Sussex Research Online

Finite element convergence for the time-dependent Joule heating problem with mixed boundary conditions

Article (Accepted Version)

Jensen, Max, Målqvist, Axel and Persson, Anna (2022) Finite element convergence for the time-dependent Joule heating problem with mixed boundary conditions. IMA Journal of Numerical Analysis, 42 (1). pp. 199-228. ISSN 0272-4979

This version is available from Sussex Research Online: http://sro.sussex.ac.uk/id/eprint/93508/

This document is made available in accordance with publisher policies and may differ from the published version or from the version of record. If you wish to cite this item you are advised to consult the publisher's version. Please see the URL above for details on accessing the published version.

Copyright and reuse:
Sussex Research Online is a digital repository of the research output of the University.

Copyright and all moral rights to the version of the paper presented here belong to the individual author(s) and/or other copyright owners. To the extent reasonable and practicable, the material made available in SRO has been checked for eligibility before being made available.

Copies of full text items generally can be reproduced, displayed or performed and given to third parties in any format or medium for personal research or study, educational, or not-for-profit purposes without prior permission or charge, provided that the authors, title and full bibliographic details are credited, a hyperlink and/or URL is given for the original metadata page and the content is not changed in any way.

http://sro.sussex.ac.uk

IMA Journal of Numerical Analysis (2020) Page 1 of 28
doi:10.1093/imanum/drn000

Finite element convergence for the time-dependent Joule heating problem with mixed boundary conditions

MAX JENSEN†,
Department of Mathematics, University of Sussex, Brighton BN1 9QH, United Kingdom

AXEL MÅLQVIST‡,
Department of Mathematical Sciences, Chalmers University of Technology and University of Gothenburg, SE-412 96 Gothenburg, Sweden

AND

ANNA PERSSON§
Department of Mathematics, KTH Royal Institute of Technology, SE-100 44 Stockholm, Sweden

[Received on 4 August 2020]

We prove strong convergence for a large class of finite element methods for the time-dependent Joule heating problem in three spatial dimensions with mixed boundary conditions on Lipschitz domains. We consider conforming subspaces for the spatial discretization and the backward Euler scheme for the temporal discretization. Furthermore, we prove uniqueness and higher regularity of the solution on creased domains and additional regularity in the interior of the domain. Due to a variational formulation with a cut-off functional the convergence analysis does not require a discrete maximum principle, permitting approximation spaces suitable for adaptive mesh refinement, responding to the the difference in regularity within the domain.

Keywords: Joule heating problem, Thermistor, Finite element convergence, Nonsmooth domains, Mixed boundary conditions, Regularity.

1. Introduction

The time-dependent Joule heating problem is a coupled non-linear elliptic-parabolic system of the form

$$\dot{u} - \Delta u = \sigma(u)|\nabla\varphi|^2, \quad \nabla\cdot\sigma(u)\nabla\varphi = 0, \tag{1.1}$$

where u denotes the temperature and φ the electric potential. It models the heat flow generated when an electric current is passed through a conductor. In applications the electric potential is typically only applied to smaller parts of the boundary, for instance through electric pads. To model such problems properly mixed boundary conditions are needed, see e.g. Henneken *et al.* (2006).

The Joule heating problem has been studied both in a theoretical context Cimatti (1992); Antontsev & Chipot (1994); Yuan & Liu (1994); Meinlschmidt *et al.* (2017), focusing on the well-posedness of

†Email: m.jensen@sussex.ac.uk
‡Email: axel@chalmers.se
§Corresponding author. Email: annaper3@kth.se

© The author 2020. Published by Oxford University Press on behalf of the Institute of Mathematics and its Applications. All rights reserved.

(1.1), and from a numerical point of view Elliott & Larsson (1995); Akrivis & Larsson (2005); Gao (2014); Li *et al.* (2014), focusing on convergence (with rate) of numerical solutions to (1.1). There are also several works on the stationary version of the problem, see, for instance, Howison *et al.* (1993); Holst *et al.* (2010); Jensen & Målqvist (2013), and references therein.

The main issue with the system (1.1) is the low regularity of the source term $\sigma(u)|\nabla\varphi|^2$. In one and two dimensions this does not lead to a problem. However, in three dimensions this term is not in H^{-1} and the problem does not fit into the classical variational framework for PDEs. In Antontsev & Chipot (1994) this issue is resolved by rewriting the source term using the equation for φ (see also Howison *et al.* (1993) for the stationary case). With this formulation existence of a solution in $L_2(H^1)$ is proved. However, to derive convergence for finite element approximations additional regularity of the solution is usually required, see Elliott & Larsson (1995); Akrivis & Larsson (2005). Typically, sufficient regularity in three dimensions cannot be proved, but needs to be assumed. To the authors' knowledge, there is no numerical analysis of this problem under more realistic assumptions on the domain (Lipschitz in three spatial dimensions) and the boundary conditions (mixed). This setting shall be the focus of this paper.

The purpose of this paper is to prove the strong convergence of finite element approximations of (1.1) on Lipschitz domains in three spatial dimensions with mixed boundary conditions. A challenge is to avoid the need for a discrete maximum principle and the associated restrictive mesh conditions, see e.g. Holst *et al.* (2010), because a direct energy argument only delivers L^1-control on the critical $|\nabla\varphi|^2$ term in (1.1). In our analysis this is achieved by introducing a variational formulation with a cut-off functional, extending Jensen & Målqvist (2013). The analysis presented in this paper covers finite element methods of any order that are conforming in space and piecewise constant in time, satisfying a backward Euler scheme. The choice of approximation spaces only needs to ensure the stability of the L^2 projection in the H^1-norm, which holds for a large class of non-uniform meshes, see Bank & Yserentant (2014).

Having arrived at only mild mesh conditions, we find the Joule heating problem with mixed boundary conditions well suited for adaptive mesh refinement. In the setting of so called *creased domains*, see Mitrea & Mitrea (2007); Brown (1994), which essentially means that the angle between the Neumann and Dirichlet part of the boundary is less than π, we prove uniqueness and additional regularity of the solution. This result combines the regularity for the Poisson equation on creased domains in Mitrea & Mitrea (2007) with the results for parabolic systems in Hieber & Rehberg (2008). We emphasize that this result holds for nonsmooth domains including re-entrant corners. The assumption of creased domains mainly affects the way the Neumann and Dirichlet boundary parts can meet. Higher regularity and uniqueness is also proven in Yuan & Liu (1994), but for smoother domains. Furthermore, the additional regularity we obtain for φ, namely $\varphi \in L_{2q/(q-3)}(W^1_q)$ for some $q > 3$, is in line with the sufficient condition for uniqueness established in Antontsev & Chipot (1994). Importantly, we can show higher regularity, in some cases C^∞, in the interior of the domain. To exploit the difference in regularity within the domain we equip the Joule heating problem with a goal functional to examine duality-based additive mesh refinement in the numerical experiments.

The paper is outlined as follows: In Section 2 we formulate the problem of interest and introduce some notation. Section 3 is devoted to the analysis of semi-discrete methods and Section 4 to fully discrete methods. In Section 5 we prove additional regularity and uniqueness of the solution. In Section 6 we present some numerical examples that confirm the convergence results and investigate adaptive mesh refinements.

2. Variational formulations and weak solutions

In this section we introduce two variational formulations, one "classical", see (2.2) below, and one based on a cut-off functional, see (2.4) below. We prove that these two are equivalent, that is, that they have the same set of solutions. The latter formulation is preferable when working with finite element discretizations of the problem, since we avoid using a discrete maximum principle, see Section 3 and Section 4.

2.1 *Problem formulation and notation*

Let D_t denote the time derivative $\frac{\partial}{\partial t}$ and $\Omega \subseteq \mathbb{R}^3$ be a domain describing the body of a conductor. Let $u\colon \Omega \times [0,T] \to \mathbb{R}$ denote the temperature inside the conductor, $\varphi\colon \Omega \times [0,T] \to \mathbb{R}$ the electric potential, and $\sigma\colon \mathbb{R} \to \mathbb{R}_+$ the electric conductivity. Furthermore, we use Γ_D^u and Γ_N^u to denote the Dirichlet and Neumann boundary for u and $\overline{\Gamma_D^u} \cup \overline{\Gamma_N^u} = \partial\Omega$. Analogously, we define Γ_D^φ and Γ_N^φ for φ. With this notation, the time-dependent Joule heating problem is given by the following nonlinear elliptic-parabolic system

$$
\begin{aligned}
D_t u - \Delta u &= \sigma(u)|\nabla\varphi|^2, && \text{in } \Omega \times (0,T), && \text{(2.1a)}\\
\nabla\cdot(\sigma(u)\nabla\varphi) &= 0, && \text{in } \Omega \times (0,T), && \text{(2.1b)}\\
u &= g_u, && \text{on } \Gamma_D^u \times (0,T), && \text{(2.1c)}\\
\varphi &= g_\varphi, && \text{on } \Gamma_D^\varphi \times (0,T), && \text{(2.1d)}\\
n\cdot\nabla u &= 0, && \text{on } \Gamma_N^u \times (0,T), && \text{(2.1e)}\\
n\cdot\nabla\varphi &= 0, && \text{on } \Gamma_N^\varphi \times (0,T), && \text{(2.1f)}\\
u(\cdot,0) &= u_0, && \text{in } \Omega. && \text{(2.1g)}
\end{aligned}
$$

Let $W_p^k(\Omega)$ denote the classical range of Sobolev spaces and define

$$
W_p^k(\Omega;\Gamma_D^u) := \{v \in W_p^k(\Omega) : v|_{\Gamma_D^u} = 0\}, \quad \text{for } k > 1/p.
$$

The space $W_p^k(\Omega;\Gamma_D^\varphi)$ is defined analogously and H^1 is used to denote W_2^1. We also use V^* for the dual space to V. Furthermore, we adopt the notation $L_p(0,T;V)$ for the Bochner space with norm

$$
\begin{aligned}
\|v\|_{L_p(0,T;V)} &= \Big(\int_0^T \|v\|_V^p \,\mathrm{d}t\Big)^{1/p}, \quad 1 \leqslant p < \infty,\\
\|v\|_{L_\infty(0,T;V)} &= \operatorname*{ess\,sup}_{0\leqslant t\leqslant T} \|v\|_V,
\end{aligned}
$$

where V is a Banach space equipped with the norm $\|\cdot\|_V$. The notation $v \in H^1(0,T;V)$ is used to denote $v, D_t v \in L_2(0,T;V)$. Finally, $C_b(\Omega)$ is the space of bounded continuous functions.

2.2 *Classical variational formulation*

To this end we make the following assumptions on the domain and the data.

(A1) $\Omega \subseteq \mathbb{R}^3$ is a bounded domain with Lipschitz boundary, $\operatorname{meas}(\Gamma_D^u) > 0$, and $\operatorname{meas}(\Gamma_D^\varphi) > 0$.

(A2) $g_u \in L_2(0,T;H^1(\Omega)) \cap H^1(0,T;H^1(\Omega)^*)$ and there are points $0 = t_0 < t_1 < \cdots < t_K = T$, $K \in \mathbb{N}$, such that

$$D_t g_u \in C_b([t_i, t_{i+1}); H^1(\Omega)^*),$$
$$g_\varphi \in C_b([t_i, t_{i+1}); W_3^1(\Omega) \cap L_\infty(\Omega)),$$

on each subinterval $[t_i, t_{i+1})$.

(A3) $u_0 \in L_2(\Omega)$.

(A4) $\sigma \in C^1(\mathbb{R})$, Lipschitz continuous, and $0 < \sigma_\circ \leqslant \sigma(x) \leqslant \sigma^\circ < \infty$, $\forall x \in \mathbb{R}$.

A weak solution to the Joule heating problem (2.1) is a pair $(u, \varphi) = (g_u + \tilde{u}, g_\varphi + \tilde{\varphi})$ such that

$$(\tilde{u}, \tilde{\varphi}) \in L_2(0,T;H^1(\Omega;\Gamma_D^u)) \cap H^1(0,T;H^1(\Omega;\Gamma_D^u)^*) \times L_2(0,T;H^1(\Omega,\Gamma_D^\varphi))$$

and for a.e. $t \in (0,T]$

$$\langle D_t u, v\rangle + \langle \nabla u, \nabla v\rangle = \langle \sigma(u)|\nabla\varphi|^2, v\rangle, \tag{2.2a}$$
$$\langle \sigma(u)\nabla\varphi, \nabla w\rangle = 0, \tag{2.2b}$$
$$\langle u(0), z\rangle = \langle u_0, z\rangle, \tag{2.2c}$$

for all $(v,w) \in W_\infty^1(\Omega;\Gamma_D^u) \times H^1(\Omega;\Gamma_D^\varphi)$ and $z \in L_2(\Omega)$, see, for instance, Cimatti (1992). Note that $\langle\cdot,\cdot\rangle$ is used to denote both the inner product in L_2 and the duality bracket. The choice of spaces guarantees $\sigma(u)|\nabla\varphi|^2 \in L_1(0,T;L_1(\Omega))$ so that the right-hand side in (2.2a) is well-defined for all $v \in W_\infty^1(\Omega;\Gamma_D^u)$.

Throughout the text we adopt the notational convention that for a function $\flat$ one understands $\tilde{\flat} = \flat - g_\varphi$ if $\flat$ is a Greek letter and $\tilde{\flat} = \flat - g_u$ if $\flat$ is a Latin letter.

REMARK 2.1 In some works, e.g. Roubíček (2013), the notion of strong (instead of weak) solution is used when the equation is satisfied almost everywhere in time.

The following lemma provides a maximum principle for $\varphi(x,t)$.

LEMMA 2.1 If (u,φ) is a solution to (2.2), then $g_\circ \leqslant \varphi(x,t) \leqslant g^\circ$ for a.e. $(x,t) \in \bar{\Omega} \times [0,T]$, where

$$g^\circ := \max_{(x,t)\in\Gamma_D^\varphi\times[0,T]} g_\varphi(x,t), \quad g_\circ := \min_{(x,t)\in\Gamma_D^\varphi\times[0,T]} g_\varphi(x,t),$$

Proof. Define $\chi = \max(0, \varphi - g^\circ) \in L_2(0,T;H^1(\Omega;\Gamma_D^\varphi))$ and choose $w = \chi(t)$ in (2.2b) and integrate from 0 to T. Then

$$0 = \int_0^T \langle \sigma(u)\nabla\varphi, \nabla\chi\rangle \,\mathrm{d}t = \int_0^T \langle \sigma(u)\nabla(\varphi - g^\circ), \nabla\chi\rangle \,\mathrm{d}t$$
$$= \int_0^T \int_{\mathrm{supp}(\chi)\cap\Omega} \sigma(u)\nabla\chi\cdot\nabla\chi \,\mathrm{d}x\,\mathrm{d}t = \int_0^T \langle \sigma(u)\nabla\chi, \nabla\chi\rangle \,\mathrm{d}t.$$

Using $\sigma(u) \geqslant \sigma_\circ$ and the Poincaré-Friedrichs inequality we get $\int_0^T \|\chi\|^2 \,\mathrm{d}t = 0$ and we deduce $\varphi \leqslant g^\circ$. A similar argument using $g_\circ$ proves $\varphi \geqslant g_\circ$. This gives $\varphi \in L_\infty(0,T;L_\infty(\Omega))$. □

In one and two spatial dimensions the formulation (2.2) is suitable for proving existence of a solution, see, e.g. Cimatti (1992); Elliott & Larsson (1995). However, because of the low regularity of the right-hand side in (2.2a) this strategy does not apply to the three dimensional setting. To overcome this difficulty, it can be proved that due to (2.1b), see, for instance, Antontsev & Chipot (1994); Howison *et al.* (1993),

$$\sigma(u)|\nabla\varphi|^2 = \nabla\cdot(\sigma(u)\varphi\nabla\varphi), \tag{2.3}$$

and from Lemma 2.1 it follows that $\nabla\cdot(\sigma(u)\varphi\nabla\varphi) \in L_2(0,T;H^1(\Omega,\Gamma_D^u)^*)$. With this right-hand side it is now possible to use Schauder's fixed point theorem to prove existence of a solution also in three dimensions.

THEOREM 2.1 There exists a solution (u,φ) to (2.2). If $\nabla\varphi \in L_{\frac{2q}{q-3}}(0,T;L_q(\Omega))$ for some $q>3$, then the solution is unique.

Proof. This follows by adapting the fixed point argument in Antontsev & Chipot (1994, Theorem 2.2) to mixed boundary conditions. The proof uses identity (2.3) and Schauder's fixed point theorem on the space $L_2(0,T;L_2(\Omega))$. More precisely, we consider the mapping $F : L_2(0,T;L_2(\Omega)) \to L_2(0,T;L_2(\Omega))$ where $y = F(s)$ is the solution to

$$\begin{aligned}\langle D_t y, v\rangle + \langle\nabla y,\nabla v\rangle &= \langle\nabla\cdot(\sigma(s)\psi\nabla\psi), v\rangle,\\ \langle\sigma(s)\nabla\psi,\nabla w\rangle &= 0\\ \langle y(0), z\rangle &= \langle u_0, z\rangle,\end{aligned}$$

for all $(v,w) \in H^1(\Omega;\Gamma_D^u)\times H^1(\Omega;\Gamma_D^\varphi)$ and $z \in L_2(\Omega)$. It is clear, via (2.3) and the fact that $W^1_\infty(\Omega;\Gamma_D^u)$ is dense in $H^1(\Omega;\Gamma_D^u)$, that a fixed point to F solves (2.2). To prove that F satisfies the conditions of Schauder's fixed point theorem on some ball B_R we may now follow Antontsev & Chipot (1994, Theorem 2.2). The mixed boundary conditions only affect the definition of the space $V \subset H^1(\Omega)$, that is, the functions in V in our case only vanish on $\Gamma_D \subset \partial\Omega$.

The uniqueness follows from Antontsev & Chipot (1994, Theorem 4.1), also by first adapting the argument to mixed boundary conditions. □

2.3 *Variational formulation with cut-off*

In this paper we are interested in proving convergence of finite element approximations. For this purpose, we propose a variational formulation based on a cut-off functional to avoid using a discrete maximum principle. The cut-off functional was introduced for the stationary problem in Jensen & Målqvist (2013), and is defined as

$$\lceil f\rceil := \min\{\max\{f+g_\varphi, a\}, b\} - g_\varphi.$$

for some fixed $a,b \in \mathbb{R}$ with $a \leqslant g_\circ$ and $b \geqslant g^\circ$. Note that min and max are taken over both space and time $\Omega\times[0,T]$ and we have $a - g_\varphi \leqslant \lceil f\rceil \leqslant b - g_\varphi$.

To introduce the new weak formulation we define

$$\begin{aligned}X &:= L_2(0,T;H^1(\Omega;\Gamma_D^u)) \cap H^1(0,T;H^1(\Omega;\Gamma_D^u)^*) \times L_2(0,T;H^1(\Omega,\Gamma_D^\varphi)),\\ Y &:= H^1(\Omega;\Gamma_D^u)\times H^1(\Omega;\Gamma_D^\varphi).\end{aligned}$$

Using these spaces, a weak solution to the system (2.1) is a pair $(u,\varphi) = (g_u + \tilde{u}, g_\varphi + \tilde{\varphi})$ such that $(\tilde{u},\tilde{\varphi}) \in X$ and for a.e. $t \in (0,T]$

$$\langle D_t u, v\rangle + \langle \nabla u, \nabla v\rangle = -\langle \sigma(u)\lceil\tilde{\varphi}\rceil\nabla\varphi, \nabla v\rangle + \langle \sigma(u)\nabla\varphi\cdot\nabla g_\varphi, v\rangle, \tag{2.4a}$$

$$\langle \sigma(u)\nabla\varphi, \nabla w\rangle = 0, \tag{2.4b}$$

$$\langle u(0), z\rangle = \langle u_0, z\rangle, \tag{2.4c}$$

for all $(v,w) \in Y$ and $z \in L_2(\Omega)$.

LEMMA 2.2 The set of solutions which satisfy (2.4) is equal to the set of solutions to (2.2). In particular, the right-hand side in (2.4a) defines an element in $L_2(0,T;(H^1(\Omega;\Gamma_D^u))^*)$.

Proof. The identity

$$\langle \sigma(u)|\nabla\varphi|^2, v\rangle = -\langle \sigma(u)\tilde{\varphi}\nabla\varphi, \nabla v\rangle + \langle \sigma(u)\nabla\varphi\cdot\nabla g_\varphi, v\rangle,$$

for $v \in W_\infty^1(\Omega)$ and a.e. $t \in [0,T]$ follows by choosing $w = (\varphi(t) - g_\varphi(t))v$ in (2.2b), see Howison *et al.* (1993, Lemma 1) for the stationary case. The definition of the cut-off functional and the maximum principle for φ in Lemma 2.1 implies that $\lceil\tilde{\varphi}\rceil = \tilde{\varphi}$. The larger space of test functions does not affect the set of solutions since $W_\infty^1(\Omega;\Gamma_D^u)$ is dense in $H^1(\Omega;\Gamma_D^u)$.

Furthermore, the right-hand side in (2.4a) satisfies the following bound

$$\begin{aligned}|-\langle \sigma(u)\lceil\tilde{\varphi}\rceil\nabla\varphi,\nabla v\rangle + \langle \sigma(u)\nabla\varphi\cdot\nabla g_\varphi, v\rangle| &\leqslant C(\sigma, g_\varphi)\|\nabla\varphi\|_{L_2(\Omega)}\|\nabla v\|_{L_2(\Omega)}\\ &\quad + \sigma^\circ\|\nabla\varphi\|_{L_2(\Omega)}\|\nabla g_\varphi\|_{L_3(\Omega)}\|v\|_{L_6(\Omega)},\end{aligned}$$

where the Sobolev embedding in $\mathbb{R}^3$ gives $\|v\|_{L_6(\Omega)} \leqslant C\|v\|_{H^1(\Omega)}$. Hence

$$\begin{aligned}&\int_0^T \|\nabla\cdot(\sigma(u)\lceil\tilde{\varphi}\rceil\nabla\varphi) + \sigma(u)\nabla\varphi\cdot\nabla g_\varphi\|^2_{H^1(\Omega;\Gamma_D^u)^*}\,\mathrm{d}t\\ &\quad\leqslant C(\sigma, g_\varphi)\big(\|\nabla\varphi\|_{L_2(0,T;L_2(\Omega))} + \|\nabla\varphi\|_{L_2(0,T;L_2(\Omega))}\|\nabla g_\varphi\|_{L_\infty(0,T;L_3(\Omega))}\big),\end{aligned}$$

where $\|\nabla g_\varphi\|_{L_\infty(0,T;L_3(\Omega))}$ is bounded due to assumption (A2), so the right-hand side defines an element in $L_2(0,T;H^1(\Omega;\Gamma_D^u)^*)$. □

3. Semidiscrete methods

In this section we analyze spatially semidiscrete Galerkin methods. We prove existence and uniqueness of semidiscrete solutions and strong convergence to a weak solution satisfying (2.4).

3.1 *Semidiscrete formulation*

Let $\{V_m^u\}_{m\in\mathbb{N}}$ and $\{V_m^\varphi\}_{m\in\mathbb{N}}$ be hierarchical families of finite-dimensional subspaces, whose unions are dense in $H^1(\Omega;\Gamma_D^u)$ and $H^1(\Omega;\Gamma_D^\varphi)$, respectively and define

$$X_m := \{v \in C(0,T;V_m^u) : v|_{[t_i,t_{i+1})} \in C^1(t_i,t_{i+1};V_m^u)\ \forall i\} \times L_\infty(0,T;V_m^\varphi).$$

Typically, V_m^u and V_m^φ are finite element spaces corresponding to a family of meshes $\{\mathscr{T}_m\}_{m\in\mathbb{N}}$. For instance, one may choose Lagrangian finite elements or conforming *hp*-finite elements, see also the numerical examples in Section 6.

We make the following additional assumption on V_m^u;

(A5) Let V_m^u be of a form such that the L_2-projection P_m onto V_m^u is stable, uniformly in m, in the H^1-norm.

In the case when V_m^u is a finite element space, we refer to Bank & Yserentant (2014) and references therein, where the H^1-stability of the L_2-projection is proved for a large class of (non-uniform) meshes in three spatial dimensions.

In the subsequent sections we let C_m denote a generic constant that depends on the discretization m, for instance, the mesh size h_m.

A semidiscrete Galerkin solution is a pair $(u_m, \varphi_m) = (g_u + \tilde{u}_m, g_\varphi + \tilde{\varphi}_m)$ such that $(\tilde{u}_m, \tilde{\varphi}_m) \in X_m$ and for a.e. $t \in (0,T]$

$$\begin{aligned} \langle D_t u_m, v\rangle + \langle \nabla u_m, \nabla v\rangle &= -\langle \sigma(u_m)\lceil\tilde{\varphi}_m\rceil \nabla\varphi_m, \nabla v\rangle \\ &\quad + \langle \sigma(u_m)\nabla\varphi_m \cdot \nabla g_\varphi, v\rangle, \end{aligned} \tag{3.1a}$$

$$\langle \sigma(u_m)\nabla\varphi_m, \nabla w\rangle = 0, \tag{3.1b}$$

$$\langle u_m(0), z\rangle = \langle u_0, z\rangle, \tag{3.1c}$$

for all $(v,w) \in V_m^u \times V_m^\varphi$ and $z \in V_m^u$.

REMARK 3.1 Recall the bound of the cut-off functional; $a - g_\varphi \leqslant \lceil\tilde{\varphi}_m\rceil \leqslant b - g_\varphi$. This uniform boundedness of $\lceil\tilde{\varphi}_m\rceil$ in m will allow us to consider the limit of $\langle \sigma(u_m)\lceil\tilde{\varphi}_m\rceil\nabla\varphi_m, \nabla v\rangle$ as $m \to \infty$ without appealing to a discrete maximum principle.

LEMMA 3.1 A solution to (3.1) fulfils the following bounds

$$\|\nabla\tilde{\varphi}_m\|_{L_\infty(0,T;L_2(\Omega))} \leqslant C(\sigma, g_\varphi), \tag{3.2}$$

$$\|\tilde{u}_m(T)\|_{L_2(\Omega)}^2 + \int_0^T \|\nabla\tilde{u}_m\|_{L_2(\Omega)}^2 \,\mathrm{d}t \leqslant C(u_0, \sigma, g_u, D_t g_u, g_\varphi), \tag{3.3}$$

$$\int_0^T \|D_t\tilde{u}_m(t)\|_{H^1(\Omega;\Gamma_D^u)^*}^2 \leqslant C(u_0, \sigma, g_u, D_t g_u, g_\varphi). \tag{3.4}$$

Proof. By choosing $w = \tilde{\varphi}_m(t)$ in (3.1b) we can prove

$$\|\nabla\tilde{\varphi}_m(t)\|_{L_2(\Omega)}^2 \leqslant C(\sigma)\|\nabla g_\varphi(t)\|_{L_2(\Omega)}^2,$$

and (3.2) follows by using (A2) and (A4).

By choosing $v = \tilde{u}_m(t)$ in (3.1a) and integrating from 0 to T we have

$$\begin{aligned} &\int_0^T \langle D_t\tilde{u}_m, \tilde{u}_m\rangle \,\mathrm{d}t + \int_0^T \|\nabla\tilde{u}_m\|_{L_2(\Omega)}^2 \,\mathrm{d}t \\ &= -\int_0^T \langle D_t g_u, \tilde{u}_m\rangle \,\mathrm{d}t - \int_0^T \langle \nabla g_u, \nabla\tilde{u}_m\rangle \,\mathrm{d}t - \int_0^T \langle \sigma(u_m)\lceil\tilde{\varphi}_m\rceil\nabla\varphi_m, \nabla\tilde{u}_m\rangle \,\mathrm{d}t \\ &\quad + \int_0^T \langle \sigma(u_m)\nabla\varphi_m \cdot \nabla g_\varphi, \tilde{u}_m\rangle \,\mathrm{d}t \\ &= I + II + III + IV. \end{aligned}$$

Using the Cauchy-Schwarz, Poincaré, and Young's (weighted) inequality we get

$$I + II \leqslant \frac{1}{4}\int_0^T \|\nabla\tilde{u}_m\|_{L_2(\Omega)}^2 \,\mathrm{d}t + C\int_0^T \|D_t g_u\|_{(H^1(\Omega;\Gamma_D^u))^*}^2 + \|\nabla g_u\|_{L_2(\Omega)}^2 \,\mathrm{d}t, \tag{3.5}$$

and

$$III + IV \leqslant \frac{1}{4}\int_0^T \|\nabla \tilde{u}_m\|^2_{L_2(\Omega)}\,\mathrm{d}t + C(\sigma, g_\varphi)\Big(\|g_\varphi\|^2_{L_\infty(0,T;W^1_3(\Omega))} + \int_0^T \|\nabla \varphi_m\|^2_{L_2(\Omega)}\,\mathrm{d}t\Big), \tag{3.6}$$

where we used Sobolev embeddings as in the proof of Lemma 2.2. We can now use (3.2) to bound the last term on the right-hand side. Finally, using (3.1c) we have

$$\begin{aligned} 2\int_0^T \langle D_t \tilde{u}_m, \tilde{u}_m\rangle\,\mathrm{d}t &= \int_0^T D_t \|\tilde{u}_m\|^2_{L_2(\Omega)} = \|\tilde{u}_m(T)\|^2_{L_2(\Omega)} - \|\tilde{u}_m(0)\|^2_{L_2(\Omega)} \\ &\geqslant \|\tilde{u}_m(T)\|^2_{L_2(\Omega)} - \|u_0\|^2_{L_2(\Omega)} - \|g_u(0)\|^2_{L_2(\Omega)}, \end{aligned}$$

and (3.3) follows.

Observe that $D_t\tilde{u}_m$ belongs to V^u_m. With P_m denoting the L_2-projection onto V^u_m we get

$$\begin{aligned} \|D_t\tilde{u}_m(t)\|_{H^1(\Omega;\Gamma^u_D)^*} &= \sup_{\substack{v\in H^1(\Omega;\Gamma^u_D)\\ v\neq 0}} \frac{\langle D_t\tilde{u}_m(t), v\rangle}{\|v\|_{H^1(\Omega)}} = \sup_{\substack{v\in H^1(\Omega;\Gamma^u_D)\\ v\neq 0}} \frac{\langle D_t\tilde{u}_m(t), P_m v\rangle}{\|v\|_{H^1(\Omega)}} \\ &\leqslant \sup_{\substack{v\in H^1(\Omega;\Gamma^u_D)\\ P_m v\neq 0}} C\frac{\langle D_t\tilde{u}_m(t), P_m v\rangle}{\|P_m v\|_{H^1(\Omega)}} = \sup_{\substack{v\in V^u_m\\ v\neq 0}} C\frac{\langle D_t\tilde{u}_m(t), v\rangle}{\|v\|_{H^1(\Omega)}}, \end{aligned}$$

where C is the H^1-norm of the L_2-projection, which is independent of m due to (A5). Now we can use bounds similar to (3.5) and (3.6) to prove (3.4). □

LEMMA 3.2 There exists a unique solution $(\tilde{u}_m, \tilde{\varphi}_m) \in X_m$ to (3.1).

Proof.

For each $\tilde{u}_m$ there is a solution $\tilde{\varphi}_m = S(\tilde{u}_m)$ of (3.1b), which defines a mapping $S : L_2(0,T;V^u_m) \to L_2(0,T;V^\varphi_m)$.

First, we assume that $g_\varphi(x,\cdot)$ and $D_t g_u(x,\cdot)$ are continuous to prove existence and uniqueness on $[0,T]$. Let $\{\lambda_i\}_{i=1}^M$ be a basis for V^u_m. Then $\tilde{u}_m = \sum_{j=1}^M \alpha_j(t)\lambda_j$ for some $\alpha(t) = (\alpha_j(t)) \in \mathbb{R}^M$. By substituting into (3.1a) we arrive at the following system of ODEs

$$MD_t\alpha(t) + K\alpha(t) = F(\alpha(t), t) + G(t), \tag{3.7}$$

where M and K denote the mass and stiffness matrices, respectively, and

$$\begin{aligned} F_i(\alpha(t),t) &= -\langle \sigma(u_m(t))\lceil\tilde{\varphi}_m(t)\rceil\nabla\varphi_m(t), \nabla\lambda_i\rangle + \langle \sigma(u_m(t))\nabla\varphi_m(t)\cdot\nabla g_\varphi(t), \lambda_i\rangle, \\ G_i(t) &= -\langle D_t g_u(t), \lambda_i\rangle - \langle \nabla g_u(t), \nabla\lambda_i\rangle, \end{aligned}$$

where $\tilde{\varphi}_m = S(\tilde{u}_m)$. The initial data is given by (3.1c) and corresponds to the equation $M\alpha(0) = b$, where $b_i = \langle u_0 - g_u(0), \lambda_i\rangle$.

Let $\tilde{u}^1_m = \sum_{j=1}^M \alpha^1_j\lambda_j$ and $\tilde{u}^2_m = \sum_{j=1}^M \alpha^2_j\lambda_j$, and $\tilde{\varphi}^1_m := S(\tilde{u}^1_m)$ and $\tilde{\varphi}^2_m := S(\tilde{u}^2_m)$. Note that

$$\|\nabla(\tilde{u}^1_m - \tilde{u}^2_m)\|_{L_2(\Omega)} \leqslant C_m\|\alpha^1 - \alpha^2\|_{\mathbb{R}^M},$$

due to the Lipschitz continuity of linear mappings in V_m^u. For $\tilde{\varphi}_m^1 - \tilde{\varphi}_m^2$ we use Strang's first Lemma Braess (2007, Chapter III) and the Lipschitz continuity of σ to get

$$\begin{aligned}\sigma_\circ \|\nabla(\varphi_m^1 - \varphi_m^2)\|_{L_2(\Omega)} &\leqslant \|(\sigma(u_m^1) - \sigma(u_m^2))\nabla\varphi_m^1\|_{L_2(\Omega)} \\ &\leqslant C\|\tilde{u}_m^1 - \tilde{u}_m^2\|_{L_2(\Omega)}\|\nabla\varphi_m^1\|_{L_\infty(\Omega)} \leqslant C_m\|\alpha^1 - \alpha^2\|_{\mathbb{R}^M},\end{aligned}$$

which means that the mapping S is Lipschitz continuous. Here we have used that the boundary data is identical for the two instances, that is, $u_m^1 - u_m^2 = \tilde{u}_m^1 - \tilde{u}_m^2$ and $\varphi_m^1 - \varphi_m^2 = \tilde{\varphi}_m^1 - \tilde{\varphi}_m^2$.

Now, for F_i we have for $i = 1, ..., M$

$$\begin{aligned}|F_i(\alpha^1(t),t) - F_i(\alpha^2(t),t)| &\leqslant |\langle\sigma(u_m^1)\lceil\tilde{\varphi}_m^1\rceil\nabla\varphi_m^1 - \sigma(u_m^2)\lceil\tilde{\varphi}_m^2\rceil\nabla\varphi_m^2, \nabla\lambda_i\rangle| \\ &\quad + |\langle(\sigma(u_m^1)\nabla\varphi_m^1 - \sigma(u_m^2)\nabla\varphi_m^2)\cdot\nabla g_\varphi, \lambda_i\rangle|.\end{aligned}$$

Note that σ and the cut-off functional $\lceil\cdot\rceil$ are Lipschitz continuous and bounded. Furthermore, the image of $S(\cdot)$ is bounded owing to Lemma 3.1. Using this, together with the fact that the product of bounded Lipschitz continuous functions is Lipschitz continuous, we prove the Lipschitz continuity of F in t;

$$\|F(\alpha^1(t),t) - F(\alpha^2(t),t)\|_{\mathbb{R}^M} \leqslant C_m\|\alpha^1(t) - \alpha^2(t)\|_{\mathbb{R}^M},$$

where C_m is a generic constant that does not depend on t.

Picard-Lindelöf's theorem gives existence and uniqueness on some maximal interval (β_1, β_2). If (β_1, β_2) is a strict subset $(0, T)$ then by Amann (1990, Theorem 7.6) either

$$\lim_{t\to\beta_1^+}\|\alpha(t)\|_{\mathbb{R}^M} = \infty, \qquad \text{or} \qquad \lim_{t\to\beta_2^-}\|\alpha(t)\|_{\mathbb{R}^M} = \infty,$$

which contradicts Lemma 3.1. Hence there exist a unique solution in $C^1(0,T;V_m^u)\times L_\infty(0,T;V_m^\varphi)$.

Finally, we consider the case when $D_t g_u(x,\cdot)$ and $g_\varphi(x,\cdot)$ have at most finitely many discontinuities as specified in (A2). We may then use Picard-Lindelöf's theorem on the sub-interval $[t_k, t_{k+1})$ with the initial data $\alpha(t_k) = \lim_{t\to t_k^-}\alpha(t)$. The existence and uniqueness on $[0,T]$ now follows by induction over k. □

3.2 *Convergence of semidiscrete solutions*

The following lemma will be used several times in the convergence analysis in the subsequent text. Recall that $\tilde{\flat} = \flat - g_\varphi$ if $\flat$ is a Greek letter and $\tilde{\flat} = \flat - g_u$ if $\flat$ is a Latin letter.

LEMMA 3.3 Consider a sequence $\{\tilde{y}_m\}_m$ which converges pointwise a.e. to $\tilde{y}$ and a sequence $\{\tilde{\psi}_m\}_m$ which converges weakly in $L_2(0,T;H^1(\Omega;\Gamma_D^\varphi))$ to $\tilde{\psi}$, and the corresponding sequences $\{y_m\}_m$ and $\{\psi_m\}_m$ that converges to y and ψ, respectively. Suppose that, for all $m\in\mathbb{N}$,

$$\int_0^T \langle\sigma(y_m)\nabla\psi_m, \nabla\tilde{\psi}\rangle\,\mathrm{d}t = \int_0^T \langle\sigma(y_m)\nabla\psi_m, \nabla\tilde{\psi}_m\rangle\,\mathrm{d}t = 0. \tag{3.8}$$

Then $\tilde{\psi}_m \to \tilde{\psi}$ strongly in $L_2(0,T;H^1(\Omega;\Gamma_D^\varphi))$ as $m\to\infty$. Furthermore, subsequences of

$$\sigma(y_m)\nabla\psi_m \quad \text{and} \quad \sigma(y_m)\lceil\tilde{\psi}_m\rceil\nabla\psi_m$$

converge strongly in $L_2(0,T;L_2(\Omega;\mathbb{R}^3))$ to $\sigma(y)\nabla\psi$ and $\sigma(y)\lceil\tilde{\psi}\rceil\nabla\psi$, respectively.

Proof. The dominated convergence theorem implies that $\sigma(y_m)\nabla\tilde{\psi} \to \sigma(y)\nabla\tilde{\psi}$ converges strongly in $L_2(0,T;L_2(\Omega;\mathbb{R}^3)) \simeq L_2((0,T)\times\Omega;\mathbb{R}^3)$, as $|\sigma(y_m)\,\partial_i\psi| \leqslant \sigma^\circ\|\nabla\psi\|$ for all i,m.

Because in $L_2(0,T;L_2(\Omega))$ the scalar product of a bounded and weakly convergent sequence and a strongly convergent sequence converges to the scalar product of the limits,

$$\begin{aligned} 0 &= \int_0^T \langle\sigma(y_m)\nabla\psi_m, \nabla\tilde{\psi}\rangle\,\mathrm{d}t = \int_0^T \langle\nabla\psi_m, \sigma(y_m)\nabla\tilde{\psi}\rangle\,\mathrm{d}t \\ &\to \int_0^T \langle\sigma(y)\nabla\psi, \nabla\tilde{\psi}\rangle\,\mathrm{d}t, \quad \text{as} \quad m\to\infty. \end{aligned} \tag{3.9}$$

In particular, $\int_0^T \langle\sigma(y)\nabla\psi, \nabla\tilde{\psi}\rangle\,\mathrm{d}t = 0$. To prove strong convergence of $\tilde{\psi}_m$ we write

$$\begin{aligned} 0 &\leqslant \int_0^T \langle\sigma(y_m)\nabla(\tilde{\psi}-\tilde{\psi}_m), \nabla(\tilde{\psi}-\tilde{\psi}_m)\rangle\,\mathrm{d}t \\ &= \int_0^T \big(\langle\sigma(y_m)\nabla\tilde{\psi}, \nabla\tilde{\psi}\rangle - 2\langle\sigma(y_m)\nabla\tilde{\psi}_m, \nabla\tilde{\psi}\rangle + \langle\sigma(y_m)\nabla\tilde{\psi}_m, \nabla\tilde{\psi}_m\rangle\big)\,\mathrm{d}t \\ &=: I + II + III. \end{aligned}$$

Using the strong convergence of $\sigma(y_m)\nabla\tilde{\psi}$ we get

$$I \to \int_0^T \langle\sigma(y)\nabla\tilde{\psi}, \nabla\tilde{\psi}\rangle\,\mathrm{d}t = -\int_0^T \langle\sigma(y)\nabla g_\varphi, \nabla\tilde{\psi}\rangle\,\mathrm{d}t,$$

and due to (3.9) we have

$$II \to -2\int_0^T \langle\sigma(y)\nabla\tilde{\psi}, \nabla\tilde{\psi}\rangle\,\mathrm{d}t = 2\int_0^T \langle\sigma(y)\nabla g_\varphi, \nabla\tilde{\psi}\rangle\,\mathrm{d}t.$$

Now, due to (3.8) the third term gives

$$III = -\int_0^T \langle\sigma(y_m)\nabla g_\varphi, \nabla\tilde{\psi}_m\rangle\,\mathrm{d}t \to -\int_0^T \langle\sigma(y)\nabla g_\varphi, \nabla\tilde{\psi}\rangle\,\mathrm{d}t,$$

since, due to the dominated convergence theorem, $\sigma(y_m)\nabla g_\varphi \to \sigma(y)\nabla g_\varphi$ strongly in $L_2(0,T;L_2(\Omega))$ and $\nabla\tilde{\psi}_m$ converges weakly. Thus, we conclude that
$I + II + III \to 0$ and, by applying a Poincaré-Friedrichs inequality, that $\tilde{\psi}_m$ converges strongly in $L_2(0,T;H^1(\Omega;\Gamma_D^\varphi))$, and, by passing to a subsequence, also pointwise a.e.

Finally, by the dominated convergence theorem in the form of Royden (1988, p. 270), $\sigma(y_{m_k})\nabla\psi_{m_k}$ and $\sigma(y_{m_k})\lceil\tilde{\psi}_{m_k}\rceil\nabla\psi_{m_k}$ converge strongly in $L_2(0,T;L_2(\Omega;\mathbb{R}^3))$,
where m_k denotes a subsequence. □

THEOREM 3.1 A subsequence of solutions $(\tilde{u}_{m_k}, \tilde{\varphi}_{m_k}) \in X_{m_k}$ of (3.1) converges strongly in X to a solution $(\tilde{u}, \tilde{\varphi})$ of (2.4).

Proof. Owing to Lemma 3.1 and the reflexivity of X, there exists a subsequence m_k and $(\tilde{y}, \tilde{\psi}) \in X$ such that

$$(\tilde{u}_{m_k}, \tilde{\varphi}_{m_k}) \rightharpoonup (\tilde{y}, \tilde{\psi}) \quad \text{in } X, \text{ as } k\to\infty.$$

Because initial conditions are L_2-projected onto V_m^u it follows that $\tilde{u}_{m_k}(0) \to \tilde{y}(0) = \tilde{u}(0)$ in $L_2(\Omega)$.

The compactness of the embedding (Aubin-Lions lemma)

$$L_2(0,T;H^1(\Omega;\Gamma_D^u))\cap H^1(0,T;H^1(\Omega;\Gamma_D^u)^*)\hookrightarrow L_2(0,T;L_2(\Omega)),$$

implies that there exist a subsequence, still denoted m_k, such that $\tilde{u}_{m_k}\to\tilde{y}$ strongly and pointwise a.e. in $L_2(0,T;L_2(\Omega))$.

Owing to Lemma 3.3 we can pass to subsequences, without change of notation, so that φ_{m_k}, $\nabla\varphi_{m_k}$, $\sigma(u_{m_k})\nabla\varphi_{m_k}$ and $\sigma(u_{m_k})\lceil\tilde{\varphi}_{m_k}\rceil\nabla\varphi_{m_k}$ converge strongly in $L_2(0,T,L_2(\Omega))$.

Now choose $v(t)\in L_2(0,T;V_{m_k}^u)$ in (3.1a). Integrating from 0 to T gives,

$$\begin{aligned}\int_0^T\langle D_t u_{m_k},v\rangle+\langle\nabla u_{m_k},\nabla v\rangle\,\mathrm{d}t=\int_0^T\big(&-\langle\sigma(u_{m_k})\lceil\tilde{\varphi}_{m_k}\rceil\nabla\varphi_{m_k},\nabla v\rangle\\&+\langle\sigma(u_{m_k})\nabla\varphi_{m_k}\cdot\nabla g_\varphi,v\rangle\big)\,\mathrm{d}t.\end{aligned}$$

Fixing v we may now let $k\to\infty$ to get that

$$\int_0^T\langle D_t y,v\rangle+\langle\nabla y,\nabla v\rangle\,\mathrm{d}t=\int_0^T-\langle\sigma(y)\lceil\tilde{\psi}\rceil\nabla\psi,\nabla v\rangle+\langle\sigma(y)\nabla\psi\cdot\nabla g_\varphi,v\rangle\,\mathrm{d}t,$$

which holds for all $v\in L_2(0,T;H^1(\Omega;\Gamma_D^u))$, since $\cup_{k\in\mathbb{N}}L_2(0,T;V_{m_k}^u)$ is dense in this space and $u_{m,k}$, $\varphi_{m,k}$ are bounded according to Lemma 3.1, see Yosida (1995, Theorem 3, p. 121). In the spirit of Lemma 3.3 we may also prove that $\tilde{\psi}$ satisfies (2.4b). This together with the convergence of initial conditions imply that the limit $(\tilde{y},\tilde{\psi})$ is a solution to (2.4).

To prove that $\{\tilde{u}_{m_k}\}_k$ converges strongly in $L_2(0,T;H^1(\Omega;\Gamma_D^u))$, we write

$$\begin{aligned}&\int_0^T\langle\nabla(\tilde{y}-\tilde{u}_{m_k}),\nabla(\tilde{y}-\tilde{u}_{m_k})\rangle\,\mathrm{d}t\\&\quad=\int_0^T\big(\langle\nabla\tilde{y},\nabla\tilde{y}\rangle-2\langle\nabla\tilde{y},\nabla\tilde{u}_{m_k}\rangle+\langle\nabla\tilde{u}_{m_k},\nabla\tilde{u}_{m_k}\rangle\big)\,\mathrm{d}t=:I+II+III.\end{aligned}$$

Then $II\to-2\int_0^T\langle\nabla\tilde{y},\nabla\tilde{y}\rangle\,\mathrm{d}t$ since $\tilde{u}_{m_k}$ converges weakly in $L_2(0,T;H^1(\Omega;\Gamma_D^u))$. For the third term we use (3.1a) to get

$$\begin{aligned}III=\int_0^T\big(&\langle D_t\tilde{u}_{m_k},\tilde{u}_{m_k}\rangle-\langle D_t g_u,\tilde{u}_{m_k}\rangle-\langle\nabla g_u,\nabla\tilde{u}_{m_k}\rangle\\&-\langle\sigma(u_{m_k})\lceil\tilde{\varphi}_{m_k}\rceil\nabla\varphi_{m_k},\nabla\tilde{u}_{m_k}\rangle+\langle\sigma(u_{m_k})\nabla\varphi_{m_k}\cdot\nabla g_\varphi,\tilde{u}_{m_k}\rangle\big)\,\mathrm{d}t\\\to\int_0^T\big(&\langle D_t\tilde{y},\tilde{y}\rangle-\langle D_t g_u,\tilde{y}\rangle-\langle\nabla g_u,\nabla\tilde{y}\rangle\\&-\langle\sigma(y)\lceil\tilde{\psi}\rceil\nabla\psi,\nabla\tilde{y}\rangle+\langle\sigma(y)\nabla\psi\cdot\nabla g_\varphi,\tilde{y}\rangle\big)\,\mathrm{d}t\end{aligned}$$

as $k\to\infty$, recalling that $D_t\tilde{u}_{m_k}$ converges weakly, u_{m_k} strongly, and the statement of Lemma 3.3. Now, (2.4a) gives $III=\int_0^T\langle\nabla\tilde{y},\nabla\tilde{y}\rangle\,\mathrm{d}t$ so that $I+II+III\to0$.

Finally, to show strong convergence of the time derivative, we note that

$$\begin{aligned}&\|D_t\tilde{y}-D_t\tilde{u}_{m_k}\|_{H^1(\Omega;\Gamma_D^u)^*}\\&\quad\leqslant\sup_{\substack{v\in H^1(\Omega;\Gamma_D^u)\\v\neq0}}\frac{\langle D_t\tilde{y}-P_{m_k}D_t\tilde{y},v\rangle}{\|v\|_{H^1(\Omega)}}+\sup_{\substack{v\in H^1(\Omega;\Gamma_D^u)\\v\neq0}}\frac{\langle P_{m_k}D_t\tilde{y}-D_t\tilde{u}_{m_k},v\rangle}{\|v\|_{H^1(\Omega)}}\\&\quad=:a_{m_k}+b_{m_k},\end{aligned}$$

where P_{m_k} is the L_2-projection onto V^u_m. It follows, due to the density of L_2 in $(H^1)^*$, that $\int_0^T a^2_{m_k}\,\mathrm{d}t \to 0$ as $k \to \infty$. Now, we use the self-adjointness of the L_2-projection, (2.4a), and (3.1a) to get $\int_0^T b^2_{m_k}\,\mathrm{d}t \to 0$ as $k \to 0$ since ∇u_{m_k}, $\sigma(u_{m_k})\lceil\tilde{\varphi}_{m_k}\rceil\nabla\varphi_{m_k}$, and $\sigma(u_{m_k})\nabla\varphi_{m_k}\cdot\nabla g_\varphi$ converge strongly in $L_2(0,T;L_2(\Omega))$. We conclude that $\|D_t\tilde{y} - D_t\tilde{u}_{m_k}\|_{L_2(0,T;H^1(\Omega;\Gamma_D^u)^*)} \to 0$.

□

COROLLARY 3.1 If the solution (u,φ) to (2.4) is unique, then the full sequence of Galerkin solutions $(u_m,\varphi_m) \in X_m$ converges.

Proof. Due to Lemma 3.1 the sequence is bounded in X. From the proof of Theorem 3.1 we deduce that any accumulation point of the sequence is a solution to (2.4) and that an accumulation point exists. If the solution to (2.4) is unique there can only be one accumulation point and, hence, the full sequence $\{(u_m,\varphi_m)\}_m$ must converge. □

4. Fully discrete methods

In this section we analyze fully discrete methods based on a backward Euler scheme in time and hierarchical families of finite dimensional subspaces, as introduced in Section 3, in space. We prove existence and uniqueness of fully discrete solutions and strong convergence to a weak solution satisfying (2.4).

4.1 *Fully discrete formulation*

Let $\{J_l\}_{l\in\mathbb{N}}$ be a family of nested partitions of the time interval $J=[0,T]$, which subordinate to the decomposition of (A2). For each partition $0=t_0<t_1<\ldots<t_N=T$ we denote the subintervals $I_n := (t_{n-1},t_n]$ and $f^n := f(t_n)$. We consider a uniform time discretization in the analysis, that is, we assume $t_n - t_{n-1} = \tau_l$ with $\tau_l = 2^{-l}T$. It simplifies some of the analysis, but it is also a requirement for the compactness argument in Walkington (2010).

Fix m and let V^u_m and V^φ_m be as in Section 3 and define the discrete space

$$\begin{aligned} X_{m,l} =& \{v(x,t) : \forall n\ \exists w \in V^u_m : v(t,\cdot) = w,\ t \in I_n\} \\ &\times \{v(x,t) : \forall n\ \exists w \in V^\varphi_m : v(t,\cdot) = w,\ t \in I_n\}. \end{aligned}$$

This means that functions in $X_{m,l}$ are piecewise constant in time and on each interval equal to a function from $V^u_m \times V^\varphi_m$. Note that $X_{m,l} \not\subseteq X_m$ since $X_{m,l}$ discontinuous in time. However, we have $X_{m,l} \subseteq L_2(0,T;V^u_m)\times L_2(0,T;V^\varphi_m)$.

We use the backward Euler scheme to define a fully discrete solution. Find a pair $(u_{m,l},\varphi_{m,l}) = (g_u+\tilde{u}_{m,l}, g_\varphi+\tilde{\varphi}_{m,l})$ such that $(\tilde{u}_{m,l},\tilde{\varphi}_{m,l}) \in X_{m,l}$ and for $n=1,\ldots,N$,

$$\begin{aligned} \left\langle \frac{u^n_{m,l}-u^{n-1}_{m,l}}{\tau_l}, v \right\rangle + \langle\nabla u^n_{m,l},\nabla v\rangle &= -\langle\sigma(u^n_{m,l})\lceil\tilde{\varphi}^n_{m,l}\rceil\nabla\varphi^n_{m,l},\nabla v\rangle \\ &\quad + \langle\sigma(u^n_{m,l})\nabla\varphi^n_{m,l}\cdot\nabla g^n_\varphi, v\rangle, \qquad (4.1a) \\ \langle\sigma(u^n_{m,l})\nabla\varphi^n_{m,l},\nabla w\rangle &= 0, \qquad (4.1b) \\ \langle u^0_{m,l}, z\rangle &= \langle u_0, z\rangle, \qquad (4.1c) \end{aligned}$$

for all $(v,w) \in V^u_m \times V^\varphi_m$ and $z \in V^u_m$, where $u^n_{m,l} = u_{m,l}(t_n)$ and $\varphi^n_{m,l} = \varphi_{m,l}(t_n)$.

LEMMA 4.1 A solution $(u_{m,l}, \varphi_{m,l})$ to (4.1) fulfils the following bounds

$$\|\nabla\tilde{\varphi}^n_{m,l}\|^2_{L_2(\Omega)} \leqslant C(\sigma, g_\varphi), \tag{4.2}$$

$$\|\tilde{u}^n_{m,l}\|^2_{L_2(\Omega)} + \int_0^{t_n} \|\nabla\tilde{u}_{m,l}\|^2_{L_2(\Omega)}\,\mathrm{d}t \leqslant C(u_0, \sigma, g_u, D_t g_u, g_\varphi), \tag{4.3}$$

$$\sum_{n=1}^{N} \|\tilde{u}^n_{m,l} - \tilde{u}^{n-1}_{m,l}\|^2_{L_2(\Omega)} \leqslant C(u_0, \sigma, g_u, D_t g_u, g_\varphi). \tag{4.4}$$

for $n = 0, ..., N$.

Proof. Choosing $w = \tilde{\varphi}^n_{m,l}$ in (4.1b) we have $\|\nabla\tilde{\varphi}^n_{m,l}\|^2_{L_2(\Omega)} \leqslant C(\sigma)\|g^n_\varphi\|_{L_2(\Omega)}$. To prove (4.3), we note that

$$\langle \tilde{u}^n_{m,l} - \tilde{u}^{n-1}_{m,l}, \tilde{u}^n_{m,l}\rangle = \frac{1}{2}\|\tilde{u}^n_{m,l}\|^2_{L_2(\Omega)} - \frac{1}{2}\|\tilde{u}^{n-1}_{m,l}\|^2_{L_2(\Omega)} + \frac{1}{2}\|\tilde{u}^n_{m,l} - \tilde{u}^{n-1}_{m,l}\|^2_{L_2(\Omega)},$$

and by choosing $v = \tau_l\tilde{u}^n_{m,l}$ in (4.1a) and summing from $n = 1$ to N

$$\begin{aligned}
&\frac{1}{2}\sum_{n=1}^{N}\left(\|\tilde{u}^n_{m,l}\|^2_{L_2(\Omega)} - \|\tilde{u}^{n-1}_{m,l}\|^2_{L_2(\Omega)} + \|\tilde{u}^n_{m,l} - \tilde{u}^{n-1}_{m,l}\|^2_{L_2(\Omega)}\right) \\
&\quad + \frac{1}{2}\int_0^T \|\nabla\tilde{u}_{m,l}\|^2_{L_2(\Omega)}\,\mathrm{d}t \\
&\quad = -\sum_{n=1}^{N}\langle g^n_u - g^{n-1}_u, \tilde{u}^n_{m,l}\rangle - \int_0^T \langle \nabla g_u, \nabla\tilde{u}_{m,l}\rangle\,\mathrm{d}t \\
&\qquad - \int_0^T \langle \sigma(u_{m,l})\lceil\tilde{\varphi}_{m,l}\rceil\nabla\varphi_{m,l}, \nabla\tilde{u}_{m,l}\rangle\,\mathrm{d}t \\
&\qquad + \int_0^T \langle \sigma(u_{m,l})\nabla\varphi_{m,l}\cdot\nabla g_\varphi, \tilde{u}_{m,l}\rangle\,\mathrm{d}t =: I + II + III + IV.
\end{aligned} \tag{4.5}$$

For the first term we get

$$\sum_{n=1}^{N}\int_{t_{n-1}}^{t_n}\langle D_t g_u, \tilde{u}^n_{m,l}\rangle\,\mathrm{d}t \leqslant C\sum_{n=1}^{N}\int_{t_{n-1}}^{t_n}\|D_t g_u\|^2_{H^1(\Omega;\Gamma^u_D)^*}\,\mathrm{d}t + \frac{1}{8}\int_0^T\|\nabla\tilde{u}_{m,l}\|^2_{L_2(\Omega)}\,\mathrm{d}t.$$

The remaining terms $II - IV$ can be estimated as in the proof of Lemma 3.1. Using the telescoping effect of the first two terms in the sum in (4.5) completes the proof. □

LEMMA 4.2 There exists a solution $(\tilde{u}_{m,l}, \tilde{\varphi}_{m,l}) \in X_{m,l}$ to (4.1). Furthermore, for any fixed m, there is an $L \in \mathbb{N}$, such that the solution $(\tilde{u}_{m,l}, \tilde{\varphi}_{m,l})$ is unique for any $l > L$.

Proof. To prove this we use the ODE setting introduced in the proof of Lemma 3.2. Let $\tilde{u}^n_{m,l} = \sum_{j=1}^{M}\alpha^n_{l,j}\lambda_j$ for $n = 1, ..., N$. Then (4.1) corresponds to finding $\alpha^n_l \in \mathbb{R}^M$ such that

$$\frac{M\alpha^n_l - M\alpha^{n-1}_l}{\tau_l} + K\alpha^n_l = F(\alpha^n_l, t_n) + G(t_n), \tag{4.6}$$

for $n = 1,..,N$, with $M\alpha_l^0 = b$, where $b_i = \langle u_0 - g_u(0), \lambda_i \rangle$, F as in (3.7), and G slightly modified as

$$G(t_n) := -\langle \frac{g_u^n - g_u^{n-1}}{\tau_l}, \lambda_i \rangle - \langle \nabla g_u^n, \nabla \lambda_i \rangle.$$

Note that (4.6) is the backward Euler discretization of the ODE (3.7).

To apply Brouwer's fixed point theorem we define the mapping $f : \mathbb{R}^M \to \mathbb{R}^M$ such that $\beta = f(\gamma)$ is the solution to the system

$$\frac{M\beta - M\alpha_l^{n-1}}{\tau_l} + K\beta = F(\gamma, t_n) + G(t_n),$$

which is equivalent to

$$\beta = (M + \tau_l K)^{-1}(M\alpha_l^{n-1} + \tau_l F(\gamma, t_n) + \tau_l G(t_n)). \tag{4.7}$$

Let $\tilde{y}$ be the function corresponding to the vector γ, that is, $\tilde{y}^n = \sum_{j=1}^M \gamma_j^n \lambda_j$. From the definitions of F_i and G_i in the proof of Lemma 3.2, and with $\tilde{\psi} = S(\tilde{y})$ it follows that

$$\begin{aligned} |F_i(\gamma, t_n)| &\leqslant C(\sigma, g_\varphi) \|\nabla \psi\|_{L_2(\Omega)} \|\nabla \lambda_i\|_{L_2(\Omega)} \\ &\quad + \sigma^\circ \|\nabla \psi\|_{L_2(\Omega)} \|g_\varphi(t_n)\|_{L_3(\Omega)} \|\lambda_i\|_{L_6(\Omega)} \leqslant C_m(\sigma, g_\varphi), \\ |G_i(t_n)| &\leqslant C_m(D_t g_u, g_u). \end{aligned}$$

Here the boundedness of $\|\nabla \psi\|_{L_2(\Omega)}$ follows from Lemma 4.1 with $\tilde{u}_{m,l}^n = \tilde{y}$. Hence, letting $B_R \in \mathbb{R}^M$ denote the ball with radius $R > 0$, it is clear that $f \colon B_R \to B_R$ if R sufficiently large. Now define $\beta_1 = f(\gamma_1), \beta_2 = f(\gamma_2) \in \mathbb{R}^M$. From (4.7) we have

$$(\beta^1 - \beta^2) = \tau_l (M + \tau_l K)^{-1}(F(\gamma_1, t_n) - F(\gamma_2, t_n)),$$

and using the Lipschitz continuity of $F(\cdot, t)$, see proof of Lemma 3.2, we get

$$\|\beta^1 - \beta^2\|_{\mathbb{R}^M} \leqslant C_L \tau_l \|(M + \tau_l K)^{-1}\|_F \|\gamma_1 - \gamma_2\|_{\mathbb{R}^M},$$

where $\|\cdot\|_F$ denotes the matrix Frobenius norm which is finite since $M + \tau_l K$ is invertible. This proves that f is continuous and the existence of a solution follows from Brouwer's fixed point theorem. Furthermore, it is clear that if τ_l is sufficiently small, or equivalently l sufficiently large, then f is a contraction on $\mathbb{R}^M$ and Banach's fixed point theorem gives uniqueness. □

4.2 *Convergence of fully discrete solutions*

To prove convergence of the fully discrete method we introduce the continuous and piecewise affine interpolant $\tilde{U}_{m,l}(t)$ of $\tilde{u}_{m,l}$

$$\tilde{U}_{m,l}(t) := \tilde{u}_{m,l}^{n-1} \frac{t_n - t}{t_n - t_{n-1}} + \tilde{u}_{m,l}^n \frac{t - t_{n-1}}{t_n - t_{n-1}}, \quad t \in I_n. \tag{4.8}$$

Note that

$$D_t \tilde{U}_{m,l}(t) = \frac{\tilde{u}_{m,l}^n - \tilde{u}_{m,l}^{n-1}}{\tau_l}, \quad t \in I_n. \tag{4.9}$$

Using (4.1a), we have for $t \in I_n$ and $v \in H^1(\Omega;\Gamma_D^u)$

$$\begin{aligned}\langle D_t\tilde{U}_{m,l}(t), v\rangle \leqslant \big(&\|u_{m,l}^n\|_{H^1(\Omega)} + C(\sigma, g_\varphi)\|\nabla\varphi_{m,l}^n\|_{L_2(\Omega)} \\ &+ \sigma^\circ\|\nabla\varphi_{m,l}^n\|_{L_2(\Omega)}\|\nabla g_\varphi^n\|_{L_3(\Omega)} + \|\partial_t g_u^n\|_{H^1(\Omega;\Gamma_D^u)^*}\big)\|v\|_{H^1(\Omega)},\end{aligned}$$

where $\partial_t g_u^n = (g_u^n - g_u^{n-1})/\tau_l$. Note that $\partial_t g_u^n = \tau_l^{-1}\int_{I_n} D_t g_u \,\mathrm{d}t$. This, together with Lemma 4.1 and (A2), gives

$$\|D_t\tilde{U}_{m,l}\|_{L_2(0,T;(V_m^u)^*)} \leqslant C(u_0, \sigma, g_u, D_t g_u, g_\varphi).$$

In addition, using (A5) as in the proof of Lemma 3.1, we get

$$\|D_t\tilde{U}_{m,l}\|_{L_2(0,T;H^1(\Omega;\Gamma_D^u)^*)} \leqslant C(u_0, \sigma, g_u, D_t g_u, g_\varphi). \tag{4.10}$$

In the analysis we also use the following reformulation of (4.1a)

$$\langle u_{m,l}^n - u_{m,l}^{n-1}, v\rangle = \langle F_{m,l}, v\rangle, \tag{4.11}$$

with

$$\begin{aligned}\langle F_{m,l}, v\rangle := &-\tau_l\langle\nabla u_{m,l}^n, \nabla v\rangle - \tau_l\langle\sigma(u_{m,l}^n)\lceil\tilde{\varphi}_{m,l}^n\rceil\nabla\varphi_{m,l}^n, \nabla v\rangle \\ &+ \tau_l\langle\sigma(u_{m,l}^n)\nabla\varphi_{m,l}^n\cdot\nabla g_\varphi^n, v\rangle, \quad \forall v \in V_m^u.\end{aligned}$$

THEOREM 4.1 A subsequence of solutions $(\tilde{u}_{m_k,l_k}, \tilde{\varphi}_{m_k,l_k}) \in X_{m_k,l_k}$ of (4.1) converges strongly in $L_2(0,T;H^1(\Omega;\Gamma_D^u)) \times L_2(0,T;H^1(\Omega;\Gamma_D^\varphi))$ to a solution $(\tilde{u},\tilde{\varphi})$ of (2.4).

Proof. From Lemma 4.1, (4.10), and the reflexivity of the spaces

$$L_2(0,T;H^1(\Omega;\Gamma_D^u)), \quad L_2(0,T;H^1(\Omega;\Gamma_D^\varphi)), \quad L_2(0,T;H^1(\Omega;\Gamma_D^u)^*)$$

there exists a subsequence such that

$$\begin{aligned}(\tilde{u}_{m_k,l_k}, \tilde{\varphi}_{m_k,l_k}) &\rightharpoonup (\tilde{y},\tilde{\psi}) && \text{in } L_2(0,T;H^1(\Omega;\Gamma_D^u)) \times L_2(0,T;H^1(\Omega;\Gamma_D^\varphi)), \\ D_t\tilde{U}_{m_k,l_k} &\rightharpoonup D_t\tilde{U} && \text{in } L_2(0,T;H^1(\Omega;\Gamma_D^u)^*),\end{aligned}$$

for some

$$(\tilde{y},\tilde{\psi}) \in L_2(0,T;H^1(\Omega;\Gamma_D^u)) \times L_2(0,T;H^1(\Omega;\Gamma_D^\varphi)),\ D_t\tilde{U} \in L_2(0,T;H^1(\Omega;\Gamma_D^u)^*).$$

The convergence of the initial conditions follows as in the semi-discrete case and we conclude $u_{m_k,l_k}^0 \to y(0) = u(0)$.

Next, we prove that $D_t\tilde{y} = \tilde{D}_t U$. First, we note that $\tilde{U}_{m_k,l_k} - \tilde{u}_{m_k,l_k} \rightharpoonup 0$ weakly in $L_2(0,T;L_2(\Omega))$. To see this, pick $\chi_{[\bar{\tau}a,\bar{\tau}b]} \otimes v : (t,x) \mapsto \chi_{[\bar{\tau}a,\bar{\tau}b]}(t)\,v(x)$, for $v \in L_2(\Omega)$, $a < b$, and some $\bar{\tau} > 0$. Now, due to Roubíček (2013, Theorem 8.9), for $\tau_{l_k} \leqslant \bar{\tau}$

$$\int_0^T \langle\tilde{U}_{m_k,l_k} - \tilde{u}_{m_k,l_k}, \chi_{[\bar{\tau}a,\bar{\tau}b]} \otimes v\rangle \,\mathrm{d}t = \frac{\tau_{l_k}}{2}\langle\tilde{u}_{m_k,l_k}^{\bar{\tau}b/\tau_{l_k}} - \tilde{u}_{m,l_k}^{\bar{\tau}a/\tau_{l_k}}, v\rangle \leqslant C\tau_{l_k} \to 0,$$

where we used (4.8) and the bounds in Lemma 4.1. Since $\tilde{U}_{m_k}$ is bounded in $L_2(0;T;L_2(\Omega))$ and functions of the form $\chi_{[\bar{\tau}a,\bar{\tau}b]}\otimes v$ are dense in $L_2(0,T;L_2(\Omega))$ this implies $\tilde{U}_{m_k,l_k}-\tilde{u}_{m_k,l_k}\rightharpoonup 0$ in $L_2(0,T;L_2(\Omega))$, see Yosida (1995, Theorem 3 p. 121). Thus we get, for $v\in C^1(0,T;H^1(\Omega;\Gamma_D^u))$ with $v(0)=v(T)=0$,

$$\int_0^T \langle D_t\tilde{U},v\rangle\,\mathrm{d}t \leftarrow \int_0^T \langle D_t\tilde{U}_{m_k,l_k},v\rangle\,\mathrm{d}t = -\int_0^T \langle \tilde{U}_{m_k,l_k},D_tv\rangle\,\mathrm{d}t \to -\int_0^T \langle \tilde{y},D_tv\rangle\,\mathrm{d}t,$$

and we conclude $D_t\tilde{y}=D_t\tilde{U}$, due to (4.10) and since $C^1(0,T;H^1(\Omega;\Gamma_D^u))$ with $v(0)=v(T)=0$ is dense in $L_2(0,T;L_2(\Omega))$, see see Yosida (1995, Theorem 3 p. 121). This implies that $(\tilde{y},\tilde{\psi})\in X$.

In Walkington (2010, Theorem 3.1) it is proved that if $\{u_{m_k,l_k}\}_k$ and $\{F_{m_k,l_k}\}_k$ in (4.11) are bounded in $L_2(0,T;H^1(\Omega;\Gamma_D^u))$ and $L_2(0,T;(V_{m_k}^u)^*)$, respectively, uniformly in k, then $\{u_{m_k,l_k}\}$ is precompact in $L_2(0,T;L_2(\Omega))$. Here, the boundedness of $\{u_{m_k,l_k}\}$ and $\{F_{m_k,l_k}\}$ follows from Lemma 4.1, (A2), and the bounds on σ and $\lceil\cdot\rceil$. Hence, there exists a subsequence, still denoted (m_k,l_k), such that $\tilde{u}_{m_k,l_k}\to\tilde{y}$ strongly and pointwise a.e. in $L_2(0,T;L_2(\Omega))$.

Owing to Lemma 3.3 we have for some subsequence, still denoted (m_k,l_k), that $\varphi_{m_k,l_k},\nabla\varphi_{m_k,l_k}$, $\sigma(u_{m_k,l_k})\nabla\varphi_{m_k,l_k}$, and $\sigma(u_{m_k,l_k})\nabla\lceil\tilde{\varphi}_{m_k,l_k}\rceil\varphi_{m_k,l_k}$, converges strongly in $L_2(0,T;L_2(\Omega))$. Using (4.1a) we get for $v\in L_2(0,T;V_m^u)$

$$\begin{aligned}&\int_0^T \langle D_t\tilde{U}_{m_k,l_k}+\partial_t g_u,v\rangle+\langle\nabla u_{m_k,l_k},\nabla v\rangle\,\mathrm{d}t\\&\quad=\int_0^T\big(-\langle\sigma(u_{m_k,l_k})\lceil\tilde{\varphi}_{m_k,l_k}\rceil\nabla\varphi_{m_k,l_k},\nabla v\rangle+\langle\sigma(u_{m_k,l_k})\nabla\varphi_{m_k,l_k}\cdot\nabla g_\varphi,v\rangle\big)\,\mathrm{d}t,\end{aligned}$$

where $\partial_t g_u^n=(g_u^n-g_u^{n-1})/\tau_l$. Keeping v fixed we get as $k\to\infty$

$$\begin{aligned}\int_0^T \langle D_t\tilde{y}+D_tg_u,v\rangle+\langle\nabla u,\nabla v\rangle\,\mathrm{d}t=\int_0^T\big(&-\langle\sigma(u_{m_k})\lceil\tilde{\varphi}_{m_k}\rceil\nabla\varphi_{m_k},\nabla v\rangle\\&+\langle\sigma(u_{m_k})\nabla\varphi_{m_k}\cdot\nabla g_\varphi,v\rangle\big)\,\mathrm{d}t,\end{aligned}$$

where we used the weak convergence of $\tilde{U}_{m_k}$ and ∇u_{m_k}, the strong convergence of $\sigma(u_{m_k,l_k})\nabla\varphi_{m_k,l_k}$ and $\sigma(u_{m_k,l_k})\nabla\lceil\tilde{\varphi}_{m_k,l_k}\rceil\varphi_{m_k,l_k}$, and $\partial_t g_u\to D_tg_u$. Now, due to the density of $\{v(x,t):v|_{I_n}\in V_m^u\}$ and the boundedness of $u_{m_k,l_k},\varphi_{m_k,l_k}$, this holds for all $v\in L_2(0,T;H^1(\Omega;\Gamma_D^u))$ Yosida (1995, Theorem 3 p. 121). Furthermore, in the spirit of Lemma 3.8 we may prove that $\tilde{\psi}$ solves (2.4b). Hence, $(\tilde{y},\tilde{\psi})$ is a solution to (2.4).

To prove that $\tilde{u}_{m_k,l_k}\to\tilde{y}$ strongly in $L_2(0,T;H^1(\Omega;\Gamma_D^u))$ we may mimic the argument in the proof of Theorem 3.1 (recall (4.9)). □

COROLLARY 4.1 If the solution (u,φ) to (2.4) is unique, then the whole sequence of fully discrete approximations $(\tilde{u}_{m,l},\tilde{\varphi}_{m,l})\in X_{m,l}$ converges.

Proof. If the solution to (2.4) is unique there can only be one accumulation point, cf. Corollary 3.1. □

REMARK 4.1 There are many other time discretizations available, see, e.g., Thomée (2006, Chapter 16) for discretizations of a parabolic nonlinear problem. For instance, one could use linearized schemes to avoid solving a nonlinear problem in each time step. In this case the existence and uniqueness result follows easily. Convergence may be deduced by comparing the linearized solution to the backward Euler solution. However, to avoid overloading the paper, we shall not analyze this here.

5. Regularity and uniqueness

In this section we prove additional regularity and uniqueness of a solution to the weak problem (2.4). For this purpose we use Theorem 5.2 below, which is based on Hieber & Rehberg (2008); Meinlschmidt *et al.* (2017), see, in particular, Hieber & Rehberg (2008, Theorem 3.1). The theory in Mitrea & Mitrea (2007) gives a setting where assumption (B5) below is satisfied, so called creased domains. We recall here the definition, cf. Brown & Mitrea (2009).

DEFINITION 5.1 (Creased domains) Since $\Omega \subset \mathbb{R}^3$ is a bounded Lipschitz domain there is a constant $M > 0$ such that for all $x \in \partial\Omega$ there exists a coordinate system (x_1, x_2, x_3), a cylinder $C_r(x) = \{(y_1, y_2, y_3) : \ |(y_1, y_2) - (x_1, x_2)| < r, |y_3 - x_3| < 2Mr\}$ for some $r > 0$, and a Lipschitz function $\xi : \mathbb{R}^2 \to \mathbb{R}$ with $\|\nabla\xi\|_\infty < M$ such that

$$C_r(x) \cap \Omega = \{(y_1, y_2, y_3) : y_3 > \xi(y_1, y_2)\} \cap C_r(x),$$
$$C_r(x) \cap \partial\Omega = \{(y_1, y_2, y_3) : y_3 = \xi(y_1, y_2)\} \cap C_r(x).$$

Let Γ_D and Γ_N be the Dirichlet and Neumann boundaries, respectively. We call Ω a *creased* domain if there is a constant $m > 0$ such that for all $x \in \overline{\Gamma_D} \cap \overline{\Gamma_N}$ one can find a Lipschitz function $\psi : \mathbb{R} \to \mathbb{R}$ and $r > 0$ such that

$$\Gamma_N \cap C_r(x) = C_r(x) \cap \partial\Omega \cap \{(x_1, x_2, x_3) : \ x_1 \geqslant \psi(x_2)\},$$
$$\Gamma_D \cap C_r(x) = C_r(x) \cap \partial\Omega \cap \{(x_1, x_2, x_3) : \ x_1 < \psi(x_2)\}$$

and

$$\tfrac{\partial\xi}{\partial x_1} > m \text{ almost everywhere where } x_1 > \psi(x_2),$$

$$\tfrac{\partial\xi}{\partial x_1} < -m \text{ almost everywhere where } x_1 < \psi(x_2).$$

Our aim is to combine the results in Hieber & Rehberg (2008) and Meinlschmidt *et al.* (2017) to obtain additional regularity for the Joule heating problem with mixed boundary conditions on creased domains. For similar settings, see Meinlschmidt *et al.* (2017, Section 3), where regularity for the Joule heating problem with pure Robin boundary conditions for the temperature and mixed boundary conditions for the potential is studied. In addition, we also prove higher regularity of the solution in the interior of the domain. We emphasize that the differences in the regularity within the domain makes the problem well suited for *h*- and *hp*-adaptive finite elements.

In Hieber & Rehberg (2008) the following type of systems are studied

$$D_t u - \Delta u = R(u), \quad \text{in } \Omega \times (0,T), \tag{5.1a}$$
$$u = g_u, \quad \text{on } \Gamma_D^u \times (0,T) \tag{5.1b}$$
$$n \cdot \nabla u = 0, \quad \text{on } \Gamma_N^u \times (0,T), \tag{5.1c}$$
$$u(\cdot, 0) = u_0, \quad \text{in } \Omega. \tag{5.1d}$$

If we define $R(u) = \sigma(u)|\nabla S(u)|^2$ such that $\varphi = S(u)$ solves

$$-\nabla \cdot (\sigma(u)\nabla\varphi) = 0, \quad \text{in } \Omega \times (0,T), \tag{5.2a}$$
$$\varphi = g_\varphi, \quad \text{on } \Gamma_D^\varphi \times (0,T) \tag{5.2b}$$
$$n \cdot \nabla\varphi = 0, \quad \text{on } \Gamma_N^\varphi \times (0,T), \tag{5.2c}$$

then (5.1) is equivalent to (2.1).

For $p > \frac{3}{2}$ and a fixed $r > \frac{4p}{2p-3}$ we consider the following assumptions.

(B1) Ω is a bounded domain with Lipschitz boundary (in the sense of Gröger, see Hieber & Rehberg (2008)), $\text{meas}(\Gamma_D^u) > 0$, and $\text{meas}(\Gamma_D^\varphi) > 0$.

(B2) $g_u \in C([0,T];W^1_{2p}(\Omega)) \cap W^1_r(0,T;L_p(\Omega))$, $\Delta g_u(t) = 0$, $t \in (0,T)$, and $g_\varphi \in L_r(0,T;W^1_{2p}(\Omega))$.

(B3) $u_0 - g_u(0) \in (L_p(\Omega), D(\Delta_p))_{1-\frac{1}{r},r}$.

(B4) $\sigma \in C^1(\mathbb{R})$, Lipschitz continuous, and $0 < \sigma_\circ \leqslant \sigma(x) \leqslant \sigma^\circ < \infty$, $\forall x \in \mathbb{R}$.

(B5) Δ is a topological isomorphism from $W^1_{2p}(\Omega;\Gamma_D^u)$ onto $W^{-1}_{2p}(\Omega;\Gamma_D^u)$ and from $W^1_{2p}(\Omega;\Gamma_D^\varphi)$ onto $W^{-1}_{2p}(\Omega;\Gamma_D^\varphi)$.

Here $D(A)$ denotes the domain of the operator A, that is

$$D(A) = \{v \in H^1(\Omega;\Gamma_D) : \exists f \in L_2(\Omega),\ -a(v,w) = \langle f, w\rangle,\ \forall w \in H^1(\Omega;\Gamma_D)\},$$

for $\Gamma_D \subseteq \partial\Omega$. The semigroup generated by A extends to a C_0-semigroup on $L_p(\Omega)$, $1 < p < \infty$, and we denote its generator by A_p, see Hieber & Rehberg (2008) and references therein. In particular, we use Δ_p for the Laplacian that maps onto $L_p(\Omega)$. Furthermore, for Banach spaces V and W forming an interpolation couple, $(V,W)_{\alpha,\beta}$ denotes the real interpolation space.

The following theorem relies on Hieber & Rehberg (2008, Theorem 3.1).

THEOREM 5.2 Let $p > \frac{3}{2}$ and $r > \frac{4p}{2p-3}$. Under the assumptions (B1)-(B5), there exists a unique solution to (2.1) satisfying

$$\tilde{u} \in W^1_r(0,T_*;L_p(\Omega)) \cap L_r(0,T_*;W^1_{2p}(\Omega;\Gamma_D^u)), \quad \tilde{\varphi} \in L_r(0,T_*;W^1_{2p}(\Omega;\Gamma_D^\varphi)), \tag{5.3}$$

for some $0 < T_* \leqslant T$.

Proof. Consider

(B6) The function $R\colon W^1_{2p}(\Omega) \to L_p(\Omega)$ is continuous.

(B7) $R(0) \in L_r(0,T;L_p(\Omega))$ and for $\beta > 0$ there exist $g_\beta \in L_r(0,T)$ such that

$$\|R(u_1) - R(u_2)\|_{L_p(\Omega)} \leqslant g_\beta(t)\|u_1 - u_2\|_{W^1_{2p}(\Omega)}, \quad t \in (0,T),$$

provided $\max(\|u_1\|_{W^1_{2p}(\Omega)}, \|u_2\|_{W^1_{2p}(\Omega)}) \leqslant \beta$.

In Hieber & Rehberg (2008, Theorem 3.1) it is proved that if the conditions (B1)-(B5) together with (B6)-(B7) on the operator R are satisfied, then there is a unique solution to (5.1) satisfying

$$u \in W^1_r(0,T_*;L_p(\Omega)) \cap L_r(0,T_*;D(\Delta_p)),$$

for some $0 < T_* \leqslant T$.

Our aim is now to prove that if $R(u) = \sigma(u)|\nabla S(u)|^2$ with $S : W^1_{2p}(\Omega) \to W^1_{2p}(\Omega)$ defined as in (5.2), then R satisfies (B6)-(B7) and we may conclude that (u,φ) solves (2.1) and (5.3) is fulfilled.

From Meinlschmidt *et al.* (2017, Corollary 3.24) it follows that the operator $(-\nabla\phi\cdot\nabla)^{-1}$ is a linear homeomorphism from $W_{2p}^{-1}(\Omega;\Gamma_D^\varphi)$ to $W_{2p}^1(\Omega;\Gamma_D^\varphi)$, if $\phi\in\mathfrak{C}$ is uniformly continuous on $\bar\Omega$, $\mathfrak{C}$ compact set in $C(\bar\Omega)$, ϕ admits a positive lower bound, and (B5) holds. It also holds that $(-\nabla\cdot\phi\nabla)^{-1}$ is Lipschitz with respect to ϕ.

In our setting, we have due to Morrey's inequality $W_{2p}^1(\Omega)\subseteq C^{0,\alpha}(\bar\Omega)$, for some $\alpha>0$, and $C^{0,\alpha}(\bar\Omega)$ embeds compactly into $C(\bar\Omega)$. This implies that $\mathfrak{C}=\{\sigma(u):u\in W_{2p}^1(\Omega)\}$ is compact and $\sigma(u)$ is uniformly continuous, since σ is Lipschitz continuous. Furthermore, $\sigma(\cdot)\geqslant\sigma_\circ>0$. Hence, given $\phi=\sigma(u)$ we deduce that there exists a unique solution $\tilde\varphi(t)\in W_{2p}^1(\Omega;\Gamma_D^\varphi)$ to (5.2). This proves that the mappings S and R are well defined in the given spaces.

Given u fixed, let $F:W_{2p}^{-1}(\Omega;\Gamma_D^\varphi)\to W_{2p}^1(\Omega;\Gamma_D^\varphi)$, such that $\psi=F(g)$ is the solution to

$$\nabla\cdot(\sigma(u)\nabla\psi)=g,\quad \psi|_{\Gamma_D^\varphi}=0.$$

By letting $G=\nabla\cdot(\sigma(u)\nabla g_\varphi)$ and using $\varphi=\tilde\varphi+g_\varphi$ it follows from (5.2a) that $\tilde\varphi=F(G)$. Since the operator F is bounded we get

$$\begin{aligned}\|\tilde\varphi\|_{W_{2p}^1(\Omega;\Gamma_D^\varphi)}&\leqslant C\|G\|_{W_{2p}^{-1}(\Omega;\Gamma_D^\varphi)}=C\sup_{w\in W_{(2p)'}^1(\Omega;\Gamma_D^\varphi)\setminus\{0\}}\frac{\langle\nabla\cdot(\sigma(u)\nabla g_\varphi),w\rangle}{\|w\|_{W_{(2p)'}^1(\Omega)}}\\&=C\sup_{w\in W_{(2p)'}^1(\Omega;\Gamma_D^\varphi)\setminus\{0\}}\frac{\langle\sigma(u)\nabla g_\varphi,\nabla w\rangle}{\|w\|_{W_{(2p)'}^1(\Omega)}}\leqslant C\|g_\varphi\|_{W_{2p}^1(\Omega)},\end{aligned}\tag{5.4}$$

where $(2p)'$ is the Hölder conjugate exponent to $2p$. Thus, $\varphi\in L_r(0,T;W_{2p}^1(\Omega))$, since $g_\varphi\in L_r(0,T;W_{2p}^1(\Omega))$, which implies $R(0)\in L_r(0,T;L_p(\Omega))$ in (B7).

For $\varphi_1=S(u_1)$ and $\varphi_2=S(u_2)$ we get

$$\begin{aligned}&\|R(u_1)-R(u_2)\|_{L_p(\Omega)}\\&\quad\leqslant\|(\sigma(u_1)-\sigma(u_2))|\nabla\varphi_1|^2\|_{L_p(\Omega)}+\|\sigma(u_2)(|\nabla\varphi_1|^2-|\nabla\varphi_2|^2\|_{L_p(\Omega)}\\&\quad\leqslant C(\|u_1-u_2\|_{L_\infty(\Omega)}\|\nabla\varphi_1\|_{L_{2p}(\Omega)}^2+\|\nabla(\varphi_1+\varphi_2)\|_{L_{2p}(\Omega)}\|\nabla(\varphi_1-\varphi_2)\|_{L_{2p}(\Omega)}).\end{aligned}$$

Using Sobolev's inequality we get $\|u_1-u_2\|_{L_\infty(\Omega)}\leqslant C\|u_1-u_2\|_{W_{2p(\Omega)}^1}$. Due to the Lipschitz property of $(-\nabla\cdot\sigma(u)\nabla)^{-1}$ we get

$$\|\tilde\varphi_1-\tilde\varphi_2\|_{W_{2p}^1(\Omega)}\leqslant C\|\sigma(u_1)-\sigma(u_2)\|_{L_\infty}\leqslant C\|u_1-u_2\|_{L_\infty}\leqslant C\|u_1-u_2\|_{W_{2p}^1(\Omega)}.$$

which proves (B7) and the continuity in (B6). Hence there exists a unique solution

$$u\in W_r^1(0,T_*;L_p(\Omega))\cap L_r(0,T_*;D(\Delta_p))$$

for some $T_*>0$.

Finally, by definition, $D(\Delta_p)$ denotes the domain such that the Laplacian maps into $L_p(\Omega)$. Let p' and $(2p)'$ be the Hölder conjugates to p and $2p$, respectively. Then

$$L_p(\Omega)=(L_{p'}(\Omega))^*,\quad W_{2p}^{-1}(\Omega;\Gamma_D^u)=(W_{(2p)'}^1(\Omega;\Gamma_D^u))^*.$$

By Sobolev's inequality we have,

$$W^1_{(2p)'}(\Omega) \subseteq L_{6p/(4p-3)}(\Omega),$$

and since $p' = p/(p-1) < 6p/(4p-3)$ when $p > 3/2$ we conclude $W^1_{(2p)'}(\Omega;\Gamma^u_D) \subseteq L_{p'}(\Omega)$, or equivalently $L_p(\Omega) \subseteq W^{-1}_{2p}(\Omega;\Gamma^u_D)$. Now, since we know that Δ is an isomorphism between $W^1_{2p}(\Omega;\Gamma^u_D)$ and $W^{-1}_{2p}(\Omega;\Gamma^u_D)$ and $L_p(\Omega) \subseteq W^{-1}_{2p}(\Omega;\Gamma^u_D)$ we deduce $D(\Delta_p) \subseteq W^1_{2p}(\Omega;\Gamma^u_D)$ and (5.3) follows. □

Note that the result is only local in time, that is, the additional regularity and uniqueness are only guaranteed up to some $T_* \leqslant T$.

We provide an example of a geometric setting for which (B5) is satisfied. Assume instead of (B1) the following

(B1') Ω is a creased domain with respect to the boundary conditions for u and φ. In addition $\text{meas}(\Gamma^u_D) > 0$ and $\text{meas}(\Gamma^\varphi_D) > 0$.

COROLLARY 5.1 Let $r > \frac{4p}{2p-3}$. Under the assumptions (B1') and (B2)-(B4), there exists $p > \frac{3}{2}$ such that there is a unique solution to (2.1) satisfying

$$\tilde{u} \in W^1_r(0,T_*;L_p(\Omega)) \cap L_r(0,T_*;W^1_{2p}(\Omega;\Gamma^u_D)), \quad \tilde{\varphi} \in L_r(0,T_*;W^1_{2p}(\Omega;\Gamma^\varphi_D)),$$

for some $0 < T_* \leqslant T$.

Proof. To see that condition (B5) is fulfilled we use the result on equations of Poisson type in Mitrea & Mitrea (2007). If Ω is a creased domain, $\Gamma_D \subseteq \partial\Omega$, and $g \in B^s_{q,q}(\Gamma_D)$, then there exists $\varepsilon \in (0, \frac{1}{2})$ such that Poisson's equation is well-posed in the spaces

$$v \in W_q^{s+\frac{1}{q}}(\Omega;\Gamma_D),\ -\Delta v = f \in \big(W_{q'}^{2-s-\frac{1}{q}}(\Omega;\Gamma_D)\big)^*,\ v|_{\Gamma_D} = g_u, \tag{5.5}$$

for $\frac{1}{q'} = 1 - \frac{1}{q}$ and $(s, \frac{1}{q}) \in \mathscr{H}_\varepsilon$ where $\mathscr{H}_\varepsilon$ is the polygon with vertices in

$$(0,0), \quad (\varepsilon,0), \quad (1, \frac{1}{2} - \varepsilon), \quad (1,1), \quad (1-\varepsilon,1), \quad (0, \frac{1}{2} + \varepsilon).$$

Choosing $s = \frac{8+\varepsilon}{12}$ and $q = \frac{12}{4-\varepsilon}$ we have for $p = \frac{6}{4-\varepsilon} > \frac{3}{2}$

$$W_q^{s+\frac{1}{q}} = W^1_{\frac{12}{4-\varepsilon}} = W^1_{2p} \text{ and } (\, W_{q'}^{2-s-\frac{1}{q}})^* = (W^1_{\frac{12}{8+\varepsilon}})^* = W^{-1}_{2p},$$

and since $W^1_{2p}(\Omega)|_{\Gamma_d} = B^s_{p,p}(\Gamma_D)$ assumption (B2) gives $g_u(t), g_\varphi(t) \in B^s_{q,q}(\Gamma^u_D)$. We conclude that (B5) holds. □

REMARK 5.1 There are other geometric settings where condition (B5) is fulfilled, see, for instance, Hieber & Rehberg (2008); Meinlschmidt *et al.* (2017); Disser *et al.* (2015).

REMARK 5.2 In Antontsev & Chipot (1994, Section 4) it is established that $\nabla\varphi \in L_{2q/(q-3)}(0,T;L_q(\Omega))$, for $q > 3$, is sufficient for a unique solution, see also Theorem 2.1. This agrees with the regularity we get in Theorem 5.2.

The next theorem shows that higher regularity achieved in the interior of the domain, cf. the stationary case Jensen & Målqvist (2013, Theorem 4.2). Here we use the notation $D(\Delta_{p,k})$ for the domain such that Δ maps into $W^k_p(\Omega)$. Note that $D(\Delta_p) = D(\Delta_{p,0})$.

THEOREM 5.3 Let $0 < T_0 < T_*$ and let Ω_0 be a relatively compact domain in Ω: $\Omega_0 \Subset \Omega$. Let (u, φ) be the solution to (2.1) such that

$$\tilde{u} \in W_r^1(0,T_*;L_p(\Omega)) \cap L_r(0,T_*;W_{2p}^1(\Omega;\Gamma_D^u)), \quad \tilde{\varphi} \in L_r(0,T_*;W_{2p}^1(\Omega;\Gamma_D^\varphi)), \tag{5.6}$$

for some $p > \frac{3}{2}$ and $r > \frac{4p}{2p-3}$. Then $\tilde{u}, \tilde{\varphi} \in L_r(T_0,T_*;W_s^2(\Omega_0))$ for all $s \in (1,\infty)$. If $\sigma \in C^\infty(\mathbb{R})$, then $\tilde{u}, \tilde{\varphi} \in L_r(T_0,T_*;C^\infty(\Omega_0))$.

Proof. Let Ω_∞ and $\{\Omega_i\}_{i=1}^\infty$ be smooth domains such that $\Omega_{i-1} \Subset \Omega_i$ and $\Omega_i \subset \Omega_\infty \Subset \Omega$, for $i = 0, 1, \ldots$. Assume, without loss of generality, that the boundary data g_u and g_φ have smooth extensions to Ω_∞ such that $g_u, D_t g_u, g_\varphi \in L_r(0,T;C^\infty(\Omega_\infty))$. Without loss of generality, we also assume $p = 6/(4-\varepsilon)$ for some $\varepsilon > 0$.

Let $\bar{\zeta}_i \in C^\infty(\Omega_i,[0,1])$, such that $\bar{\zeta}_i|_{\partial\Omega_i} = 0$ and $\bar{\zeta}_i|_{\Omega_{i-1}} = 1$. Furthermore, let $\{T_i\}_{i=1}^\infty$ and T_0 be positive numbers such that $T_i < T_{i-1}$, and $0 < T_i < T_0 < T_*$ for all i. Define $\eta_i(t) \in C^\infty([T_i,T_*],[0,1])$ such that $\eta_i(T_i) = 0$ and $\eta_i|_{[T_{i-1},T_0]} = 1$. Let $\zeta_i := \eta_i\bar{\zeta}_i$, then $(\zeta_i\tilde{u}, \zeta_i\tilde{\varphi})$ satisfies the following system in $\Omega_i \times (T_i, T_*)$.

$$\begin{aligned} D_t(\zeta_i\tilde{u}) - \Delta(\zeta_i\tilde{u}) = {} & \zeta_i\sigma(u)|\nabla\varphi|^2 - \zeta_i(D_t g_u - \Delta g_u) + \tilde{u}D_t\zeta_i - 2\nabla\zeta_i\cdot\nabla\tilde{u} \\ & - \tilde{u}\Delta\zeta_i, \end{aligned} \tag{5.7a}$$

$$\Delta(\zeta_i\tilde{\varphi}) = \zeta_i\frac{\sigma'(u)}{\sigma(u)}\nabla u\cdot\nabla\varphi - \zeta_i\Delta g_\varphi - 2\nabla\zeta_i\cdot\nabla\tilde{\varphi} - \tilde{\varphi}\Delta\zeta_i, \tag{5.7b}$$

with homogeneous Dirichlet conditions and zero initial data. Note that we have used

$$\nabla\cdot(\sigma(u)\nabla\varphi) = 0 \Leftrightarrow \Delta\varphi = \frac{\sigma'(u)}{\sigma(u)}\nabla u\cdot\nabla\varphi,$$

in the second equation. Because of the assumed regularity in (5.6) and the smoothness of ζ_i, g_u, and g_φ, the right-hand sides in (5.7) are in $L_r(T_i,T_*;L_p(\Omega_i))$.

There exists a unique solution in $W_p^2(\Omega_i)$ to Poisson's equation with homogeneous Dirichlet boundary conditions, if the domain is smooth and the right-hand side in $L_p(\Omega_i)$, $1 < p < \infty$, see e.g. Gilbarg & Trudinger (2001, Theorem 9.15). We conclude that, for a fixed t, $\zeta_i(t)\tilde{\varphi}(t) \in W_p^2(\Omega_i)$. We may now use elliptic regularity in L_p, see Gilbarg & Trudinger (2001, Lemma 9.17), to deduce

$$\|\zeta_i\tilde{\varphi}\|_{W_p^2(\Omega_i)} \leqslant C\|\zeta_i\frac{\sigma'(u)}{\sigma(u)}\nabla u\cdot\nabla\varphi - \zeta_i\Delta g_\varphi - 2\nabla\zeta_i\cdot\nabla\tilde{\varphi} - \tilde{\varphi}\Delta\zeta_i\|_{L_p(\Omega_i)}.$$

The regularity in time of the right hand side implies $\zeta_i\tilde{\varphi} \in L_r(T_i,T_*;W_p^2(\Omega_i))$. Thus $\tilde{\varphi} \in L_r(T_{i-1},T_*;W_p^2(\Omega_{i-1}))$, since $\zeta_i = 1$ on $[T_{i-1},T_*]\times\Omega_{i-1}$.

For the parabolic equation we use the theory for maximal L_p-regularity with homogeneous Dirichlet boundary conditions on smooth domains, see, e.g., Hieber & Prüss (1997, Theorem 3.1). If the right-hand side is in $L_r(0,T;L_p(\Omega))$ and the initial data is zero, then the solution belongs to $L_r(0,T;D(\Delta_p)) \cap W_r^1(0,T;L_p(\Omega))$. From the results on Poisson's equation we deduce $D(\Delta_p) \subset W_p^2(\Omega_i)$ and thus $\tilde{u} \in L_r(T_{i-1},T_*;W_p^2(\Omega_{i-1})) \cap W_r^1(T_{i-1},T_*;L_p(\Omega_{i-1}))$.

From the Sobolev inequality we have $W_p^2(\Omega_{i-1}) \subseteq W_{3p/(3-p)}^1(\Omega_{i-1})$. Using that $2p = 12/(4-\varepsilon)$ for some $\varepsilon > 0$ we get $3p/(3-p) = 12/(4-2\varepsilon)$. Hence, we can substitute ε by 2ε, pass from i to $i-1$ and repeat the argument. Note that if $12/(4-\varepsilon)$ becomes negative, the right-hand side is in $L_\infty(\Omega_i)$. By induction $\tilde{u}, \tilde{\varphi} \in L_r(T_0,T_*;W_s^2(\Omega_0))$ for any $s \in (1,\infty)$.

Now assume $\sigma \in C^\infty(\mathbb{R})$. A solution to Poisson's equation on a smooth domain is in $W_p^{k+2}(\Omega_i)$ if the right-hand side is in $W_p^k(\Omega_i)$, see e.g. Gilbarg & Trudinger (2001, Theorem 9.19). By applying Leibniz's rule it is clear that there is an s' such that the right-hand sides in (5.7) belongs to $L_r(T_i, T_*; W_s^k(\Omega_i))$ if $\tilde{\varphi}, \tilde{u} \in L_r(T_i, T_*; W_{s'}^{k+1}(\Omega_i))$. Hence, we may perform induction over k and pass from i to $i-1$, to achieve $\tilde{u}, \tilde{\varphi} \in L_r(T_0, T_*; W_s^{k+2}(\Omega_0))$, for any $k, s > 1$. This implies $\tilde{u}, \tilde{\varphi} \in L_r(T_0, T_*; C^\infty(\Omega_0))$. □

6. Numerical Examples

In this section we consider four different examples. The first two are designed to test the convergence rates for different settings. In the first example we choose the domain and the data such that the exact solution is known. To achieve this we add a function $f(x,t)$ to the right-hand side in (2.4a) and consider non-zero Neumann data for ϕ, see Subsection 6.1 below. For the second example we consider a setting that does not fulfil the creased domain conditions. For this problem we expect low regularity and reduced convergence rates. Finally, in the last two examples we test a goal oriented adaptivity method.

In all cases we consider a continuous, piecewise affine finite element discretization. We let $\{\mathscr{T}_m\}_m$ denote a family of uniform triangulations of the domain such that $h_{m+1} = 2^{-1}h_m$, $h_0 \in \mathbb{R}$, where h_m is the maximal mesh size on $\mathscr{T}_m$. With this notation we may define

$$V_m^u := \{v \in H^1(\Omega; \Gamma_D^u) \cap C^0(\bar{\Omega}) : v|_K \text{ is a polynomial of degree} \leqslant 1, \forall K \in \mathscr{T}_h\},$$
$$V_m^\varphi := \{v \in H^1(\Omega; \Gamma_D^\varphi) \cap C^0(\bar{\Omega}) : v|_K \text{ is a polynomial of degree} \leqslant 1, \forall K \in \mathscr{T}_h\}.$$

For the time discretization, we let $\tau_l = 2^{-l}T$ and the fully discrete space $X_{m,l}$ is defined as in Section 4.

In the first two experiments we keep the time step proportional to the mesh size in each refinement. That is, we consider spaces of the form $X_{k,k}$, for $k = 1, 2, 3, \ldots$. This means that if the solution has sufficient regularity, then we expect at most linear convergence rate in the norm $L_2(0,T;H^1(\Omega))$, see also Elliott & Larsson (1995); Målqvist & Stillfjord (2017); Gao (2014).

All computations are made using the FEniCS software Logg *et al.* (2012).

6.1 *Example 1*

We let $T = 0.1$, Ω be the unit cube, $\Gamma_D^u = \partial\Omega$, and $\Gamma_D^\varphi = \partial\Omega \setminus \{x_3 = 0 \text{ or } x_3 = 1\}$. To construct an example where the exact solution is known, we consider non-zero Neumann data g_N for φ and an additional function f in the right-hand side of (2.4a). We get

$$\langle D_t u, v\rangle + \langle \nabla u, \nabla v\rangle = -\langle \sigma(u)\lceil\tilde{\varphi}\rceil \nabla\varphi, \nabla v\rangle + \langle \sigma(u)\nabla\varphi \cdot \nabla g_\varphi, v\rangle + \langle f, v\rangle,$$
$$\langle \sigma(u)\nabla\varphi, \nabla w\rangle = \langle g_N, w\rangle_{\Gamma_N^\varphi},$$

where $\langle\cdot,\cdot\rangle_{\Gamma_N^\varphi}$ denotes integration on the boundary Γ_N^φ.

Letting $g_u = t, g_\varphi = x_2, g_N = -1 + 2x_2$, $\sigma = 1$, and

$$f = 2(x_1x_2(1-x_1)(1-x_2) + x_1x_3(1-x_1)(1-x_3) + x_2x_3(1-x_2)(1-x_3)),$$
$$u_0 = x_1(1-x_1)x_2(1-x_2)x_3(1-x_3),$$

the exact solution is given by $u = x_1(1-x_1)x_2(1-x_2)x_3(1-x_3) + t$ and $\varphi = x_2$. Note that $\varphi = \tilde{\varphi} + g_\varphi$ and thus $\tilde{\varphi} = 0$. In our setting, the approximations $\varphi_{m,l}$ are all close to zero and, hence, we omit to plot the error for φ below.

We compute the finite element approximation on meshes with tetrahedra of maximal diameter $h = 2^{-k}\sqrt{3}$ and time step size $\tau = 2^{-k}T = 2^{-k}0.1$ for $k = 1,..,6$. With this refinement, the finest approximation ($k = 6$) is computed on a mesh with 274625 nodes. The error in $L_2(0,T;H^1(\Omega))$ is approximated using Simpson's rule in time on each interval I_n and the FEniCS function `errornorm` in space. The relative error is depicted in Figure 1. The convergence rate is approximately linear, which is expected for sufficiently regular problems.

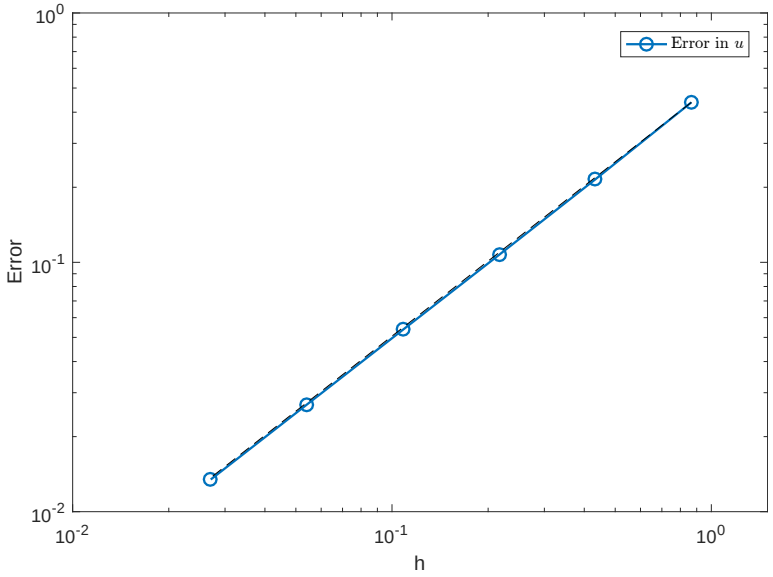
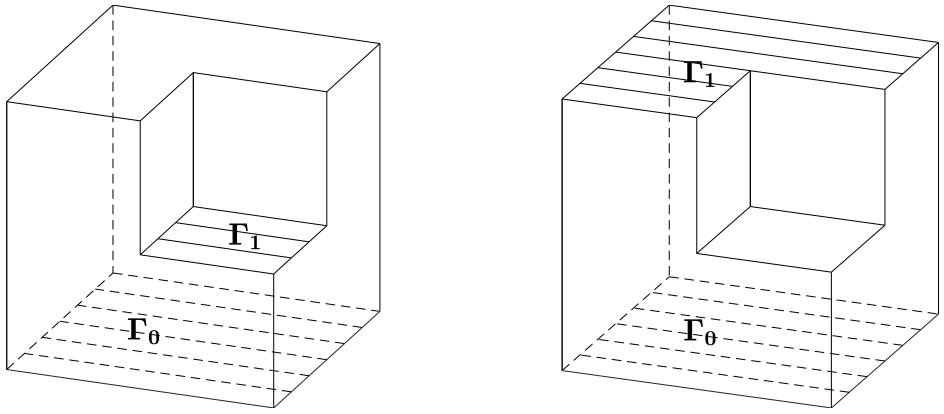


FIG. 1: Relative errors for the temperature u (blue o) of Example 1 plotted against the mesh size h. The dashed line is Ch.

6.2 *Example 2*

We let $T = 0.1$ and Ω be the Fichera cube depicted in Figure 2 (left). We consider non-creased boundary conditions by imposing Dirichlet conditions on the striped areas, Γ_0 and Γ_1 in the figure (left), and homogeneous Neumann conditions on the remaining parts. On Γ_0 we set $g_u = 0$ and $g_\varphi = 10$ and on Γ_1 we set $g_u = g_\varphi = 0$. Furthermore, we let $\sigma(u) = 2^{-1}(\pi - \arctan(u))$ and $u_0 = 0$.

We compute the finite element approximation on meshes with tetrahedra of maximal diameter $h = 2^{-(k-1)}\sqrt{2}$ and time step size $\tau = 2^{-k}T = 2^{-k}0.1$ for $k = 1,..,5$. Since the exact solution is not known, the approximations are compared to a reference solution computed for $k = 6$ corresponding to a mesh with 471233 nodes. The relative error in $L_2(0,T;H^1(\Omega))$-norm is plotted in Figure 3. We have convergence, but not with order one. This is due to the low regularity in the vicinity of the edges where the Dirichlet and Neumann boundaries meet with an angle greater than π, that is, where the creased domain condition fails.



(a) Non-creased (b) Creased

FIG. 2: Two settings of the Fichera cube with centre in the origin. Dirichlet boundary conditions are imposed on the striped areas (Γ_0 and Γ_1).

6.3 *Example 3*

We continue in the setting of the Fichera cube as in Example 2, but with a choice of boundary conditions that fulfil the creased domain condition. We choose Γ_0 and Γ_1 as in Figure 2 (right), with $g_u = 0$ and $g_\varphi(x,t) = 2x_2(x_2+1)+5$ on both Γ_0 and Γ_1.

The aim is to utilize the observation that the solution has higher regularity in the interior of the domain, see Theorem 5.3, and that the problem thus is suitable for h-adaptive finite elements. In this example we use a goal-oriented approach for the mesh refinement, which is supported for stationary problems in the FEniCS software, see Rognes & Logg (2013).

We summarize the goal-oriented procedure here, and refer to Rognes & Logg (2013) and references therein for details. Consider a nonlinear variational problem; find $u \in V$ such that

$$F(u,v) = 0, \quad \forall v \in \hat{V}, \tag{6.1}$$

and the corresponding finite element problem; find $u_h \in V_h$ such that

$$F(u_h,v) = 0, \quad \forall v \in \hat{V}_h, \tag{6.2}$$

for some triangulation $\mathscr{T}_h$ and appropriate finite element space $V_h \subset V$, $\hat{V}_h \subset \hat{V}$. Let $\mathscr{M} : V \to \mathbb{R}$ denote a linear goal functional and define the dual problem; find $z \in V^*$ such that

$$\overline{F'}^*(z,v) = \mathscr{M}(v), \quad \forall v \in \hat{V}^*,$$

where $\hat{V}^* = V_0 = \{v - w : v,w \in V\}$ and $V^* = \hat{V}$. The bilinear form $\overline{F'}^*$ denotes the following average of the Fréchet derivative F' of F,

$$\overline{F'}(\cdot,\cdot) = \int_0^1 F'(su + (1-s)u_h;\cdot,\cdot)\,\mathrm{d}s,$$

and by the chain rule we have $\overline{F'}(u-u_h,\cdot) = F(u,\cdot) - F(u_h,\cdot)$.

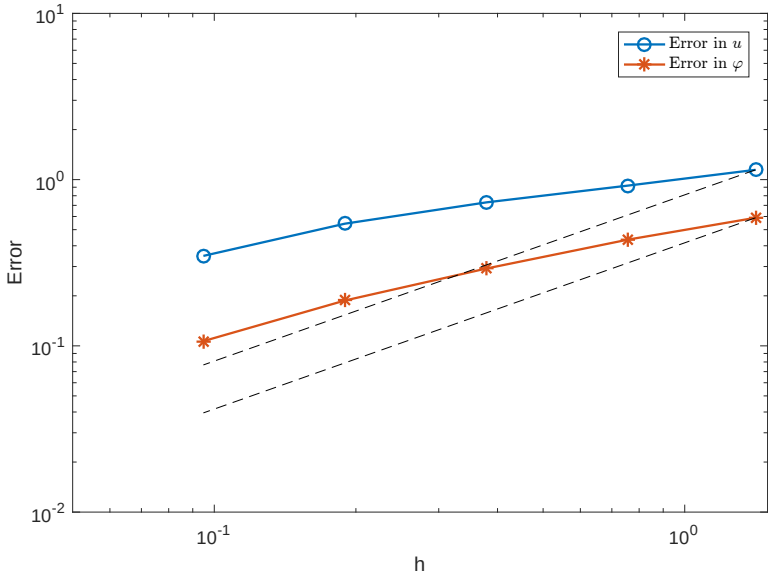


FIG. 3: Relative errors for the temperature u (blue $\circ$) and the potential φ (red $*$) of Example 2 plotted against the mesh size h. The dashed line is Ch.

Using the definition of the dual problem we may now express the error in the goal functional as

$$\begin{aligned}\mathscr{M}(u)-\mathscr{M}(u_h)=\mathscr{M}(u-u_h)&=\overline{F'}^{*}(z,u-u_h)=\overline{F'}(u-u_h,z)\\&=F(u,z)-F(u_h,z)=-F(u_h,z)=:r(z),\end{aligned}$$

where $r(z)$ denotes the residual. The residual can be decomposed into local contributions from each cell $T\in\mathscr{T}_h$

$$r(v)=\sum_{T\in\mathscr{T}_h}r_T(v)=\sum_{T\in\mathscr{T}_h}\left(\int_T R_T v\,\mathrm{d}x+\int_{\partial T}R_{\partial T}v\,\mathrm{d}s\right),$$

where R_T and $R_{\partial T}$ are the cell and facet residuals. In Rognes & Logg (2013, Theorem 4.1) it is proved that the error indicators $R_T, R_{\partial T}$ can be determined by solving a set of local problems on each cell T.

The procedure of computing the error indicators R_T and $R_{\partial T}$ and refining the mesh accordingly is performed in FEniCS by using `solve` together with the goal functional and a given tolerance. In our case, the fully discrete problem (4.1) is a stationary problem of the form (6.1) with

$$\begin{aligned}F((u^n_{m,l},\varphi^n_{m,l}),(v,w)):=&\langle\frac{u^n_{m,l}-u^{n-1}_{m,l}}{\tau_l},v\rangle+\langle\nabla u^n_{m,l},\nabla v\rangle\\&+\langle\sigma(u^n_{m,l})\lceil\tilde{\varphi}^n_{m,l}\rceil\nabla\varphi^n_{m,l},\nabla v\rangle\\&-\langle\sigma(u^n_{m,l})\nabla\varphi^n_{m,l}\cdot\nabla g^n_{\varphi},v\rangle+\langle\sigma(u^n_{m,l})\nabla\varphi^n_{m,l},\nabla w\rangle.\end{aligned}$$

In each time step the error indicators are computed and the mesh refined. Note that the refined mesh is reused in the next time step and additionally refined if needed.

In this example we choose $\mathscr{M}(u) = \int_\Omega u \,\mathrm{d}x$. The initial data remains the same as in Example 2. We choose to have fixed (small) time step $\tau = 2^{-6}T$ in this experiment, since the the spatial error is the main concern here. The relative error in the goal functional for $h = 2^{-4}\sqrt{2}$ compared to the reference solution, here denoted u_{ref} and computed on mesh with 471233 nodes, is

$$\frac{\max_{0\leqslant n\leqslant N} |\mathscr{M}(u_h^n) - \mathscr{M}(u_{\text{ref}}^n)|}{\max_{0\leqslant n\leqslant N} |\mathscr{M}(u_{\text{ref}}^n)|} \approx 0.0254.$$

We note that our uniform mesh of size $h = 2^{-4}\sqrt{2}$ corresponds to 7985 nodes. Using the goal oriented adaptivity, denoted u_{ad} below, we get

$$\frac{\max_{0\leqslant n\leqslant N} |\mathscr{M}(u_{\text{ad}}^n) - \mathscr{M}(u_{\text{ref}}^n)|}{\max_{0\leqslant n\leqslant N} |\mathscr{M}(u_{\text{ref}}^n)|} \approx 0.0282,$$

already for 1628 nodes.

This example indicates that the problem is suitable for h-adaptive finite elements and motivates a further analysis of a posteriori methods for the Joule heating problem, which will be considered in later works.

6.4 *Example 4*

In this example, we use the non-creased Fichera cube as in Example 2, see Figure 2 (left). The aim is to investigate the use of goal-oriented adaptivity for non-creased domains. We emphasize that, in this setting, Theorem 5.3 is not directly applicable.

As in Example 3 we choose $\mathscr{M}(u) = \int_\Omega u \,\mathrm{d}x$. The initial and boundary data remain the same as in Example 2 and the time step is $\tau = 2^{-6}T$. The error in the goal functional for $h = 2^{-5}\sqrt{2}$ compared to the reference solution is

$$\frac{\max_{0\leqslant n\leqslant N} |\mathscr{M}(u_h^n) - \mathscr{M}(u_{\text{ref}}^n)|}{\max_{0\leqslant n\leqslant N} |\mathscr{M}(u_{\text{ref}}^n)|} \approx 0.0271.$$

Here $h = 2^{-5}\sqrt{2}$ corresponds to 60513 nodes. For the goal oriented adaptivity we get

$$\frac{\max_{0\leqslant n\leqslant N} |\mathscr{M}(u_{\text{ad}}^n) - \mathscr{M}(u_{\text{ref}}^n)|}{\max_{0\leqslant n\leqslant N} |\mathscr{M}(u_{\text{ref}}^n)|} \approx 0.0254,$$

for 6560 nodes.

This example indicates that the goal oriented adaptivity is applicable also in non-creased domain settings. However, it is still an open problem to show that the solution to such a problem enjoys the appropriate regularity to be suitable for h-adaptivity.

Acknowledgement

The authors acknowledge the hospitality of the Hausdorff Research Institute for Mathematics in Bonn, where parts of this paper were written.

Funding

Axel Målqvist's work is supported by the Swedish Research Council and The Göran Gustafsson Foundation for Research in Natural Sciences and Medicine.

REFERENCES

AKRIVIS, G. & LARSSON, S. (2005) Linearly implicit finite element methods for the time-dependent Joule heating problem. *BIT*, **45**, 429–442.

AMANN, H. (1990) *Ordinary differential equations*. De Gruyter Studies in Mathematics, vol. 13. Walter de Gruyter & Co., Berlin, pp. xiv+458. An introduction to nonlinear analysis, Translated from the German by Gerhard Metzen.

ANTONTSEV, S. N. & CHIPOT, M. (1994) The thermistor problem: existence, smoothness uniqueness, blowup. *SIAM J. Math. Anal.*, **25**, 1128–1156.

BANK, R. E. & YSERENTANT, H. (2014) On the H^1-stability of the L_2-projection onto finite element spaces. *Numer. Math.*, **126**, 361–381.

BRAESS, D. (2007) *Finite elements*, third edn. Cambridge University Press, Cambridge, pp. xviii+365. Theory, fast solvers, and applications in elasticity theory, Translated from the German by Larry L. Schumaker.

BROWN, R. (1994) The mixed problem for Laplace's equation in a class of Lipschitz domains. *Comm. Partial Differential Equations*, **19**, 1217–1233.

BROWN, R. M. & MITREA, I. (2009) The mixed problem for the Lamé system in a class of Lipschitz domains. *J. Differential Equations*, **246**, 2577–2589.

CIMATTI, G. (1992) Existence of weak solutions for the nonstationary problem of the joule heating of a conductor. *Ann. Mat. Pura Appl. (4)*, **162**, 33–42.

DISSER, K., KAISER, H.-C. & REHBERG, J. (2015) Optimal Sobolev regularity for linear second-order divergence elliptic operators occurring in real-world problems. *SIAM J. Math. Anal.*, **47**, 1719–1746.

ELLIOTT, C. M. & LARSSON, S. (1995) A finite element model for the time-dependent Joule heating problem. *Math. Comp.*, **64**, 1433–1453.

GAO, H. (2014) Optimal error analysis of Galerkin FEMs for nonlinear Joule heating equations. *J. Sci. Comput.*, **58**, 627–647.

GILBARG, D. & TRUDINGER, N. S. (2001) *Elliptic partial differential equations of second order*. Classics in Mathematics. Springer-Verlag, Berlin, pp. xiv+517. Reprint of the 1998 edition.

HENNEKEN, V. A., TICHEM, M. & SARRO, P. M. (2006) In-package mems-based thermal actuators for micro-assembly. *Journal of Micromechanics and Microengineering*, **16**, S107.

HIEBER, M. & PRÜSS, J. (1997) Heat kernels and maximal L^p-L^q estimates for parabolic evolution equations. *Comm. Partial Differential Equations*, **22**, 1647–1669.

HIEBER, M. & REHBERG, J. (2008) Quasilinear parabolic systems with mixed boundary conditions on nonsmooth domains. *SIAM J. Math. Anal.*, **40**, 292–305.

HOLST, M. J., LARSON, M. G., MÅLQVIST, A. & SÖDERLUND, R. (2010) Convergence analysis of finite element approximations of the Joule heating problem in three spatial dimensions. *BIT*, **50**, 781–795.

HOWISON, S. D., RODRIGUES, J. F. & SHILLOR, M. (1993) Stationary solutions to the thermistor problem. *J. Math. Anal. Appl.*, **174**, 573–588.

JENSEN, M. & MÅLQVIST, A. (2013) Finite element convergence for the Joule heating problem with mixed boundary conditions. *BIT*, **53**, 475–496.

LI, B., GAO, H. & SUN, W. (2014) Unconditionally optimal error estimates of a Crank-Nicolson Galerkin method for the nonlinear thermistor equations. *SIAM J. Numer. Anal.*, **52**, 933–954.

LOGG, A., MARDAL, K.-A., WELLS, G. N. *et al.* (2012) *Automated Solution of Differential Equations by the Finite Element Method*. Springer.

MÅLQVIST, A. & STILLFJORD, T. (2017) Finite element convergence analysis for the thermoviscoelastic joule heating problem. *BIT Numerical Mathematics*, **57**, 787–810.

MEINLSCHMIDT, H., MEYER, C. & REHBERG, J. (2017) Optimal Control of the Thermistor Problem in Three Spatial Dimensions, Part 1: Existence of Optimal Solutions. *SIAM J. Control Optim.*, **55**, 2876–2904.

MITREA, I. & MITREA, M. (2007) The Poisson problem with mixed boundary conditions in Sobolev and Besov spaces in non-smooth domains. *Trans. Amer. Math. Soc.*, **359**, 4143–4182 (electronic).

ROGNES, M. E. & LOGG, A. (2013) Automated goal-oriented error control I: Stationary variational problems. *SIAM J. Sci. Comput.*, **35**, C173–C193.

ROUBÍČEK, T. (2013) *Nonlinear partial differential equations with applications*. International Series of Numerical Mathematics, vol. 153, second edn. Birkhäuser/Springer Basel AG, Basel, pp. xx+476.

ROYDEN, H. L. (1988) *Real analysis*, third edn. Macmillan Publishing Company, New York, pp. xx+444.

THOMÉE, V. (2006) *Galerkin Finite Element Methods for Parabolic Problems*. Springer Series in Computational Mathematics, second edn. Berlin: Springer-Verlag, pp. xii–370.

WALKINGTON, N. J. (2010) Compactness properties of the DG and CG time stepping schemes for parabolic equations. *SIAM J. Numer. Anal.*, **47**, 4680–4710.

YOSIDA, K. (1995) *Functional analysis*. Classics in Mathematics. Springer-Verlag, Berlin, pp. xii+501. Reprint of the sixth (1980) edition.

YUAN, G. W. & LIU, Z. H. (1994) Existence and uniqueness of the C^{α} solution for the thermistor problem with mixed boundary value. *SIAM J. Math. Anal.*, **25**, 1157–1166.