu'(\theta))\phi\nabla\mathbf{u}_1 + \nu'(\theta)\phi(\nabla\mathbf{u}_1-\nabla\mathbf{u})\big) \\
&\qquad\qquad + P\big(((\mathbf{u}_1-\mathbf{u})\cdot\nabla)\mathbf{w} + (\mathbf{w}\cdot\nabla)(\mathbf{u}_1-\nabla\mathbf{u}_2)\big).
\end{aligned}$$

Consequently, arguing as in the previous estimates, the following is found:

$$\begin{aligned}
&\|DM^{(1)}(\mathbf{u}_1,\theta_1,v_1)(\mathbf{w},\phi,\bar{v}) - DM^{(1)}(\mathbf{u},\theta,v)(\mathbf{w},\phi,\bar{v})\|_{L^2(0,T;H)} \\
&\qquad \le C(\|\theta_1-\theta\|_{W_\theta}\|\mathbf{w}\|_{W_\mathbf{u}} + \|\mathbf{u}_1\|_{W_\mathbf{u}}\|\theta_1-\theta\|_{W_\theta}\|\phi\|_{W_\theta} \\
&\qquad\quad + \|\mathbf{u}_1-\mathbf{u}\|_{W_\mathbf{u}}\|\phi\|_{W_\theta} + \|\mathbf{u}_1-\mathbf{u}\|_{W_\mathbf{u}}\|\phi\|_{W_\theta}) \\
&\qquad \le C(1+\|\mathbf{u}_1\|_{W_\mathbf{u}}t)\|(\mathbf{u}_1,\theta_1,v_1)-(\mathbf{u},\theta,v)\|_W\|(\mathbf{w},\phi,\bar{v})\|_W.
\end{aligned}$$

On the other hand,

$$\begin{aligned}
&DM^{(2)}(\mathbf{u}_1,\theta_1,v_1)(\mathbf{w},\phi,\bar{v}) - DM^{(2)}(\mathbf{u},\theta,v)(\mathbf{w},\phi,\bar{v}) \\
&\qquad\qquad = (\mathbf{u}_1-\mathbf{u})\cdot\nabla\phi + \mathbf{w}\cdot(\nabla\theta_1-\nabla\theta),
\end{aligned}$$

whence

$$\begin{aligned}\|DM^{(2)}(\mathbf{u}_1,\theta_1,v_1)(\mathbf{w},\phi,\bar v) - DM^{(2)}(\mathbf{u},\theta,v)(\mathbf{w},\phi,\bar v)\|_{L^\infty(0,T;L^2(\Omega))}&\\ \le C(\|\mathbf{u}_1-\mathbf{u}\|_{W_\mathbf{u}}\|\phi\|_{W_\theta} + \|\theta_1-\theta\|_{W_\theta}\|\mathbf{w}\|_{W_\theta})&\\ \le C\|(\mathbf{u}_1,\theta_1,v_1)-(\mathbf{u},\theta,v)\|_W\|(\mathbf{w},\phi,\bar v)\|_W&\end{aligned}$$

and, also,

$$\begin{aligned}\|\nabla\big(DM^{(2)}(\mathbf{u}_1,\theta_1,v_1)(\mathbf{w},\phi,\bar v)\big) - \nabla(DM^{(2)}(\mathbf{u},\theta,v)(\mathbf{w},\phi,\bar v))\|_{L^2(0,T;L^2(\Omega))}&\\ \le C(\|\mathbf{u}_1-\mathbf{u}\|_{W_\mathbf{u}}\|\phi\|_{W_\theta} + \|\theta_1-\theta\|_{W_\theta}\|\mathbf{w}\|_{W_\theta})&\\ \le C\|(\mathbf{u}_1,\theta_1,v_1)-(\mathbf{u},\theta,v)\|_W\|(\mathbf{w},\phi,\bar v)\|_W.&\end{aligned}$$

Finally, since estimates of the same kind are trivial for $DM^{(3)}$ and $DM^{(4)}$, we see that

$$\begin{aligned}\|(DM(\mathbf{u}_1,\theta_1,v_1) - DM(\mathbf{u},\theta,v))(\mathbf{w},\phi,\bar v)\|_{\widetilde W}&\\ \le C(1+\|\mathbf{u}_1\|_{W_\mathbf{u}})\|(\mathbf{u}_1,\theta_1,v_1)-(\mathbf{u},\theta,v)\|_W\|(\mathbf{w},\phi,\bar v)\|_W&\end{aligned}$$

for any $(\mathbf{u},\theta,v), (\mathbf{u}_1,\theta_1,v_1), (\mathbf{w},\phi,\bar v) \in W$. (Recall that $\widetilde W = L^2(0,T;H)\times W_c\times W_{ic}$.)

Therefore, DM is continuous and the lemma is proved. □

We also have the following result:

Lemma 3.4. *At any* $(\mathbf{u},\theta,v)\in W$*, the linear operator* $DM(\mathbf{u},\theta,v): W\mapsto \widetilde W$ *given by* (7)–(10) *is onto.*

Proof. We have to prove that, for any $(\mathbf{u},\theta,v)\in W$ and $(\psi_1,\psi_2,\psi_3,\psi_4)\in\widetilde W$, there exists $(\mathbf{w},\phi,\bar v)\in W$ such that

$$\begin{aligned}&\mathbf{w}_t - P\big(\mathrm{div}(\nu(\theta)\nabla\mathbf{w} + \nu'(\theta)\phi\nabla\mathbf{u}) + (\mathbf{u}\cdot\nabla)\mathbf{w} + (\mathbf{w}\cdot\nabla)\mathbf{u}\big) = \psi_1,\\ &\phi_t - k\Delta\phi + \mathbf{u}\cdot\nabla\phi + \mathbf{w}\cdot\nabla\theta - \bar v = \psi_2,\\ &\mathbf{w}|_{t=0} = \psi_3,\\ &\phi|_{t=0} = \psi_4.\end{aligned}\tag{11}$$

We take $\bar v = -\psi_2$ in these equations and proceed to find the corresponding $(\mathbf{w},\phi)$.

The existence of a solution of the previous problem can be deduced in a standard way: we can use the spectral Faedo-Galerkin method, i.e., the Faedo-Galerkin method determined by the eigenfunctions of the Stokes operator A and the eigenfunctions of the Laplace-Dirichlet operator $-\Delta$ as a basis to find $\mathbf{w}$ and ϕ. The local existence in time of the approximate solutions is then a consequence of usual existence results for ordinary differential equations.

Then, we establish *a priori* estimates for these approximate solutions to ensure that they exist globally in time and that an appropriate subsequence converges to a solution of the original equations in the required functional spaces.

Since most arguments to complete the proof are standard, in the sequel we will just establish the needed uniform estimates. To ease the notation, since the formal computations are the same, we will present the estimates working directly on (11). We will pay special attention to some specific points where we have to be careful.

We start by multiplying the first equation by $\mathbf{w}$, integrating over Ω and proceeding as usual to obtain

$$\begin{aligned}\frac{d}{dt}\|\mathbf{w}(t)\|^2_{L^2(\Omega)} &+ \nu_0\|\nabla\mathbf{w}\|^2_{L^2(\Omega)} \\ &\le C\big(\|\mathbf{u}\|^2_{H^2(\Omega)}\|\nabla\phi\|^2_{L^2(\Omega)} + \|\mathbf{u}\|^2_{H^2(\Omega)}\|\mathbf{w}\|^2_{L^2(\Omega)} + \|\psi_1\|^2_{L^2(\Omega)}\big). \quad (12)\end{aligned}$$

Now, we multiply the second equation in (11) by ϕ, we integrate over Ω, and proceed as usual to obtain

$$\frac{d}{dt}\|\phi(t)\|^2_{L^2(\Omega)} + k\|\nabla\phi\|^2_{L^2(\Omega)} \le C\|\theta\|^2_{H^2(\Omega)}\|\nabla\mathbf{w}\|^2_{L^2(\Omega)}. \quad (13)$$

Next, we multiply the second equation in (11) by $-\Delta\phi$. (Recall that we are using the *spectral* Faedo-Galerkin method.) Again, integrating in Ω we deduce that

$$\begin{aligned}\frac{d}{dt}\|\nabla\phi(t)\|^2_{L^2(\Omega)} &+ k\|\Delta\phi\|^2_{L^2(\Omega)} \\ &\le C\big(\|\mathbf{u}\|^2_{H^2(\Omega)}\|\nabla\phi\|^2_{L^2(\Omega)} + \|\theta\|^2_{L^\infty(0,T;H^2(\Omega))}\|\nabla\mathbf{w}\|^2_{L^2(\Omega)}\big). \quad (14)\end{aligned}$$

By adding (13) to (14) and to (12) multiplied by a constant $1/\beta$ such that $\beta\nu_0^{-1}\|\theta\|_{L^\infty(0,T;H^2(\Omega))}$ is sufficiently small, we see that

$$\begin{aligned}\frac{d}{dt}(\|\mathbf{w}(t)\|^2_{L^2(\Omega)} &+ \|\phi(t)\|^2_{L^2(\Omega)} + \|\nabla\phi(t)\|^2_{L^2(\Omega)}) \\ &+ \bar{D}(\|\Delta\mathbf{w}\|^2_{L^2(\Omega)} + \|\nabla\phi\|^2_{L^2(\Omega)} + \|\Delta\phi\|^2_{L^2(\Omega)}) \\ &\le F(\|\mathbf{w}\|^2_{L^2(\Omega)} + \|\phi\|^2_{L^2(\Omega)} + \|\nabla\phi\|^2_{L^2(\Omega)}) + C\|\psi_1\|^2_{L^2(\Omega)},\end{aligned}$$

where $F = C(\|\mathbf{u}\|^2_{H^2(\Omega)} + \|\theta\|^2_{H^2(\Omega))})$.

Since F is integrable, by using Gronwall's inequality in the last inequality, we deduce the following uniform estimates:

$$\begin{gathered}\|\mathbf{w}\|_{L^\infty(0,T;L^2(\Omega))} + \|\nabla\mathbf{w}\|_{L^2(0,T;L^2(\Omega))} \le C, \\ \|\phi\|_{L^\infty(0,T;L^2(\Omega))} + \|\nabla\phi\|_{L^\infty(0,T;L^2(\Omega))} + \|\Delta\phi\|_{L^2(0,T;L^2(\Omega))} \le C.\end{gathered} \quad (15)$$

We now proceed to find higher order estimates. To this end, we recall that we are actually working with spectral approximations.

Since the eigenfunctions are invariant under powers of Δ, $\Delta^2\phi$ belongs to the appropriate approximation subspace.

Recall that we have assumed that $\partial\Omega \in C^3$. We can multiply the second equation in (11) by $\Delta^2\phi$ and integrate over Ω, which leads in the usual way to the inequality

$$\frac{d}{dt}\|\Delta\phi\|^2_{L^2(\Omega)} + k\|\nabla\Delta\phi\|^2_{L^2(\Omega)} \leq C\|\mathbf{u}\|^2_{H^2(\Omega)}\|\Delta\phi\|^2_{L^2(\Omega)} + \tilde{C}\|\theta\|^2_{L^\infty(0,T;H^2(\Omega))}\|A\mathbf{w}\|^2_{L^2(\Omega)}. \quad (16)$$

Here, we have used the well known fact that $\|\mathbf{w}\|_{H^2(\Omega)} \leq C\|A\mathbf{w}\|_{L^2(\Omega)}$.

To obtain higher order estimates for $\mathbf{w}$, we have to be a little more careful. Therefore, we will describe this with more detail. We start by rewriting the first equation in (11) in the form

$$\mathbf{w}_t - P\big(\nu(\theta)\Delta\mathbf{w} + \nu'(\theta)(\nabla\theta\cdot\nabla)\mathbf{w} + \nu''(\theta)\phi(\nabla\theta\cdot\nabla)\mathbf{u} + \nu'(\theta)(\nabla\phi\cdot\nabla)\mathbf{u} + \nu'(\theta)\phi\,\Delta\mathbf{u} + (\mathbf{u}\cdot\nabla)\mathbf{w} + (\mathbf{w}\cdot\nabla)\mathbf{u}\big) = \psi_1.$$

Then, we multiply this equation by $A\mathbf{w}$ and we integrate the resulting identity over Ω. Next, by using the Helmholtz decomposition $-\Delta\mathbf{w} = A\mathbf{w}+\nabla\eta$ for a suitable η, we obtain

$$\begin{aligned}\frac{1}{2}\frac{d}{dt}\|\nabla\mathbf{w}\|^2_{L^2(\Omega)} + \nu_0\|A\mathbf{w}\|^2_{L^2(\Omega)} \leq{}& -\int_\Omega \nu(\theta)\nabla\eta\cdot A\mathbf{w} \\ &+ \int_\Omega \nu'(\theta)(\nabla\theta\cdot\nabla)\mathbf{w}\cdot A\mathbf{w} + \int_\Omega \nu''(\theta)\phi(\nabla\theta\cdot\nabla)\mathbf{u}\cdot A\mathbf{u} \\ &+ \int_\Omega \nu'(\theta)(\nabla\phi\cdot\nabla)\mathbf{u}\cdot A\mathbf{w} + \int_\Omega \nu'(\theta)\phi\Delta\mathbf{u}\cdot A\mathbf{w} \\ &+ \int_\Omega (\mathbf{u}\cdot\nabla)\mathbf{w}\cdot A\mathbf{w} + \int_\Omega (\mathbf{w}\cdot\nabla)\mathbf{u}\cdot A\mathbf{w} \\ &+ \int_\Omega \psi_1\cdot A\mathbf{w}. \end{aligned} \quad (17)$$

We have to estimate each of the terms in the right hand side of the last inequality. We observe that the first one can be written in the form

$$\begin{aligned}\int_\Omega \nu(\theta)\nabla\eta\cdot A\mathbf{w} &= \int_\Omega \nabla(\nu(\theta)\eta)\cdot A\mathbf{w} - \int_\Omega \nu'(\theta)\nabla\theta\,\eta\,A\mathbf{w} \\ &= -\int_\Omega \nu'(\theta)\eta\nabla\theta\cdot A\mathbf{w},\end{aligned}$$

since $A\mathbf{w}$ and $\nabla(\nu(\theta)\eta)$ are orthogonal in L^2. Thus, by using interpolation results, proposition 1.2, and the fact that $\|\eta\|_{H^1(\Omega)} \leq C\|A\mathbf{w}\|_{L^2(\Omega)}$, we deduce the following for any $\epsilon > 0$:

$$\begin{aligned}\left|\int_\Omega \nu(\theta)\nabla\eta\cdot A\mathbf{w}\right| &\le \int_\Omega |\nu'(\theta)||\nabla\theta|\,|\eta|\,|A\mathbf{w}| \\ &\le C\|\nabla\theta\|_{L^4(\Omega)}\|\eta\|_{L^4(\Omega)}\|A\mathbf{w}\|_{L^2(\Omega)} \\ &\le C\|\nabla\theta\|_{L^4(\Omega)}\|\eta\|^{1/4}_{L^2(\Omega)}\|\eta\|^{3/4}_{H^1(\Omega)}\|A\mathbf{w}\|_{L^2(\Omega)} \\ &\le C\|\nabla\theta\|_{L^4(\Omega)}(C_\epsilon\|\mathbf{w}\|_{L^2(\Omega)}+\epsilon\|A\mathbf{w}\|_{L^2(\Omega)})^{1/4}C\|A\mathbf{w}\|^{7/4}_{L^2(\Omega)} \\ &\le C_\epsilon\|\nabla\theta\|_{L^4(\Omega)}\|\mathbf{w}\|^{1/4}_{L^2(\Omega)}\|A\mathbf{w}\|^{7/4}_{L^2(\Omega)}+\epsilon\|\nabla\theta\|_{L^4(\Omega)}\|A\mathbf{w}\|^2_{L^2(\Omega)} \\ &\le C_\epsilon\|\theta\|^8_{L^\infty(0,T;H^2(\Omega))}\|\mathbf{w}\|^2_{L^2(\Omega)} \\ &\quad+\epsilon\|A\mathbf{w}\|^2_{L^2(\Omega)}+\epsilon\|\theta\|_{L^\infty(0,T;H^2(\Omega))}\|A\mathbf{w}\|^2_{L^2(\Omega)}.\end{aligned}$$

From well known interpolation results, we also find that

$$\begin{aligned}\left|\int_\Omega \nu'(\theta)(\nabla\theta\cdot\nabla)\mathbf{w}\cdot A\mathbf{w}\right| &\le C\int_\Omega|\nabla\theta||\nabla\mathbf{w}||A\mathbf{w}| \\ &\le \|\nabla\theta\|_{L^4(\Omega)}\|\nabla\mathbf{w}\|_{L^4(\Omega)}\|A\mathbf{w}\|_{L^2(\Omega)} \\ &\le C_\epsilon\|\theta\|^2_{L^\infty(0,T;H^2(\Omega))}\|\nabla\mathbf{w}\|^2_{L^2(\Omega)}+\epsilon\|A\mathbf{w}\|^2_{L^2(\Omega)}.\end{aligned}$$

Also,

$$\begin{aligned}\left|\int_\Omega \nu''(\theta)\phi(\nabla\theta\cdot\nabla)\mathbf{u}\cdot A\mathbf{u}\right| &\le C\int_\Omega|\nabla\theta|\,|\phi|\,|\nabla\mathbf{u}||A\mathbf{u}| \\ &\le C\|\nabla\theta\|_{L^4(\Omega)}\,\|\phi\|_{L^\infty(\Omega)}\,\|\nabla\mathbf{u}\|_{L^4(\Omega)}\|A\mathbf{u}\|_{L^2(\Omega)} \\ &\le C\|\theta\|_{H^2(\Omega)}\,\|\Delta\phi\|_{L^2(\Omega)}\,\|\mathbf{u}\|_{H^2(\Omega)}\|A\mathbf{u}\|_{L^2(\Omega)} \\ &\le C_\epsilon\|\theta\|^2_{L^\infty(0,T;H^2(\Omega))}\|\mathbf{u}\|^2_{H^2(\Omega)}\,\|\Delta\phi\|^2_{L^2(\Omega)}+\epsilon\|A\mathbf{u}\|^2_{L^2(\Omega)}.\end{aligned}$$

Next,

$$\begin{aligned}\left|\int_\Omega \nu'(\theta)(\nabla\phi\cdot\nabla)\mathbf{u}\cdot A\mathbf{w}\right| &\le C\int_\Omega|\nabla\phi||\nabla\mathbf{u}||A\mathbf{w}| \\ &\le C\|\nabla\phi\|_{L^4(\Omega)}\|\nabla\mathbf{u}\|_{L^4(\Omega)}\|A\mathbf{w}\|_{L^2(\Omega)} \\ &\le C\|\Delta\phi\|_{L^2(\Omega)}\|\mathbf{u}\|_{H^2(\Omega)}\|A\mathbf{w}\|_{L^2(\Omega)} \\ &\le C_\epsilon\|\mathbf{u}\|^2_{H^2(\Omega)}\|\Delta\phi\|^2_{L^2(\Omega)}+\epsilon\|A\mathbf{w}\|^2_{L^2(\Omega)},\end{aligned}$$

$$\begin{aligned}\left|\int_\Omega \nu'(\theta)\phi\Delta\mathbf{u}\cdot A\mathbf{w}\right| &\le C\int_\Omega|\phi||\Delta\mathbf{u}||A\mathbf{w}| \\ &\le C\|\phi\|_{L^\infty(\Omega)}\|\Delta\mathbf{u}\|_{L^2(\Omega)}\|A\mathbf{w}\|_{L^2(\Omega)} \\ &\le C\|\Delta\phi\|_{L^2(\Omega)}\|\mathbf{u}\|_{H^2(\Omega)}\|A\mathbf{w}\|_{L^2(\Omega)} \\ &\le C_\epsilon\|\mathbf{u}\|^2_{H^2(\Omega)}\|\Delta\phi\|^2_{L^2(\Omega)}+\epsilon\|A\mathbf{w}\|^2_{L^2(\Omega)},\end{aligned}$$

$$\begin{aligned}\left|\int_\Omega (\mathbf{u}\cdot\nabla)\mathbf{w}\cdot A\mathbf{w}\right| &\le \int_\Omega |\mathbf{u}||\nabla\mathbf{w}||A\mathbf{w}|\\ &\le \|\mathbf{u}\|_{L^4(\Omega)}\|\nabla\mathbf{w}\|_{L^4(\Omega)}\|A\mathbf{w}\|_{L^2(\Omega)}\\ &\le C\|\mathbf{u}\|_{H^1(\Omega)}\|\nabla\mathbf{w}\|_{L^2(\Omega)}^{1/4}\|A\mathbf{w}\|_{L^2(\Omega)}^{7/4}\\ &\le C_\epsilon\|\mathbf{u}\|^8_{L^\infty(0,T;H^1(\Omega))}\|\nabla\mathbf{w}\|^2_{L^2(\Omega)}+\epsilon\|A\mathbf{w}\|^2_{L^2(\Omega)}\end{aligned}$$

and

$$\begin{aligned}\left|\int_\Omega (\mathbf{w}\cdot\nabla)\mathbf{u}\cdot A\mathbf{w}\right| &\le \int_\Omega |\mathbf{w}||\nabla\mathbf{u}||A\mathbf{w}|\\ &\le \|\mathbf{w}\|_{L^4(\Omega)}\|\nabla\mathbf{u}\|_{L^4(\Omega)}\|A\mathbf{w}\|_{L^2(\Omega)}\\ &\le C\|\nabla\mathbf{w}\|_{L^2(\Omega)}\|\mathbf{u}\|_{H^2(\Omega)}\|A\mathbf{w}\|_{L^2(\Omega)}\\ &\le C_\epsilon\|\mathbf{u}\|^2_{H^2(\Omega)}\|\nabla\mathbf{w}\|^2_{L^2(\Omega)}+\epsilon\|A\mathbf{w}\|^2_{L^2(\Omega)}.\end{aligned}$$

Finally,

$$\left|\int_\Omega \psi_1\cdot A\mathbf{w}\right| \le \|\psi_1\|_{L^2(\Omega)}\|A\mathbf{w}\|_{L^2(\Omega)} \le C_\epsilon\|\psi_1\|^2_{L^2(\Omega)}+\epsilon\|A\mathbf{w}\|^2_{L^2(\Omega)}.$$

From these estimates and (17), choosing ϵ sufficiently small, we obtain

$$\begin{aligned}&\frac{d}{dt}\|\nabla\mathbf{w}\|^2_{L^2(\Omega)}+\nu_0\|A\mathbf{w}\|^2_{L^2(\Omega)}\\ &\qquad\le C\Big(\|\theta\|^8_{L^\infty(0,T;H^2(\Omega))}\|\mathbf{w}\|^2_{L^2(\Omega)}\\ &\qquad\quad+\|\theta\|^2_{L^\infty(0,T;H^2(\Omega))}\big(\|\mathbf{w}\|^2_{L^2(\Omega)}+\|\mathbf{u}\|^2_{H^2(\Omega)}\,\|\Delta\phi\|^2_{L^2(\Omega)}\big)\\ &\qquad\quad+\|\mathbf{u}\|^2_{H^2(\Omega)}\|\Delta\phi\|^2_{L^2(\Omega)}+\|\mathbf{u}\|^8_{L^\infty(0,T;H^1(\Omega))}\|\nabla\mathbf{w}\|^2_{L^2(\Omega)}\\ &\qquad\quad+\|\mathbf{u}\|^2_{H^2(\Omega)}\|\nabla\mathbf{w}\|^2_{L^2(\Omega)}+\|\psi_1\|^2_{L^2(\Omega)}\Big).\end{aligned}$$

Multiplying this last inequality by a constant $1/\beta$ such that $\beta\nu_0^{-1}\|\theta\|^2_{L^\infty(0,T;H^2(\Omega))}$ is small enough, adding the result to (16), simplifying and rearranging the resulting terms, we see that

$$\begin{aligned}\frac{d}{dt}\big(\|\Delta\phi\|^2_{L^2(\Omega)}+\frac{1}{\beta}\|\nabla\mathbf{w}\|^2_{L^2(\Omega)}\big)+C\big(\|\nabla(\Delta\phi)\|^2_{L^2(\Omega)}+\|A\mathbf{w}\|^2_{L^2(\Omega)}\big)&\\ \le G_1\big(\|\Delta\phi\|^2_{L^2(\Omega)}+\|\nabla\mathbf{w}\|^2_{L^2(\Omega)}\big)+G_2,&\qquad(18)\end{aligned}$$

where

$$G_1 = C[(1+\|\theta\|^2_{L^\infty(0,T;H^2(\Omega))})\|\mathbf{u}\|^2_{H^2(\Omega)}+\|\theta\|^8_{L^\infty(0,T;H^2(\Omega))}+\|\mathbf{u}\|^8_{L^\infty(0,T;H^1(\Omega))}]$$

and

$$G_2 = C\|\psi_1\|^2_{L^2(\Omega)}+C\|\theta\|^8_{L^\infty(0,T;H^2(\Omega))}\|\mathbf{w}\|^2_{L^2(\Omega)}.$$

Observe that G_1 and G_2 are positive integrable functions in $(0,T)$, due to the properties of $\mathbf{u}$, θ and ψ and the estimates for $\mathbf{w}$ in (15). Hence, from (18) and Gronwall's lemma, we finally obtain the following estimates:

$$\|\phi\|_{L^\infty(0,T;H^2(\Omega))} + \|\phi\|_{L^2(0,T;H^3(\Omega))} \leq C,$$
$$\|\mathbf{w}\|_{L^\infty(0,T;H^1(\Omega))} + \|\mathbf{w}\|_{L^2(0,T;H^2(\Omega))} \leq C.$$

Using these estimates and the second equation in (11), we easily get that

$$\|\mathbf{w}_t\|_{L^2(0,T;L^2(\Omega))} + \|\phi_t\|_{L^\infty(0,T;L^2(\Omega))} \leq C.$$

As mentioned before, once these estimates are proved for the spectral Faedo-Galerkin approximations, it is easy to pass to the limit and obtain the existence of a strong solution for (11) in W, as desired. (Recall that we fixed $\bar{v} = -\psi_2$.) □

The previous proof of the existence of a solution $(\mathbf{w},\phi)$ starts from the choice $\bar{v} = -\psi_2$. With only very few and minor modifications, w would have obtained he same result starting from another $\bar{v} \in W_c$. Moreover, using the estimates we have established in the previous proof, it is easy to see that, for any fixed $\bar{v}$, the solution is unique.

For future reference, let us emphasize this assertion as follows:

Lemma 3.5. *For any $(\mathbf{u},\theta,v) \in W$, $\bar{v} \in W_c$, and $(\psi_1,\psi_2,\psi_3,\psi_4) \in \widetilde{W}$, there exists an unique $(\mathbf{w},\phi) \in W_{\mathbf{u}} \times W_\theta$ that solves* (11)*. Moreover, such $(\mathbf{w},\phi)$ satisfies the following estimates:*

$$\begin{aligned}\|\mathbf{w}\|_{L^\infty(0,T;H)} + \|\mathbf{w}\|_{L^2(0,T;V)} + \|\phi\|_{L^\infty(0,T;H_0^1(\Omega))} + \|\phi\|_{L^2(0,T;H^2(\Omega))}\\ \leq C(\|\bar{v}\|_{L^2(0,T;L^2(\Omega))} + \|\psi_1\|_{L^2(0,T,H)} + \|\psi_2\|_{L^2(0,T;L^2(\Omega))}\\ + \|\psi_3\|_{W_{ic\mathbf{u}}} + \|\psi_4\|_{W_{ic\theta}}),\end{aligned} \quad (19)$$

$$\begin{aligned}\|\mathbf{w}\|_{W_{\mathbf{u}}} + \|\phi\|_{W_\theta}\\ \leq C(\|\bar{v}\|_{W_c} + \|\psi_1\|_{L^2(0,T;H)} + \|\psi_2\|_{W_c} + \|\psi_3\|_{W_{ic\mathbf{u}}} + \|\psi_4\|_{W_{ic\theta}}).\end{aligned}$$

The next result is immediate:

Lemma 3.6. *The functional $J : W \mapsto \mathbb{R}$ is C^1 and its derivative at $(\mathbf{u},\theta,v)$ in the direction $(\mathbf{w},\phi,\bar{v})$ is given by*

$$DJ(\mathbf{u},\theta,v)(\mathbf{w},\phi,\bar{v}) = \alpha_1 \iint_{\omega_{\mathbf{u}}\times(0,T)} (\mathbf{u}-\mathbf{u}_d)\cdot\mathbf{w} + \alpha_2 \iint_{\omega_\theta\times(0,T)} (\theta-\theta_d)\phi + \mu \iint_Q v\bar{v}.$$

We will also need the following lemma:

Lemma 3.7. *For any $\mathbf{u} \in W_{\mathbf{u}}$ and $\theta \in W_\theta$, there exists a unique solution of* (5) *in the sense of transposition. In other words, there exists exactly one $(q, \zeta) \in L^2(0,T;H) \times L^2(0,T;L^2(\Omega))$ such that*

$$\iint_Q (q \cdot \psi_1 + \zeta \cdot \psi_2) = -\alpha_1 \iint_{\omega_{\mathbf{u}} \times (0,T)} (\mathbf{u} - \mathbf{u}_d) \cdot \mathbf{w} - \alpha_2 \iint_{\omega_\theta \times (0,T)} (\theta - \theta_d)\phi \tag{20}$$

for any $\psi_1 \in L^2(0,T;H)$ and $\psi_2 \in W_c$, where $(\mathbf{w}, \phi)$ is the unique couple in $W_{\mathbf{u}} \times W_\theta$ satisfying

$$\begin{cases} \mathbf{w}_t - P\left(\operatorname{div}(\nu(\theta)\nabla \mathbf{w} + \nu'(\theta)\phi\nabla\mathbf{u}) + (\mathbf{u}\cdot\nabla)\mathbf{w} + (\mathbf{w}\cdot\nabla)\mathbf{u}\right) = \psi_1, \\ \phi_t - k\Delta\phi + \mathbf{u}\cdot\nabla\phi + \mathbf{w}\cdot\nabla\theta = \psi_2, \\ \mathbf{w}|_{t=0} = 0, \\ \phi|_{t=0} = 0. \end{cases}$$

Proof. It will suffice to prove that the linear functional $\ell : L^2(0,T;H) \times W_c \mapsto \mathbb{R}$ defined by the right hand side of (20) is continuous. But, in accordance with (19), we have

$$\begin{aligned} |\ell(\psi_1,\psi_2)| &\le \alpha_1(\|\mathbf{u}\|_{L^2(0,T;H)} + \|\mathbf{u}_d\|_{L^2(0,T;L^2(\omega_{\mathbf{u}}))})\|\mathbf{w}\|_{L^2(0,T;H)} \\ &\quad + \alpha_2(\|\theta\|_{L^2(0,T;L^2(\Omega)} + \|\theta_d\|_{L^2(0,T;L^2(\omega_\theta))})\|\phi\|_{L^2(0,T;L^2(\Omega))} \\ &\le C(\|\mathbf{u}\|_{L^2(0,T;H)} + \|\mathbf{u}_d\|_{L^2(0,T;L^2(\omega_{\mathbf{u}}))})\|\psi_1\|_{L^2(0,T;H)} \\ &\quad + C(\|\theta\|_{L^2(0,T;L^2(\Omega)} + \|\theta_d\|_{L^2(0,T;L^2(\omega_\theta))})\|\psi_2\|_{L^2(0,T;L^2(\Omega))}. \end{aligned}$$

Thus, Riesz's theorem guarantees the existence of a unique (q, ζ) in $L^2(0,T;H) \times L^2(0,T;L^2(\Omega))$ satisfying (20). □

In fact, the solution by transposition of (5) is more regular, as we now show:

Lemma 3.8. *For any $\mathbf{u} \in W_{\mathbf{u}}$ and $\theta \in W_\theta$, the couple (q, ζ) furnished by lemma 3.7 satisfies $q \in L^\infty(0,T;V) \cap L^2(0,T;H^2(\Omega)^N)$, $\zeta \in L^\infty(0,T;L^2(\Omega)) \cap L^2(0,T;H^1(\Omega))$ and solves* (5) *in the following sense:*

$$\begin{aligned} \iint_Q q\cdot\mathbf{w}_t + \iint_Q \big(\nabla q \cdot \nabla \mathbf{w} + \nu'(\theta)\phi(\nabla q\cdot\nabla)\mathbf{u}\big) \\ + \iint_Q q \cdot \big((\mathbf{u}\cdot\nabla)\mathbf{w} + (\mathbf{w}\cdot\nabla)\mathbf{u}\big) \\ + \iint_Q \zeta\phi_t + k\iint_Q \nabla\zeta\cdot\nabla\phi + \iint_Q \zeta(\mathbf{u}\cdot\nabla\phi + \mathbf{w}\cdot\nabla\theta) \\ = -\alpha_1 \iint_{\omega_{\mathbf{u}}\times(0,T)} (\mathbf{u}-\mathbf{u}_d)\cdot\mathbf{w} - \alpha_2 \iint_{\omega_\theta\times(0,T)} (\theta-\theta_d)\phi, \end{aligned} \tag{21}$$

for any $\mathbf{w} \in L^2(0,T;V)$ such that $\mathbf{w}_t \in L^2(0,T;L^2(\Omega))$ and $\mathbf{w}|_{t=0} = 0$ and any $\phi \in L^2(0,T;H_0^1(\Omega)$ such that $\phi_t \in L^2(0,T;L^2(\Omega)$ and $\phi|_{t=0} = 0$.

Proof. As in the proof of lemma 3.4, we first introduce a spectral Faedo-Galerkin method. As there, the local in time existence of approximate solutions is a consequence of standard existence results for ordinary differential equations. Then, one proceeds to find *a priori* estimates for these approximations to ensure that they exist globally in time and we have, at least for a subsequence, convergence to a solution.

Again, since the arguments to complete the proof are standard, in the sequel we just present the necessary estimates. Since the computations are formally the same, we will show these estimates working directly with (5).

We start as follows.

For each i, we multiply the first equation in (5) by $q^{(i)}$, we add the resulting equations and we argue as usual, which yields the following for each $\epsilon > 0$:

$$-\frac{d}{dt}\|q\|^2_{L^2(\Omega)} + \nu_0\|\nabla q\|^2_{L^2(\Omega)} \leq C\big(\|\mathbf{u}\|^2_{H^2(\Omega)}\|q\|^2_{L^2(\Omega)} + \|\mathbf{u}\|^2_{L^2(\Omega)} + \|\mathbf{u}_d\|^2_{L^2(\omega_u)}\big) + C_\epsilon\|\theta\|^2_{H^2(\Omega)}\|q\|_{L^2(\Omega)} + \epsilon\|\nabla\zeta\|^2_{L^2(\Omega)}. \quad (22)$$

Now, we multiply the second equation in (5) by ζ and proceed as before, to obtain

$$-\frac{d}{dt}\|\zeta\|^2_{L^2(\Omega)} + k\|\nabla\zeta\|^2_{L^2(\Omega)} \leq C\big(\|\mathbf{u}\|^2_{H^2(\Omega)}\|\nabla q\|^2_{L^2(\Omega)} + \|\theta\|^2_{L^2(\Omega)} + \|\theta_d\|^2_{L^2(\omega_{\mathbf{u}})}\big). \quad (23)$$

Next, for each i we multiply the first equation in (5) by $-(Aq)^{(i)}$ (the i-th component of $-Aq$) and we add the resulting identities. We use the Helmholtz decomposition to write $-\Delta q = Aq + \nabla\bar{\eta}$ for a suitable $\bar{\eta}$ and then we proceed as in lemma 3.4 to estimate the terms with this "artificial" pressure $\bar{\eta}$, using proposition 1.2. After some work, using interpolation to estimate all the appearing terms, we find

$$-\frac{d}{dt}\|\nabla q\|^2_{L^2(\Omega)} + \nu_0\|Aq\|^2_{L^2(\Omega)} \leq C\Big(\big(\|\theta\|^8_{L^\infty(0,T;H^2(\Omega))} + \|\mathbf{u}\|^8_{L^\infty(0,T;H^1(\Omega))}\big)\|q\|^2_{L^2(\Omega)} + \|\mathbf{u}\|^2_{L^2(\Omega)} + \|\mathbf{u}_d\|^2_{L^2(\omega_u)} + \|\theta\|^2_{L^\infty(0,T;H^2(\Omega))}\|\nabla\zeta\|_{L^2(\Omega)}\Big). \quad (24)$$

By adding (22), (24), and (23) multiplied by a sufficiently large constant, after some simplification and grouping, we obtain

$$\begin{aligned}
-\frac{d}{dt}&\big(\|q\|^2_{L^2(\Omega)} + \|\zeta\|^2_{L^2(\Omega)} + \|\nabla q\|^2_{L^2(\Omega)}\big) \\
&+ \nu_0\|\nabla q\|^2_{L^2(\Omega)} + C_1\|\nabla\zeta\|^2_{L^2(\Omega)} + \nu_0\|Aq\|^2_{L^2(\Omega)} \\
&\leq C\Big(\|\mathbf{u}\|^2_{H^2(\Omega)}\|\nabla q\|^2_{L^2(\Omega)} \\
&\quad + \big(\|\theta\|^8_{L^\infty(0,T;H^2(\Omega))} + \|\mathbf{u}\|^8_{L^\infty(0,T;H^1(\Omega))} + \|\mathbf{u}\|^2_{H^2(\Omega)}\big)\|q\|^2_{L^2(\Omega)} \\
&\quad + \|\mathbf{u}\|^2_{L^2(\Omega)} + \|u_d\|^2_{L^2(\omega_{\mathbf{u}})} + \|\theta\|^2_{L^2(\Omega)} + \|\theta_d\|^2_{L^2(\omega_{\mathbf{u}})}\Big) \\
&\quad + C_\epsilon\|\theta\|^2_{H^2(\Omega)}\|q\|_{L^2(\Omega)} + \epsilon\|\nabla\zeta\|^2_{L^2(\Omega)}.
\end{aligned}$$

By taking $\epsilon > 0$ sufficiently small, we finally obtain

$$
\begin{aligned}
-\frac{d}{dt}\big(\|q\|^2_{L^2(\Omega)} + \|\zeta\|^2_{L^2(\Omega)} + \|\nabla q\|^2_{L^2(\Omega)}\big)& \\
+ C\big(\|\nabla q\|^2_{L^2(\Omega)} + \|\nabla \zeta\|^2_{L^2(\Omega)} + \|Aq\|^2_{L^2(\Omega)}\big)& \\
\leq G_3\big(\|q\|^2_{L^2(\Omega)} + \|\zeta\|^2_{L^2(\Omega)} + \|\nabla q\|^2_{L^2(\Omega)}\big) + G_4&,
\end{aligned}
$$

where

$$G_3 = C(\|\theta\|^8_{L^\infty(0,T;H^2(\Omega))} + \|\theta\|^2_{H^2(\Omega)} + \|\mathbf{u}\|^8_{L^\infty(0,T;H^1(\Omega))} + \|\mathbf{u}\|^2_{H^2(\Omega)})$$

and

$$G_4 = C(\|u\|^2_{L^2(\Omega)} + \|\mathbf{u}_d\|^2_{L^2(\omega_\mathbf{u})} + \|\theta\|^2_{L^2(\Omega)} + \|\theta_d\|^2_{L^2(\omega_\mathbf{u})}).$$

With the help of Gronwall's inequality, we now obtain the following estimates for the spectral approximations (q_n, ζ_n), where C is of course independent of n:

$$
\begin{aligned}
\|q_n\|_{L^\infty(0,T;,V)} + \|q_n\|_{L^2(0,T;,H^2(\Omega))} \leq C,\\
\|\zeta_n\|_{L^\infty(0,T;,L^2(\Omega))} + \|\zeta_n\|_{L^2(0,T;,H^1(\Omega))} \leq C.
\end{aligned}
\tag{25}
$$

From (25), it is easy to obtain estimates for $q_{n,t}$ and $\zeta_{n,t}$ in suitable dual spaces. Using this, again (25) and the well known Aubin-Lions' lemma, we can extract a subsequence of (q_n, ζ_n) converging in a suitable sense to (q, ζ) and we can pass to the limit to deduce that (21) is satisfied by (q, ζ). □

Next, we describe the formalism of Dubovitskii and Milyutin as applied to our specific problem.

First of all, we associate to any $(\mathbf{u}, \theta, v) \in W$ the cone of decreasing directions of the functional J:

$$
\begin{aligned}
DC(J; \mathbf{u}, \theta, v) = \{\, (\mathbf{w}, \phi, \bar{v}) \in W \,:\, \exists \epsilon > 0 \quad \text{such that}& \\
J(\mathbf{u}, \theta, v) + \lambda(\mathbf{w}, \phi, \bar{v}) < J(\mathbf{u}, \theta, v) \quad \forall \lambda \in (0, \epsilon] \,\}&.
\end{aligned}
$$

In view of this definition and lemma 3.6, we have:

Lemma 3.9. *The cone of decreasing directions of J at $(\mathbf{u}, \theta, v)$ is given by*

$$DC(J; \mathbf{u}, \theta, v) = \{\, (\mathbf{w}, \phi, \bar{v}) \in W \,:\, DJ(\mathbf{u}, \theta, v)(\mathbf{w}, \phi, \bar{v}) < 0 \,\}.$$

The corresponding dual cone is

$$[DC(J; \mathbf{u}, \theta, v)]^* = \{\, -\lambda DJ(\mathbf{u}, \theta, v) \,:\, \lambda \geq 0 \,\}.$$

Now, we introduce the cone of feasible directions of $\mathcal{U}$ at $(\mathbf{u}, \theta, v)$. By definition, it is given by

$$
\begin{aligned}
FC(\mathcal{U}; \mathbf{u}, \theta, v) = \{\, (\mathbf{w}, \phi, \bar{v}) \in W : \exists \epsilon > 0 \quad \text{such that}& \\
(\mathbf{u}, \theta, v) + \lambda(\mathbf{w}, \phi, \bar{v}) \in W_\mathbf{u} \times W_\theta \times \mathcal{U} \quad \forall \lambda \in (0, \epsilon] \,\}&.
\end{aligned}
$$

Since $\mathcal{U}$ is a convex set with nonempty interior, we have:

Lemma 3.10. *The cone of feasible directions of $\mathcal{U}$ at $(\mathbf{u}, \theta, v)$ is given by*

$$FC(\mathcal{U}; \mathbf{u}, \theta, v) = W_{\mathbf{u}} \times W_\theta \times \{\lambda(\bar{v} - v) \,:\, \bar{v} \in \operatorname{int}\mathcal{U},\ \lambda > 0\}.$$

Its dual cone is given by

$$[FC(\mathcal{U}; \mathbf{u}, \theta, v)]^* = \{\, (0, 0, h) \,:\, h \in W_c' \text{ is a support functional for } \mathcal{U} \text{ at } v \,\}.$$

Finally, let us consider the cone of tangent directions of $\mathcal{M}$ at $(\mathbf{u}, \theta, v)$, where

$$\mathcal{M} = \{\, (\mathbf{u}, \theta, v) \in W \,:\, M(\mathbf{u}, \theta, v) = 0 \,\}.$$

This is defined as follows:

$$\begin{aligned} TC(\mathcal{M}; \mathbf{u}, \theta, v) = \Big\{ & (\mathbf{w}, \phi, \bar{v}) \in W \,:\, \exists \lambda_n, (\mathbf{u}_n, \theta_n, v_n) \\ & \text{for } n = 1, 2, \dots, \text{ with } \lambda_n \to 0+,\ (\mathbf{u}_n, \theta_n, v_n) \in \mathcal{M}, \\ & \lim_{n \to +\infty} \frac{1}{\lambda_n} [(\mathbf{u}_n, \theta_n, v_n) - (\mathbf{u}, \theta, v)] = (\mathbf{w}, \phi, \bar{v}) \Big\}. \end{aligned}$$

From lemma 3.3, we know that M is a C^1 mapping (and in particular is strictly differentiable). Lemma 3.4 guarantees that $DM(\mathbf{u}, \theta, v)$ is onto for each $(\mathbf{u}, \theta, v) \in W$. Thus, from Lyusternik's theorem (theorem 1.1) we obtain the following:

Lemma 3.11. *The cone of tangent directions of $\mathcal{M}$ at $(\mathbf{u}, \theta, v)$ is given by*

$$TC(\mathcal{M}; \mathbf{u}, \theta, v) = \{\, (\mathbf{w}, \phi, \bar{v}) \in W \,:\, DM(\mathbf{u}, \theta, v)(\mathbf{w}, \phi, \bar{v}) = 0 \,\}.$$

Consequently, if f belongs to the dual cone $[TC(\mathcal{M}; \mathbf{u}, \theta, v)]^*$, then $f(\mathbf{w}, \phi, \bar{v}) = 0$ for any $(\mathbf{w}, \phi, \bar{v}) \in W$ such that $DM(\mathbf{u}, \theta, v)(\mathbf{w}, \phi, \bar{v}) = 0$.

3.1. Proof of theorem 3.1

Let $(\mathbf{u}, \theta, v)$ be an optimal solution of problem (4). Then, from the results implied by the Dubovitskii-Milyutin formalism, we know that

$$DC(J; \mathbf{u}, \theta, v) \cap FC(\mathcal{U}; \mathbf{u}, \theta, v) \cap TC(\mathcal{M}; \mathbf{u}, \theta, v) = \emptyset;$$

see Girsanov [10], Flett [9].

Accordingly, here must exist $f_1 \in [DC(J; \mathbf{u}, \theta, v)]^*$, $f_2 \in [FC(\mathcal{U}; \mathbf{u}, \theta, v)]^*$ and $f_3 \in [TC(\mathcal{M}; \mathbf{u}, \theta, v))]^*$, not simultaneously zero, such that the following Euler-Lagrange equation is satisfied:

$$f_1 + f_2 + f_3 = 0.$$

Now, let us choose $\bar{v} \in W_c$ arbitrarily and let $(\mathbf{w}, \phi) \in W_{\mathbf{u}} \times W_\theta$ be the unique strong solution of

$$\begin{cases} \mathbf{w}_t - P\big(\operatorname{div}(\nu(\theta)\nabla\mathbf{w} + \nu'(\theta)\phi\nabla\mathbf{u}) + (\mathbf{u}\cdot\nabla)\mathbf{w} + (\mathbf{w}\cdot\nabla)\mathbf{u}\big) = 0, \\ \phi_t - k\Delta\phi + \mathbf{u}\cdot\nabla\phi + w\cdot\nabla\theta = \bar{v}, \\ \mathbf{w}|_{t=0} = 0, \\ \phi|_{t=0} = 0, \end{cases} \tag{26}$$

furnished by corollary 3.5.

Then $(\mathbf{w}, \phi, \bar{v}) \in TC(\mathcal{M}; \mathbf{u}, \theta, v)$ and, in view of lemma 3.11, $f_3(\mathbf{w}, \phi, \bar{v}) = 0$. Therefore, the Euler-Lagrange equation implies

$$(f_1 + f_2)(\mathbf{w}, \phi, \bar{v}) = 0 \tag{27}$$

From lemmas 3.9 and 3.10, we know that

$$f_1(\mathbf{w}, \phi, \bar{v}) = -\lambda DJ(\mathbf{u}, \theta, v)(\mathbf{w}, \theta, \bar{v}) \tag{28}$$

for some $\lambda \geq 0$ and

$$f_2(\mathbf{w}, \phi, \bar{v}) = h(\bar{v}) \tag{29}$$

for some $h \in W_c'$.

Observe that λ cannot be zero. Otherwise, we would have $f_1 = 0$ and, from (27), we would conclude that $f_2(\mathbf{w}, \phi, \bar{v}) = h(\bar{v}) = 0$; since $\bar{v} \in W_c$ was arbitrary, we would have $h = 0$ and also $f_2 = 0$, in contradiction with Dubovistskii-Milyutin theorem.

Thus, we must have $\lambda > 0$ and, without loss of generality, we can assume that $\lambda = 1$.

Now, from (27), (28) with $\lambda = 1$, (29), and lemma 3.6, we deduce that

$$h(\bar{v}) = \alpha_1 \iint_{\omega_{\mathbf{u}}\times(0,T)} (\mathbf{u} - \mathbf{u}_d)\cdot\mathbf{w} + \alpha_2 \iint_{\omega_\theta\times(0,T)} (\theta - \theta_d)\phi + \mu \iint_Q u\bar{v}, \tag{30}$$

where we recall that $(\mathbf{w}, \phi)$ is the solution of (26) associated to $\bar{v}$.

Let (q, ζ) be the solution of the adjoint system (5), which exists by lemma 3.7. Let us take $\psi_1 = 0$ and $\psi_2 = -\bar{v}$ in (5). Then

$$\iint_Q \zeta\bar{v} = -\alpha_1 \iint_{\omega_{\mathbf{u}}\times(0,T)} (\mathbf{u} - \mathbf{u}_d)\cdot\mathbf{w} - \alpha_2 \iint_{\omega_\theta\times(0,T)} (\theta - \theta_d)\phi.$$

From this identity and (30), we find that

$$h(\bar{v}) = \iint_Q (-\zeta + \mu v)\bar{v}.$$

Finally, taking into account that $\bar{v} \in W_c$ was arbitrary and h is a support functional for $\mathcal{U}$, we obtain (6).

This ends the proof. □

4. Some additional remarks

4.1. Other cost functionals

The problem considered in the previous sections can be generalized in several ways. For instance, we can consider locally supported in space controls, that is, controls that act on the system only through a small part ω_c of Ω. In this case, the previous results hold with some obvious modifications: we have again the existence of an optimal control, as well as optimality conditions similar to those in theorem 3.1.

The modifications are just the following: where v appears, replace it by $v\chi_{\omega_c}$; the associated minimum principle (6) is replaced by

$$\iint_Q (-\zeta + \mu v)(\bar{v} - v) \leq 0 \quad \forall \bar{v} \in \mathcal{U}_\ell, \quad v \in \mathcal{U}_\ell,$$

where

$$\mathcal{U}_\ell = \{\, v \in W_{c\ell} \, : \, \|v\|_{W_{c\ell}} \leq \delta/2 \,\}$$

and

$$W_{c\ell} = L^2(0,T; H_0^1(\omega_c)) \cap L^\infty(0,T; L^2(\omega_c)).$$

Results of the same kind also hold for (1) together with other cost functionals (again, we can also assume here that the controls are localized in space). For instance, this is the case for

$$\widetilde{J}(\mathbf{u},\theta,v) = \frac{\alpha}{2}\int_{\omega_{\mathbf{u}}} |\mathbf{u}(T) - \mathbf{u}_d(T)|^2 + \frac{\beta}{2}\int_{\omega_\theta} |\theta(T) - \theta_d(T)|^2 + \frac{\mu}{2}\iint_{\omega_c \times (0,T)} |v|^2.$$

On the other hand, notice that the smallness condition in the definition of $\mathcal{U}$ can be replaced by another assumption imposing the smallness of T. In fact, given $R > 0$ (not necessarily small), the set of admissible controls can be $\mathcal{U}_R = \{\, v \in W_c \, : \, \|v\|_{W_c} \leq R \,\}$. Then the previous results for the associated optimal problem hold whenever $T \leq T^*$, where T^* (which depends on R, $\|f\|_{W_c}$, $\|\mathbf{h}\|_{L^2(0,T;L^2(\Omega))}$, $\|\mathbf{u}_0\|_{W_{ic\mathbf{u}}}$ and $\|\theta_0\|_{W_{ic\theta}}$) is the time for which the solutions exist.

4.2. Other state equations

In the case of nonhomogeneous boundary conditions, a similar analysis applies with appropriate regularity and smallness conditions.

For more general Boussinesq models (for instance, with temperature-depending thermal conductivity $k = k(\theta)$) and Dirichlet boundary conditions as in (1), related control problems can be considered. The situation is unclear. Indeed, the local existence of a strong solution (i.e., a result like proposition 1.3) is needed, but this does not seem obvious. (Nevertheless, when the boundary conditions are of the Neumann kind, the situation is hopeful. Indeed, in this case, the existence result in [6] can be used; this is presently under investigation.)

Very probably, similar results are also true in the context of boundary controls.

4.3. Weak versus strong solutions

One may wonder why not to consider weak solutions of the generalized Boussinesq model instead of strong ones. This seems natural since the existence of weak solutions requires weaker conditions than the strong ones.

One may conceive that one of the difficulties in doing so is the lack of uniqueness of weak solutions in the three-dimensional case. We will say more about the uniqueness of the control-to-state mapping latter on, but now we want to stress another point. From the theoretical perspective of the methodology we used in this work, such uniqueness may not be an essential issue; in principle, the Dubovitskii-Milyoutin formalism can handle situations when the control-to-state mapping is not single-valued, or when the set of admissible control is not convex, etc. if one is capable of computing the required cones and dual cones.

In the present problem, we think that one of the main difficulties in using weak solutions comes from the fact that, whatever the methodology used to find the optimality conditions is, one always ends up with a linearized partial differential system with coefficients that may depend on the possible optimal solutions. And in our case, differently from the Navier-Stokes or classical Boussineq equations, when one is considering just weak solutions, the nonlinear behavior of the viscosity introduces in this system a term having a coefficient whose regularity is only L^1. Thus, although linear, such system is very difficult to handle in a rigorous way and we do not know at present whether results similar to theorems 2.1 and 3.1 hold when we consider *weak* (and not necessarily strong) solutions.

On the other hand, in the case of strong solutions, such coefficient has better regularity, and the linearized system can be handled. This was the main reason we had to work with strong solutions.

Unfortunately, the known results concerning existence of strong solutions for the generalized Boussinesq model require more regularity of the forcing terms besides being just in a certain L^p-space, and such extra regularity implies the uniqueness of such strong solutions as well. And for global existence they also require certain smallness assumptions like those imposed in proposition 1.3.

Since our control acts as a forcing term, in order to guarantee that the associated solution (when it exists) is strong, we had to impose regularity conditions on the space of controls. To guarantee existence, one option is to impose smallness conditions as in proposition 1.3 on the control set, and this was our choice in this work.

4.4. Small and large controls

We can adopt another viewpoint, considering the possibility of handling large controls and trying to explore the fact that an optimal solution must be special in some sense, since it minimizes a cost functional. This is usually done by considering a minimizing sequence, as in the proof of theorem 2.1 and using additional estimates for this sequence that come from the functional itself. In this way, one tries to get rid off the previous smallness conditions on the controls. However, this is not possible with the functional that we considered in this work.

One could consider, for instance, different functionals having terms that give information on L^r-norms, with r large enough, of the state or the control variables. We will come again to this possibility shortly, when we consider the issue of uniqueness. However, in our case this does not appear to lead to better results. Other possibilities with better chances of success rely on cost functionals including terms with information on certain L^r-norms of the gradient of state or the control variables. Such an additional information on a minimizing sequence could be used to improve the estimates and this could lead to the existence of an optimal solution associated to a not necessarily small control. These changes in the functional would lead to different optimality conditions, which are presently under investigation.

4.5. Some comments on uniqueness

In our analysis, we have used hypotheses $(\mathbf{H_5})$–$(\mathbf{H_9})$ to ensure the uniqueness of the state associated to a control. Thus, the control set was given by (3). This makes the control-to-state mapping $v \mapsto (\mathbf{u}, \theta)$ single-valued.

In this way we are constrained to work with small controls.

In the framework of the Navier-Stokes or classical Boussinesq system, there are other ways to ensure uniqueness. For instance, let us assume that $\nu(\sigma) \equiv \nu_0$ and, instead of (2), the cost functional is given by

$$K(\mathbf{u}, \theta, v) = \frac{\alpha_1}{2}\left(\iint_{\omega_{\mathbf{u}}\times(0,T)} |\mathbf{u} - \mathbf{u}_d|^s\right)^{1/r} + \frac{\alpha_2}{2}\iint_{\omega_\theta\times(0,T)} |\theta - \theta_d|^2 + \frac{\mu}{2}\iint_Q |v|^2,$$

where

$$\frac{2}{r} + \frac{3}{s} \leq 1, \quad s > 3.$$

Then a control v provides at most one state $(\mathbf{u}, \theta)$ such that $K(\mathbf{u}, \theta, v) < +\infty$; for more details, see for instance [3].

Unfortunately, the situation is less favorable for the generalized Boussinesq system we have considered in this paper. Indeed, to our knowledge, it is not clear which are the minimal regularity conditions that must be imposed to $(\mathbf{u}, \theta)$ in order to ensure uniqueness for varying ν.

Notice that, in fact, we can still do something without requiring uniqueness. Indeed, let us assume that $\mathcal{U}$ is a non-empty closed convex set with

$$\mathcal{U} \subset W_c = L^2(0,T; H^1(\Omega)) \cap L^\infty(0,T; L^2(\Omega)).$$

Then, the arguments and results we have presented in section 3 to deduce the optimality system hold again: if $(\mathbf{u}, \theta, v) \in \mathcal{Q}$ is an optimal solution of(4), there are variables $\mathbf{q} \in L^2(0,T;H)$ and $\zeta \in L^2(0,T;L^2(\Omega))$ that solve in the transposition sense the adjoint system (5) and one has

$$\iint_Q (-\zeta + \mu v)(\bar{v} - v) \leq 0 \quad \forall \bar{v} \in \mathcal{U}, \quad v \in \mathcal{U}.$$

Actually, the difficulty arises only when we try to prove the existence of an optimal solution. (Again, with suitable modifications of the cost functional, the existence of optimal triplets $(\mathbf{u}, \theta, v) \in \mathcal{Q}$ can be established.)

These questions will be analyzed in a forthcoming paper.

Acknowledgements. The authors are grateful to the "Fundação de Apoio à Pesquisa do Estado de São Paulo" (FAPESP) and the Spanish "Ministerio de Educación y Ciencia" (MEC), for having supported visits of the second author to the Department of Mathematics of the IMECC-Unicamp, Brazil, during part of the period of this research (grant FAPESP 2005/50705-4). They are grateful to the anonymous referee whose suggestions helped to improve this work.

References

[1] F. Abergel and R. Temam, *On some control problems in fluid mechanics*, Theoret. Comput. Fluid Dyn. **1** (1990), 303–325.

[2] R. A. Adams, *Sobolev spaces*, Pure and Applied Mathematics, vol. 65, Academic Press, New York-London, 1975.

[3] E. Casas, *The Navier-Stokes equations coupled with the heat equation: analysis and control*, Control Cybernet. **23** (1994), no. 4, 605–620.

[4] E. Casas, J.-P. Raymond, and H. Zidani, *Pontryagin's principle for local solutions of control problems with mixed control-state constraints*, SIAM J. Control Optim. **39** (2000), no. 4, 1182–1203.

[5] V. Alexéev, S. Fomine, and V. Tikhomirov, *Commande optimale*, Mir, Moscow, 1982.

[6] B. Climent-Ezquerra, F. Guillén-González, and M. A. Rojas-Medar, *Regular time-reproductive solutions for generalized Boussinesq model with Neumann boundary conditions for the temperature*, to appear.

[7] J. I. Díaz and G. Galiano, *Existence and uniqueness of solutions of the Boussinesq system with nonlinear thermal diffusion*, Topol. Methods Nonlinear Anal. **11** (1998), no. 1, 59–82.

[8] P. G. Drazin and W. H. Reid, *Hydrodynamic stability*, Cambridge Monographs on Mechanics and Applied Mathematics, Cambridge University Press, Cambridge, 1981.

[9] T. M. Flett, *Differential analysis*: *Differentiation, differential equations and differential inequalities*, Cambridge University Press, Cambridge, 1980.

[10] I. V. Girsanov, *Lectures on mathematical theory of extremum problems*, Lecture Notes in Economics and Mathematical Systems, vol. 67, Springer-Verlag, Berlin, 1972.

[11] T. Hishida, *Existence and regularizing properties of solutions for the nonstationary convection problem*, Funkcial. Ekvac. **34** (1991), no. 3, 449–474.

[12] H.-C. Lee and B. C. Shin, *Piecewise optimal distributed controls for 2D Boussinesq equations*, Math. Methods Appl. Sci. **23** (2000), no. 3, 227–254.

[13] S. Li and G. Wang, *The time optimal control of the Boussinesq equations*, Numer. Funct. Anal. Optim. **24** (2003), no. 1-2, 163–180.

[14] J.-L. Lions, *Optimal control of systems governed by partial differential equations*, Springer-Verlag, New York, 1971.

[15] ———, *Control of distributed singular systems*, Gautier-Villars, Paris, 1985.

[16] S. A. Lorca and J. L. Boldrini, *Stationary solutions for generalized Boussinesq models*, J. Differential Equations **124** (1996), no. 2, 389–406.

[17] ———, *The initial value problem for a generalized Boussinesq model*: *Regularity and global existence of strong solutions*, Fourth Workshop on Partial Differential Equations — Part II (Rio de Janeiro, 1995), Matemática Contemporánea, vol. 11, 1996.

[18] ———, *The initial value problem for a generalized Boussinesq model*, Nonlinear Anal. **36** (1999), no. 4, 457–480.

[19] H. Morimoto, *Nonstationary Boussinesq equations*, J. Fac. Sci. Univ. Tokyo Sect. IA Math. **39** (1992), no. 1, 61–75.

[20] K. Ōeda, *On the initial value problem for the heat convection equation of Boussinesq approximation in a time-dependent domain*, Proc. Japan Acad. Ser. A Math. Sci. **64** (1988), no. 5, 143–146.

[21] T. N. Shilkin, *Partial regularity of weak solutions of the stationary 3D Boussinesq system*, J. Math. Sci. (N. Y.) **123** (2004), no. 6, 4668–4677.

[22] M. Shinbrot and W. P. Kotorynski, *The initial value problem for a viscous heat-conducting fluid*, J. Math. Anal. Appl. **45** (1974), 65–96.

[23] E. Zabrodzki, *On the initial-boundary value problem for the system of equations describing a motion of incompressible, viscous and heat conducting fluid with data in L^p*, Bull. Polish Acad. Sci. Math. **33** (1985), no. 9-10, 503–510.