Available online at www.sciencedirect.com
ScienceDirect
J. Math. Anal. Appl. 323 (2006) 891–912
Journal of MATHEMATICAL ANALYSIS AND APPLICATIONS
www.elsevier.com/locate/jmaa

Semidiscrete approximations of optimal Robin boundary control problems constrained by semilinear parabolic PDE

K. Chrysafinos [a], M.D. Gunzburger [b,*], L.S. Hou [c]

[a] Institut für Numerische Simulation, Universität at Bonn, Wegelerstr 6, Bonn 53115, Germany
[b] School of Computational Science, Florida State University, Tallahassee, FL 32306, USA
[c] Department of Mathematics, Iowa State University, Ames, IA 50011, USA

Received 11 October 2005
Available online 1 December 2005
Submitted by William F. Ames

Abstract

An optimal Robin boundary control problem associated with semilinear parabolic partial differential equations is considered. Existence of an optimal solution is proved and an optimality system of equations is derived. Semidiscrete finite element approximations of the optimality system are defined and error estimates are obtained.

Keywords: Robin boundary control; Finite element methods; Parabolic equations

1. Introduction

We consider an optimal boundary control problem for the semilinear parabolic partial differential equation (PDE)

$$\partial_t u - \operatorname{div}\big[A(\mathbf{x})\nabla u\big] + \phi(u) + b(t,\mathbf{x})u = f \quad \text{in } (0,T)\times\Omega \tag{1.1}$$

* Corresponding author.
E-mail addresses: chrysafinos@ins.uni-bonn.edu (K. Chrysafinos), gunzburg@csit.fsu.edu (M.D. Gunzburger), hou@math.iastate.edu (L.S. Hou).

0022-247X/$ – see front matter
doi:10.1016/j.jmaa.2005.10.053

with the initial condition

$$u(0) = u_0 \quad \text{in } \Omega \tag{1.2}$$

and the Robin boundary condition

$$u + \lambda^{-1}\big[A(\mathbf{x})\nabla u\big] \cdot \mathbf{n} = g \quad \text{on } (0, T) \times \partial\Omega. \tag{1.3}$$

Here, Ω is a two-dimensional convex polygon with a boundary $\partial\Omega$, λ a positive constant, $[0, T]$ a time interval, f a function in $L^2(0, T; H^1(\Omega)^*)$, u_0 a function in $L^2(\Omega)$, $b(t, \mathbf{x})$ a continuous function on $[0, T] \times \overline{\Omega}$, and $A(\mathbf{x})$ a $C^1(\overline{\Omega})$, symmetric-matrix-valued function that is uniformly positive definite. Also, $\phi \in C^2(\mathbb{R})$ satisfies the following monotonicity and growth conditions:

$$\begin{aligned} &\phi'(s) \geqslant 0 \quad \forall s \in \mathbb{R}, \qquad \big|\phi'(s)\big| \leqslant C_3\big(1 + |s|^{r_2 - 1}\big) \quad \forall s \in \mathbb{R}, \\ &s\phi(s) \geqslant C_1|s|^{r_1} \quad \text{and} \quad \big|\phi(s)\big| \leqslant C_2\big(1 + |s|^{r_2}\big) \quad \forall s \in \mathbb{R}, \end{aligned} \tag{1.4}$$

where $r_1 \in (2, 4)$ and $r_2 \in (1, 3)$. The Robin boundary control g belongs to $L^2((0, T) \times \partial\Omega)$ and the control objective is to track a global target state $U(t, \mathbf{x})$ in $(0, T) \times \Omega$. We formulate the control problem as follows: minimize the cost functional

$$\mathcal{J}(u, g) = \frac{1}{2}\int_0^T \int_\Omega |u - U|^2 \, d\mathbf{x}\, dt + \frac{\gamma}{2}\int_0^T \int_{\partial\Omega} |g|^2 \, ds\, dt \tag{1.5}$$

subject to the initial boundary value problem (1.1)–(1.3).

In Section 2, we prove the existence of a solution for (1.1)–(1.3), establish the existence of an optimal solution, and derive an optimality system of equations which the optimal solution must satisfy. In Section 3, we define semidiscrete finite element approximations of the optimality system, quote the Brezzi–Rappaz–Raviart (BRR) theory for the approximation of a class of nonlinear problems, and apply that theory to derive error estimates for the semidiscrete approximations of the optimality system.

Some remarks about the literature are in order. Extremal problems for linear parabolic PDE with nonsmooth Dirichlet boundary control were analyzed mathematically and numerically in [18]. In [17], a Robin boundary control problem for a linear parabolic PDE is studied. The objective of the control problem in [17] was to determine the minimal time for the controlled state to reach within a specified distance from the desired state. The convergence and error estimates for semidiscrete finite element approximations were studied in [17]. The Robin control used in [17] was of the separation of variable type and the domain was assumed to be of class C^∞ as required by elliptic regularity results. Several results are proposed and analyzed in the continuous setting for a conjugate gradient method for solving an optimal control problem constrained by a linear parabolic PDE in [28]. The problem studied in [28] involved a terminal-state tracking functional and a Neumann boundary control. A key idea of [28] is to formulate the optimal solution as a solution to a system of two sequentially-coupled initial value problems; as a result, the methods of [28] applies only to terminal-state functionals tracking functionals. In [23], error estimates for the fully-discrete approximation of a Neumann boundary control problem associated to a homogeneous linear parabolic equation are presented.

In [19] (see also relevant work in [18]), Dirichlet and Neumann control problems are considered for linear homogeneous parabolic PDEs. Several results concerning analysis and finite element approximations are presented based on semigroup techniques. In [26,27], nonlinear boundary controls are used to minimize a general functional that can handle terminal, normal, and matching norms. The size of the control is constrained and additional regularity on the controls is needed. A nonstandard weak form for the PDE is used for which mild solutions are defined as Bochner integral solutions.

For boundary optimal control problems having states constrained by elliptic partial differential equations, there has been much progress with respect to both analyses and the finite element approximations; see, e.g., [7,12–16,19]. Finally, a posteriori error estimates for a distributed optimal control problems governed by parabolic PDEs are studied in [20]. For an overview of theoretical results in control theory, we refer the reader to [8,19], while for related references and applications of flow control see [11].

It appears that little work has been done in case of semidiscrete error estimates for the optimality system arising from boundary optimal control problems for semilinear parabolic PDEs. The main difficulty consists of the lack of sufficient techniques to "uncouple" the optimality system, in particular in presence nonlinearities.

The optimality systems arising from boundary optimal control problems are usually coupled and in order to uncouple the primal and dual variable, we will use the theory of Brezzi–Rappaz–Raviart (BRR theory) which requires the availability of error estimates under minimal regularity assumptions. The main advantage of this methodology is that it enables the derivation of estimates of arbitrary order in the natural energy norms for all involved variables (primal, dual and control) and it can handle nonlinearities as well as nonhomogeneous data. In addition, using the BRR theory, we will be able to avoid the semigroup machinery.

Furthermore, the use of a Robin-type boundary control may alleviate the difficulties arising from the nonhomogeneous boundary condition.

We note that although we confine our study to Robin-type boundary controls, semidiscrete approximations of Neumann and distributed control problems can be treated in a similarly, and in the latter case, more easily.

2. Existence of an optimal solution and an optimality system of equations

In this section we formulate and analyze the control problem in an appropriate mathematical framework.

Throughout, C denotes a generic constant whose value changes with context. We freely make use of the standard Sobolev spaces $H^s(\Omega)$ and $H^s(\partial\Omega)$ for $s \in \mathbb{R}$ with norms $\|\cdot\|_s$ and $\|\cdot\|_{s,\partial\Omega}$, respectively.

For $p \in [1, \infty]$, an interval $(a, b) \subset \mathbb{R}$, and a Banach space B with norm $\|\cdot\|_B$, we denote by $L^p(a, b; B)$ the set of measurable functions $v : (a, b) \to B$ such that $\int_a^b \|v(t)\|_B^p \, dt < \infty$. The norm on $L^p(a, b; B)$ for $p \in [1, \infty)$ is defined by

$$\|v\|_{L^p(a,b;B)} = \left(\int_a^b \|v(t)\|_B^p \, dt \right)^{1/p} \quad \forall v \in L^p(a, b; B),$$

and for $p = \infty$ by

$$\|v\|_{L^\infty(a,b;B)} = \operatorname*{ess\,sup}_{(a,b)} \|v(t)\|_B \quad \forall v \in L^\infty(a, b; B).$$

We denote by $C([a,b];B)$ the set of all continuous functions $v:[a,b]\to B$ with the norm $\|v\|_{C([a,b];B)}=\max_{t\in[a,b]}\|v(t)\|_B$. For real numbers $s\geqslant 0$ and $p\geqslant 1$, the space $H^s(a,b;B)$ is defined as follows. First,

$$H^s(\mathbb{R};B)=\bigl\{v\in L^2(\mathbb{R};B):\ |\tau|^s\hat{v}\in L^2(\mathbb{R};B)\bigr\}$$

endowed with the norm

$$\|v\|_{H^s(\mathbb{R};B)}=\biggl(\int_{\mathbb{R}}\bigl\|v(t)\bigr\|_B^2\,dt+\int_{\mathbb{R}}|\tau|^{2s}\bigl\|\hat{v}(\tau)\bigr\|_B^2\,d\tau\biggr)^{1/2},$$

where $\hat{v}$ is the temporal Fourier transform of v:

$$\hat{v}(\tau)=\int_{\mathbb{R}}e^{-2i\pi t\tau}v(t)\,dt.$$

Then, we set

$$H^s(a,b;B)=\bigl\{v=\tilde{v}|_{[a,b]}:\ \tilde{v}\in H^s(\mathbb{R};B)\bigr\}$$

with the norm

$$\|v\|_{H^s(a,b;B)}=\inf_{\substack{\tilde{v}\in H^s(\mathbb{R};B)\\ \tilde{v}|_{[a,b]}=v}}\|\tilde{v}\|_{H^s(\mathbb{R};B)}\quad\forall v\in H^s(a,b;B).$$

A function $v=v(t,\mathbf{x})\in H^s(a,b;B)$ for some spatial function space B is often simply written as $v(t)$. Further discussions of the Banach-space-valued Sobolev spaces $H^s(a,b;B)$ may be found in [6,22,24]. In particular, we will use the following function spaces involving time: $H^r(0,T;H^s(\Omega))$ and $H^r(0,T;H^s(\partial\Omega))$ for $r,s\in\mathbb{R}$.

We also introduce the solution space $\mathcal{W}(0,T)$ for (1.1)–(1.3) as

$$\mathcal{W}(0,T)\equiv L^2\bigl(0,T;H^1(\Omega)\bigr)\cap H^1\bigl(0,T;H^1(\Omega)^*\bigr)$$

with the norm defined by

$$\|v\|^2_{\mathcal{W}(0,T)}=\|v\|^2_{L^2(0,T;H^1(\Omega))}+\|\partial_t v\|^2_{L^2(0,T;H^1(\Omega)^*)}\quad\forall v\in\mathcal{W}(0,T).$$

Given an $f\in L^2(0,T;H^1(\Omega)^*)$ and a $g\in L^2(0,T;H^{-1/2}(\partial\Omega))$, a function $u\in\mathcal{W}(0,T)$ is said to be a (weak) solution of (1.1)–(1.3) if

$$\bigl\langle\partial_t u(t),v\bigr\rangle+a\bigl(u(t),v\bigr)+\lambda\bigl[u(t),v\bigr]_{\partial\Omega}=\bigl\langle f(t),v\bigr\rangle+\lambda\bigl\langle g(t),v\bigr\rangle_{\partial\Omega}\quad\forall v\in H^1(\Omega),\ \text{a.e. }t,\tag{2.1}$$

and

$$u(0)=u_0\quad\text{in }L^2(\Omega),\tag{2.2}$$

where $\langle\cdot,\cdot\rangle$ denotes the duality pairing between $H^1(\Omega)^*$ and $H^1(\Omega)$ and $\langle\cdot,\cdot\rangle_{\partial\Omega}$ the duality pairing between $H^{-1/2}(\partial\Omega)$ and $H^{1/2}(\partial\Omega)$, see [6]. Equation (2.2) makes sense thanks to the following lemma.

Lemma 2.1. *Suppose $v\in L^2(0,T;H^1(\Omega))\cap H^1(0,T;H^1(\Omega)^*)$. Then,*

$$v\in C\bigl([0,T];L^2(\Omega)\bigr)\quad\textit{and}\quad\frac{d}{dt}\bigl\|v(t)\bigr\|_0^2=2\bigl\langle v'(t),v(t)\bigr\rangle\quad\textit{a.e. }t\in(0,T).\tag{2.3}$$

The proofs for Lemma 2.1 and the next lemma are identical to that of [6, Theorem 3, pp. 287–288] and are omitted.

Lemma 2.2. *Suppose* $v \in L^2(0,T;H^{1/2}(\partial\Omega)) \cap H^1(0,T;H^{-1/2}(\partial\Omega))$. *Then,*

$$v \in C\big([0,T];L^2(\partial\Omega)\big) \quad \text{and} \quad \frac{d}{dt}\|v(t)\|^2_{0,\partial\Omega} = 2\langle v'(t), v(t)\rangle_{\partial\Omega} \quad \text{a.e. } t \in (0,T). \tag{2.4}$$

The solution of (2.1)–(2.2) should be sought in $L^2(0,T;H^1(\Omega)) \cap H^1(0,T;H^1(\Omega)^*) \cap L^{r_2}(0,T\times\Omega)$. The following lemma indicates that the requirement for the solution to belong to $L^{r_2}(0,T\times\Omega)$ is redundant.

Lemma 2.3. *The embedding*

$$L^2\big(0,T;H^1(\Omega)\big) \cap H^1\big(0,T;H^1(\Omega)^*\big) \hookrightarrow L^4\big(0,T;L^4(\Omega)\big)$$

is continuous.

Proof. Using Sobolev embedding theorems and interpolation theory, we have

$$\|w(t)\|_{L^4(\Omega)} \leqslant C\|w(t)\|_{1/2} \leqslant C\|w(t)\|_0^{1/2}\|w(t)\|_1^{1/2}.$$

Thus,

$$\int_0^T \|w(t)\|^4_{L^4(\Omega)}\,dt \leqslant C\int_0^T \|w(t)\|_0^2\|w(t)\|_1^2\,dt \leqslant C\|w\|^2_{L^\infty(0,T;L^2(\Omega))}\int_0^T \|w(t)\|_1^2\,dt$$

so that by virtue of Lemma 2.1 we obtain

$$\int_0^T \|w(t)\|^4_{L^4(\Omega)}\,dt \leqslant C\|w\|^4_{\mathcal{W}(0,T)}.$$

This completes the proof. □

The optimal control problem described in Section 1 can now be stated precisely as follows:

$$\begin{aligned}&\text{seek a pair } (u,g) \in \mathcal{W}(0,T) \times L^2\big(0,T;L^2(\partial\Omega)\big) \text{ such that}\\ &\text{functional (1.5) is minimized subject to (2.1)–(2.2).}\end{aligned} \tag{2.5}$$

We will prove the existence of a solution for the PDE problem (2.1)–(2.2), establish the existence of an optimal solution for (2.5), and derive an optimality system of equations.

2.1. Existence and uniqueness of solutions of the PDE problem

Theorem 2.4. *Suppose* $f \in L^2(0,T;H^1(\Omega)^*)$, $u_0 \in L^2(\Omega)$, *and* $g \in L^2(0,T;H^{-1/2}(\partial\Omega))$. *Then, there exists a unique* $u \in L^2(0,T;H^1(\Omega)) \cap H^1(0,T;H^1(\Omega)^*)$ *satisfying* (2.1)–(2.2) *and*

$$\begin{aligned}&\|u\|_{L^2(0,T;H^1(\Omega))} + \|\partial_t u\|_{L^2(0,T;H^1(\Omega)^*)}\\ &\quad \leqslant C\big(\|f\|_{L^2(0,T;H^1(\Omega)^*)} + \|u_0\|_0 + \|g\|_{L^2(0,T;H^{-1/2}(\partial\Omega))}\big).\end{aligned} \tag{2.6}$$

Proof. The proof follows standard Galerkin techniques, see, e.g., [6,21]. We choose an *orthogonal* basis $\{e_k\}_{k=1}^{\infty}$ of $H^1(\Omega)$ such that $\{e_k\}_{k=1}^{\infty}$ is *orthonormal* in $L^2(\Omega)$. For each $m = 1, 2, \ldots,$ we set $V_m = \operatorname{span}\{e_1, \ldots, e_m\}$ and let $u_m = \sum_{k=1}^{m} d_k^{(m)} e_k \in H^1(0, T; V_m)$ be a solution of

$$\begin{aligned}
&\langle u_m', v\rangle + a[u_m, v] + \lambda[u_m, v]_{\partial\Omega} + \big[\phi(u_m), v\big] + \big[b(t, \mathbf{x})u_m, v\big] = \langle f, v\rangle + \lambda\langle g, v\rangle_{\partial\Omega}\\
&\quad \forall v \in V_m,\ \text{a.e. } t \in (0, T),\\
&\big[u_m(0), v\big] = [u_0, v] \quad \forall v \in V_m.
\end{aligned} \tag{2.7}$$

The existence of such u_m can be easily proved based on the assumptions on ϕ. Using Gronwall's inequality, we obtain the following energy estimate:

$$\begin{aligned}
&\max_{[0,T]} \big\|u_m(t)\big\|_0 + \|u_m\|_{L^2(0,T;H^1(\Omega))} + \big\|u_m'\big\|_{L^2(0,T;H^1(\Omega)^*)}\\
&\quad \leqslant C\big(\|f\|_{L^2(0,T;H^1(\Omega)^*)} + \|g\|_{L^2(0,T;H^{-1/2}(\partial\Omega))}\big)
\end{aligned} \tag{2.8}$$

for $m = 1, 2, \ldots.$ Thus, we may extract a subsequence of $\{u_m\}_{m=1}^{\infty}$, still denoted by $\{u_m\}_{m=1}^{\infty}$, such that

$$\begin{aligned}
&u_m \rightharpoonup u \quad \text{weakly in } L^2\big(0, T; H^1(\Omega)\big),\\
&u_m' \rightharpoonup \partial_t u \quad \text{weakly in } L^2\big(0, T; H^1(\Omega)^*\big),\\
&u_m|_{\partial\Omega} \rightharpoonup u|_{\partial\Omega} \quad \text{weakly in } L^2\big(0, T; H^{1/2}(\partial\Omega)\big),
\end{aligned}$$

and

$$u_m \to u \quad \text{strongly in } L^2\big(0, T; L^2(\Omega)\big),$$

where the strong convergence result follows directly from a well-known compact embedding result; see, e.g., [24]. Also, Lemma 2.3 and (2.8) imply that $\{\|u_m\|_{L^4(0,T;L^4(\Omega))}\}$ is uniformly bounded. Thus, by passing to the limit in (2.7), we see that u satisfies (2.1)–(2.2). Moreover, the solution of (2.1)–(2.2) is unique. Passing to the limit in (2.8) yields (2.6). □

2.2. Existence of a solution to the optimal Robin control problem

Theorem 2.5. *There exists a pair* $(u, g) \in \mathcal{W}(0, T) \times L^2(0, T; L^2(\partial\Omega))$ *that minimizes* (1.5) *subject to* (2.1)–(2.2).

Proof. Set $\mathcal{U}_{\mathrm{ad}} = \{(\tilde u, \tilde g) \in \mathcal{W}(0, T) \times L^2(0, T; L^2(\partial\Omega))\colon\ (\tilde u, \tilde g) \text{ satisfies (2.1)–(2.2)}\}$. $\mathcal{U}_{\mathrm{ad}}$ is obviously nonempty because of Theorem 2.4.

Let $\{(u_n, g_n)\} \subset \mathcal{U}_{\mathrm{ad}}$ be a minimizing sequence, i.e.,

$$\lim_{n\to\infty} \mathcal{J}(u_n, g_n) = \inf_{(\tilde u, \tilde g)\in\mathcal{U}_{\mathrm{ad}}} \mathcal{J}(\tilde u, \tilde g), \tag{2.9}$$

$$\begin{aligned}
&\langle \partial_t u_n, v\rangle + a[u_n, v] + \lambda[u_n, v]_{\partial\Omega} + \big[\phi(u_n), v\big] + \big[b(t, \mathbf{x})u_n, v\big] = \langle f, v\rangle + \lambda[g_n, v]_{\partial\Omega}\\
&\quad \forall v \in H^1(\Omega),\ \text{a.e. } t,
\end{aligned} \tag{2.10}$$

and

$$u_n(0) = u_0 \quad \text{in } \Omega. \tag{2.11}$$

Equation (2.9) implies that $\|g_n\|_{L^2(0,T;L^2(\partial\Omega))} \leqslant C$ for all n. Then, using the estimate (2.6), we deduce $\|u_n\|_{\mathcal{W}(0,T)} \leqslant C$ for all n. Hence, we may extract a subsequence, still denoted by $\{(u_n, g_n)\}$, that satisfies the following convergence properties:

$$g_n \rightharpoonup g \quad \text{weakly in } L^2\big(0,T; L^2(\partial\Omega)\big),$$
$$u_n \rightharpoonup u \quad \text{weakly in } L^2\big(0,T; H^1(\Omega)\big),$$
$$\partial_t u_n \rightharpoonup \partial_t u \quad \text{weakly in } L^2\big(0,T; H^1(\Omega)^*\big),$$
$$u_n|_{\partial\Omega} \rightharpoonup u|_{\partial\Omega} \quad \text{weakly in } L^2\big(0,T; H^{1/2}(\partial\Omega)\big),$$

and

$$u_n \to u \quad \text{strongly in } L^2\big(0,T; L^2(\Omega)\big) \text{ and a.e.}$$

These convergence relations allow us to pass to the limit in (2.10)–(2.11) to show that (u, g) satisfies (2.1)–(2.2).

Using the weak lower semicontinuity of the functional $\mathcal{J}$, we have

$$\mathcal{J}(u,g) \leqslant \lim_{n\to\infty} \mathcal{J}(u_n, g_n) = \inf_{(\tilde{u},\tilde{g})\in\mathcal{U}_{\mathrm{ad}}} \mathcal{J}(\tilde{u},\tilde{g})$$

so that

$$\mathcal{J}(u,g) = \inf_{(\tilde{u},\tilde{g})\in\mathcal{U}_{\mathrm{ad}}} \mathcal{J}(\tilde{u},\tilde{g}).$$

This shows that (u, g) is an optimal solution. □

2.3. *An optimality system of equations*

We will use the Lagrange multiplier principle to derive an optimality system of equations that the optimal solutions must satisfy.

We consider an abstract minimization problem. Let X_1 and X_2 be two Banach spaces. Let $J : X_1 \to \mathbb{R}$ be a functional and $F : X_1 \to X_2$ be a mapping. We seek a $z \in X_1$ such that

$$J(z) = \inf_{u\in\mathcal{U}_{\mathrm{ad}}} J(u), \tag{2.12}$$

where

$$\mathcal{U}_{\mathrm{ad}} = \big\{u \in X_1:\ F(u) = 0\big\}.$$

The Lagrange functional for the minimization problem (2.12) is defined by

$$\mathcal{L}(z, q, \lambda_0) = \lambda_0 J(z) + \big\langle F(z), q\big\rangle \tag{2.13}$$

for all $z \in X_1$, $\lambda_0 \in \mathbb{R}$, and $q \in X_2^*$, where X_2^* is the dual space of X_2 and $\langle\cdot,\cdot\rangle$ denotes the duality pairing between X_2 and X_2^*. We quote a standard abstract Lagrange principle in the following particular form (see [1,8]).

Theorem 2.6. *Let z be a solution of* (2.12). *Assume that the mappings J and F are continuously differentiable and that the image of the operator $F'(z) : X_1 \to X_2$ is closed. Then, there exist $q \in X_2^*$ and $\lambda_0 \in \mathbb{R}$ such that $(q, \lambda_0) \neq (0, 0)$, i.e., q and λ_0 do not vanish simultaneously,*

$$\big\langle \mathcal{L}_z(z, q, \lambda_0), \eta\big\rangle = 0 \quad \forall \eta \in X_1, \tag{2.14}$$

and

$$\lambda_0 \geqslant 0, \tag{2.15}$$

where $\mathcal{L}_z(\cdot,\cdot,\cdot)$ *denotes the Fréchet derivative of* $\mathcal{L}$ *with respect to the first argument. Furthermore, if* $F'(z): X_1 \to X_2$ *is an epimorphism, then* $\lambda_0 \neq 0$ *and* λ_0 *can be taken as* 1.

Theorem 2.7. *Let* $(u,g) \in \mathcal{W}(0,T) \times L^2(0,T;L^2(\partial\Omega))$ *denote an optimal solution for* (2.5). *Then, there exists* $\xi \in \mathcal{W}(0,T)$ *such that*

$$-\langle \partial_t \xi, \eta\rangle + a[\xi,\eta] + \lambda[\xi,\eta]_{\partial\Omega} + \left[\phi'(u)\xi,\eta\right] + \left[b(t,\mathbf{x})\xi,\eta\right] = \langle u-U,\eta\rangle$$
$$\forall \eta \in H^1(\Omega), \ \text{a.e. } t, \tag{2.16}$$

$$\xi(T) = 0 \quad \text{in } \Omega, \tag{2.17}$$

and

$$\lambda\xi + \gamma g = 0 \quad \text{on } (0,T)\times\partial\Omega. \tag{2.18}$$

Proof *(Sketch).* We define

$$X_1 = \mathcal{W}(0,T) \times L^2\big(0,T;L^2(\partial\Omega)\big),$$
$$X_2 = L^2\big(0,T;H^1(\Omega)^*\big) \times L^2(\Omega) \times L^2\big(0,T;H^{-1/2}(\partial\Omega)\big).$$

The variable z is understood as the pair $z=(u,g)$, the functional $J(z)$ is defined as

$$J(z) \equiv \frac{1}{2}\int_0^T \int_\Omega |u-U|^2\,d\mathbf{x}\,dt + \frac{\gamma}{2}\int_0^T\int_{\partial\Omega} |g|^2\,ds\,dt$$

and the operator $F: X_1 \to X_2$ as the operator related to the constraint equations (1.1)–(1.3), i.e.,

$$F(u,g) = \big(\partial_t u - \operatorname{div}\big[A(\mathbf{x})\nabla u\big] + \phi(u) + b(t,\mathbf{x})u,\ u(0)-u_0,\ u|_{\partial\Omega} + \lambda^{-1}\big[A(\mathbf{x})\nabla u\big]\cdot\mathbf{n} - g\big).$$

The operator $F'(u): X_1 \to X_2$ can be defined as

$$F'(u)(\bar{u},\bar{g}) = \big(\partial_t \bar{u} - \operatorname{div}\big[A(\mathbf{x})\nabla \bar{u}\big] + \phi'(u)\bar{u} + b(t,\mathbf{x})\bar{u},\ \bar{u}(0)-\bar{u}_0,\ \bar{u}|_{\partial\Omega} + \lambda^{-1}\big[A(\mathbf{x})\nabla \bar{u}\big]\cdot\mathbf{n} - \bar{g}\big).$$

Note that the constraints can be expressed as $F(u,g) = (f,0,0)$. The range of operator $F'(u)$ is closed. It remains to show that it is also an epimorphism. First, observe that the operator is continuous due to the embedding $\mathcal{W}(0,T) \subset C(0,T;L^2(\Omega))$ and a well-known trace theorem (see, e.g., [22]). Then, for each $(\bar{f},\bar{u}_0,\bar{g}_1) \in X_2$ we need to show that there exists a solution $(\bar{u},0) \in X_1$ of system

$$\partial_t \bar{u} - \operatorname{div}\big[A(\mathbf{x})\nabla\bar{u}\big] + \phi'(u)\bar{u} + b(t,\mathbf{x})\bar{u} = \bar{f}, \qquad \bar{u}(0) = \bar{u}_0,$$
$$\bar{u}|_{\partial\Omega} + \lambda^{-1}\big[A(\mathbf{x})\nabla\bar{u}\big]\cdot\mathbf{n} = \bar{g}_1.$$

This is true due to growth assumptions on $\phi'(u)$, the linearity of the equation and the (chosen) boundary condition (see, e.g., Theorem 2.4). Therefore, we may apply Theorem 2.6 to conclude that there exist $q \in L^2(0,T;H^1(\Omega))$, $\lambda_0 \in \mathbb{R}$, such that q,λ_0 do not vanish simultaneously and

$$\big\langle \mathcal{L}_z(z,q,\lambda_0),\eta\big\rangle = 0 \quad \forall \eta \in X_1,$$

where the Lagrangian is defined as

$$\begin{aligned}
&\mathcal{L}\big((u,g),\xi\big)\\
&\quad=\frac{1}{2}\int_0^T\int_\Omega |u-U|^2\,d\mathbf{x}\,dt+\frac{\gamma}{2}\int_0^T\int_{\partial\Omega}|g|^2\,ds\,dt\\
&\qquad+\int_0^T\Big(\langle u,\partial_t\xi\rangle-a[u,\xi]-\big[\phi(u),\xi\big]-\lambda\big[u(t),\xi\big]_{\partial\Omega}+\langle f,\xi\rangle+\lambda\langle g,\xi\rangle_{\partial\Omega}\Big)\,dt\\
&\qquad-\big[u(T),\xi(T)\big]+\big[u_0,\xi(0)\big].
\end{aligned}$$

Note that we denote by ξ the multiplier q and that λ_0 can be taken as 1 in the above definition. Combining the last two equalities, taking the derivative of the Lagrangian with respect to the first argument and using standard techniques from Calculus of Variations (see, e.g., [8, Section 2]), we conclude that there exists $\xi\in\mathcal{W}(0,T)$ such that (2.16)–(2.18) hold. (The fact $\partial_t\xi\in L^2(0,T;H^1(\Omega)^*)$ can be easily deduced by the linearity of our equation.) □

Theorem 2.7 implies that an optimal solution (u,g) must satisfy the system formed by (2.1)–(2.2) and (2.16)–(2.18). Using (2.18), we can eliminate g in (2.1) and obtain the reduced optimality system of equations:

$$\begin{aligned}
&\langle\partial_t u,v\rangle+a[u,v]+\lambda[u,v]_{\partial\Omega}+\big[\phi(u),v\big]+\big[b(t,\mathbf{x})u,v\big]=\langle f,v\rangle-\lambda^2\gamma^{-1}[g,v]_{\partial\Omega}\\
&\quad\forall v\in H^1(\Omega),\ \text{a.e. } t,
\end{aligned}\tag{2.19}$$

(2.2), and (2.16)–(2.17).

3. Semidiscrete approximations of the optimality system

Let V_h be a family of finite element subspaces of $H^1(\Omega)$ defined over a family of regular triangulations of Ω. The parameter h denotes the largest grid size for a given triangulation. We assume that V_h satisfies the following approximation properties:

(i) for every $v\in H^s(\Omega)$,

$$\inf_{v_h\in V_h}\|v-v_h\|_s\to 0\quad\text{as } h\to 0,\ s=-1,0,1;\tag{3.1}$$

(ii) there exists a constant $C>0$ such that for every $v\in H^{r+1}(\Omega)$ and every $r\in[s-1,k]$,

$$\inf_{v_h\in V_h}\|v-v_h\|_s\leqslant Ch^{r+1-s}\|v\|_{r+1},\quad s=-1,0,1,\tag{3.2}$$

where $k\geqslant 1$ is a positive integer that is usually determined by the order of the piecewise polynomials used to define V_h.

We also assume that finite element triangulations are uniformly regular so that the following inverse inequality holds:

$$\|v_h\|_1\leqslant Ch^{-1}\|v_h\|_0\quad\forall v_h\in V_h.\tag{3.3}$$

For detailed discussions of the properties (3.1)–(3.3) and constructions of the finite element spaces with these properties, see, e.g., [4].

We denote by P_h the $L^2(\Omega)$ projection from $L^2(\Omega)$ onto V_h, namely, for each $v \in L^2(\Omega)$,

$$\big(P_h v - v, w^h\big) = 0 \quad \forall w^h \in V_h. \tag{3.4}$$

As a consequence of (3.3), we have

$$\|P_h v\|_1 \leqslant C\|v\|_1 \quad \forall v \in H_0^1(\Omega); \tag{3.5}$$

see [3,25].

The semidiscrete finite element approximations of the optimality system (2.19), (2.2) and (2.16)–(2.17) are defined as follows: seek $u_h \in H^1(0,T;V_h)$ and $\xi_h \in H^1(0,T;V_h)$ such that

$$\begin{cases} \langle \partial_t u_h(t), v\rangle + a[u_h(t), v_h] + \lambda[u_h(t), v_h]_{\partial\Omega} + [\phi(u_h(t)), v_h] + [b(t,\mathbf{x})u_h(t), v_h] \\ \quad = \langle f(t), v_h\rangle - \lambda^2\gamma^{-1}\langle \xi_h(t), v_h\rangle_{\partial\Omega} \quad \forall v_h \in V_h, \text{ a.e. } t \in [0,T], \\ u_h(0) = P_h u_0, \\ -\langle \partial_t \xi_h(t), \eta_h\rangle + a[\xi_h(t), v_h] + \lambda[\xi_h(t), \eta_h]_{\partial\Omega} + [\phi'(u_h(t))\xi_h(t), \eta_h] \\ \quad + [b(t,\mathbf{x})\xi_h(t), \eta_h] = [u(t) - U(t), \eta_h] \quad \forall \eta_h \in V_h, \text{ a.e. } t \in [0,T], \\ \xi_h(T) = 0, \end{cases} \tag{3.6}$$

where P_h is the $L^2(\Omega)$ projection onto V_h defined by (3.4).

The goal of this section is to prove the following error estimate for the semidiscrete approximations of the optimality system:

$$\begin{aligned} \|u - u_h\|_{\mathcal{W}(0,T)} + \|\xi - \xi_h\|_{\mathcal{W}(0,T)} \leqslant Ch^r\big(&\|u\|_{L^2(0,T;H^{r+1}(\Omega))} + \|\xi\|_{L^2(0,T;H^{r+1}(\Omega))} \\ &+ \|\partial_t u\|_{L^2(0,T;H^{r-1}(\Omega))} + \|\partial_t u\|_{L^2(0,T;H^{r-1}(\Omega))}\big) \end{aligned}$$

provided $u, \xi \in L^2(0,T;H^{r+1}(\Omega)) \cap H^1(0,T;H^{r-1}(\Omega))$ for some $r \in [0,k]$.

We will use the approximation theory of Brezzi–Rappaz–Raviart (BRR) to prove this error estimate.

3.1. *Quotation of the Brezzi–Rappaz–Raviart theory for a class of nonlinear problems*

The Brezzi–Rappaz–Raviart theory [2,5,9] implies that the error of approximations of solutions of certain classes of nonlinear problem is basically the same as the error of approximations of related linear problems. We quote the relevant results here.

Consider the following type of nonlinear problems on a Banach space $\mathcal{X}$: we seek a $\psi \in \mathcal{X}$ such that

$$\psi + \mathcal{T}\mathcal{G}(\psi) = 0, \tag{3.7}$$

where $\mathcal{Y}$ is another Banach space, $\mathcal{T} \in \mathcal{L}(\mathcal{Y};\mathcal{X})$, and $\mathcal{G}$ is a C^2 mapping from $\mathcal{X}$ into $\mathcal{Y}$. We say that $\psi \in \mathcal{X}$ is a *regular solution* of (3.7) if (3.7) holds and $\psi + \mathcal{T}\mathcal{G}_\psi(\psi)$ is an isomorphism from $\mathcal{X}$ into $\mathcal{X}$. Here, $\mathcal{G}_\psi(\cdot)$ denotes the Fréchet derivative of $\mathcal{G}(\cdot)$. We assume that there exists another Banach space $\mathcal{Z}$, contained in $\mathcal{Y}$, with continuous imbedding, such that

$$\mathcal{G}_\psi(\psi) \in \mathcal{L}(\mathcal{X};\mathcal{Z}) \quad \forall \psi \in \mathcal{X}. \tag{3.8}$$

Approximations are defined by introducing a subspace $\mathcal{X}^h \subset \mathcal{X}$ and an approximating operator $\mathcal{T}^h \in \mathcal{L}(\mathcal{Y};\mathcal{X}^h)$: we seek $\psi^h \in \mathcal{X}^h$ such that

$$\psi^h + \mathcal{T}^h \mathcal{G}(\psi^h) = 0. \tag{3.9}$$

Concerning the linear operator $\mathcal{T}^h$, we assume the approximation properties

$$\lim_{h \to 0} \|(\mathcal{T}^h - \mathcal{T})\omega\|_{\mathcal{X}} = 0 \quad \forall \omega \in \mathcal{Y}, \tag{3.10}$$

and

$$\lim_{h \to 0} \|\mathcal{T}^h - \mathcal{T}\|_{\mathcal{L}(\mathcal{Z};\mathcal{X})} = 0. \tag{3.11}$$

Note that whenever the imbedding $\mathcal{Z} \hookrightarrow \mathcal{Y}$ is compact, (3.11) follows from (3.10) and moreover, (3.8) implies that the operator $\mathcal{T}\mathcal{G}_\psi(\psi) \in \mathcal{L}(\mathcal{X};\mathcal{X})$ is compact.

One of the results in [2] is the following theorem. In the statement, $D^2\mathcal{G}$ represents the second Fréchet derivatives of $\mathcal{G}$.

Theorem 3.1. *Let $\mathcal{X}$ and $\mathcal{Y}$ be Banach spaces. Assume that $\mathcal{G}$ is a C^2 mapping from $\mathcal{X}$ into $\mathcal{Y}$ and that $D^2\mathcal{G}$ is bounded on all bounded sets of $\mathcal{X}$. Assume that* (3.8), (3.10)*, and* (3.11) *hold and that $\psi \in \mathcal{X}$ is a regular solution of* (3.7)*. Then, for $h \leqslant h_0$ small enough, there exists a unique $\psi^h \in \mathcal{X}^h$ such that ψ^h is a regular solution of* (3.9)*. Moreover, there exists a positive constant C independent of h such that*

$$\|\psi^h - \psi\|_{\mathcal{X}} \leqslant C\|(\mathcal{T}^h - \mathcal{T})\mathcal{G}(\psi)\|_{\mathcal{X}}. \tag{3.12}$$

3.2. Recasting the optimality system and its semidiscrete approximations into the BRR framework

We set $\mathcal{X} = \mathcal{W}(0,T) \times \mathcal{W}(0,T)$ and $\mathcal{Y} = L^2(0,T;H^1(\Omega)^*) \times L^2(0,T;H^{-1/2}(\partial\Omega)) \times L^2(\Omega) \times L^2(0,T;H^1(\Omega)^*) \times L^2(\Omega)$. We define the linear operator $\mathcal{T}:\mathcal{Y} \to \mathcal{X}$ as follows: $(\tilde{u},\tilde{\xi}) = \mathcal{T}(\tilde{f},\tilde{g},\tilde{u}_0,\tilde{\zeta},\tilde{\xi}_T)$ for $(\tilde{u},\tilde{\xi}) \in \mathcal{X}$ and $(\tilde{f},\tilde{g},\tilde{u}_0,\tilde{\zeta},\tilde{\xi}_T) \in \mathcal{Y}$ if and only if

$$\begin{cases} \langle \partial_t \tilde{u}(t), v\rangle + a[\tilde{u}(t), v] = \langle \tilde{f}(t), v\rangle + \lambda\langle \tilde{g}(t), v\rangle_{\partial\Omega} \quad \forall v \in H_0^1(\Omega), \text{ a.e. } t \in [0,T], \\ \tilde{u}(0) = \tilde{u}_0 \quad \text{in } L^2(\Omega), \\ -\langle \partial_t \tilde{\xi}(t), \eta\rangle + a[\tilde{\xi}(t), \eta] = \langle \tilde{\zeta}(t), \eta\rangle \quad \forall \eta \in H_0^1(\Omega), \text{ a.e. } t \in [0,T], \\ \tilde{\xi}(T) = \tilde{\xi}_T \quad \text{in } L^2(\Omega). \end{cases}$$

We define the nonlinear operator $\mathcal{G}:\mathcal{X} \to \mathcal{Y}$ by

$$\mathcal{G}(\tilde{u},\tilde{\xi}) = \left(-f + \phi(\tilde{u}) + b(t,\mathbf{x})\tilde{u}, \lambda^2\gamma\tilde{\xi}, u_0, \tilde{u} - U, 0\right) \quad \forall (\tilde{u},\tilde{\xi}) \in \mathcal{X},$$

where $f \in L^2(0,T;H^1(\Omega)^*)$, $u_0 \in L^2(\Omega)$ and $U \in L^2(0,T;L^2(\Omega))$ are the prescribed (fixed) data in (2.1). Clearly, the optimality system (2.19), (2.2) and (2.16)–(2.17) may be written as

$$(u,\xi) = -\mathcal{T}\mathcal{G}(u,\xi),$$

i.e., the optimality system is recast into the form of (3.7).

We set

$$\mathcal{X}_h = H^1(0,T;V_h) \times H^1(0,T;V_h)$$

and define the discrete operator $\mathcal{T}_h:\mathcal{Y} \to \mathcal{X}_h$ as follows: $(\tilde{u}_h,\tilde{\xi}_h) = \mathcal{T}_h(\tilde{f},\tilde{g},\tilde{u}_0,\tilde{\zeta},\tilde{\xi}_T)$ for $(\tilde{u}_h,\tilde{\xi})_h \in \mathcal{X}$ and $(\tilde{f},\tilde{g},\tilde{u}_0,\tilde{\zeta},\tilde{\xi}_T) \in \mathcal{Y}$ if and only if

$$\begin{cases} \langle \partial_t \tilde{u}_h(t), v_h\rangle + a[\tilde{u}_h(t), v_h] = \langle \tilde{f}(t), v_h\rangle + \lambda\langle \tilde{g}(t), v_h\rangle_{\partial\Omega} \quad \forall v_h \in V_h, \text{ a.e. } t \in [0,T], \\ \tilde{u}_h(0) = P_h\tilde{u}_0, \\ -\langle \partial_t \tilde{\xi}_h(t), \eta_h\rangle + a[\tilde{\xi}_h(t), \eta_h] = \langle \tilde{\zeta}(t), \eta_h\rangle \quad \forall \eta_h \in V_h, \text{ a.e. } t \in [0,T], \\ \tilde{\xi}_h(T) = P_h\tilde{\xi}_T, \end{cases}$$

where $P_h : L^2(\Omega) \to V_h$ is the $L^2(\Omega)$ projection. Comparing the definitions of $\mathcal{T}_h$ and $\mathcal{G}$ with (3.6), we see that (3.6) can be written as

$$(u_h, \xi_h) = -\mathcal{T}_h\mathcal{G}(u_h, \xi_h),$$

i.e., the semidiscrete optimality system is recast into the form of (3.9).

In order to apply BRR theory to derive error estimates for the semidiscrete approximations of the optimality system, we study in Section 3.3 the semidiscrete approximations of the associated linear problem. Then, in Section 3.4, we establish some embedding and trace results that will help us to choose an appropriate space $\mathcal{Z}$ in Theorem 3.1.

3.3. *The linear boundary value problem and its semidiscrete approximations*

We consider semidiscrete approximations of the linear Robin boundary value problem:

$$\partial_t u - \operatorname{div}\big[A(\mathbf{x})\nabla u\big] = f \quad \text{in } Q \equiv (0,T)\times\Omega,$$
$$u(0) = u_0 \quad \text{in } \Omega,$$

and

$$u + \lambda^{-1}\big[A(\mathbf{x})\nabla u\big]\cdot\mathbf{n} = g \quad \text{on } (0,T)\times\partial\Omega.$$

We will work with the following weak formulation: seek $u \in \mathcal{W}(0,T)$ such that

$$\big\langle \partial_t u(t), v\big\rangle + a\big[u(t), v\big] + \lambda\big[u(t), v\big]_{\partial\Omega} = \big\langle f(t), v\big\rangle + \lambda\big\langle g(t), v\big\rangle_{\partial\Omega}$$
$$\forall v \in H^1(\Omega), \text{ a.e. } t, \tag{3.13}$$

and

$$u(0) = u_0 \quad \text{in } L^2(\Omega). \tag{3.14}$$

Lemma 3.2. *The integration by parts formula*

$$\int_0^t \big[g(s), \partial_t v(s)\big]_{\partial\Omega}\, ds = \big[g(s), v(s)\big]\big|_0^t - \int_0^t \big\langle \partial_t g(s), v(s)\big\rangle\, ds$$

holds for $g \in L^2(0,T; H^{1/2}(\partial\Omega)) \cap H^1(0,T; H^{-1/2}(\partial\Omega))$ and $v \in H^1(0,T; H^{1/2}(\partial\Omega))$.

Proof. The formula obviously holds for $g, v \in H^1(0,T; H^{1/2}(\partial\Omega))$. Using the denseness of $H^1(0,T; H^{1/2}(\partial\Omega))$ in $L^2(0,T; H^{1/2}(\partial\Omega)) \cap H^1(0,T; H^{-1/2}(\partial\Omega))$, we easily complete the proof. □

Theorem 3.3. *Suppose $g \in L^2(0,T; H^{-1/2+\theta}(\partial\Omega)) \cap H^\theta(0,T; H^{-1/2}(\partial\Omega))$, $u_0 \in H^\theta(\Omega)$, and $f \in L^2(0,T; H^{1-\theta}(\Omega)^*)$ for some $\theta \in [0,1]$. Then, there exists a unique $u \in L^2(0,T; H^{1+\theta}(\Omega)) \cap H^1(0,T; H^{1-\theta}(\Omega)^*)$ satisfying* (3.13)–(3.14) *and*

$$\|u\|_{L^2(0,T;H^{1+\theta}(\Omega))} + \|\partial_t u\|_{L^2(0,T;H^{1-\theta}(\Omega)^*)} \leqslant C\big(\|f\|_{L^2(0,T;H^{1-\theta}(\Omega)^*)} + \|u_0\|_\theta + \|g\|_{L^2(0,T;H^{-1/2+\theta}(\partial\Omega))} + \|g\|_{H^\theta(0,T;H^{-1/2}(\partial\Omega))}\big). \tag{3.15}$$

Proof. We only need to examine the cases $\theta = 0$ and $\theta = 1$, thanks to interpolation theorems.

For the case $\theta = 0$, we state without proof the following results (this case may be treated in a way similar to [6, §7.1.2, Theorems 1–3], thus detailed justifications are omitted). We choose an orthogonal basis $\{e_k\}_{k=1}^\infty$ of $H^1(\Omega)$ such that $\{e_k\}_{k=1}^\infty$ is orthonormal in $L^2(\Omega)$. For each $m = 1, 2, \ldots$, we set $V_m = \operatorname{span}\{e_1, \ldots, e_m\}$ and let $u_m = \sum_{k=1}^m d_k^{(m)} e_k \in C([0,T]; V_m)$ be the unique solution of

$$\begin{aligned}&\big[u_m', v\big] + a[u_m, v] + \lambda[u_m, v]_{\partial\Omega} = \langle f, v\rangle + \lambda[g, v]_{\partial\Omega} \quad \forall v \in V_m, \text{ a.e. } t \in (0,T),\\ &\big[u_m(0), v\big] = [u_0, v] \quad \forall v \in V_m.\end{aligned} \tag{3.16}$$

The following energy estimates hold:

$$\max_{[0,T]}\big\|u_m(t)\big\|_0 + \|u_m\|_{L^2(0,T;H^1(\Omega))} + \big\|u_m'\big\|_{L^2(0,T;H^1(\Omega)^*)} \leqslant C\big(\|f\|_{L^2(0,T;H^1(\Omega)^*)} + \|u_{0,m}\|_0 + \|g\|_{L^2(0,T;H^{-1/2}(\partial\Omega))}\big) \tag{3.17}$$

for $m = 1, 2, \ldots$. Thus, we may extract a subsequence of $\{u_m\}_{m=1}^\infty$, still denoted by $\{u_m\}_{m=1}^\infty$, such that

$$\begin{aligned}&u_m \rightharpoonup u \quad \text{weakly in } L^2\big(0,T; H^1(\Omega)\big) \quad \text{and}\\ &u_m' \rightharpoonup u' \quad \text{weakly in } L^2\big(0,T; H^1(\Omega)^*\big).\end{aligned} \tag{3.18}$$

By passage to the limits in (3.16) we see that u satisfies (3.13)–(3.14). Moreover, the solution to (3.13)–(3.14) is unique. Also, passage to the limits in (3.17) yields

$$\|u\|_{L^2(0,T;H^1(\Omega))} + \|\partial_t u\|_{L^2(0,T;H^1(\Omega)^*)} \leqslant C\big(\|f\|_{L^2(0,T;H^1(\Omega)^*)} + \|u_0\|_0 + \|g\|_{L^2(0,T;H^{-1/2}(\partial\Omega))}\big). \tag{3.19}$$

For the case $\theta = 1$, we have $g \in L^2(0,T; H^{1/2}(\partial\Omega)) \cap H^1(0,T; H^{-1/2}(\partial\Omega))$, $u_0 \in H^1(\Omega)$ and $f \in L^2(0,T; L^2(\Omega))$. Let $\{u_m\}_{m=1}^\infty$ be the sequence defined in the case of $\theta = 0$ and u denotes the weak limit of u_m in the sense of (3.18). Setting $v = u_m'(t)$ in (3.16), we obtain

$$\begin{aligned}&\big\|u_m'(t)\big\|_0^2 + \frac{1}{2}\frac{d}{dt}a\big[u_m(t), u_m(t)\big] + \frac{\lambda}{2}\frac{d}{dt}\big\|u_m(t)\big\|_{0,\partial\Omega}^2\\ &\quad = \big[f(t), u_m'(t)\big] + \lambda\big[g(t), u_m'(t)\big]_{\partial\Omega}\\ &\quad \leqslant \frac{1}{2}\big\|f(t)\big\|_0^2 + \frac{1}{2}\big\|u_m'(t)\big\|_0^2 + \lambda\big[g(t), u_m'(t)\big]_{\partial\Omega}.\end{aligned} \tag{3.20}$$

Transferring the $\|u_m'(t)\|_0^2/2$ term on the right-hand side to the left-hand side and integrating from 0 to t, we are led to

$$\frac{1}{2}\int_0^t \big\|u_m'(s)\big\|_0^2\,ds + \frac{1}{2}a\big[u_m(t), u_m(t)\big] - \frac{1}{2}a\big[u_m(0), u_m(0)\big] + \frac{\lambda}{2}\big\|u_m(t)\big\|_{0,\partial\Omega}^2 - \frac{\lambda}{2}\big\|u_m(0)\big\|_{0,\partial\Omega}^2$$

$$\leqslant \frac{1}{2}\int_0^t \|f(s)\|_0^2\,ds + \lambda\int_0^t \big[g(s), u_m'(s)\big]_{\partial\Omega}\,ds$$

$$= \frac{1}{2}\int_0^t \|f(s)\|_0^2\,ds + \lambda\big[g(t), u_m(t)\big]_{\partial\Omega} - \lambda\big[g(0), u_m(0)\big]_{\partial\Omega} - \lambda\int_0^t \big\langle \partial_t g(s), u_m(s)\big\rangle_{\partial\Omega}\,ds$$

$$\leqslant \frac{1}{2}\int_0^t \|f(s)\|_0^2\,ds + \frac{C}{\epsilon}\big\|g(t)\big\|_{-1/2,\partial\Omega}^2 + \epsilon\big\|u_m(t)\big\|_{1/2,\partial\Omega}^2 + C\big\|g(0)\big\|_{-1/2,\partial\Omega}^2$$

$$+ C\big\|u_m(0)\big\|_{1/2,\partial\Omega}^2 + C\int_0^t \big\|g'(s)\big\|_{-1/2,\partial\Omega}^2\,ds + C\int_0^t \big\|u_m(s)\big\|_{1/2,\partial\Omega}^2\,ds. \tag{3.21}$$

Since $\|u_m(t)\|_{1/2,\partial\Omega} \leqslant C\|u_m(t)\|_1$ and $a[u_m(t), u_m(t)] \geqslant C_a\|u_m(t)\|_1^2$, we may fix an ϵ such that

$$\epsilon\big\|u_m(t)\big\|_{1/2,\partial\Omega}^2 \leqslant \frac{C_a}{4}\big\|u_m(t)\big\|_1^2 \leqslant \frac{1}{4}a\big[u_m(t), u_m(t)\big].$$

By [6, Theorem 2, p. 286], we have $H^1(0,T; H^{-1/2}(\partial\Omega)) \hookrightarrow C([0,T]; H^{-1/2}(\partial\Omega))$ with the estimate

$$\big\|g(t)\big\|_{-1/2,\partial\Omega} \leqslant C\|g\|_{L^2(0,T;H^{-1/2}(\partial\Omega))} + C\|\partial_t g\|_{L^2(0,T;H^{-1/2}(\partial\Omega))} \quad \forall t \in [0,T].$$

Also,

$$\big\|u_m(0)\big\|_{0,\partial\Omega} \leqslant \big\|u_m(0)\big\|_{1/2,\partial\Omega} \leqslant C\big\|u_m(0)\big\|_1 \leqslant C\|u_0\|_1.$$

Substituting the last three relations into (3.21), we obtain

$$\frac{1}{2}\int_0^t \big\|u_m'(s)\big\|_0^2\,ds + \frac{C_a}{4}\big\|u_m(t)\big\|_1^2 + \frac{\lambda}{2}\big\|u_m(t)\big\|_{0,\partial\Omega}^2$$

$$\leqslant \frac{1}{2}\int_0^t \|f(s)\|_0^2\,ds + C\|u_0\|_1^2 + C\|g\|_{L^2(0,T;H^{1/2}(\partial\Omega))}^2 + C\|\partial_t g\|_{L^2(0,T;H^{-1/2}(\partial\Omega))}^2$$

$$+ C\int_0^t \big\|u_m(s)\big\|_1^2\,ds. \tag{3.22}$$

Applying Gronwall's inequality to (3.22), we deduce

$$\big\|u_m(t)\big\|_1^2 \leqslant \frac{1}{2}\int_0^t \|f(s)\|_0^2\,ds + C\|u_0\|_1^2 + C\|g\|_{L^2(0,T;H^{1/2}(\partial\Omega))}^2 + C\|g\|_{H^1(0,T;H^{-1/2}(\partial\Omega))}^2. \tag{3.23}$$

Combining (3.23) with (3.22), we obtain

$$\frac{1}{2}\|u_m'(t)\|_0^2\,dt \leqslant \frac{1}{2}\int_0^t \|f(s)\|_0^2\,ds + C\|u_0\|_1^2 + C\|g\|^2_{L^2(0,T;H^{1/2}(\partial\Omega))} + C\|g\|^2_{H^1(0,T;H^{-1/2}(\partial\Omega))}. \tag{3.24}$$

Equation (3.24) implies that for a subsequence we have $u_m' \rightharpoonup \partial_t u$ in $L^2(\Omega)$ so that $u \in H^1(0,T;L^2(\Omega))$ and

$$\frac{1}{2}\|u'(t)\|_0^2\,ds \leqslant \frac{1}{2}\int_0^t \|f(s)\|_0^2\,ds + C\|u_0\|_1^2 + C\|g\|^2_{L^2(0,T;H^{1/2}(\partial\Omega))} + C\|g\|^2_{H^1(0,T;H^{-1/2}(\partial\Omega))}. \tag{3.25}$$

Thus, for almost every t, $u(t)$ is the solution of the elliptic problem

$$-\operatorname{div}\big[A(\mathbf{x})\nabla u(t)\big] = f(t) - u_t(t) \quad \text{in } \Omega,$$
$$\big[A(\mathbf{x})\nabla u(t)\big]\big|_{\partial\Omega} = \lambda\big[g(t) - u(t)\big] \quad \text{on } \partial\Omega.$$

Elliptic regularity for the Neumann boundary value problem on a convex polygon (see [9,10]) yields

$$\|u(t)\|_2 \leqslant C\big(\|f(t)\|_0 + \|u_t(t)\|_0 + \|g(t)\|_{1/2,\partial\Omega}\big).$$

Combining the last estimate with (3.25), we obtain

$$\|u\|^2_{L^2(0,T;H^2(\Omega))} + \|u'\|^2_{L^2(0,T;L^2(\Omega))} \leqslant \|f\|_{L^2(0,T;L^2(\Omega))} + C\|u_0\|_1^2 + C\|g\|^2_{L^2(0,T;H^{1/2}(\partial\Omega))} + C\|g\|^2_{H^1(0,T;H^{-1/2}(\partial\Omega))}. \tag{3.26}$$

Interpolations between (3.26) and (3.19) yield (3.15). □

We next derive error estimates for semidiscrete approximations of the linear boundary value problem. Let V_h be a family of finite element subspaces of $H^1(\Omega)$ introduced in Section 2.1. The semidiscrete finite element approximations of (3.13)–(3.14) are defined as follows: seek $u_h \in H^1(0,T;V_h)$ such that

$$\begin{cases} \langle \partial_t u_h(t), v\rangle + a[u_h(t), v_h] + \lambda[u_h(t), v_h]_{\partial\Omega} = \langle f(t), v_h\rangle + \lambda\langle g(t), v_h\rangle_{\partial\Omega} \\ \quad \forall v_h \in V_h,\ \text{a.e. } t \in [0,T], \\ u_h(0) = P_h u_0, \end{cases} \tag{3.27}$$

where P_h is the $L^2(\Omega)$ projection onto V_h. Similar to [3] we may prove:

Theorem 3.4. *Assume that $g \in L^2(0,T;H^{-1/2}(\partial\Omega))$, $f \in L^2(0,T;H^1(\Omega)^*)$ and $u_0 \in L^2(\Omega)$. Let $u \in \mathcal{W}(0,T)$ be the solution of* (3.13)–(3.14) *and let $u_h \in H^1(0,T;V_h)$ be the solution of* (3.27). *Then,*

$$\|u - u_h\|_{\mathcal{W}(0,T)} \to 0 \quad \text{as } h \to 0.$$

If, in addition, $g \in L^2(0,T;H^{-1/2+\theta}(\partial\Omega)) \cap H^\theta(0,T;H^{-1/2}(\partial\Omega))$, $f \in L^2(0,T;H^{1-\theta}(\Omega)^*)$ and $u_0 \in H^\theta(\Omega)$ for some $\theta \in [0,1]$, then

$$\begin{aligned}\|u-u_h\|_{\mathcal{W}(0,T)} &\leqslant Ch^\theta\big(\|u\|_{L^2(0,T;H^{1+\theta}(\Omega))}+\|\partial_t u\|_{L^2(0,T;H^{1-\theta}(\Omega)^*)}\big)\\ &\leqslant Ch^\theta\big(\|f\|_{L^2(0,T;H^{1-\theta}(\Omega)^*)}+\|u_0\|_\theta+\|g\|_{L^2(0,T;H^{-1/2+\theta}(\partial\Omega))}\\ &\quad +\|g\|_{H^\theta(0,T;H^{-1/2}(\partial\Omega))}\big).\end{aligned}$$

Moreover, if $u\in L^2(0,T;H^{r+1}(\Omega))\cap H^1(0,T;H^{r-1}(\Omega))$ *for some* $r\in[0,k]$*, then*

$$\|u-u_h\|_{\mathcal{W}(0,T)}\leqslant Ch^r\big(\|u\|_{L^2(0,T;H^{r+1}(\Omega))}+\|\partial_t u\|_{L^2(0,T;H^{r-1}(\Omega))}\big).$$

Remark 3.5. In the special case of $\theta=1/2$, Theorem 3.4 says that if $g\in L^2(0,T;L^2(\partial\Omega))\cap H^{1/2}(0,T;H^{-1/2}(\partial\Omega))$, then $u\in L^2(0,T;H^{3/2}(\Omega))\cap H^1(0,T;H^{1/2}(\Omega)^*)$ and an $\mathcal{O}(h^{1/2})$ error estimate holds. In other words, $g\in L^2(0,T;L^2(\partial\Omega))$ alone is not sufficient to guarantee an $\mathcal{O}(h^{1/2})$ error estimate in the norm of the solution space $\mathcal{W}(0,T)$.

3.4. Embedding and trace theorems

We will establish embedding and trace theorems for $\mathcal{W}(0,T)$ and an embedding-like result for the product of functions in $\mathcal{W}(0,T)$.

Lemma 3.6 *(An embedding theorem for* $\mathcal{W}(0,T)$*).*

(i) *For every* $\epsilon>0$*, the continuous embedding*

$$\mathcal{W}(0,T)\hookrightarrow H^{1/2-\epsilon}\big(0,T;L^2(\Omega)\big)$$

holds and

$$\|w\|_{H^{1/2-\epsilon}(0,T;L^2(\Omega))}\leqslant C\|w\|_{\mathcal{W}(0,T)},$$

where C may depend on ϵ*.*

(ii) *For every* $\sigma\in(0,1/4)$*, the continuous embedding*

$$\mathcal{W}(0,T)\hookrightarrow H^\sigma\big(0,T;H^{1/2+\epsilon}(\Omega)\big)$$

holds and

$$\|w\|_{H^\sigma(0,T;H^{1/2+\epsilon}(\Omega))}\leqslant C\|w\|_{\mathcal{W}(0,T)},$$

where $0<\epsilon<(1-4\sigma)/2$ *and C may depend on* σ*.*

Proof. Let $\epsilon>0$ and $w\in\mathcal{W}(0,T)$ be given. We define E_0w to be the extension of w onto $\mathbb{R}$ by zero outside $(0,T)$, i.e., $E_0w=w$ for $t\in(0,T)$ and $E_0w=0$ otherwise. Let $\widehat{E_0w}$ denote the temporal Fourier transform of E_0w. It is easily verified that

$$2i\pi\tau\widehat{E_0w}(\tau)=\widehat{E_0\partial_t w}(\tau)+w(0)-w(0)e^{-2i\pi\tau T}$$

(a similar relation was used in [24, Theorem 2.3, Eq. (2.41), pp. 187–188].) By taking the $H^1(\Omega)^*$–$H^1(\Omega)$ duals against $\widehat{E_0w}(\tau)$, we obtain

$$\begin{aligned}&2\pi\tau\big\|\widehat{E_0w}(\tau)\big\|_0^2\\ &\quad\leqslant\big\|\widehat{E_0\partial_t w}(\tau)\big\|_{H^1(\Omega)^*}\big\|\widehat{E_0w}(\tau)\big\|_1+C\tau^{-1}\big\|w(0)\big\|_0^2+C\tau^{-1}\big\|w(T)\big\|_0^2+\pi\tau\big\|\widehat{E_0w}(\tau)\big\|_0^2.\end{aligned}$$

By virtue of the continuous embedding $\mathcal{W}(0,T)\hookrightarrow C([0,T];L^2(\Omega))$ the last estimate reduces to

$$\pi\tau\left\|\widehat{E_0w}(\tau)\right\|_0^2 \leqslant C\left\|\widehat{E_0\partial_t w}(\tau)\right\|_{H^1(\Omega)^*}^2 + C\left\|\widehat{E_0w}(\tau)\right\|_1^2 + C\tau^{-1}\|w\|_{\mathcal{W}(0,T)}^2.$$

Thus, if $|\tau| \geqslant 1$, we have

$$\pi\tau^{1-2\epsilon}\left\|\widehat{E_0w}(\tau)\right\|_0^2 \leqslant C\left\|\widehat{E_0\partial_t w}(\tau)\right\|_{H^1(\Omega)^*}^2 + C\left\|\widehat{E_0w}(\tau)\right\|_1^2 + C\tau^{-1-2\epsilon}\|w\|_{\mathcal{W}(0,T)}^2$$

so that

$$\int\limits_{|\tau|\geqslant 1} \tau^{1-2\epsilon}\left\|\widehat{E_0w}(\tau)\right\|_0^2 d\tau \leqslant C\int\limits_{|\tau|\geqslant 1}\left\|\widehat{E_0\partial_t w}(\tau)\right\|_{H^1(\Omega)^*}^2 d\tau + C\int\limits_{|\tau|\geqslant 1}\left\|\widehat{E_0w}(\tau)\right\|_1^2 d\tau + C\|w\|_{\mathcal{W}(0,T)}^2.$$

Also, it is evident that

$$\int\limits_{|\tau|<1} \tau^{1-2\epsilon}\left\|\widehat{E_0w}(\tau)\right\|_0^2 d\tau \leqslant C\int\limits_{|\tau|<1}\left\|\widehat{E_0w}(\tau)\right\|_0^2 d\tau \leqslant C\int\limits_{|\tau|<1}\left\|\widehat{E_0w}(\tau)\right\|_1^2 d\tau.$$

Combining the last two relations and using the Parseval equality, we obtain

$$\begin{aligned}\int\limits_{\mathbb{R}} \tau^{1-2\epsilon}\left\|\widehat{E_0w}(\tau)\right\|_0^2 d\tau &\leqslant C\int\limits_{\mathbb{R}}\left\|\widehat{E_0\partial_t w}(\tau)\right\|_{H^1(\Omega)^*}^2 d\tau + C\int\limits_{\mathbb{R}}\left\|\widehat{E_0w}(\tau)\right\|_1^2 d\tau + C\|w\|_{\mathcal{W}(0,T)}^2 \\ &= C\|\partial_t w\|_{L^2(0,T;H^1(\Omega)^*)}^2 + C\|w\|_{L^2(0,T;H^1(\Omega))}^2 + C\|w\|_{\mathcal{W}(0,T)}^2 \leqslant C\|w\|_{\mathcal{W}(0,T)}^2.\end{aligned}$$

Hence,

$$\|w\|_{H^{1/2-\epsilon}(0,T;L^2(\Omega))}^2 \leqslant \int\limits_{\mathbb{R}} \tau^{1-2\epsilon}\left\|\widehat{E_0w}(\tau)\right\|_0^2 d\tau \leqslant C\|w\|_{\mathcal{W}(0,T)}^2. \tag{3.28}$$

Let $\sigma \in (0, 1/4)$ and $w \in \mathcal{W}(0,T)$ be given. We wish to prove $w \in H^\sigma(0,T;H^{1/2+\epsilon}(\Omega))$ where $0 < \epsilon < (1-4\sigma)/2$. By interpolation, we have

$$\left\|\widehat{E_0w}(\tau)\right\|_{1/2+\epsilon} \leqslant C\left\|\widehat{E_0w}(\tau)\right\|_0^{1/2-\epsilon}\left\|\widehat{E_0w}(\tau)\right\|_1^{1/2+\epsilon}, \quad \text{a.e. } \tau \in \mathbb{R},$$

so that

$$\begin{aligned}\int\limits_{\mathbb{R}} |\tau|^{2\sigma}\left\|\widehat{E_0w}(\tau)\right\|_{1/2+\epsilon}^2 d\tau &\leqslant C\int\limits_{\mathbb{R}} |\tau|^{2\sigma}\left\|\widehat{E_0w}(\tau)\right\|_0^{1-2\epsilon}\left\|\widehat{E_0w}(\tau)\right\|_1^{1+2\epsilon} d\tau \\ &\leqslant C\left(\int\limits_{\mathbb{R}} |\tau|^{4\sigma/(1-2\epsilon)}\left\|\widehat{E_0w}(\tau)\right\|_0^2 d\tau\right)^{(1-2\epsilon)/2}\left(\int\limits_{\mathbb{R}}\left\|\widehat{E_0w}(\tau)\right\|_1^2 d\tau\right)^{(1+2\epsilon)/2}.\end{aligned}$$

Since $0 < 2\sigma/(1-2\epsilon) < 1/2$, it follows from (3.28) that

$$\int\limits_{\mathbb{R}} |\tau|^{4\sigma/(1-2\epsilon)}\left\|\widehat{E_0w}(\tau)\right\|_0^2 d\tau \leqslant C\|w\|_{\mathcal{W}(0,T)}^2.$$

Combining the last two estimates and applying the Parseval equality, we obtain

$$\int_{\mathbb{R}} |\tau|^{2\sigma} \big\| \widehat{E_0 w}(\tau) \big\|^2_{1/2+\epsilon} \, d\tau \leqslant C \|w\|^{(1-2\epsilon)}_{\mathcal{W}(0,T)} \|w\|^{(1+2\epsilon)}_{L^2(0,T;H^1(\Omega))} \leqslant C \|w\|^2_{\mathcal{W}(0,T)}.$$

Hence,

$$\|w\|^2_{H^\sigma(0,T;H^{1/2+\epsilon}(\Omega))} \leqslant \int_{\mathbb{R}} |\tau|^{2\sigma} \big\| \widehat{E_0 w}(\tau) \big\|^2_{1/2+\epsilon} \, d\tau \leqslant C \|w\|^2_{\mathcal{W}(0,T)}.$$

This completes the proof. □

The following theorem concerning the trace of $\mathcal{W}(0,T)$ is a direct consequence of the continuous embedding $\mathcal{W}(0,T) \hookrightarrow H^s(0,T;H^{1/2+\epsilon}(\Omega))$ and the well-known trace estimate $\|w\|_{\epsilon,\partial\Omega} \leqslant C\|w\|_{1/2+\epsilon}$ for every $\epsilon > 0$.

Theorem 3.7. *For every $\sigma \in (0, 1/4)$, the continuous embedding*

$$\mathcal{W}(0,T)|_{\partial\Omega} \hookrightarrow H^\sigma\big(0,T;H^\epsilon(\partial\Omega)\big)$$

holds with

$$\|w\|_{H^\sigma(0,T;H^\epsilon(\partial\Omega))} \leqslant C\|w\|_{\mathcal{W}(0,T)} \quad \forall w \in \mathcal{W}(0,T),$$

where $0 < \epsilon < (1-4\sigma)/2$ and C may depend on σ.

Next we prove an embedding-like result for the product of functions in $L^4(0,T;L^4(\Omega))$. This result also holds for functions in $\mathcal{W}(0,T)$ since $\mathcal{W}(0,T) \hookrightarrow L^4(0,T;L^4(\Omega))$.

Theorem 3.8. *If $1 < p < 3$ and $w, v \in L^4(0,T;L^4(\Omega))$, then there exists $\sigma > 0$ such that $(w^{p-1}v) \in L^2(0,T;H^{1-\sigma}(\Omega)^*)$ and*

$$\big\| w^{p-1} v \big\|_{L^2(0,T;H^{1-\sigma}(\Omega)^*)} \leqslant C\|v\|_{\mathcal{W}(0,T)} \|w\|^{p-1}_{L^4(0,T;L^4(\Omega))} < \infty. \tag{3.29}$$

Proof. The case $p < 1 \leqslant 2$ can be handled easily. Let $2 < p < 3$, and note that

$$\int_\Omega |w|^{p-1} |v\phi| \, dx \leqslant \big\| |w|^{p-1} \big\|_{L^{q_1}(\Omega)} \|v\|_{L^{q_2}(\Omega)} \|\phi\|_{L^{q_3}(\Omega)},$$

where $1/q_1 + 1/q_2 + 1/q_3 = 1$. Choose $q_2 = 2$, $q_1 = (2+\epsilon)$, $\epsilon > 0$. Then, $1/q_3 = (1/2) - 1/(2+\epsilon)$, i.e., $q_3 = (2+\epsilon)/\epsilon$ and

$$\int_\Omega |w|^{p-1} |v\phi| \, dx \leqslant \big\| |w|^{p-1} \big\|_{L^{(2+\epsilon)}(\Omega)} \|v\|_{L^2(\Omega)} \|\phi\|_{L^{(2+\epsilon)/\epsilon}(\Omega)}.$$

After integrating from 0 to T, standard algebraic manipulations lead to

$$\int_0^T \int_\Omega |w|^{p-1} |v\phi| \, dx$$

$$\leqslant \int_0^T \left(\int_\Omega |w|^{(p-1)(2+\epsilon)}\right)^{1/(2+\epsilon)} \|v\|_{L^2} \|\phi\|_{L^{(2+\epsilon)/\epsilon}}$$

$$= \int_0^T \|w\|^{p-1}_{L^{(p-1)(2+\epsilon)}(\Omega)} \|v\|_{L^2(\Omega)} \|\phi\|_{L^{(2+\epsilon)/\epsilon}(\Omega)}$$

$$\leqslant \left(\int_0^T \|w\|^{(p-1)k_1}_{L^{(p-1)(2+\epsilon)}(\Omega)}\right)^{1/k_1} \left(\int_0^T \|v\|^{k_2}_{L^2(\Omega)}\right)^{1/k_2} \left(\int_0^T \|\phi\|^{k_3}_{L^{(2+\epsilon)/\epsilon}(\Omega)}\right)^{1/k_3},$$

where $1/k_1 + 1/k_2 + 1/k_3 = 1$. Choose $k_1 = k_3 = 2$ and $k_2 = \infty$ to obtain

$$\int_0^T \int_\Omega |w|^{p-1} |v\phi|\, dx \leqslant \|w\|^{p-1}_{L^{2(p-1)}(0,T;L^{(p-1)(2+\epsilon)}(\Omega))} \|v\|_{L^\infty(0,T;L^2(\Omega))} \|\phi\|_{L^2(0,T;L^{(2+\epsilon)/\epsilon}(\Omega))}.$$

Note that in $\mathbb{R}^2$ the following embedding is valid:

$$H^{1-\sigma} \subset L^{2/\sigma}, \quad 0 < \sigma.$$

Now, let $\sigma, \epsilon > 0$ be chosen to satisfy

$$\frac{2}{\sigma} = \frac{2+\epsilon}{\epsilon} \quad \text{so that} \quad L^{2/\sigma}(\Omega) \equiv L^{(2+\epsilon)/\epsilon}(\Omega).$$

Therefore,

$$\int_0^T \int_\Omega |w|^{p-1} |v\phi|\, dx \leqslant \|w\|^{p-1}_{L^{2(p-1)}(0,T;L^{(p-1)(2+\epsilon)}(\Omega))} \|v\|_{L^\infty(0,T;L^2(\Omega))} \|\phi\|_{L^2(0,T;L^{2/\sigma}(\Omega))}. \tag{3.30}$$

Using $\frac{2}{\sigma} = \frac{2+\epsilon}{\epsilon}$, we compute

$$2\epsilon = 2\sigma + \sigma\epsilon \quad \text{or} \quad \epsilon = (2\sigma)/(2-\sigma).$$

Then, substituting this value of ϵ into (3.30), we obtain

$$\int_0^T \int_\Omega |w|^{p-1} |v\phi|\, dx \leqslant \|w\|^{p-1}_{L^{2(p-1)}(0,T;L^{(p-1)(4/(2-\sigma))}(\Omega))} \|v\|_{L^\infty(0,T;L^2(\Omega))} \|\phi\|_{L^2(0,T;L^{2/\sigma}(\Omega))}. \tag{3.31}$$

Here, we have used $2 + \epsilon = 2 + (2\sigma)/(2-\sigma) = (4 - 2\sigma + 2\sigma)/(2-\sigma) = 4/(2-\sigma)$.

Note that the above relation indicates that if we choose $\sigma \to 0$, then $4/(2-\sigma) \to 2$ which, for fixed $p < 3$, results in $(p-1)(4/(2-\sigma)) \to 2(p-1)$ as $\sigma \to 0$.

In particular, let $2 < p < 3$ and choose σ_0 such that $(p-1) \times \frac{4}{2-\sigma_0} = 4$, i.e., $p = 3 - \sigma_0$, and note that $2(p-1) = 2(2-\sigma_0)$. Therefore, (3.31) gives

$$\int_0^T\int_\Omega |w|^{p-1}|v\phi|\,dx \leqslant \|w\|^{p-1}_{L^{2(2-\sigma_0)}(0,T;L^4(\Omega))}\|v\|_{L^\infty(0,T;L^2(\Omega))}\|\phi\|_{L^2(0,T;L^{2/\sigma_0}(\Omega))}$$

or, equivalently, using the embedding $H^{1-\sigma_0}\subset L^{2/\sigma_0}$, we obtain

$$\int_0^T\int_\Omega |w|^{p-1}|v\phi|\,dx \leqslant \|w\|^{p-1}_{L^{2(2-\sigma_0)}(0,T;L^4(\Omega))}\|v\|_{L^\infty(0,T;L^2(\Omega))}\|\phi\|_{L^2(0,T;H^{1-\sigma_0}(\Omega))}.$$

Taking the supremum over $\phi\in L^2(0,T;H^{1-\sigma_0}(\Omega))$, we obtain the desired estimate. □

Remark 3.9. The embedding of Theorem 3.8 is also valid for every $\sigma_1\leqslant\sigma$ and moreover,

$$\begin{aligned}\big\|w^{p-1}v\big\|_{L^2(0,T;H^{1-\sigma_1}(\Omega)^*)} &\leqslant C\big\|w^{p-1}v\big\|_{L^2(0,T;H^{1-\sigma}(\Omega)^*)}\\ &\leqslant C\|v\|_{\mathcal{W}(0,T)}\|w\|^{p-1}_{L^4(0,T;L^4(\Omega))}<\infty.\end{aligned}$$

Therefore, combining the results of Theorems 3.7 and 3.8, we can always assume that $0<\sigma<(1/4)$.

3.5. Error estimates for semidiscrete approximations of the optimality system

Let $\mathcal{X}$, $\mathcal{Y}$, $\mathcal{T}$, $\mathcal{G}$, $\mathcal{X}_h$, and $\mathcal{T}_h$ be defined as in Section 3.2 where we recasted the optimality system and its semidiscrete approximations into the abstract forms (3.7) and (3.9), respectively. We now proceed to verify all assumptions of Theorem 3.1.

Given $1<p<3$ we choose $\sigma=\sigma_0$, such that $p=3-\sigma_0$ similar to Theorem 3.8. Set

$$\begin{aligned}\mathcal{Z}&=L^2\big(0,T;H^{1-\sigma_0}(\Omega)^*\big)\times\big[L^2\big(0,T;H^{-1/2+\sigma_0}(\partial\Omega)\big)\cap H^{\sigma_0}\big(0,T;H^{-1/2}(\partial\Omega)\big)\big]\\ &\quad\times H^{\sigma_0}(\Omega)\times L^2\big(0,T;H^{1-\sigma_0}(\Omega)^*\big)\times H^{\sigma_0}(\Omega)\end{aligned}$$

with the obvious graph norm. We denote the Fréchet derivative of $\mathcal{G}(u,\xi)$ with respect to (u,ξ) by $D\mathcal{G}(u,\xi)$; then we find that for $(u,\xi)\in\mathcal{X}$,

$$D\mathcal{G}(\tilde u,\tilde\xi)\cdot(v,\eta)=\big(\phi'(\tilde u)(v)+b(t,\mathbf{x})v,\lambda^2\gamma\eta,0,v,0\big)\quad\forall(\tilde u,\tilde\xi)\in\mathcal{X}. \tag{3.32}$$

Proposition 3.10 *(Verification of (3.8)). Suppose that* $1<p<3$. *Then, there exists* $C>0$ *such that*

$$\big\|D\mathcal{G}(\tilde u,\tilde\xi)\big\|_{\mathcal{L}(\mathcal{X},\mathcal{Z})}\leqslant C\big(1+\|\tilde u\|^{p-1}_{\mathcal{W}(0,T)}\big)\infty. \tag{3.33}$$

Proof. Let $(\tilde u,\tilde\xi),(v,\eta)\in\mathcal{X}=\mathcal{W}(0,T)\times\mathcal{W}(0,T)$ be given. For any p such that $1<p<3$, we can choose σ_0 as in Theorem 3.8. Therefore, for the pair p,σ_0 using Theorems 3.7 and 3.8 we obtain:

$$\begin{aligned}&\big\|D\mathcal{G}(\tilde u,\tilde\xi)\cdot(v,\eta)\big\|^2_{\mathcal{Z}}\\ &\quad=\big\|\big(\phi'(\tilde u)(v)+b(t,\mathbf{x})v,\lambda^2\gamma\eta,0,v,0\big)\big\|^2_{\mathcal{Z}}\\ &\quad=\big\|\phi'(\tilde u)(v)+b(t,\mathbf{x})v\big\|^2_{L^2(0,T;H^{1-\sigma_0}(\Omega)^*)}+\big\|\lambda^2\gamma\eta\big\|^2_{L^2(0,T;H^{-1/2+\sigma_0}(\partial\Omega))}\\ &\qquad+\big\|\lambda^2\gamma\eta\big\|^2_{H^{\sigma_0}(0,T;H^{-1/2}(\partial\Omega))}+\|0\|_{H^{\sigma_0}(\Omega)}+\|v\|^2_{L^2(0,T;H^{1-\sigma_0}(\Omega)^*)}+\|0\|_{H^{\sigma_0}(\Omega)}\end{aligned}$$

$$\leqslant C\|\tilde{u}^{p-1}v\|^2_{L^2(0,T;H^{1-\sigma_0}(\Omega)^*)} + C\|\eta\|^2_{\mathcal{W}(0,T)} + \|0\|^2_{\mathcal{W}(0,T)}$$
$$\leqslant C\big(1 + \|\tilde{u}\|^{2-\sigma_0}_{\mathcal{W}(0,T)}\big)\big(\|v\|^2_{\mathcal{W}(0,T)} + \|\eta\|^2_{\mathcal{W}(0,T)}\big).$$

Taking the supremum over all $(v,\eta) \in \mathcal{W}(0,T) \times \mathcal{W}(0,T)$ with $\|v\|^2_{\mathcal{W}} + \|\eta\|^2_{\mathcal{W}(0,T)} = 1$, we arrive at (3.33). □

Proposition 3.11 *(Verification of continuous differentiability of $\mathcal{G}$). $\mathcal{G}$ is twice continuously differentiable and $D^2\mathcal{G}$ is bounded on all bounded sets of $\mathcal{X}$.*

Proof. The second Fréchet derivative of $\mathcal{G}$ is defined by

$$D^2\mathcal{G}(\tilde{u},\tilde{\xi}) \cdot \big((v,\eta),(w,\zeta)\big) = \big(\phi''(\tilde{u})(v,w), 0, 0, 0, 0\big)$$
$$\forall (v,\eta),(w,\zeta) \in \mathcal{X} = \mathcal{W}(0,T) \times \mathcal{W}(0,T).$$

Similar to the proof of Proposition 3.10 we may show that $D^2\mathcal{G}$ is well defined, continuous and bounded on every bounded set of $\mathcal{X}$. □

Proposition 3.12 *(Verification of (3.10) and (3.11)). For every $(\tilde{f},\tilde{g},\tilde{u}_0,\tilde{\zeta},\tilde{\xi}_T) \in \mathcal{Y}$,*

$$\lim_{h\to 0}\big\|(\mathcal{T}-\mathcal{T}_h)(\tilde{f},\tilde{g},\tilde{u}_0,\tilde{\zeta},\tilde{\xi}_T)\big\|_{\mathcal{X}} \to 0. \tag{3.34}$$

Moreover,

$$\|\mathcal{T}-\mathcal{T}_h\|_{\mathcal{L}(\mathcal{Z},\mathcal{X})} \to 0 \quad \text{as } h \to 0. \tag{3.35}$$

Proof. Note that the definition of $\mathcal{T}$ consists of two *uncoupled* linear Robin boundary value problems and the definition of $\mathcal{T}_h$ consists of two *uncoupled* semidiscrete approximations of linear Robin boundary value problems. An application of Theorem 3.4 yields (3.34) and the error estimate

$$\|\tilde{u}-\tilde{u}_h\|_{\mathcal{W}(0,T)} + \|\tilde{u}-\tilde{u}_h\|_{\mathcal{W}(0,T)}$$
$$\leqslant Ch^{\sigma_0}\big(\|\tilde{f}\|_{L^2(0,T;H^{1-\sigma_0}(\Omega)^*)} + \|\tilde{g}\|_{L^2(0,T;H^{-1/2+\sigma_0}(\partial\Omega))} + \|\tilde{g}\|_{H^{\sigma_0}(0,T;H^{-1/2}(\partial\Omega))}$$
$$+ \|\tilde{u}_0\|_{\sigma_0} + \|\tilde{\zeta}\|_{L^2(0,T;H^{1-\sigma_0}(\Omega)^*)} + \|\tilde{\xi}_T\|_{\sigma_0}\big)$$
$$\leqslant Ch^{\sigma_0}\big\|(\tilde{f},\tilde{g},\tilde{u}_0,\tilde{\zeta},\tilde{\xi}_T)\big\|_{\mathcal{Z}}$$

for every $(\tilde{f},\tilde{g},\tilde{u}_0,\tilde{\zeta},\tilde{\xi}_T) \in \mathcal{Z}$. Hence, (3.35) follows from the last error estimate. □

Through Propositions 3.10–3.12 we have verified all assumptions of Theorem 3.1. Thus, by that theorem we obtain the following results.

Theorem 3.13. *Assume that $u_0 \in L^2(\Omega)$, $f \in L^2(0,T;H^1(\Omega)^*)$, $U \in L^2((0,T)\times\Omega))$ and $\Upsilon \in L^2(\Omega)$. Let $(u,\xi) \in \mathcal{W}(0,T)\times\mathcal{W}(0,T)$ be the solution of the optimality system* (2.19), (2.2) *and* (2.16)–(2.17). *Let $(u_h,\xi_h) \in H^1(0,T;V_h)\times H^1(0,T;V_h)$ be the solution of the semidiscrete optimality system* (3.6). *Then*

$$\|u-u_h\|_{\mathcal{W}(0,T)} + \|\xi-\xi_h\|_{\mathcal{W}(0,T)} \to 0 \quad \text{as } h \to 0.$$

If, in addition, $u,\xi \in L^2(0,T;H^{r+1}(\Omega)) \cap H^1(0,T;H^{r-1}(\Omega))$ for some $r \in [0,k]$, then

$$\|u-u_h\|_{\mathcal{W}(0,T)}+\|\xi-\xi_h\|_{\mathcal{W}(0,T)}\leqslant Ch^r\big(\|u\|_{L^2(0,T;H^{r+1}(\Omega))}+\|\xi\|_{L^2(0,T;H^{r+1}(\Omega))}+\|\partial_t u\|_{L^2(0,T;H^{r-1}(\Omega))}+\|\partial_t u\|_{L^2(0,T;H^{r-1}(\Omega))}\big).$$

References

[1] V. Alekseev, V. Tikhomirov, S. Fomin, Optimal Control, Consultants Bureau, New York, 1987.

[2] F. Brezzi, J. Rappaz, P.-A. Raviart, Finite dimensional approximation of nonlinear problems. Part I: Branches of nonsingular solutions, Numer. Math. 36 (1980) 1–25.

[3] K. Chrysafinos, L. Hou, Error estimates for semidiscrete finite element approximations of linear and semilinear parabolic equations under minimal regularity assumptions, SIAM J. Numer. Anal. 40 (2002) 282–306.

[4] P. Ciarlet, The Finite Element Method for Elliptic Problems, North-Holland, Amsterdam, 1978.

[5] M. Crouzeix, J. Rappaz, On Numerical Approximations in Bifurcation Theory, Masson/Springer-Verlag, Paris/Berlin, 1990.

[6] L. Evans, Partial Differential Equations, Amer. Math. Soc., Providence, RI, 1998.

[7] D. French, J. King, Approximation of an elliptic control problem by the finite element method, Numer. Funct. Anal. Optim. 12 (1991) 299–314.

[8] A. Fursikov, Optimal Control for Distributed Systems. Theory and Applications, Amer. Math. Soc., Providence, RI, 2000.

[9] V. Girault, P.-A. Raviart, Finite Element Methods for Navier–Stokes Equations, Springer-Verlag, New York, 1986.

[10] P. Grisvard, Singularities in Boundary Value Problems, Mason, Paris, 1992.

[11] M. Gunzburger, Perspectives in Flow Control and Optimization, Advances in Design and Control, SIAM, Philadelphia, 2003.

[12] M. Gunzburger, L. Hou, T. Svobodny, Analysis and finite element approximation of optimal control problems for the stationary Navier–Stokes equations with distributed and Neumann controls, Math. Comp. 57 (1991) 123–151.

[13] M. Gunzburger, L. Hou, T. Svobodny, Analysis and finite element approximation of optimal control problems for the stationary Navier–Stokes equations with Dirichlet controls, RAIRO Model. Math. Anal. Numer. 25 (1991) 711–748.

[14] M. Gunzburger, L. Hou, Treating inhomogeneous essential boundary conditions in finite element methods and the calculation of boundary stresses, SIAM J. Numer. Anal. 29 (1992) 390–424.

[15] M. Gunzburger, L. Hou, Finite-dimensional approximation of a class of constrained nonlinear optimal control problems, SIAM. J. Control Optim. 34 (1996) 1001–1043.

[16] L. Hou, S. Ravindran, A penalized Neumann control approach for solving an optimal Dirichlet control problem for the Navier–Stokes equations, SIAM J. Control Optim. 36 (1998) 1795–1814.

[17] G. Knowles, Finite element approximation of parabolic time optimal control problems, SIAM J. Control Optim. 20 (1982) 414–427.

[18] I. Lasiecka, Rietz–Galerkin approximation of the time optimal boundary control problem for parabolic systems with Dirichlet boundary data, SIAM J. Control Optim. 30 (1984) 477–500.

[19] I. Lasiecka, R. Triggiani, Control Theory for Partial Differential Equations, Cambridge Univ. Press, 2000.

[20] W.-B. Liu, N. Yan, A posteriori error estimates for optimal control problems governed by parabolic equations, Numer. Math. 93 (2003) 497–521.

[21] J.-L. Lions, Quelques méthodes de résolution des problèmes aux limites non linéaires, Dunod, Paris, 1968.

[22] J.L. Lions, E. Magenes, Non Homogeneous Boundary Value Problems and Applications, Springer-Verlag, Berlin, 1972.

[23] K. Malanowski, Convergence of approximations vs. regularity of solutions for convex, control-constrained optimal control problems, Appl. Math. Optim. 8 (1981) 69–95.

[24] R. Temam, Navier–Stokes Equations, Amer. Math. Soc. Chelsea Publishing, Providence, RI, 2001.

[25] V. Thomee, Galerkin Finite Element Methods for Parabolic Equations, Springer-Verlag, Berlin, 1997.

[26] F. Troltzsch, Semidiscrete Ritz–Galerkin approximation of nonlinear parabolic boundary control problems, Internat. Ser. Numer. Math. 111 (1993) 57–68.

[27] F. Troltzsch, Semidiscrete Ritz–Galerkin approximation of nonlinear parabolic boundary control problems – Strong convergence of optimal controls, Appl. Math. Optim. 29 (1994) 309–329.

[28] R. Winther, Initial value methods for parabolic control problems, Math. Comp. 34 (1980) 115–125.