THESIS FOR THE DEGREE OF DOCTOR OF PHILOSOPHY

The Finite Element Method for Fractional Order Viscoelasticity and the Stochastic Wave Equation

Fardin Saedpanah

CHALMERS | UNIVERSITY OF GOTHENBURG

Department of Mathematical Sciences
Chalmers University of Technology and University of Gothenburg
Gothenburg, Sweden 2009

The Finite Element Method for Fractional Order Viscoelasticity and the Stochastic Wave Equation
Fardin Saedpanah
ISBN 978-91-7385-273-9

Doktorsavhandlingar vid Chalmers Tekniska högskola
Ny serie nr 2954
ISSN 0346-718X

Department of Mathematical Sciences
Division of Mathematics
Chalmers University of Technology and University of Gothenburg
SE-412 96 Göteborg
Sweden
Telephone +46 (0)31 772 1000

Printed at the Department of Mathematical Sciences
Göteborg, Sweden 2009

The Finite Element Method for Fractional Order Viscoelasticity and the Stochastic Wave Equation

Fardin Saedpanah

Department of Mathematical Sciences
Chalmers University of Technology
and University of Gothenburg

Abstract

This thesis can be considered as two parts. In the first part a hyperbolic type integro-differential equation with weakly singular kernel is considered, which is a model for dynamic fractional order viscoelasticity. In the second part, the finite element approximation of the linear stochastic wave equation is studied. The link between these two equations is that they are both treated as perturbations of the linear wave equation.

Our study in the first part comprises investigating well-posedness of the model, and the analysis of the finite element approximation of the solution of the model problem. The equation, with homogeneous mixed Dirichlet and Neumann boundary conditions, is reformulated as an abstract Cauchy problem, and existence, uniqueness and regularity are verified in the context of linear semigroup theory. From a practical viewpoint, the problems with mixed homogeneous Dirichlet and non-homogeneous Neumann boundary conditions are of special importance. Therefore, the Galerkin method is used to prove existence, uniqueness and regularity of the solution of this type of problem. Then two variants of the continuous Galerkin finite element method are applied to the model problem. Stability properties of the discrete and the continuous problem are investigated. These are then used to obtain optimal order a priori estimates and global a posteriori error estimates. In a general framework, a space-time cellwise a posteriori error representation is also presented. The theory is illustrated by an example.

The second part concerns the study of the semidiscrete finite element approximation of the linear stochastic wave equation with additive noise in a semigroup framework. Optimal error estimates for the deterministic problem are obtained under minimal regularity assumptions. These are used to prove strong convergence estimates for the stochastic problem. The theory presented here applies to multi-dimensional domains and correlated noise. Numerical examples illustrate the theory.

Keywords: finite element method, continuous Galerkin method, linear viscoelasticity, fractional calculus, fractional order viscoelasticity, weakly singular kernel, stability, a priori error estimate, a posteriori error estimate, stochastic wave equation, additive noise, Wiener process, strong convergence.

Dissertation

This thesis consists of a short review and four papers:

Paper I: *The continuous Galerkin method for an integro-differential equation modeling dynamic fractional order viscoelasticity.*
IMA Journal of Numerical Analysis, to appear (with Stig Larsson).

Paper II: *Existence and uniqueness of the solution of an integro-differential equation with weakly singular kernel.*
Preprint 2009:16, Department of Mathematical Sciences, Chalmers University of Technology and University of Gothenburg.

Paper III: *A continuous space-time finite element method for an integro-differential equation modeling dynamic fractional order viscoelasticity.*
Preprint 2009:17, Department of Mathematical Sciences, Chalmers University of Technology and University of Gothenburg.

Paper IV: *Finite element approximation for the linear stochastic wave equation with additive noise.*
Preprint 2009:18, Department of Mathematical Sciences, Chalmers University of Technology and University of Gothenburg (with Mihály Kovács and Stig Larsson).

Contributions to co-authored papers

Paper I:
Took part in the theoretical developments.
Did a large part of the writing.
Carried out the coding and numerical experiments.

Paper IV:
Proved the error estimates for the deterministic problem.
Took part in the proof of the main theorem for the stochastic problem.
Did a large part of the writing.
Carried out the coding and numerical experiments.

Acknowledgements

First of all I would like to thank my supervisor Professor Stig Larsson for teaching me and for his excellent guidance and overall support. His supervision and support with great ideas in research is invaluable. Also I give thanks to my assistant advisor Professor Ivar Gustafsson.

I must express my appreciation to Professor Mohammad Asadzadeh for his guidance and support during the while that I have been here. I would like to thank Professor Mikael Enelund for his constructive comments when he was the opponent of my licentiate degree. I also thank Professor Thomas Ericsson for his excellent guidance on numerical linear algebra. I am grateful to Professor Milena Racheva for very nice and helpful discussion on fractional order viscoelasticity models. Thanks to Professor Mihály Kovács for interesting lectures on stochastic PDEs and for his help as a coauthor for the paper on the stochastic wave equation.

I would like to thank the staff in the department of Mathematical Sciences, in particular, Marie Kühn and Camilla Nygren. I want to thank my colleagues and friends for creating a pleasant work environment. In particular Ali Mesforush, Fredrik Lindgren, Karin Kraft, Christoffer Cromvik, Hasan Almanasreh, Iulia Pop, and David Heintz.

I would like to thank my parents and my sisters for giving me such a great support and encouragement in life.

Last, but not least, I am very thankful to my wife and my daughter for their constant support, patience and kindness. I run short of words and I dedicate this work to my family Kajal and Ainaz.

Fardin Saedpanah
Gothenburg, April 2009

Contents

1 Introduction

The fractional viscoelastic model, i.e., the linear viscoelastic model with fractional order operators in the constitutive equations, is capable of describing the behaviour of many viscoelastic matrials by using only a few parameters. Therefore the fractional order viscoelastic models have attracted considerable attention in the last decades.

The finite flement method (FEM) is a numerical technique for finding approximate solutions of partial differential equations (PDE) as well as of integral equations. This method originated from the need for solving complex elasticity and structural analysis problems in civil and aeronautical engineering. The FEM allows detailed visualization of where structures bend or twist, and indicates the distribution of stresses and displacements. Generally, FEM is the method of choice in all types of analysis in structural mechanics, i.e., solving for deformation and stresses in solid bodies or dynamics of structures.

This thesis can be considered as two parts. In the first part we study a hyperbolic type integro-differential equation with weakly singular kernel, modeling dynamic fractioal order viscoelasticity. Our study in the first part comprises investigating well-posedness of the model, and the analysis of the finite element approximation of the solution of the model problem consisting of implementation, stability, a priori and a posteriori analysis. We have collected the obtained results in the first three appended papers. The second part is devoted to studying the finite element approximation of the linear stochastic wave equation. This resulted in one paper, Paper IV. These two parts might seem different, but we consider the main subject as studying the wave problem with two types of perturbations; the first case is a perturbation with a memory term, similar to the fractional order viscoelasticity model, and the second case is a perturbation with a noise in the load term, which is the linear stochastic wave equation.

In the sequel, after providing some basic concepts from fractional calculus, we explain the derivation of the fractional viscoelasticity model, that is the main model for the first part of this work. Then in §4 we provide some materials from Paper I and Paper II, where we prove well-posedness of the model problem. In §5 we use the classical wave equation on a bounded domain to highlight the main ideas for implementation of the continuous Galerkin mehtod, and the corresponding analysis. These have been used in Paper I, Paper III, and Paper IV. We devote §6 to explain the main feature of the linear stochastic wave equation and the results from Paper IV. We discuss some works which have been done in the past in §7. Finally we summerize the appended papers.

2 Fractional calculus

Generalization have always been an interesting subject in mathematics. One example is the continuous gamma function which interpolates between the factorials. Another one is the fractional differential/integral operators which interpolates between integer order differential/integral operators. In fact analytic continuation of the gamma function for $x \leq 0$ plays an important role when we construct the theory of fractional order differential/integral operators from the corresponding integer order operators. In the following we describe the main ideas of these generalizations.

A brief historical overview of the development of fractional calculus is given by Ross [55]. The text books Oldham and Spanier [48] and Samko et al. [58] are concerned with the definitions and the properties of fractional order differential/integral operators. A survey of the many different applications which have emerged from fractional calculus is given in Podlubny [50].

In contrast to the term "fractional" the fractional order exponent can be irrational and even complex. However, in this context we take it to be real.

2.1 Gamma function

A comprehensive definition of the the gamma function $\Gamma(x)$ is that provided by the Euler limit

$$\Gamma(x) = \lim_{N \to \infty} \frac{N! N^x}{x(x+1)(x+2) \cdots (x+N)},$$

but the integral transform definition

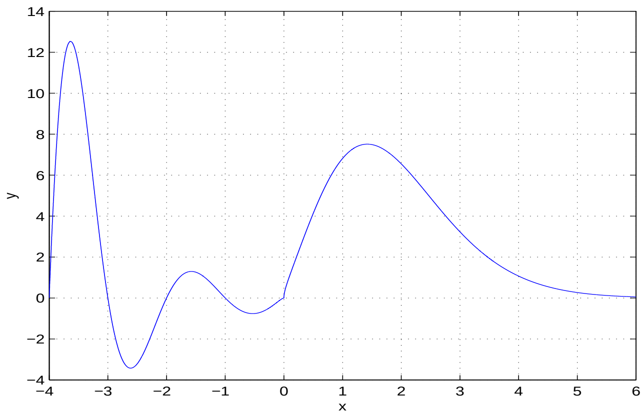
$$\Gamma(x) = \int_0^\infty t^{x-1} e^{-t}\, dt, \quad x > 0,$$

is often more useful, altough it is restricted to $x > 0$. Integration by parts then leads to the recurrence relationship $\Gamma(x+1) = x\Gamma(x)$, that is also a simple consequence of the Euler definition. Having $\Gamma(1) = 1$, for a positive integer n, we have the factorial $n! = n(n-1) \cdots 2 \cdot 1$. Rewriting the recurrence relationship as $\Gamma(x-1) = \Gamma(x)/(x-1)$, shows that $\Gamma(0)$ is infinite, as is $\Gamma(-1)$ and the value of gamma function at all negative integers. However, ratios of gamma functions of negative integers are finite, that is, for positive integers $n \leq N$

$$\frac{\Gamma(-n)}{\Gamma(-N)} = (-N)(-N+1) \cdots (-n-2)(-n-1) = (-1)^{N-1} \frac{N!}{n!}.$$

The reciprocal $1/\Gamma(x)$ of the gamma function is single-valued and finite for all x. The figure below shows a graph of this function described by

$$\frac{1}{\Gamma(x)} \sim \frac{x^{1/2-x}}{\sqrt{2\pi}} e^x, \quad x \to \infty.$$



In the generalization of the integer order differential/integral operators we will use the gamma function expression

$$\frac{\Gamma(j-q)}{\Gamma(-q)\Gamma(j+1)}, \tag{2.1}$$

where j is a nonnegative integer and q may takes any value. This can be expressed as a polynomial in q, in terms of Stirling numbers $S_j^{(m)}$,

$$\frac{\Gamma(j-q)}{\Gamma(-q)\Gamma(j+1)} = \frac{(-1)^j}{j!} \sum_{m=0}^{j} S_j^{(m)} q^m,$$

and establishes that (2.1) is finite and single-valued for all finite values of q and j, see Oldham and Spanier [48]. Though, we will use a binomial coefficient expression

$$\frac{\Gamma(j-q)}{\Gamma(-q)\Gamma(j+1)} = \binom{j-q-1}{j} = (-1)^j \binom{q}{j}. \tag{2.2}$$

We remark that one other way of extension is to split the gamma function in the form

$$\Gamma(x) = \int_0^1 t^{x-1} e^{-t}\, dt + \int_1^\infty t^{x-1} e^{-t}\, dt.$$

The first term can be evaluated by using the series expansion for the exponential function, and the second integral defines an entire function, see Podlubny [50]. That is, bringing these together we have

$$\Gamma(z) = \sum_{k=0}^{\infty} \frac{(-1)^k}{k!} \frac{1}{k+z} + \text{entire function},$$

and, indeed $\Gamma(z)$ has only simple poles at the points $z = -n$, $n = 0, 1, 2, \ldots$.

2.2 Fractional differential/integral operators

The first definition that can be offered is the one that is regarded as the most fundamental definition. That is, it involves the fewest restrictions on the functions to which it applies and avoid explicit use of the notions of ordinary derivative and integral. To this end, we first recall a unification of two notions which are usually presented separately in classical analysis: derivate of integer order $n \geq 0$ and n-fold integral. Using the basic definition of the nth derivative and the n-fold integrals, that are, respectively, the limit of difference quotient and the limit of Riemann sum, we have

$$D^n f(t) = \frac{d^n f}{dt^n}(t) = \lim_{N\to\infty} \left(\frac{t}{N}\right)^{-n} \sum_{j=0}^{N} (-1)^j \binom{n}{j} f\left(t - j\frac{t}{N}\right),$$

$$D^{-n} f(t) = \frac{d^{-n} f}{dt^{-n}}(t) = \lim_{N\to\infty} \left(\frac{t}{N}\right)^{n} \sum_{j=0}^{N} (-1)^j \binom{j+n-1}{j} f\left(t - j\frac{t}{N}\right),$$

where n is a nonnegative integer. In general we need to consider a lower limit a that is a number smaller than t, and in that case we would have $\frac{t-a}{N}$ instead. We use $a = 0$ to ease the notation. Now recalling equation (2.2) we have a unified definition, for any integer number q,

$$D^q f(t) = \lim_{N\to\infty} \left(\frac{t}{N}\right)^{-q} \sum_{j=0}^{N} (-1)^j b_j f\left(t - j\frac{t}{N}\right), \quad \text{with } b_j = \frac{\Gamma(j-q)}{\Gamma(-q)\Gamma(j+1)}.$$

Letting q be an arbitrary number we obtain the Grünwald's formula of the fractional order differential/integral operators, that was introduced by Grünwald (1867) and later extended by Post (1930). A suitable truncation of the Grünwald's definition is often used for numerical approximation of fractional derivatives and integrals, see, e.g., Podlubny [50].

A frequently encountered definition of an integral of fractional order is via an integral transform called the Riemann-Liouville integral. To motivate this definition we recall Cauchy's formula for repeated integration

$$D^{-n}f(t) = \int_0^t \int_0^{t_{n-1}} \cdots \int_0^{t_1} f(t_0)\, dt_0 \cdots dt_{n-2}\, dt_{n-1}$$
$$= \frac{1}{(n-1)!} \int_0^t \frac{f(s)}{(t-s)^{1-n}}\, ds, \quad n = 1, 2, \cdots ,$$

with $D^0 f(t) = f(t)$. Replacing the integer number n with the real number α and the discrete factorial $(n-1)!$ with the continuous gamma function Γ, the Riemann-Liouville fractional integral is obtained

$$D^{-\alpha}f(t) = \frac{1}{\Gamma(\alpha)} \int_0^t \frac{f(s)}{(t-s)^{1-\alpha}}\, ds, \quad \alpha > 0, \tag{2.3}$$

where α is the order of integration. Note that the convolution kernel $\frac{1}{\Gamma(\alpha)t^{1-\alpha}}$ is singular but integrable.

From the Riemann-Liouville fractional integral (2.3) at least two definitions of fractional differentiation can be formulated. The most common definition is the Riemann-Liouville fractional derivative. Formally, the same definition can be used for fractional differentiation of order α by making the replacement $-\alpha \to \alpha$ in (2.3), that is,

$$D^{\alpha}f(t) = \frac{1}{\Gamma(-\alpha)} \int_0^t \frac{f(s)}{(t-s)^{1+\alpha}}\, ds, \quad \alpha > 0.$$

The convolution integral above is in general divergent and needs to be interpreted in the sense of its regularization. A convergent expression for the fractional derivative operator is obtained by splitting the derivative operator into an integer order derivative and a fractional integral operator. That is for $\alpha \geq 0$,

$$D^{\alpha}f(t) = D^n D^{\alpha-n} f(t) = \frac{1}{\Gamma(n-\alpha)} \frac{d^n}{dt^n} \left[\int_0^t \frac{f(s)}{(t-s)^{\alpha-n+1}}\, ds \right],$$

where n is the integer that satisfies $n-1 < \alpha \leq n$. Another related definition of fractional differentiation is the so-called Caputo derivative, see, e.g., Caputo [13] and Podlubny [50],

$$D_*^{\alpha}f(t) = D^{\alpha-n} D^n f(t) = \frac{1}{\Gamma(n-\alpha)} \left[\int_0^t \frac{(D^n f)(s)}{(t-s)^{\alpha-n+1}}\, ds \right],$$

We note that fractional order operators are defined through convolution integrals and are therefore, unlike integer order derivatives, nonlocal operators. They depend on all function values from its lower limit $t = 0$ up to the evaluation point $s = t$.

Recalling the fact that the Riemann-Liouville definition of the fractional differential/integral operator is equivalent to the Grünwald's definition, we compare briefly the definitions of the fractional derivative by Riemann-Liouville and Caputo, see Oldham and Spanier [48] and Podlubny [50]. The Riemann-Liouville derivative of a constant function $f(t) = c,\ t \geq 0,\ f(t) = 0,\ t < 0$ is $D^\alpha c = ct^{-\alpha}/\Gamma(1-\alpha)$ while the Caputo fractional derivative of the same function is $D_*^\alpha c = 0$. This shows an unusual property of the Riemann-Liouville derivative, namely the derivative of a constant is a function of t. Fractional derivatives often appear in fractional differential equations. By taking the Laplace transform of a fractional derivative it is possible to identify the initial conditions that should be specified. The Laplace transform of the Riemann-Liouville derivative and the Caputo derivative are, respectively,

$$\mathcal{L}[D^\alpha f(t)](s) = s^\alpha \mathcal{L}[f(t)](s) - \sum_{k=0}^{n-1} s^k (D^{\alpha-k-1} f)(0^+),$$
$$\mathcal{L}[D_*^\alpha f(t)](s) = s^\alpha \mathcal{L}[f(t)](s) - \sum_{k=0}^{n-1} s^{\alpha-k-1} (D^k f)(0^+).$$

Here the last terms reveal the initial conditions. For the Riemann-Liouville definition we obtain initial conditions on integrals of fractional order, while for the Caputo derivative we obtain initial conditions on the function itself and its derivatives. The latter conditions are of course easier to interpret physically. To choose the appropriate Laplace transform formula, it is important to understand which type of initial conditions must be used.

For the constitutive models of viscoelasticity to be consistent with the second law of thermodynamics the fractional exponent must be between zero and one, see Adolfsson et al. [6]. We therefore restrict the exponent to $\alpha \in (0, 1]$. The Riemann-Liouville derivative then takes the form

$$D^\alpha f(t) = \frac{1}{\Gamma(1-\alpha)} \frac{d}{dt} \left[\int_0^t \frac{f(s)}{(t-s)^\alpha} \, ds \right]. \tag{2.4}$$

Then a simple rearrangement of the operators, yields the following relation between the Riemann-Liouville and the Caputo definitions

$$D^\alpha f(t) = D_*^\alpha f(t) + \frac{t^{-\alpha} f(0)}{\Gamma(1-\alpha)}, \quad \alpha \in (0, 1].$$

2.3 Mittag-Leffler function

The exponential function, e^z, plays an important role in the theory of integer order differential equations. Its one-parameter generalization, the function defined by

$$E_\alpha(z) = \sum_{k=0}^{\infty} \frac{z^k}{\Gamma(1+\alpha k)},$$

was introduced by Mittag-Leffler (1903, 1904, 1905) and was investigated by several authors among whom Wiman (1905), Pollard (1948), Humbert (1953). For $\alpha > 0$, $E_\alpha(z)$ is the simplest entire function of order $1/\alpha$ Phragmén (1904). The two-parameter function of the Mittag-Leffler type

$$E_{\alpha,\beta}(z) = \sum_{k=0}^{\infty} \frac{z^k}{\Gamma(\beta+\alpha k)}, \quad \alpha > 0,\, \beta > 0,$$

that plays an important role in the fractional calculus, was in fact introduced by Agarwal (1953). It was studied by Humbert and Agarwal (1953), but they used the same notation and name as for the one-parameter Mittag-Leffler function, see Podlubny [50] and Bateman [25] for references. It is noted that $E_\alpha = E_{\alpha,1}$ and $E_{1,1}(z) = e^z$.

The Laplace transform of the Mittag-Leffler function of order $\alpha > 0$ is

$$\mathcal{L}\big(E_\alpha(at^\alpha)\big) = \frac{s^{\alpha-1}}{s^\alpha - a}, \quad s > |a|^{1/\alpha}.$$

Indeed, for $s > a$, using the series expansion of the exponential function we have

$$\frac{1}{s-a} = \int_0^\infty e^{-st} e^{at}\, dt = \sum_{k=0}^{\infty} \frac{a^k}{k!} \int_0^\infty e^{-st} t^k\, dt = \sum_{k=0}^{\infty} \frac{a^k}{k!} \frac{k!}{s^{k+1}} = \sum_{k=0}^{\infty} \frac{a^k}{s^{k+1}}.$$

Then similarly for the Mittag-Leffler function we obtain

$$\begin{aligned}\mathcal{L}\big(E_\alpha(at^\alpha)\big) &= \sum_{k=0}^{\infty} \frac{a^k}{\Gamma(\alpha k+1)} \int_0^\infty e^{-st} t^{\alpha k}\, dt = \sum_{k=0}^{\infty} \frac{a^k}{\Gamma(\alpha k+1)} \frac{\Gamma(\alpha k+1)}{s^{\alpha k+1}} \\ &= \sum_{k=0}^{\infty} \frac{a^k}{s^{\alpha k+1}} = s^{\alpha-1} \sum_{k=0}^{\infty} \frac{a^k}{(s^\alpha)^{k+1}} = \frac{s^{\alpha-1}}{s^\alpha - a}.\end{aligned}$$

We now bring two definitions which are important in the theory of integro-differential equations, and therefore in viscoelasticity, see Bateman [25], Renardy et al. [53], and Widder [68].

Definition 2.1. A function $b \in \mathcal{C}^\infty(0,\infty)$ is called completely monotone if

$$(-1)^k \frac{d^k}{dt^k} b(t) \geq 0, \quad k = 0,\, 1,\, \ldots\, .$$

Pollard (1948) proved that the Mittag-Leffler function $E_\alpha(-z)$ is completely monotone for $z \geq 0$ if $\alpha \in [0,1]$, that is,

$$(-1)^k \frac{d^k}{dt^k} E_\alpha(-z) \geq 0, \quad k = 0,\, 1,\, \ldots.$$

Definition 2.2. A function $b \in L_{1,loc}[0,\infty)$ is called of positive type if for all $T > 0$ and $\varphi \in \mathcal{C}([0,T])$,

$$\int_0^T \int_0^t b(t-s)\varphi(t)\varphi(s)\, ds\, dt \geq 0. \tag{2.5}$$

The integral above may vanish without having φ identically zero. For example, for an arbitrary continuous function g, let $b(t,s) = g(t)g(s)$. Then we have only to choose $\varphi(t)$ orthogonal to $g(t)$ on $(0,T)$. A positive definite function then is a positive type function such that the integral (2.5) can vanish if and only if $\varphi \equiv 0$. We recall some important properties of the positive type functions.

The definition of a positive type function is not easy to check. Therefore, using the transform techniques one can show that $b \in L_1(0,\infty)$ is of positive type if and only if

$$\operatorname{Re} \hat{b}(i\omega) = \int_0^\infty b(t)\cos(\omega t)\, dt \geq 0, \quad \forall \omega \in \mathbb{R}, \tag{2.6}$$

where $\hat{b}$ denotes the Laplace transform of b.

Remark 2.1. From the viewpoint of applications to viscoelasticity, it is useful to know that a sufficient condition for (2.6) to hold is a certain type of sign conditions. That is, if $b \in L_1(0,\infty) \cap \mathcal{C}^2(0,\infty)$ and

$$(-1)^k \frac{d^k}{dt^k} b(t) \geq 0, \quad \forall t > 0,\ k = 0,\, 1,\, 2, \tag{2.7}$$

then b is a positive type function, that is b satisfies (2.6). Consequently, any completely monotone function $b \in L_{1,loc}(0,\infty)$ is of positive type.

We note that the function $b(t) = e^{-t}\cos(t)$, satisfies (2.6) but not (2.7).

Remark 2.2. Assume b is a positive type function (kernel), and A is a selfadjoint, positive definite operator on a Hilbert space of functions. Let $a(u,v) = (Au,v)$ be a corresponding bilinear form, for sufficiently smooth functions u, v, and $\{(\lambda_j,\varphi_j)\}_{j=1}^\infty$ be the corresponding eigenpairs. Then, for any $T > 0$ and $u \in \mathcal{C}\big([0,T];\mathcal{D}(A^{1/2})\big)$, we have

$$\int_0^T \int_0^t b(t-s)a(u(s),u(t))\, ds\, dt = \sum_{j=1}^\infty \lambda_j \int_0^T \int_0^t b(t-s)u_j(s)u_j(t)\, ds\, dt,$$

where $u_j = (u, \varphi_j)$. Since each of these integrals is positive by the positive definiteness of the kernel b, we conclude,

$$\int_0^T \int_0^t b(t-s)a(u(s), u(t))\, ds\, dt \geq 0.$$

3 Fractional order linear viscoelasticity

Linear viscoelasticity in combination with fractional order operators, i.e., the fractional order viscoelastic model, have attracted considerable attention in the last decades. The fractional order viscoelastic model is capable of describing the behaviour of many viscoelastic materials.

A perfectly elastic material does not exist since in reality: inelasticity is always present. This inelasticity leads to energy dissipation or damping. Therefore, for a wide class of materials it is not sufficient to use an elastic constitutive model to capture the mechanical behaviour. In order to replace extensive experimental tests by numerical simulations there is a need for an accurate material model. Therefore viscoelastic constitutive models have frequently been used to simulate the time dependent behaviour of polymeric materials. The classical linear viscoelastic models that use integer order time derivatives in the constitutive laws, require an excessive number of parameters to accurately predict observed material behaviour.

Bagley and Torvik [9] used fractional derivatives to construct stress-strain relationships for viscoelastic materials. The advantage of this approach is that very few empirical parameters are required (two elastic constants, one relaxation constant and the fractional order exponent).

When this fractional derivative model of viscoelasticity is incorporated directly into the structural equations a time differential equation of non-integer order higher than two is obtained. One consequence of this is that initial conditions of fractional order higher than one are required. The problems with initial conditions of fractional order have been discussed by Enelund and Olsson [24] and also by Bagley [8] and by Beyer and Kempfle [12]. To avoid the difficulties with fractional order initial conditions some alternative formulations of the fractional derivative viscoelastic model are used in structural modeling. The first form, that we will use, is based on a convolution integral formulation with a singular kernel of Mittag-Leffler type (see [24], [21] and [4]). The second form involves fractional integral operators rather than fractional derivative operators (see [20]). And the third form uses internal variables, see [22], [23] and [1].

We recall that a fractional order differential operator is not a local operator, i.e., the derivative depends on the whole history of the function.

This increases the complexity of mathematical analysis and the numerical computations of fractional order viscoelastic models.

For extensive overviews, analysis of the fractional order viscoelastic models, the hereditary theory of linear viscoelasticity and the history of linear viscoelasticity the reader is referred to [1], [6], [19], and [56].

3.1 Convolution integral formulation

Let $\sigma_{ij}, \epsilon_{ij}$ and u_i denote, respectively, the usual stress tensor, strain tensor and displacement vector. We recall that the linear strain tensor is defined by,

$$\epsilon_{ij} = \frac{1}{2}\Big(\frac{\partial u_i}{\partial x_j} + \frac{\partial u_j}{\partial x_i}\Big).$$

With the decompositions

$$s_{ij} = \sigma_{ij} - \tfrac{1}{3}\sigma_{kk}\delta_{ij}, \quad e_{ij} = \epsilon_{ij} - \tfrac{1}{3}\epsilon_{kk}\delta_{ij},$$

the constitutive equations are formulated as, see Bagley and Torvik [9],

$$\begin{aligned} s_{ij}(t) + \tau_1^{\alpha_1} D_t^{\alpha_1} s_{ij}(t) &= 2G_\infty e_{ij}(t) + 2G\tau_1^{\alpha_1} D_t^{\alpha_1} e_{ij}(t), \\ \sigma_{kk}(t) + \tau_2^{\alpha_2} D_t^{\alpha_2} \sigma_{kk}(t) &= 3K_\infty \epsilon_{kk}(t) + 3K\tau_2^{\alpha_2} D_t^{\alpha_2} \epsilon_{kk}(t), \end{aligned} \tag{3.1}$$

with initial conditions

$$s_{ij}(0+) = 2Ge_{ij}(0+), \quad \sigma_{kk}(0+) = 3K\epsilon_{kk}(0+),$$

meaning that the initial response follows Hooke's elastic law. Here G, K are the instantaneous (unrelaxed) moduli, and G_∞, K_∞ are the long-time (relaxed) moduli. Note that we have two relaxation times, $\tau_1,\ \tau_2 > 0$, and fractional orders of differentiation, $\alpha_1,\ \alpha_2 \in (0,1)$, where the fractional order derivative is defined by (2.4). The constitutive equations (3.1) can be solved for σ by means of Laplace transformation, see Enelund and Olsson [24]:

$$\begin{aligned} s_{ij}(t) &= 2G\Big(e_{ij}(t) - \frac{G - G_\infty}{G}\int_0^t f_1(t-s)e_{ij}(s)\,ds\Big), \\ \sigma_{kk}(t) &= 3K\Big(\epsilon_{kk}(t) - \frac{K - K_\infty}{K}\int_0^t f_2(t-s)\epsilon_{kk}(s)\,ds\Big), \end{aligned}$$

where

$$f_i(t) = -\frac{d}{dt}E_{\alpha_i}\Big(-\Big(\frac{t}{\tau_i}\Big)^{\alpha_i}\Big), \tag{3.2}$$

and E_{α_i} is the Mittag-Leffler function of order α_i, defined in §2.3. We make the simplifying assumption (synchronous viscoelasticity):

$$\alpha = \alpha_1 = \alpha_2, \quad \tau = \tau_1 = \tau_2, \quad f = f_1 = f_2.$$

Then we may define a parameter γ, a kernel β, and the Lamé constants μ, λ,

$$\gamma = \frac{G - G_\infty}{G} = \frac{K - K_\infty}{K}, \quad \beta(t) = \gamma f(t), \quad \mu = G, \quad \lambda = K - \tfrac{2}{3}G,$$

and the constitutive equations become

$$\begin{aligned} \sigma_{ij}(t) &= \Big(2\mu\epsilon_{ij}(t) + \lambda\epsilon_{kk}(t)\delta_{ij}\Big) - \int_0^t \beta(t-s)\Big(2\mu\epsilon_{ij}(s) + \lambda\epsilon_{kk}(s)\delta_{ij}\Big)\,ds \\ &= (\sigma_0)_{ij}(t) - \int_0^t \beta(t-s)(\sigma_0)_{ij}(s)\,ds. \end{aligned}$$

Note that the viscoelastic part of the model contains only three parameters:

$$0 < \gamma < 1, \quad 0 < \alpha < 1, \quad \tau > 0.$$

The kernel is weakly singular:

$$\begin{aligned} \beta(t) &= -\gamma \frac{d}{dt} E_\alpha\Big(-\Big(\frac{t}{\tau}\Big)^\alpha\Big) = \gamma \frac{\alpha}{\tau}\Big(\frac{t}{\tau}\Big)^{-1+\alpha} E'_\alpha\Big(-\Big(\frac{t}{\tau}\Big)^\alpha\Big) \\ &\approx C t^{-1+\alpha}, \quad t \to 0, \end{aligned}$$

and we note the properties

$$\begin{aligned} \beta(t) &\geq 0, \\ \|\beta\|_{L_1(\mathbb{R}^+)} &= \int_0^\infty \beta(t)\,dt = \gamma\big(E_\alpha(0) - E_\alpha(\infty)\big) = \gamma < 1. \end{aligned}$$

Various properties (e.g., regularity and convergence) of the memory kernel function β have been investigated in Enelund and Olsson [24].

The equations of motion now become:

$$\begin{aligned} &\rho u_{i,tt} - \sigma_{ij,j} = f_i, && \text{in } \Omega, \\ &u_i = 0, && \text{on } \Gamma_\mathrm{D}, \\ &\sigma_{ij} n_j = g_i, && \text{on } \Gamma_\mathrm{N}. \end{aligned}$$

Considering also initial values for displacement u and velocity u_t, this can be written as

$$
\begin{aligned}
&\rho u_{tt}(x,t) - \nabla\cdot\sigma_0(u;x,t) && \\
&\qquad + \int_0^t \beta(t-s)\nabla\cdot\sigma_0(u;x,s)\,ds = f(x,t) && \text{in } \Omega\times I,\\
&u(x,t) = 0 && \text{on } \Gamma_D\times I,\\
&\sigma(u;x,t)\cdot n(x) = g(x,t) && \text{on } \Gamma_N\times I,\\
&u(x,0) = u^0(x) && \text{in } \Omega,\\
&u_t(x,0) = v^0(x) && \text{in } \Omega,
\end{aligned}
\tag{3.3}
$$

which is a Volterra type integro-differential equation.

We should mention that there are also models with exponential kernels, smooth kernels, which describe polymeric materials, e.g., natural and synthetic rubber. The drawback with this kind of model is that it requires a large number of exponential kernels to describe the behaviour of the materials. This is the reason for introducing kernels of Mittag-Leffler type or fractional operators. In [24] and [6] Enelund and Adolfsson have shown that the classical viscoelastic model based on exponential kernels can describe the same viscoelastic behaviour as the fractional model if the number of kernels tend to infinity.

To motivate the derivation of the dual (adjoint) problem, we denote $Au = -\nabla\cdot\sigma_0(u)$ and $u_1 = u$, $u_2 = u_t = u_{1,t}$. Then the strong form (3.3) can be written as

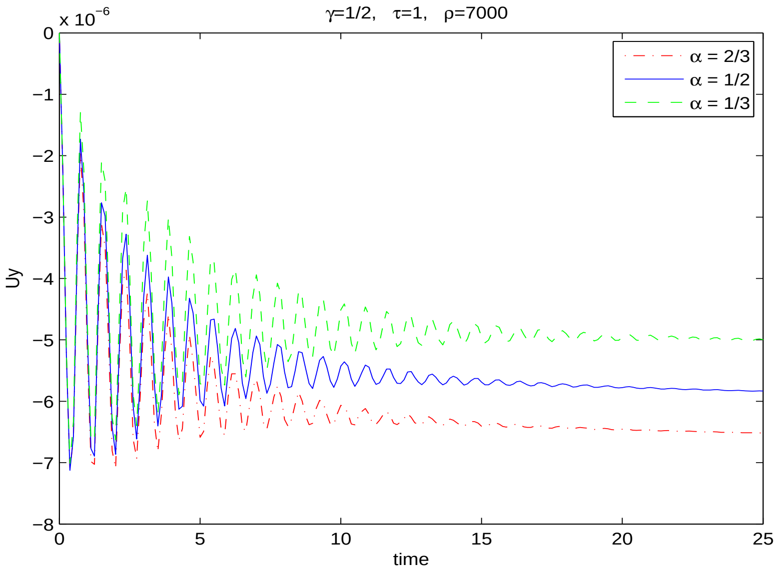
$$
\begin{pmatrix} A & 0 \\ 0 & \rho I \end{pmatrix}\begin{pmatrix} u_{1,t} \\ u_{2,t} \end{pmatrix} + \begin{pmatrix} 0 & -A \\ \left(I - \int_0^t \beta(t-s)ds\right)A & 0 \end{pmatrix}\begin{pmatrix} u_1 \\ u_2 \end{pmatrix} = \begin{pmatrix} 0 \\ f \end{pmatrix},
$$

that lead us to the dual form

$$
-\begin{pmatrix} A & 0 \\ 0 & \rho I \end{pmatrix}\begin{pmatrix} \phi_{1,t} \\ \phi_{2,t} \end{pmatrix} + \begin{pmatrix} 0 & \left(I - \int_t^T \beta(s-t)ds\right)A \\ -A & 0 \end{pmatrix}\begin{pmatrix} \phi_1 \\ \phi_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}.
$$

For a precise discussion see, the appended Paper III.

The two limits of the fractional exponent α, that is, $\alpha = 0$, $\alpha = 1$ describe two different mechanical behaviour. For the first case $\alpha = 0$, by the constitutive equation (3.1), there is no convolution term in the model, that is, there is no dissipation of energy and therefore no damping. While for the other limit $\alpha = 1$, recalling (3.2) and the fact that $E_1(t) = e^t$, we expect strong damping. These are illustrated in the figure below, see also the appended Paper I.



4 Well-Posedness of the model problem

The main tools, that we have used to prove existence, uniqueness, and regularity of the solution of the model problem (3.3), are the theory of semigroup of linear operators, and the Galerkin method. In the next section we bring some materials (without proofs) for the semigroup approach to provide the main idea that have been used in Paper I. For details on the Galerkin method, that is the main tool for Paper II, we refer to the existing references, e.g., Evans [27], or Dautray and Lions [17].

4.1 Semigroups of linear operators

Semigroup theory is the abstract study of first-order ordinary differential equations with values in Banach spaces, driven by linear, but possibly unbounded, operators. This approach provides an elegant alternative to some of the well-posedness theory for evolution equations that is one of the many applications that the theory has in different branches of analysis (such as harmonic analysis, approximation theory and many other subjects). In this section we outline the basics of the theory, without any proof, and present as well the Lumer-Phillips theorem, which will be used in §2 of the appended Paper I.

Troughout this section we let X denote a real Banach space.

For more complete and advanced details of the theory and its application to partial differential equations one may refer to Pazy [49] and Evans [27].

4.1.1 Definitions and properties

Definition 4.1. A one parameter family $T(t)$, $0 \leq t < \infty$, of bounded linear operators from the Banach space X to X is a *semigroup of bounded linear operators* on X if

$$(i)\, T(0) = I, \quad (I \text{ is the identity operator on } X),$$
$$(ii)\, T(t+s) = T(t)T(s), \quad \text{for every } t,\, s \geq 0 \quad \text{(the semigroup property)}.$$

Definition 4.2. The linear operator $\mathcal{A}$ defined by

$$\mathcal{A}x = \lim_{t \searrow 0} \frac{T(t)x - x}{t} = \frac{d^+T(t)x}{dt}\,|_{t=0} \quad \text{for} \quad x \in \mathcal{D}(\mathcal{A}),$$

is the *(infinitesimal) generator* of the semigroup $T(t)$, where $\mathcal{D}(\mathcal{A})$ is the domain of $\mathcal{A}$ defined by

$$\mathcal{D}(\mathcal{A}) = \Big\{x \in X : \lim_{t \searrow 0} \frac{T(t)x - x}{t} \text{ exists}\Big\}.$$

Definition 4.3. A semigroup $T(t)$, $0 \leq t < \infty$, of bounded linear operators on X is a *strongly continuous* semigroup if

$$\lim_{t \searrow 0} T(t)x = x \quad \forall x \in X.$$

A strongly continuous semigroup of bounded linear operators on X will be called a C_0 semigroup. If moreover $\|T(t)\| \leq 1$ for $t \geq 0$ it is called a C_0 *semigroup of contractions.*

Lemma 4.1. Let the linear operator $\mathcal{A}$ be the generator of a C_0 semigroup $T(t)$. Then for $x \in \mathcal{D}(\mathcal{A})$, $T(t)x \in \mathcal{D}(\mathcal{A})$ and

$$\frac{d}{dt}T(t)x = \mathcal{A}T(t)x.$$

Definition 4.4. For every $x \in X$ we define the *duality set* $F(x) \subset X^*$ by

$$F(x) = \big\{x^* : x^* \in X^* \text{ and } \langle x^*, x\rangle = \|x\|^2 = \|x^*\|^2\big\},$$

where X^* denotes the dual of X. And we note that by Hahn-Banach theorem $F(x) \neq \varnothing$ for every $x \in X$.

Definition 4.5. A linear operator $\mathcal{A}$ is *dissipative* if for every $x \in \mathcal{D}(\mathcal{A})$ there is $x^* \in F(x)$ such that $\langle \mathcal{A}x, x^*\rangle \leq 0$.

Theorem 4.1. Let $\mathcal{A}$ be dissipative with $R(I - \mathcal{A}) = X$. If X is reflexive then $\mathcal{D}(\mathcal{A})$ is dense in X, i.e., $\overline{\mathcal{D}(\mathcal{A})} = X$.

We use the first part of the following theorem in Paper I.

Theorem 4.2. (*Lumer-Phillips*). Let $\mathcal{A}$ be a linear operator with dense domain $\mathcal{D}(\mathcal{A})$ in X.

(a) If $\mathcal{A}$ is dissipative and there is a $\lambda > 0$ such that $R(\lambda I - \mathcal{A}) = X$, then $\mathcal{A}$ is the infinitesimal generator of a C_0 semigroup of contractions on X.

(b) If $\mathcal{A}$ is the infinitesimal generator of a C_0 semigroup of contractions on X, then $R(\lambda I - \mathcal{A}) = X$ for all $\lambda > 0$ and $\mathcal{A}$ is dissipative. Moreover, for every $x \in \mathcal{D}(\mathcal{A})$ and every $x^* \in F(x)$, $\langle \mathcal{A}x, x^* \rangle \leq 0$.

4.1.2 The abstract Cauchy problem

Let $\mathcal{A}$ be a linear operator from $\mathcal{D}(\mathcal{A}) \subset X$ into X. Given $x \in X$ the *abstract Cauchy problem* for $\mathcal{A}$ with initial data x consists of finding a solution $u(t)$ to the initial value problem

$$\begin{aligned} &\frac{d}{dt}u(t) = \mathcal{A}u(t) + f(t), \quad t > 0, \\ &u(0) = x, \end{aligned} \tag{4.1}$$

where $f : [0, T) \to X$. And by a solution we mean an X-valued function $u(t)$ such that $u(t)$ is continuous for $t \geq 0$, continuously differentiable and $u(t) \in \mathcal{D}(\mathcal{A})$ for $t > 0$ and (4.1) is satisfied. Note that since $u(t) \in \mathcal{D}(\mathcal{A})$ for $t > 0$ and u is continuous at $t = 0$, (4.1) cannot have a solution for $x \notin \overline{\mathcal{D}(\mathcal{A})}$.

From Lemma 4.1 it is clear that if $\mathcal{A}$ is the (infinitesimal) generator of a C_0 semigroup $T(t)$, the abstract Cauchy problem (4.1) when $f = 0$ has a solution, namely, $u(t) = T(t)x$, for every $x \in \mathcal{D}(\mathcal{A})$. So $T(t)$ is called the *operator solution*. It can be proved that the solution is unique. We recall a useful corollary that provides sufficient conditions to have a solution of the initial value problem (4.1).

Corollary 4.1. Let $\mathcal{A}$ be the infinitesimal generator of a C_0 semigroup $T(t)$. Let $f \in L_1(0, T; X)$ be continuous on $(0, T)$. If $f(s) \in \mathcal{D}(\mathcal{A})$ for $0 < s < T$ and $\mathcal{A}f(s) \in L_1(0, T; X)$ then for every $x \in \mathcal{D}(\mathcal{A})$ the initial value problem (4.1) has a solution on $[0, T)$.

Definition 4.6. A function u which is differentiable almost everywhere on $[0, T]$ such that $u' \in L_1(0, T; X)$ is called a *strong solution* of (4.1) if $u(0) = x$ and $u'(t) = \mathcal{A}u(t) + f(t)$ a.e. on $[0, T]$.

In the following lemmas we find the sufficient assumptions under which we get a unique strong solution of (4.1).

Corollary 4.2. If $\mathcal{A}$ generates a C_0 semigroup $T(t)$, f is differentiable a.e. on $[0,T]$ and $f' \in L_1(0,T;X)$ then for every $x \in \mathcal{D}(\mathcal{A})$ the initial value problem (4.1) has a unique strong solution.

In general, the Lipschitz continuity of f on $[0,T]$ is not sufficient to assure the existence of a strong solution of (4.1) for $x \in \mathcal{D}(\mathcal{A})$. However, if X is reflexive, for instance a Hilbert space, and f is Lipschitz continuous on $[0,T]$ that is

$$\|f(t_1) - f(t_2)\| \leq C|t_1 - t_2| \quad \text{for} \quad t_1,\, t_2 \in [0,T],$$

then by a classical result f is differentiable a.e. and $f' \in L_1(0,T;X)$. Therefore Lemma 4.3 implies the following.

Corollary 4.3. Let X be a reflexive Banach space and $\mathcal{A}$ generates a C_0 semigroup $T(t)$ on X. If f is Lipschitz continuous on $[0,T]$ then for every $x \in \mathcal{D}(\mathcal{A})$ the initial value problem (4.1) has a unique strong solution u on $[0,T]$ given by the variation of constants formula

$$u(t) = T(t)x + \int_0^t T(t-s)f(s)\,ds.$$

4.1.3 Application to partial differential equations

One important application of the theory of linear semigroups is analysing partial differential equations, e.g., well-posedness and numerical solution. In order to reformulate a PDE into a first-order ordinary differential equation, an abstract Cauchy problem, we need to construct suitable spaces and a suitable linear operator $\mathcal{A}$. It should be noticed in the previous sections that an important property for a linear operator $\mathcal{A}$ is to generate a C_0 semigroup (of contractions) of $T(t)$. This is what we have done in Paper I, inspired by Fabiano and Ito [28], to prove well-posedness and regularity properties.

To make a ready reference for §2 of the appended Paper I, we correspond the important corollaries and theorems in this draft with the main ones in Pazy [49] as follows:

here	Theorem 4.1	Theorem 4.2	Corollary 4.1	Corollary 4.2
$\Updownarrow$	$\Updownarrow$	$\Updownarrow$	$\Updownarrow$	$\Updownarrow$
[49]	Theorem 1.4.6	Theorem 1.4.3	Corollary 4.2.6	Corollary 4.2.10

5 Deterministic wave equation: numerical method

The model problem in this manuscript, i.e., (3.3), is a hyperbolic type integro-differential equation, and numerical analysis of such a problem is inherent from the numerical analysis of the hyperbolic problems. Therefore we consider the wave equation on a bounded domain, as a typical hyperbolic problem, to explain the main ideas of the numerical methods and the corresponding error analysis, that we applied to the main model (3.3).

In §5.2 part **(a)** we explain the main idea for the derivation of optimal order a priori error estimates that require minimal regularity assumptions. This will be used to prove strong convergence of the semidiscrete finite element approximation of the linear stochastic wave equation, see §6 and Paper IV.

5.1 Wave equation and variational formulations

Let us consider the wave equation

$$\begin{aligned} &\ddot{u} - \Delta u = f && \text{in } \Omega \times (0,T), \\ &u = 0 && \text{on } \partial\Omega \times (0,T), \\ &u(0) = u^0, \ \dot{u}(0) = v^0 && \text{in } \Omega, \end{aligned} \tag{5.1}$$

where $\Omega \subset \mathbb{R}^d$, $d = 1, 2, 3$, is a polygonal, convex bounded domain. We denote time derivatives by '$\cdot$' in the sequel.

We recall the Lebesgue space $L_2 = L_2(\Omega)$, with the usual L_2-inner product $(\cdot,\cdot)$ and L_2-norm $\|\cdot\|$, and the classical Sobolev spaces $H^s = H^s(\Omega)$, for s a positive integer. We also use $H_0^1 = H_0^1(\Omega)$ and its dual H^{-1}. We let $A = -\Delta$ with domain $\mathcal{D}(A) = H^2 \cap H_0^1$ which is a selfadjoint, positive definite unbounded operator. We define the Hilbert spaces $\dot{H}^l = \mathcal{D}(A^{l/2})$, for $l \in \mathbb{R}$, with norm

$$\|v\|_l = \|A^{l/2}v\| = \sqrt{(A^l v, v)} = \Big(\sum_{j=1}^{\infty} \lambda_j^l (v,\varphi_j)^2\Big)^{1/2}, \quad \forall v \in \dot{H}^l,$$

where $\{(\lambda_j,\varphi_j)\}_{j=1}^{\infty}$ are the eigenpairs of the operator A. Having the homogeneous Dirichlet boundary condition, we recall the elliptic regularity

$$\|u\|_{H^2} \le C\|Au\|, \quad u \in \mathcal{D}(A).$$

We recall the conservation law for the wave equation, that is for $t \in [0,T]$,

$$\|\dot{u}(t)\|^2 + \|u(t)\|_1^2 = \|v^0\|^2 + \|u^0\|_1^2.$$

An alternative representation of the wave equation (5.1) is by a "velocity-displacement" formulation which is obtained by introducing a new velocity variable. Using the notation $u_1 = u$, $u_2 = \dot{u}$, at least two variational forms can be formulated.

The first variational formulation is: find $\boldsymbol{u} = (u_1, u_2) \in \mathcal{V}$, such that,

$$\begin{aligned} &\mathcal{B}(\boldsymbol{u}, \boldsymbol{v}) = \mathcal{L}(\boldsymbol{v}), \quad \forall \boldsymbol{v} = (v_1, v_2) \in \mathcal{W}, \\ &u_1(0) = u^0, \;\; u_2(0) = v^0, \end{aligned} \tag{5.2}$$

where the bilinear form $\mathcal{B}(\cdot,\cdot)$ and the linear form $\mathcal{L}(\cdot)$ are defined by

$$\begin{aligned} \mathcal{B}(\boldsymbol{u}, \boldsymbol{v}) &= \int_0^T \big\{a(\dot{u}_1, v_1) - a(u_2, v_1) + (\dot{u}_2, v_2) + a(u_1, v_2)\big\}\, dt, \\ \mathcal{L}(\boldsymbol{v}) &= \int_0^T (f, v_2)\, dt. \end{aligned}$$

The second variational formulation is: find $\boldsymbol{u} = (u_1, u_2) \in \hat{\mathcal{V}}$, such that,

$$\mathcal{A}(\boldsymbol{u}, \boldsymbol{v}) = \mathcal{F}(\boldsymbol{v}), \quad \forall \boldsymbol{v} = (v_1, v_2) \in \hat{\mathcal{W}}, \tag{5.3}$$

where the bilinear form $\mathcal{A}(\cdot,\cdot)$ and the linear form $\mathcal{F}(\cdot)$ are defined by

$$\begin{aligned} \mathcal{A}(\boldsymbol{u}, \boldsymbol{v}) &= \int_0^T \big\{(\dot{u}_1, v_1) - (u_2, v_1) + (\dot{u}_2, v_2) + a(u_1, v_2)\big\}\, dt, \\ &\quad + (u_1(0), v_1(0)) + (u_2(0), v_2(0)), \\ \mathcal{F}(\boldsymbol{v}) &= \int_0^T (f, v_2)\, dt + (u^0, v_1(0)) + (v^0, v_2(0)). \end{aligned}$$

Here the function spaces $\mathcal{V}$, $\mathcal{W}$, $\hat{\mathcal{V}}$, $\hat{\mathcal{W}}$ are appropriately chosen to adjust the variational formulations, see Eriksson et al. [26], and Bangerth and Rannacher [11]. We note the main differences in the variational formulations above. In (5.2) the new velocity variable is enforced in the H^1 sense and the initial conditions are considered separately. While in (5.3) the velocity variable is enforced in the L_2 sense as well as the initial data.

The first formulation were investigated in the works by Hulbert [36], Johnson [38], and Li and Wiberg [40]. The second formulation, but without the weak enforcement of the initial values, and discretizations thereof were investigated in Hulbert and Hughes [37], Bales and Lasiecka [10], and French and Peterson [31].

5.2 The continuous Galerkin method

The continuous Galerkin (cG) method is a finite element technique which provides time discretizations for evolution problems using approximation spaces of continuous piecewise polynomial functions. This approach is particularly appropriate for wave problems as it retains discrete version of the important energy conservation properties provided by the initial/boundary value problem being approximated, see French and Schaeffer [30]. Computations and analyses have shown this is specially useful in the approximation of solutions to nonlinear wave problems, see, e.g., Glassey [34], Glassey and Schaeffer [33] or Strauss and Vazquez [64]. Another advantage of the continuous time Galerkin approach is that cG methods of any desired order of accuracy are easily formulated. We also note that the cG method follows the so-called Rothe approach of first discretizing in time and then in space on each discrete time level. This has the advantage of having the freedom to choose the spatial mesh differently at each time level.

We review two variant of the cG method in the following, for references see Eriksson et al. [26] for the first approach and Bangerth and Rannacher [11] for the second one.

(a): Let $0 = t_0 < t_1 < \cdots < t_n < \cdots < t_N = T$ be any given partition of the time interval $[0, T]$. To each subinterval $I_n = (t_{n-1}, t_n)$ of length $k_n = t_n - t_{n-1}$, we associate a triangulation $\mathcal{T}_h^n$ of Ω and the corresponding finite element space V_h^n of continuous piecewise linear polynomials that vanish on the boundary $\partial\Omega$ (that is the mesh is adjusted to fit $\partial\Omega$). We also define the spaces, for $q = 0,\ 1$,

$$W^{(q)} = \big\{ w : w|_{\Omega \times I_n} = w^n \in W_n^{(q)},\ n = 1, \ldots, N \big\},$$

where,

$$W_n^{(q)} = \big\{ w : w(x,t) = \sum_{i=0}^{q} w_i(x) t^i,\ w_i \in V_h^n \big\}.$$

This means that $w \in W^{(q)}$ may be discontinuous at time levels t_n, and $w \in W^{(0)}$ is piecewise constant in time.

The cG method is: find $U = (U_1, U_2) \in \big(W^{(1)}\big)^2$ such that,

$$\begin{aligned} &\mathcal{B}(U, V) = \mathcal{L}(V), \quad \forall V = (V_1, V_2) \in \big(W^{(0)}\big)^2, \\ &U_{1,n-1}^{+} = \mathcal{R}_{h,n} U_{1,n-1}^{-}, \quad U_{2,n-1}^{+} = \mathcal{P}_{h,n} U_{2,n-1}^{-}, \\ &U_{1,0}^{-}(0) = u^0,\ U_{2,0}^{-}(0) = v^0, \end{aligned} \tag{5.4}$$

where $\mathcal{R}_h, \mathcal{P}_h$ are, respectively, the usual elliptic and orthogonal projections. This means that cG is a Petrov-Galerkin method having different trial and test function spaces. An important property that is used in the error analysis is the Galerkin orthogonality

$$\mathcal{B}(U - \boldsymbol{u}, V) = 0, \quad \forall V \in \big(W^{(0)}\big)^2, \tag{5.5}$$

where $\boldsymbol{u}$ is the solution of the weak formulation (5.2).

When $\mathcal{T}_h^{n-1} = \mathcal{T}_h^n$, $n = 2, \ldots, N$, that is, the spatial mesh does not change between time levels, we conclude the discrete conservation property

$$\|U_1(t_N)\|_1^2 + \|U_2(t_N)\|^2 = \|U_1(0)\|_1^2 + \|U_2(0)\|^2.$$

We note that when we coarsen the mesh from the time level t_{n-1} to the next time level t_n, that is $\mathcal{T}_h^{n-1} \supset \mathcal{T}_h^n$, we use the projections to handle the degrees of freedom, by transforming the information from a time level to the next one. Obviously, in this case U_i are discontinuous at $t = t_{n-1}$. While when $\mathcal{T}_h^{n-1} \subset \mathcal{T}_h^n$, that is the spatial mesh is allowed to be refined or unchanged, then $U_{i,n-1}^+ = U_{i,n-1}^-$ ($n = 1, \ldots, N$, $i = 1, 2$), and the initial conditions reduce to $U_1(0) = u_h^0 = \mathcal{R}_{h,1}u^0$, $U_2(0) = v_h^0 = \mathcal{P}_{h,1}v^0$. In fact the initial data $U_1(0) = \mathcal{R}_{h,1}u^0$, $U_2(0) = \mathcal{P}_{h,1}v^0$ are the natural choice for a finite element method based on the weak formulation (5.2).

In our analysis (stability of the discrete equation and a priori error analysis), we consider the case $\mathcal{T}_h^{n-1} \subset \mathcal{T}_h^n$, and a slightly more general problem. That is to find $U = (U_1, U_2) \in \big(W^{(1)}\big)^2$ such that,

$$\begin{aligned} &\mathcal{B}(U, V) = \hat{\mathcal{L}}(V), \quad \forall V = (V_1, V_2) \in \big(W^{(0)}\big)^2, \\ &U_1(0) = u_h^0, \; U_2(0) = v_h^0, \end{aligned} \tag{5.6}$$

where

$$\hat{\mathcal{L}}(V) = \int_0^T \Big(a(f_1, V_1) + (f, V_2) \Big) \, dt.$$

This means that we consider an extra load term f_1 for the equation regarding the velocity variable. Then in an standard way we obtain a stability estimate of the form

$$|||U(t_N)|||_{h,l} \leq C_s \Big(|||U^0|||_{h,l} + \int_0^T |||(f_1, f_2)|||_{h,l} \, dt \Big), \tag{5.7}$$

where $|||\cdot|||_{h,l}$ is a certain discrete norm and depends on $l \in \mathbb{R}$, see Paper I and Paper IV.

Then we split the error $e = U - \boldsymbol{u}$ in the form

$$e = (U - \pi_{hk}\boldsymbol{u}) + (\pi_{hk}\boldsymbol{u} - \boldsymbol{u}) = \theta + \omega, \tag{5.8}$$

where π_{hk} is a combination of interpolation and projections, to be chosen appropriately, depending on the quantity of interest. Since ω can be estimated by the classical results from the approximation theory, we need to estimate θ. To this end, using the Galerkin orthogonality (5.5), we show that θ satisfies the problem (5.6) with certain functions f_1, f_2, such that we can use the stability estimates (5.7) with suitable choice of l. Then from the relation (5.8) and estimates of ω, from the approximation theory, we obtain optimal a priori error estimates for the quantity of interest.

We introduced this new technique to obtain optimal order a priori error estimates with minimal regularity requirement in Paper IV, where we study discretization in the spatial variable, see Remark 4.6 there for a comparison with earlier works. We used the same idea for the full discretization of the fractional order viscoelasticity model (3.3) in Paper I. We note that for a priori error analysis based on duality argument we do not need the general problem (5.6).

(b): An important feature of the cG method, as a Rothe approach, is that we have the freedom to choose the computational mesh differently at each time level. In the cG method, just presented, the projections are used to handle changing of the spatial mesh, since the meshes are associated with the space-time slabs $\Omega \times I_n$. In the second approach the functions are kept continuous by means of 'hanging nodes'.

Let $0 = t_0 < t_1 < \cdots < t_{n-1} < t_n < \cdots < t_N = T$ be a partition of the time interval $[0, T]$. To each discrete time level t_n we associate a triangulation $\mathbb{T}_h^n$ of Ω and a corresponding finite element space $\mathbb{V}_h^n$ consisting of continuous piecewise linear polynomials. For each time subinterval $I_n = (t_{n-1}, t_n)$ of length $k_n = t_n - t_{n-1}$, we define intermediate triangulation $\bar{\mathbb{T}}_h^n$ which is composed of mutually finest meshes of the neighboring meshes $\mathbb{T}_h^n$, $\mathbb{T}_h^{n-1}$ defined at discrete time levels t_n, t_{n-1}, respectively. Correspondingly, we define the finite element spaces $\bar{\mathbb{V}}_h^n$ consisting of continuous piecewise linear polynomials. This construction is used in order to allow continuity in time of the trial functions when the meshes change with time. Hence we obtain a decomposition of each slab $\Omega^n = \Omega \times I_n$ into space-time cells $K^n = K \times I_n$, $K \in \bar{\mathbb{T}}_h^n$ (prisms, for example, in case of $\Omega \subset \mathbb{R}^2$). Then, the trial and the test function spaces for the discrete form are, respectively,

$$
\begin{aligned}
\mathcal{V}_{hk} = \Big\{ U = (U_1, U_2) : \ & U \text{ continuous in } \Omega \times [0, T], \\
& U(x,t)|_{I_n} \text{ linear in } t, \\
& U(\cdot, t_n) \in (\mathbb{V}_h^n)^2, U(\cdot, t)|_{I_n} \in (\bar{\mathbb{V}}_h^n)^2 \Big\}, \\
\mathcal{W}_{hk} = \Big\{ V = (V_1, V_2) : \ & V(\cdot, t) \text{ continuous in } \Omega, \\
& V(\cdot, t)|_{I_n} \in (\mathbb{V}_h^n)^2, \\
& V(x,t)|_{I_n} \text{ piecewise constant in } t \Big\}.
\end{aligned}
\tag{5.9}
$$

We note that global continuity of the trial functions in $\mathcal{V}_{hk}$ requires the use of hanging nodes if the spatial mesh changes across a time level, see Carey and Oden [14], and Svensson [65] for numerical aspects of hanging nodes. The cG method is then read: find $U \in \mathcal{V}_{hk}$ such that,

$$\mathcal{A}(U, W) = \mathcal{F}(W), \quad \forall W \in \mathcal{W}_{hk}. \tag{5.10}$$

To avoid making this manuscript too long, we ignore the details of the analysis for the cG method (5.10). For details on a priori error estimates one can consult, e.g., Johnson [38] and Eriksson et al. [26] for the main ideas, and for a posteriori error analysis see Bangerth and Rannacher [11].

We formulated the same cG method for the model problem (3.3) in Paper III. The error analysis are based on the duality argument. For a priori error analysis we follow the main idea of duality based analysis, from Johnson [38]. For a posteriori error analysis we use the general framework presented by Bangerth and Rannacher [11], where they used the Dual Weighted Residual method.

6 Stochastic wave equation

In Paper IV we study the finite element approximation of the linear stochastic wave equation driven by additive noise,

$$\begin{aligned} &\mathrm{d}\dot{u} - \Delta u \, \mathrm{d}t = \mathrm{d}W && \text{in } \mathcal{D} \times (0, \infty), \\ &u = 0 && \text{on } \partial\mathcal{D} \times (0, \infty), \\ &u(\cdot, 0) = u_0, \, \dot{u}(\cdot, 0) = v_0 && \text{in } \mathcal{D}, \end{aligned} \tag{6.1}$$

where $\mathcal{D} \in \mathbb{R}^d$, $d = 1, 2, 3$, is a bounded convex polygonal domain with boundary $\partial\mathcal{D}$, and $\{W(t)\}_{t\geq 0}$ is a $L_2(\mathcal{D})$-valued Wiener process adapted to a filtered probability space $(\Omega, \mathcal{F}, \mathbf{P}, \{\mathcal{F}_t\}_{t\geq 0})$ with respect to the normal filtration $\{\mathcal{F}_t\}_{t\geq 0}$, and we let u_0, v_0 be $\mathcal{F}_0$-measurable random variables. For examples of stochastic wave equation and their applications see, e.g., Walsh [66], Chow [15], Allen [7], Martin [44] and references therein.

For a rigorous meaning to the infinite dimensional Wiener process W, stochastic integral, and the definition of a weak solution of (6.1) we refer to §2 in the appended Paper IV, and references therein.

Here, for simplicity, we set the initial values $u_0 = v_0 = 0$, and we describe the main feature of the presented work in Paper IV. We use the semigroup framework of Da Prato and Zabczyk [51] in which the weak solution of (6.1) is represented as a stochastic convolution

$$u(t) = \int_0^t A^{-1/2} \sin(sA^{1/2}) \, \mathrm{d}W(s).$$

Here, where we have set $u_0 = v_0 = 0$, recalling the linear operator $A = -\Delta$ with $D(A) = H^2(\mathcal{D}) \cap H_0^1(\mathcal{D})$, $v(t) = A^{-1/2}\sin(tA^{1/2})f$ is the solution of

$$\begin{aligned} &\ddot{v} + Av = 0 && \text{in } \mathcal{D} \times (0, \infty), \\ &v(0) = 0,\, \dot{v}(0) = f && \text{in } \mathcal{D}. \end{aligned} \tag{6.2}$$

Let Q and $\|\cdot\|_{\mathrm{HS}}$ denote, respectively, the covariance operator of W and the Hilbert Schmidt norm. We show that if, for some $\beta \geq 0$,

$$\|A^{(\beta-1)/2}Q^{1/2}\|_{\mathrm{HS}} < \infty, \tag{6.3}$$

then we have the spatial regularity of order β, that is,

$$\mathbf{E}\big(\|u(t)\|^2_{\dot{H}^\beta}\big) \leq C(t)\|A^{(\beta-1)/2}Q^{1/2}\|_{\mathrm{HS}},$$

where $\dot{H}^\beta = D(A^{\beta/2})$ and $\mathbf{E}$ denotes the expected value of a random variable. In particular, if $\mathrm{Tr}(Q) < \infty$ (spatially correlated noise), then we may take $\beta = 1$. On the other hand if $Q = I$ (uncorrelated noise), then $\beta < 1/2$, $d = 1$.

We discretize (6.1) in the spatial variables with a standard piecewise linear finite element method, and we show strong convergence estimates in various norms. In particular, denoting the maximal mesh size by h and the approximate solution by u_h, we have,

$$\Big(\mathbf{E}\big(\|u_h(t) - u(t)\|^2\big)\Big)^{1/2} \leq C(t)h^{\frac{2}{3}\beta}\|A^{(\beta-1)/2}Q^{1/2}\|_{\mathrm{HS}}, \tag{6.4}$$

if (6.3) holds for some $\beta \in [0, 3]$.

As a comparison, we recall from Yan [69] that for the stochastic heat equation, if (6.3) holds for some $\beta \in [0, 2]$, we have the spatial regularity,

$$\Big(\mathbf{E}\big(\|u(t)\|^2_{\dot{H}^\beta}\big)\Big)^{1/2} \leq C\|A^{(\beta-1)/2}Q^{1/2}\|_{\mathrm{HS}},$$

and the error estimate

$$\Big(\mathbf{E}\big(\|u_h(t) - u(t)\|^2\big)\Big)^{1/2} \leq Ch^\beta\|A^{(\beta-1)/2}Q^{1/2}\|_{\mathrm{HS}}.$$

Here the order of regularity coincide with the order of convergence. We note that surprisingly the stochastic heat and wave equation have both the same order of regularity.

The main tools for the proof of (6.4) are the Itô-isometry and error estimates for the deterministic problem (6.2) with minimal regularity assumptions. One way to obtain optimal order (a priori) error estimate with minimum regularity assumptions has already been explained in §5.2 in this introduction. Having the error estimates

$$\|v_h(t) - v(t)\| \le C(t)h^2\|f\|_{\dot{H}^2}, \tag{6.5}$$

$$\|v_h(t) - v(t)\| \le C\|f\|_{\dot{H}^{-1}},$$

we have, by interpolation,

$$\|v_h(t) - v(t)\| \le C(t)h^{\frac{2}{3}\beta}\|v\|_{\dot{H}^2}, \quad \beta \in [0,3].$$

When we specialize to $Q = I$, $d = 1$, we must have $\beta < 1/2$ and thus the order of strong convergence is $O(h^\alpha)$, $\alpha < 1/3$. This is in agreement with Quer-Sardanyons and Sanz-Solé [52], where spatial discretization of (6.1) is studied for $d = 1$ and with space-time white noise ($Q = I$). They used a standard difference scheme with uniform mesh size h. We note that the order of convergence is less than the order of regularity which is $\beta < 1/2$. However, it is known that in (6.5) $\|f\|_{\dot{H}^2}$ can not be replaced by $\|f\|_{\dot{H}^{2-\epsilon}}$ for any small positive number ϵ, see Paper IV, Remark 4.4. Therefore $O(h^\alpha)$, $\alpha < 1/3$ is the best that one can expect. The framework, that we present, applies to multiple dimensions and spatial correlated noise.

7 Earlier works

7.1 Integro-differential equations

A lot of work have been done during the last decades regarding well-posedness of the fractional order linear viscoelasticity and also several methods have been investigated to solve these kinds of models numerically. We try to give just some references to earlier works, but it does not seem to be possible to give a complete list.

Thomée and McLean [46] have proved the existence, uniqueness and regularity of the solution of a reformed model of (3.3) by means of Fourier series. One can also see [18] where Desch and Fašanga have used the context of analytic semigroups in terms of interpolation spaces. An abstract Volterra equation, as an abstract model for equations of linear viscoelasticity, has been studied in Dafermos [16].

For more details on numerical methods such as semidiscretization in time or space and the relevant methods that have been used, namely discontinuous Galerkin approximation as well as first and second-order backward difference methods in time or continuous Galerkin approximation in space, we refer to, e.g., [2], [3], [5], [41], [46], [47], [54], [61], and [63].

Numerical methods for quasistatic viscoelasticity problems, i.e., $\rho\ddot{u} \approx 0$, have been studied, e.g., in [3] and [62] where basically they have used discontinuous Galerkin approximation in time and continuous Galerkin approximation in space. A dynamic model for viscoelasticity based on internal variables has been studied in [54]. A posteriori analysis of temporal finite element approximation of a parabolic type problem and discontinuous Galerkin finite element approximation of a quasi-static linear viscoelasticity problem has been studied, respectively, in [4] and [62]. For more references one may see [54].

The drawback of the fractional order viscoelastic models is that the whole strain history must be saved and included in each time step that is due to the non-locality of the fractional order differential operators. The most commonly used algorithms for this integration are based on Lubich convolution quadrature [42] for fractional order operators. One example of the application of this approach to integro-differential equations with a memory term is in [43]. The Lubich convolution quadrature requires uniformly distributed time steps or alternatively logarithmically distributed time steps as outlined in [29]. These are cumbersome restrictions because it is not possible to use adaptivity and goal oriented error estimation. Some efficient numerical algorithms to overcome the mentioned problem of Lubich convolution quadrature can be found in [59] and [60]. Also sparse quadrature as a possible way to overcome the problem with the growing amount of data, that has to be stored and used in each time step, has been studied in [2], [3], and [47]. In this approach variable time steps can be used.

7.2 Evolution stochastic PDEs

An extensive work on stochastic parabolic PDEs, their applications and their numerical approximation can be found in the literature, see for example, [15], [32], [35], [39], [51], [66], [70], and and references therein.

For analysis of the stochastic wave equation and the properties we refer to e.g., [15], [51], [52], [67] for references. However the numerical analysis of the stochastic wave equation is less studied, see [45], [52], [57], [67], for existing results. In particular these works do not deal with multiple dimensions or correlated noise.

8 Summary of the appended papers

8.1 Paper I

The continuous Galerkin method for an integro-differential equation modeling dynamic fractional order viscoelasticity, IMA Journal of Numerical

Analysis, to appear (with Stig Larsson).

In this paper, first we consider the model problem (3.3) with homogeneous boundary conditions, and we prove existence and uniqueness in the context of linear semigroup theory. We also study regularity properties as well, and we observe that the regularity of high order is limited due to the mixed boundary conditions and singularity of the convolution kernel. We then formulate a continuous Galerkin method of degree one for the model problem, and we investigate the stability properties. Using stability estimates, we then prove optimal order a priori error estimates for the displacement and the velocity at the temporal nodal points. This shows that the error estimates for the semidiscretization in spatial variables are optimal in $L_\infty(L_2)$ and $L_\infty(H^1)$ for the displacement and the velocity, respectively. At the end we illustrate the theory by an example.

8.2 Paper II

Existence and uniqueness of the solution of an integro-differential equation with weakly singular kernel, preprint 2009:16.

The abstract framework presented in Paper I does not admit the mixed homogeneous Dirichlet and non-homogeneous Neumann boundary conditions. In this paper we consider the model problem (3.3) with mixed homogeneous Dirichlet and non-homogeneous Neumann boundary conditions. Then we prove existence and uniqueness of the solution, in the weak sense, based on the Galerkin method. We then prove regularity estimates and we discuss the limitations for higher regularity.

8.3 Paper III

A continuous space-time finite element method for an integro-differential equation modeling dynamic fractional order viscoelasticity, preprint 2009:17.

In this paper, which is based on a weak formulation of the model problem (3.3), we present a simple proof of stability for the primal and the dual problems. We then formulate a continuous Galerkin method of degree one, and we obtain optimal order a priori error estimate by a duality argument. We also present a posteriori error representation in terms of space-time cells, such that it can be used for adaptive strategies based on the dual weighted residual method. We then prove a weighted global a posteriori error estimate. Due to some restrictions for such weighted global estimate, we prove a global a posteriori estimate which has less restrictions.

8.4 Paper IV

Finite element approximation for the linear stochastic wave equation with additive noise, preprint 2009:18, (with Mihály Kovács and Stig Larsson).

In this paper we study the semidiscrete finite element approximation of the linear stochastic wave equation with additive noise in a semigroup framework. We then obtain optimal error estimates of the deterministic problem under minimal regularity assumptions. We use these to prove strong convergence estimates for the stochastic problem. The theory presented here applies to multi-dimensional domains and spatially correlated noise. We illustrate the theory by numerical examples.

References

[1] K. Adolfsson, *Models and numerical procedures for fractional order viscoelasticity*, PhD thesis, Chalmers University of Technology, Göteborg, Sweden, 2003.

[2] K. Adolfsson, M. Enelund, and S. Larsson, *Adaptive discretization of an integro-differential equation with a weakly singular convolution kernel*, Comput. Methods Appl. Mech. Engrg. **192** (2003), 5285–5304.

[3] ———, *Adaptive discretization of fractional order viscoelasticity using sparse time history*, Comput. Methods Appl. Mech. Engrg. **193** (2004), 4567–4590.

[4] K. Adolfsson, M. Enelund, and S. Larsson, *Space-time discretization of an integro-differential equation modeling quasi-static fractional-order viscoelasticity*, J. Vib. Control **14** (2008), 1631–1649.

[5] K. Adolfsson, M. Enelund, S. Larsson, and M. Racheva, *Discretization of integro-differential equations modeling dynamic fractional order viscoelasticity*, LNCS **3743** (2006), 76–83.

[6] K. Adolfsson, M. Enelund, and P. Olsson, *On the fractional order model of viscoelasticity*, Mechanics of Time-Dependent Materials **9** (2005), 15–34.

[7] E. Allen, *Modeling with Itô Stochastic Differential Equations*, Mathematical Modelling: Theory and Applications, vol. 22, Springer, 2007.

[8] R. L. Bagley, *On the fractional order initial value problem and its engineering applications*, Proceedings of the international conference of Fractional Calculus and Its Applications, Tokyo, Japan (1990), 12–22.

[9] R. L. Bagley and P. J. Torvik, *Fractional calculus–a different approach to the analysis of viscoelastically damped structures*, AIAA J. **21** (1983), 741–748.

[10] L. Bales and I. Lasiecka, *Continuous finite elements in space and time for the nonhomogeneous wave equation*, Comput. Math. Appl. **27** (1994), 91–102.

[11] W. Bangerth and R. Rannacher, *Adaptive Finite Element Methods for Differential Equations*, Lectures in Mathematics ETH Zürich, Birkhäuser Verlag, 2003.

[12] H. Beyer and S. Kempfle, *Definition of physically consistent damping laws with fractional derivatives*, Zeitschrift für Angewandte Mathematik und Mechanik **75** (1995), 623–635.

[13] M. Caputo, *Linear models of dissipation whose Q is almost frequency independent. II*, Fract. Calc. Appl. Anal. **11** (2008), 4–14.

[14] G. F. Carey and J. T. Oden, *Finite Elements. Vol. III*, The Texas Finite Element Series, III, Prentice Hall Inc., 1984, Computational aspects.

[15] P. Chow, *Stochastic Partial Differential Equations*, Chapman & Hall/CRC Applied Mathematics and Nonlinear Science Series, Chapman & Hall/CRC, Boca Raton, FL, 2007.

[16] C. M. Dafermos, *An abstract Volterra equation with applications to linear viscoelasticity*, J. Differential Equations **7** (1970), 554–569.

[17] R. Dautray and J. L. Lions, *Mathematical Analysis and Numerical Methods for Science and Technology. Vol. 5*, Springer-Verlag, 1992.

[18] W. Desch and E. Fašanga, *Stress obtained by interpolation methods for a boundary value problem in linear viscoelasticity*, J. Differential Equations **217** (2005), 282–304.

[19] M. Enelund, *Fractional calculus and linear viscoelasticity in structural dynamics*, PhD thesis, Chalmers University of Technology, Göteborg, Sweden, 1996.

[20] M. Enelund, A. Fenander, and P. Olsson, *A fractional integral formulation of constitutive equations of viscoelasticity*, AIAA J. **35** (1997), 1356–1362.

[21] M. Enelund and B. L. Josefson, *Time domain FE-analysis of viscoelastic structures having constitutive relations involving fractional derivatives*, Proceedings 37th Structures, Structural Dynamics and Materials Conference, Salt Lake City. UT, USA, AIAA Washington DC (1996), 685–694.

[22] M. Enelund and G. A. Lesieutre, *Time domain modeling of damping using anelastic displacement fields and fractional calculus*, International J. Solids Structures **36** (1999), 4447–4472.

[23] M. Enelund, L. Mähler, K. Runesson, and B. L. Josefson, *Unified formulation and integration of the standard linear viscoelastic solid with integer and fractional order rate laws*, International J. Solids Structures **36** (1999), 2417–2442.

[24] M. Enelund and P. Olsson, *Damping described by fading memory-analysis and application to fractional derivative models*, International J. Solids Structures **36** (1999), 939–970.

[25] A. Erdélyi, W. Magnus, F. Oberhettinger, and F. G. Tricomi, *Higher Transcendental Functions. Vol. III*, Robert E. Krieger Publishing Co. Inc., 1981, Based on notes left by Harry Bateman, Reprint of the 1955 original.

[26] K. Eriksson, D. Estep, P. Hansbo, and C. Johnson, *Computational Differential Equations*, Cambridge University Press, 1996.

[27] L. C. Evans, *Partial Differential Equations*, American Mathematical Society, Providence, Rhode Island, 1998.

[28] R. H. Fabiano and K. Ito, *Semigroup theory and numerical approximation for equations in linear viscoelasticity*, SIAM J. Math. Anal. **21** (1990), 374–393.

[29] N. J. Ford and A. C. Simpson, *The numerical solution of fractional differential equations: speed versus accuracy*, Numer. Algorithms **26** (2001), 333–346.

[30] D. A. French and J. W. Schaeffer, *Continuous finite element methods which preserve energy properties for nonlinear problems*, Appl. Math. Comput. **39** (1990), 271–295.

[31] D.A. French and T.E. Peterson, *A continuous space-time finite element method for the wave equation*, Math. Comp. **65** (1996), 491–506.

[32] M. Geissert, M. Kovács, and S. Larsson, *Rate of weak convergence of the finite element method for the stochastic heat equation with additive noise*, BIT, to appear.

[33] R. Glassey and J. Schaeffer, *Convergence of a second-order scheme for semilinear hyperbolic equations in* $2+1$ *dimensions*, Math. Comp. **56** (1991), 87–106.

[34] R. T. Glassey, *Convergence of an energy-preserving scheme for the Zakharov equations in one space dimension*, Math. Comp. **58** (1992), 83–102.

[35] I. Gyöngy, *Approximations of stochastic partial differential equations*, Stochastic partial differential equations and applications (Trento, 2002), Lecture Notes in Pure and Appl. Math., vol. 227, Dekker, 2002, pp. 287–307.

[36] G. M. Hulbert, *Time finite element methods for structural dynamics*, Internat. J. Numer. Methods Engrg. **33** (1992), 307–331.

[37] G. M. Hulbert and T. J. R. Hughes, *Space-time finite element methods for second-order hyperbolic equations*, Comput. Methods Appl. Mech. Engrg. **84** (1990), 327–348.

[38] C. Johnson, *Discontinuous Galerkin finite element methods for second order hyperbolic problems*, Comput. Methods Appl. Mech. Engrg. **107** (1993), 117–129.

[39] M. Kovács, S. Larsson, and F. Lindgren, *Strong convergence of the finite element method with truncated noise for semilinear parabolic stochastic equations*, Numer. Algorithms, to appear.

[40] X. D. Li and N. E. Wiberg, *Implementation and adaptivity of a space-time finite element method for structural dynamics*, Comput. Methods Appl. Mech. Engrg. **156** (1998), 211–229.

[41] Y. Lin, V. Thomée, and L. B. Wahlbin, *Ritz-volterra projections to finite-element spaces and application to integro-differential and related equations*, SIAM J. Numer. Anal. **28** (1991), 1047–1070.

[42] C. Lubich, *Convolution quadrature and discretized operational calculus I*, Numeriche Mathematik **52** (1988), 129–145.

[43] Ch. Lubich, I. H. Sloan, and V. Thomée, *Nonsmooth data error estimates for approximations of an evolution equation with a positive-type memory term*, Math. Comp. **65** (1996), 1–17.

[44] A. Martin, *Hyperbolic stochastic partial differential equations: Small balls and simulation; propagation of singularities.*, PhD Thesis, Aachen, Shaker Verlag. Neuherberg: Institut f. Biomathematik u. Biometrie, GSF Forschungszentrum f. Umwelt u. Gesundheit, 2002.

[45] A. Martin, S.M. Prigarin, and G. Winkler, *Exact and fast numerical algorithms for the stochastic wave equation*, Int. J. Comput. Math. **80** (2003), 1535–1541.

[46] W. McLean and V. Thomée, *Numerical solution of an evolution equation with a positive-type memory term*, J. Austral. Math. Soc. Ser. B **35** (1993), 23–70.

[47] W. McLean, V. Thomée, and L. B. Wahlbin, *Discretization with variable time steps of an evolution equation with a positive-type memory term*, J. of Comput. Appl. Math. **69** (1996), 49–69.

[48] K. B. Oldham and J. Spanier, *The Fractional Calculus*, Academic Press, New york and London, 1974.

[49] A. Pazy, *Semigroups of Linear Operators and Applications to Partial Differential Equations*, Springer-Verlag, 1983.

[50] I. Podlubny, *Fractional Differential Equations*, Academic Press, San diego, CA, 1999.

[51] G. Da Prato and J. Zabczyk, *Stochastic Equations in Infinite Dimensions*, Encyclopedia of Mathematics and its Applications, vol. 44, Cambridge University Press, 1992.

[52] L. Quer-Sardanyons and M. Sanz-Solé, *Space semi-discretisations for a stochastic wave equation*, Potential Anal. **24** (2006), 303–332.

[53] M. Renardy, W. J. Hrusa, and J.—A. Nohel, *Mathematical Problems in Viscoelasticity*, Pitman Monographs and Surveys in Pure and Applied Mathematics, vol. 35, Longman Scientific & Technical, 1987.

[54] B. Riviére, S. Shaw, and J. R. Whiteman, *Discontinuous Galerkin finite element methods for dynamic linear solid viscoelasticity problems*, Numer. Methods Partial Differential Equations **23** (2007), 1149–1166.

[55] B. Ross, *A Brief History and Exposition of the Fundamental Theory of Fractional Calculus*, Lecture notes in Mathematics, vol. 457, Springer-Verlag, 1975.

[56] Y. A. Rossikhin and M. V. Shitikova, *Applications of fractional calculus to dynamic problems of linear and nonlinear hereditary mechanics of solids*, Appl. Mech. Rev. **15-67** (1997), 704–718.

[57] C. Roth, *Weak approximations of solutions of a first order hyperbolic stochastic partial differential equation*, Monte Carlo Methods Appl. **13** (2007), 117–133.

[58] S. G. Samko, A. A. Kilbas, and O. I. Marichev, *Fractional Integrals and Derivatives*, Gordon and Breach Science Publishers, 1993.

[59] A. Schädle, M. López-Fernández, and Ch. Lubich, *Fast and oblivious convolution quadrature*, SIAM J. Sci. Comput. **28** (2006), 421–438.

[60] ______, *Adaptive, fast, and oblivious convolution in evolution equations with memory*, SIAM J. Sci. Comput. **30** (2008), 1015–1037.

[61] S. Shaw and J. R. Whiteman, *Applications and numerical analysis of partial differential Volterra equations: a brief survey*, Comput. Methods Appl. Mech. Engrg. **150** (1997), 397–409.

[62] ______, *A posteriori error estimates for space-time finite element approximation of quasistatic hereditary linear viscoelasticity problems*, Comput. Methods Appl. Mech. Engrg. **193** (2004), 5551–5572.

[63] I. H. Sloan and V. Thomée, *Time discretization of an integro-differential equation of parabolic type*, SIAM J. Numer. Anal. **23** (1986), 1052–1061.

[64] W. Strauss and L. Vazquez, *Numerical solution of a nonlinear Klein-Gordon equation*, J. Comput. Phys. **28** (1978), 271–278.

[65] E. D. Svensson, *Multigrid methods on adaptively refined triangulations: Practical considerations*, Preprint 2006:23, Dept. of Mathematics, Chalmers University of Technology and University of Gothenburg.

[66] J.B. Walsh, *An introduction to stochastic partial differential equations*, École d'été de probabilités de Saint-Flour, XIV—1984, Lecture Notes in Math., vol. 1180, Springer, 1986, pp. 265–439.

[67] ______, *On numerical solutions of the stochastic wave equation*, Illinois J. Math. **50** (2006), 991–1018 (electronic).

[68] D. V. Widder, *The Laplace Transform*, Princeton Mathematical Series, v. 6, Princeton University Press, 1941.

[69] Y. Yan, *Semidiscrete Galerkin approximation for a linear stochastic parabolic partial differential equation driven by an additive noise*, BIT **44** (2004), 829–847.

[70] ———, *Galerkin finite element methods for stochastic parabolic partial differential equations*, SIAM J. Numer. Anal. **43** (2005), 1363–1384 (electronic).

THE CONTINUOUS GALERKIN METHOD FOR AN INTEGRO-DIFFERENTIAL EQUATION MODELING DYNAMIC FRACTIONAL ORDER VISCOELASTICITY

STIG LARSSON AND FARDIN SAEDPANAH

ABSTRACT. We consider a fractional order integro-differential equation with a weakly singular convolution kernel. The equation with homogeneous mixed Dirichlet and Neumann boundary conditions is reformulated as an abstract Cauchy problem, and well-posedness is verified in the context of linear semigroup theory. Then we formulate a continuous Galerkin method for the problem, and we prove stability estimates. These are then used to prove a priori error estimates. The theory is illustrated by a numerical example.

1. INTRODUCTION

Bagley and Torvik [5] have proved that fractional order operators are very suitable for modelling viscoelastic materials. The basic equations of the viscoelastic dynamic problem, with surface loads, can be written in the strong form,

$$
\begin{aligned}
&\rho \ddot{\boldsymbol{u}}(\boldsymbol{x},t) - \nabla \cdot \boldsymbol{\sigma}_0(\boldsymbol{u};\boldsymbol{x},t) \\
&\qquad + \int_0^t b(t-s) \nabla \cdot \boldsymbol{\sigma}_1(\boldsymbol{u};\boldsymbol{x},s)\, ds = \boldsymbol{f}(\boldsymbol{x},t) && \text{in } \Omega \times (0,T), \\
&\boldsymbol{u}(\boldsymbol{x},t) = 0 && \text{on } \Gamma_D \times (0,T), \\
&\boldsymbol{\sigma}(\boldsymbol{u};\boldsymbol{x},t) \cdot \boldsymbol{n}(\boldsymbol{x}) = \boldsymbol{g}(\boldsymbol{x},t) && \text{on } \Gamma_N \times (0,T), \\
&\boldsymbol{u}(\boldsymbol{x},0) = \boldsymbol{u}^0(\boldsymbol{x}) && \text{in } \Omega, \\
&\dot{\boldsymbol{u}}(\boldsymbol{x},0) = \boldsymbol{v}^0(\boldsymbol{x}) && \text{in } \Omega,
\end{aligned}
\tag{1.1}
$$

(throughout this text we use '$\cdot$' to denote '$\frac{\partial}{\partial t}$') where $\boldsymbol{u}$ is the displacement vector, ρ is the (constant) mass density, $\boldsymbol{f}$ and $\boldsymbol{g}$ represent, respectively, the

Date: April 14, 2009.

1991 *Mathematics Subject Classification.* 65M60, 45K05.

Key words and phrases. finite element, continuous Galerkin, linear viscoelasticity, fractional calculus, weakly singular kernel, stability, a priori error estimate.

Research supported in part by the Swedish Foundation for Strategic Research through GMMC, the Gothenburg Mathematical Modelling Centre.

volume and surface loads, $\boldsymbol{\sigma}_0$ and $\boldsymbol{\sigma}_1$ are the stresses according to

$$\begin{aligned}\boldsymbol{\sigma}(t) &= \boldsymbol{\sigma}_0(t) - \int_0^t b(t-s)\boldsymbol{\sigma}_1(s)\,ds, \quad \text{with} \\ \boldsymbol{\sigma}_0(t) &= 2\mu_0\boldsymbol{\epsilon}(t) + \lambda_0 \mathrm{Tr}(\boldsymbol{\epsilon}(t))\mathbf{I}, \quad \boldsymbol{\sigma}_1(t) = 2\mu_1\boldsymbol{\epsilon}(t) + \lambda_1 \mathrm{Tr}(\boldsymbol{\epsilon}(t))\mathbf{I},\end{aligned} \tag{1.2}$$

where $\lambda_0 > \lambda_1 > 0$ and $\mu_0 > \mu_1 > 0$ are elastic constants of Lamé type, $\boldsymbol{\epsilon}$ is the strain which is defined through the usual linear kinematic relation $\boldsymbol{\epsilon} = \frac{1}{2}(\nabla \boldsymbol{u} + (\nabla \boldsymbol{u})^T)$, and b is the convolution kernel

$$b(t) = -\frac{d}{dt}\Big(\mathrm{E}_\alpha(-(t/\tau)^\alpha)\Big) = \frac{\alpha}{\tau}\Big(\frac{t}{\tau}\Big)^{\alpha-1}\mathrm{E}'_\alpha\Big(-\big(\frac{t}{\tau}\big)^\alpha\Big) \approx Ct^{-1+\alpha}, \ t \to 0. \tag{1.3}$$

Here $\tau > 0$ is the relaxation time and $\mathrm{E}_\alpha(z) = \sum_{k=0}^\infty \frac{z^k}{\Gamma(1+\alpha k)}$ is the Mittag-Leffler function of order $\alpha \in (0,1)$, $\gamma = \frac{\mu_1}{\mu_0} = \frac{\lambda_1}{\lambda_0} < 1$, so that $\boldsymbol{\sigma}_1 = \gamma\boldsymbol{\sigma}_0$ and we define $\beta(t) = \gamma b(t)$. The convolution kernel is weakly singular and $\beta \in L_1(0,\infty)$ with $\int_0^\infty \beta(t)\,dt = \gamma$. We introduce the function

$$\xi(t) = \gamma - \int_0^t \beta(s)\,ds = \int_t^\infty \beta(s)\,ds, \tag{1.4}$$

which is decreasing with $\xi(0) = \gamma$, $\lim_{t\to\infty} \xi(t) = 0$, so that $0 < \xi(t) \le \gamma$.

We let $\Omega \subset \mathbb{R}^d$, $d = 2, 3$, be a bounded polygonal domain with boundary $\Gamma = \Gamma_D \cup \Gamma_N$, where Γ_D and Γ_N are disjoint and $\mathrm{meas}(\Gamma_D) \neq 0$. We introduce the function spaces $H = L_2(\Omega)^d$, $H_{\Gamma_N} = L_2(\Gamma_N)^d$, and $V = \{\boldsymbol{v} \in H^1(\Omega)^d : \boldsymbol{v}|_{\Gamma_D} = \mathbf{0}\}$. We denote the norms in H and H_{Γ_N} by $\|\cdot\|$ and $\|\cdot\|_{\Gamma_N}$, respectively, and we equip V with the inner product $a(\cdot,\cdot)$ and norm $\|\boldsymbol{v}\|_V^2 = a(\boldsymbol{v},\boldsymbol{v})$, where (with the usual summation convention)

$$a(\boldsymbol{v},\boldsymbol{w}) = \int_\Omega \big(2\mu_0\epsilon_{ij}(\boldsymbol{v})\epsilon_{ij}(\boldsymbol{w}) + \lambda_0\epsilon_{ii}(\boldsymbol{v})\epsilon_{jj}(\boldsymbol{w})\big)\,dx, \quad \boldsymbol{v},\boldsymbol{w} \in V, \tag{1.5}$$

which is a coercive bilinear form on V. Setting $A\boldsymbol{u} = -\nabla\cdot\boldsymbol{\sigma}_0(\boldsymbol{u})$ with $\mathcal{D}(A) = H^2(\Omega)^d \cap V$ such that $a(\boldsymbol{u},\boldsymbol{v}) = (A\boldsymbol{u},\boldsymbol{v})$ for sufficiently smooth $\boldsymbol{u},\boldsymbol{v} \in V$, we can write the weak form of the equation of motion as: Find $\boldsymbol{u}(t) \in V$ such that $\boldsymbol{u}(0) = \boldsymbol{u}^0$, $\dot{\boldsymbol{u}}(0) = \boldsymbol{v}^0$, and

$$\begin{aligned}\rho(\ddot{\boldsymbol{u}}(t),\boldsymbol{v}) + a(\boldsymbol{u}(t),\boldsymbol{v}) - \int_0^t \beta(t-s)a(\boldsymbol{u}(s),\boldsymbol{v})\,ds \\ = (\boldsymbol{f}(t),\boldsymbol{v}) + (\boldsymbol{g}(t),\boldsymbol{v})_{\Gamma_N}, \quad \forall \boldsymbol{v} \in V,\ t \in (0,T),\end{aligned} \tag{1.6}$$

with $(\boldsymbol{g}(t),\boldsymbol{v})_{\Gamma_N} = \int_{\Gamma_N} \boldsymbol{g}(t)\cdot\boldsymbol{v}\,dS$. For more details see [1], [2], [3], [4] and references therein.

Defining $\boldsymbol{u}_1 = \boldsymbol{u}$ and $\boldsymbol{u}_2 = \dot{\boldsymbol{u}}$ we write (1.6) as: Find $\boldsymbol{u}_1(t),\, \boldsymbol{u}_2(t) \in V$ such that $\boldsymbol{u}_1(0) = \boldsymbol{u}^0$, $\boldsymbol{u}_2(0) = \boldsymbol{v}^0$, and

(1.7)
$$\begin{aligned} &a\big(\dot{\boldsymbol{u}}_1(t), \boldsymbol{v}_1\big) - a\big(\boldsymbol{u}_2(t), \boldsymbol{v}_1\big) = 0, \\ &\rho(\dot{\boldsymbol{u}}_2(t), \boldsymbol{v}_2) + a(\boldsymbol{u}_1(t), \boldsymbol{v}_2) - \int_0^t \beta(t-s) a(\boldsymbol{u}_1(s), \boldsymbol{v}_2)\, ds \\ &\qquad\qquad = (\boldsymbol{f}(t), \boldsymbol{v}_2) + (\boldsymbol{g}(t), \boldsymbol{v}_2)_{\Gamma_N}, \quad \forall \boldsymbol{v}_1, \boldsymbol{v}_2 \in V,\, t \in (0,T). \end{aligned}$$

In the next section, using (1.6) with $\Gamma_N \neq \varnothing$, $\boldsymbol{g} = 0$ or $\Gamma_N = \varnothing$, we reformulate the problem as an abstract Cauchy problem and prove well-posedness. We also discuss the regularity and obtain some regularity estimates. In §3 we use (1.7) to formulate a continuous Galerkin method based on linear polynomials both in time and space. Then in §4 we show stability estimates for the continuous Galerkin method, and in §5 we use them to prove a priori error estimates that are optimal in $L_\infty(L_2)$ and $L_\infty(H^1)$. Finally, in §6, we illustrate the theory by computing the approximate solutions of (1.1) in a simple but realistic numerical example.

There is an extensive literature on finite element methods for partial differential equations with memory, see, e.g., [1], [7], [8], [9], [10]. The present work extends previous works, e.g., [2], [3], [15], on quasi-static fractional order viscoelasticity ($\rho\ddot{\boldsymbol{u}} \approx 0$) to the dynamic case. The paper [4] also deals with the dynamic case but considers only spatial discretization. A dynamic model for viscoelasticity based on internal variables is studied in [12]. The memory term generates a growing amount of data that has to be stored and used in each time step. This can be dealt with by introducing "sparse quadrature" in the convolution term [16]. For a different approach based on "convolution quadrature", see [13], [14].

The main result in the present work are derived under rather restrictive assumptions, $\Gamma_N = \varnothing$ or $\Gamma_N \neq \varnothing$, $\boldsymbol{g} = 0$, which guarantee the global regularity needed for the a priori error analysis. Also our results do not admit adaptive meshes. In general such global regularity is not present, which calls for adaptive methods based on a posteriori error analysis. We plan to address these issues in future work.

2. Existence and uniqueness

In this section, using the theory of linear operator semigroups, we show that there is a unique solution of (1.6), with pure Dirichlet boundary condition, that is, $\Gamma_N = \varnothing$, or with homogeneous mixed Dirichlet-Neumann boundary condition, that is, $\boldsymbol{g} = 0$, $\Gamma_N \neq \varnothing$. The theory presented here does not admit the term $(\boldsymbol{g}, \boldsymbol{v})_{\Gamma_N} \neq 0$ in (1.6). We then investigate the regularity in the case of homogeneous Dirichlet boundary condition, that is, $\Gamma_N = \varnothing$. The techniques are adapted from [6].

We consider the strong form of (1.6), for any fixed $T > 0$, that is,

$$\rho \ddot{\boldsymbol{u}}(t) + A\boldsymbol{u}(t) - \int_0^t \beta(t-s) A\boldsymbol{u}(s)\, ds = \boldsymbol{f}(t), \quad t \in (0,T), \tag{2.1}$$

with the initial conditions

$$\boldsymbol{u}(0) = \boldsymbol{u}^0 \in \mathcal{D}(A), \ \ \dot{\boldsymbol{u}}(0) = \boldsymbol{v}^0 \in V. \tag{2.2}$$

We extend $\boldsymbol{u}$ by $\boldsymbol{u}(t) = \boldsymbol{h}(t)$ for $t < 0$ with $\boldsymbol{h}$ to be chosen. By adding $-\int_{-\infty}^0 \beta(t-s) A\boldsymbol{h}(s)\, ds$ to both sides of (2.1), changing the variables in the convolution terms and defining $\boldsymbol{w}(t,s) = \boldsymbol{u}(t) - \boldsymbol{u}(t-s)$, we get

$$\rho \ddot{\boldsymbol{u}}(t) + \tilde{\gamma} A\boldsymbol{u}(t) + \int_0^\infty \beta(s) A\boldsymbol{w}(t,s)\, ds = \boldsymbol{f}(t) - \int_t^\infty \beta(s) A\boldsymbol{h}(t-s)\, ds, \tag{2.3}$$

where $\tilde{\gamma} = 1 - \gamma = 1 - \int_0^\infty \beta(s)\, ds > 0$.

2.1. **An abstract Cauchy problem.** We choose $\boldsymbol{h}(t) = \boldsymbol{u}^0$ in (2.3), so that

$$\rho \ddot{\boldsymbol{u}}(t) + \tilde{\gamma} A\boldsymbol{u}(t) + \int_0^\infty \beta(s) A\boldsymbol{w}(t,s)\, ds = \tilde{\boldsymbol{f}}(t), \tag{2.4}$$

where,

$$\boldsymbol{w}(t,s) = \begin{cases} \boldsymbol{u}(t) - \boldsymbol{u}(t-s), & s \in [0,t], \\ \boldsymbol{u}(t) - \boldsymbol{u}^0, & s \in [t,\infty), \end{cases}$$

and, in view of (1.4),

$$\tilde{\boldsymbol{f}}(t) = \boldsymbol{f}(t) - \xi(t) A\boldsymbol{u}^0. \tag{2.5}$$

Then we reformulate (2.4) as an abstract Cauchy problem and prove well-posedness.

We set $\boldsymbol{v} = \rho \dot{\boldsymbol{u}}$ and define the Hilbert spaces

$$\begin{aligned} W &= L_{2,\beta}\big((0,\infty); V\big) = \Big\{ \boldsymbol{w} : \|\boldsymbol{w}\|_W^2 = \rho \int_0^\infty \beta(s) \|\boldsymbol{w}(s)\|_V^2\, ds < \infty \Big\}, \\ Z &= V \times H \times W \\ &= \Big\{ \boldsymbol{z} = (\boldsymbol{u}, \boldsymbol{v}, \boldsymbol{w}) : \|\boldsymbol{z}\|_Z^2 = \tilde{\gamma}\rho \|\boldsymbol{u}\|_V^2 + \|\boldsymbol{v}\|^2 + \|\boldsymbol{w}\|_W^2 < \infty \Big\}. \end{aligned}$$

We also define the linear operator $\mathcal{A}$ on Z such that, for $\boldsymbol{z} = (\boldsymbol{u}, \boldsymbol{v}, \boldsymbol{w})$,

$$\mathcal{A}\boldsymbol{z} = \Big(\frac{1}{\rho}\boldsymbol{v} \, , \, -A\Big(\tilde{\gamma}\boldsymbol{u} + \int_0^\infty \beta(s) \boldsymbol{w}(s)\, ds \Big) \, , \, \frac{1}{\rho}\boldsymbol{v} - D\boldsymbol{w} \Big),$$

with domain of definition

$$\mathcal{D}(\mathcal{A}) = \Big\{(\boldsymbol{u}, \boldsymbol{v}, \boldsymbol{w}) \in Z : \boldsymbol{v} \in V,\, \tilde{\gamma}\boldsymbol{u} + \int_0^\infty \beta(s)\boldsymbol{w}(s)\, ds \in \mathcal{D}(A), \boldsymbol{w} \in \mathcal{D}(D)\Big\},$$

where

$$D\boldsymbol{w} = \frac{d}{ds}\boldsymbol{w} \quad \text{with} \quad \mathcal{D}(D) = \{\boldsymbol{w} \in W : D\boldsymbol{w} \in W \text{ and } \boldsymbol{w}(0) = 0\}.$$

Therefore, a solution of (2.1) with (2.2) satisfies the abstract Cauchy problem

$$\begin{aligned} &\dot{\boldsymbol{z}}(t) = \mathcal{A}\boldsymbol{z}(t) + F(t), \quad 0 < t < T, \\ &\boldsymbol{z}(0) = \boldsymbol{z}^0, \end{aligned} \tag{2.6}$$

where $F(t) = (0\,,\, \tilde{\boldsymbol{f}}(t)\,,\, 0)$ and $\boldsymbol{z}^0 = \big(\boldsymbol{u}^0,\, \rho\boldsymbol{v}^0,\, 0\big)$, since

$$\boldsymbol{w}(0, s) = \boldsymbol{u}(0) - \boldsymbol{u}(-s) = \boldsymbol{u}(0) - h(-s) = \boldsymbol{u}^0 - \boldsymbol{u}^0 = 0. \tag{2.7}$$

We also note that $\boldsymbol{w}(t, 0) = \boldsymbol{u}(t) - \boldsymbol{u}(t) = 0$, so that $\boldsymbol{w}(t, \cdot) \in \mathcal{D}(D)$.

A function $\boldsymbol{z}$ which is differentiable a.e. on $[0, T]$ with $\dot{\boldsymbol{z}} \in L_1((0, T); Z)$ is called a *strong solution* of the initial value problem (2.6) if $\boldsymbol{z}(0) = \boldsymbol{z}^0$, $\boldsymbol{z}(t) \in \mathcal{D}(\mathcal{A})$, and $\dot{\boldsymbol{z}}(t) = \mathcal{A}\boldsymbol{z}(t) + F(t)$ a.e. on $[0, T]$.

Lemma 1. *Let $\boldsymbol{z} = (\boldsymbol{u}, \boldsymbol{v}, \boldsymbol{w})$ be a strong solution of* (2.6) *with $\boldsymbol{z}^0 = \big(\boldsymbol{u}^0,\, \rho\boldsymbol{v}^0,\, 0\big)$. Then $\boldsymbol{u}$ is a solution of* (2.1) *with initial conditions* (2.2).

Proof. For the components of the strong solution $\boldsymbol{z}$ of (2.6), we have

$$\begin{aligned} \dot{\boldsymbol{u}}(t) &= \frac{1}{\rho}\boldsymbol{v}(t), \quad t \in (0, T), \\ \dot{\boldsymbol{v}}(t) &= -A\Big(\tilde{\gamma}\boldsymbol{u}(t) + \int_0^\infty \beta(s)\boldsymbol{w}(t, s)\, ds\Big) + \tilde{\boldsymbol{f}}(t), \quad t \in (0, T), \\ \dot{\boldsymbol{w}}(t, s) &= \frac{1}{\rho}\boldsymbol{v}(t) - D\boldsymbol{w}(t, s), \quad s \in (0, \infty),\, t \in (0, T). \end{aligned}$$

The first equation and $\boldsymbol{z}^0 = \big(\boldsymbol{u}^0,\, \rho\boldsymbol{v}^0,\, 0\big)$ imply the initial conditions (2.2). The first and third equations mean that $\boldsymbol{w}$ satisfies the first order PDE

$$\frac{\partial}{\partial t}\boldsymbol{w} + \frac{\partial}{\partial s}\boldsymbol{w} = \frac{\partial}{\partial t}\boldsymbol{u}.$$

Besides, since $\boldsymbol{w}(0, \cdot) = 0$ and $\boldsymbol{w}(t, \cdot) \in \mathcal{D}(D)$ we have the boundary conditions $\boldsymbol{w}(0, s) = 0$ and $\boldsymbol{w}(t, 0) = 0$. Hence $\boldsymbol{w}(t, s) = \boldsymbol{u}(t) - \boldsymbol{u}(t-s)$, $0 \le s \le t$, and $\boldsymbol{w}(t, s) = \boldsymbol{u}(t) - \boldsymbol{u}^0 = \boldsymbol{u}(t) - \boldsymbol{u}(t-s)$, $0 \le t \le s$. This and the fact that (2.4) is obtained from the first two equations, imply that $\boldsymbol{u}$ satisfies (2.1) a.e. on $[0, T]$ by backward calculations from (2.3). □

Theorem 1. *Assume that $\Gamma_N = \varnothing$ or $\Gamma_N \neq \varnothing$ and $\boldsymbol{g} = 0$. There is a unique solution $\boldsymbol{u} = \boldsymbol{u}(t)$ of (2.1)–(2.2) for all $\boldsymbol{u}^0 \in \mathcal{D}(A)$ and $\boldsymbol{v}^0 \in V$, if $\boldsymbol{f} : [0,T] \to H$ is Lipschitz continuous. Moreover, for some $C = C(\tilde{\gamma}, \rho, T)$, we have the regularity estimate*

$$\|\boldsymbol{u}(t)\|_V + \|\dot{\boldsymbol{u}}(t)\| \le C\Big(\|\boldsymbol{u}^0\|_{H^2} + \|\boldsymbol{v}^0\| + \int_0^t \|\boldsymbol{f}\|\, ds\Big), \quad t \in [0,T]. \tag{2.8}$$

Proof. For any $\boldsymbol{u}^0 \in \mathcal{D}(A)$ and $\boldsymbol{v}^0 \in V$, we have $\boldsymbol{z}^0 = \big(\boldsymbol{u}^0, \boldsymbol{v}^0, 0\big) \in \mathcal{D}(\mathcal{A})$. We first show that F in (2.6) is differentiable a.e. on $[0,T]$ and $\dot{F} \in L_1([0,T];Z)$. We then show that the linear operator $\mathcal{A}$ is the infinitesimal generator of a C_0 semigroup $e^{t\mathcal{A}}$ on Z. These prove that there is a unique strong solution of (2.6) by [11, Corollary 4.2.10], and the proof of the first part is then complete by Lemma 1. Finally we prove (2.8).

1. By assumption $\boldsymbol{f}$ is Lipschitz continuous on $[0,T]$. Hence $\boldsymbol{f}$ is differentiable a.e. on $[0,T]$ and $\dot{\boldsymbol{f}} \in L_1((0,T);H)$, since H is a Hilbert space. Since $\dot{\xi}(t) = -\beta(t)$ by (1.4), from (2.5) we get

$$\dot{\tilde{\boldsymbol{f}}}(t) = \dot{\boldsymbol{f}}(t) + A\boldsymbol{u}^0\beta(t),$$

which shows that $\tilde{\boldsymbol{f}}$ is differentiable a.e. on $[0,T]$. Thus F is differentiable a.e. on $[0,T]$ and $\dot{F} \in L_1((0,T);Z)$.

2. We use the Lumer-Philips Theorem [11] to show that $\mathcal{A}$ generates a C_0 semigroup of contractions on Z. To this end we first show that $\mathcal{A}$ is dissipative. For $\boldsymbol{z} = (\boldsymbol{u}, \boldsymbol{v}, \boldsymbol{w}) \in \mathcal{D}(\mathcal{A})$ we have

$$\begin{aligned}\langle \mathcal{A}\boldsymbol{z}, \boldsymbol{z}\rangle_Z &= \tilde{\gamma} a(\boldsymbol{v}, \boldsymbol{u}) - \Big(A\big(\tilde{\gamma}\boldsymbol{u} + \int_0^\infty \beta(s)\boldsymbol{w}(s)\, ds\big), \boldsymbol{v}\Big) + \Big(\frac{1}{\rho}\boldsymbol{v} - D\boldsymbol{w}, \boldsymbol{w}\Big)_W \\ &= -\rho \int_0^\infty \beta(s) a(D\boldsymbol{w}(s), \boldsymbol{w}(s))\, ds = -\frac{1}{2}\rho \int_0^\infty \beta(s) D\|\boldsymbol{w}(s)\|_V^2\, ds.\end{aligned}$$

To prove that the last term is non-positive, and hence $\mathcal{A}$ is dissipative, we consider for $\epsilon > 0$,

$$\begin{aligned}\int_\epsilon^\infty \beta(s) D\|\boldsymbol{w}(s)\|_V^2\, ds &= \lim_{M\to\infty} \int_\epsilon^M \beta(s) D\|\boldsymbol{w}(s)\|_V^2\, ds \\ &= \lim_{M\to\infty} \Big(\beta(M)\|\boldsymbol{w}(M)\|_V^2 - \beta(\epsilon)\|\boldsymbol{w}(\epsilon)\|_V^2 - \int_\epsilon^M \beta'(s)\|\boldsymbol{w}(s)\|_V^2\, ds\Big) \\ &\ge -\beta(\epsilon)\|\boldsymbol{w}(\epsilon)\|_V^2,\end{aligned}$$

because $\beta'(s) < 0$ and $\lim_{M\to\infty} \beta(M)\|\boldsymbol{w}(M)\|_V^2 = 0$, since $\boldsymbol{w} \in W$, that is, $\int_0^\infty \beta(s)\|\boldsymbol{w}(s)\|_V^2\, ds < \infty$. Since $\boldsymbol{w}(\epsilon) = \int_0^\epsilon D\boldsymbol{w}(s)\, ds$, by the Cauchy-Schwarz inequality we have

$$\|\boldsymbol{w}(\epsilon)\|_V^2 \leq \Big(\int_0^\epsilon \|D\boldsymbol{w}(s)\|_V \, ds \Big)^2 \leq \int_0^\epsilon \frac{1}{\beta(s)} \, ds \int_0^\epsilon \beta(s) \|D\boldsymbol{w}(s)\|_V^2 \, ds,$$

and consequently we get

$$\beta(\epsilon)\|\boldsymbol{w}(\epsilon)\|_V^2 \leq \int_0^\epsilon \frac{\beta(\epsilon)}{\beta(s)} \, ds \int_0^\epsilon \beta(s)\|D\boldsymbol{w}(s)\|_V^2 \, ds \leq \epsilon \frac{1}{\rho} \|D\boldsymbol{w}\|_W^2,$$

since $\beta(\epsilon) \leq \beta(s)$ and $(\boldsymbol{u},\, \boldsymbol{v},\, \boldsymbol{w}) \in \mathcal{D}(\mathcal{A})$ implies $D\boldsymbol{w} \in W$. Therefore

$$\langle \mathcal{A}\boldsymbol{z}, \boldsymbol{z} \rangle_Z \leq \lim_{\epsilon \to 0^+} \frac{\epsilon}{2} \|D\boldsymbol{w}\|_W^2 = 0,$$

and $\mathcal{A}$ is dissipative.

Next we show that $R(I - \mathcal{A}) = Z$. To see this, for arbitrary $(\boldsymbol{\phi}, \boldsymbol{\psi}, \boldsymbol{\theta}) \in Z$ we must find a unique $(\boldsymbol{u}, \boldsymbol{v}, \boldsymbol{w}) \in \mathcal{D}(\mathcal{A})$ such that $(I - \mathcal{A})(\boldsymbol{u}, \boldsymbol{v}, \boldsymbol{w}) = (\boldsymbol{\phi}, \boldsymbol{\psi}, \boldsymbol{\theta})$, that is,

$$\begin{aligned} &\boldsymbol{u} - \frac{1}{\rho}\boldsymbol{v} = \boldsymbol{\phi}, \\ &\boldsymbol{v} + A\Big(\tilde{\gamma}\boldsymbol{u} + \int_0^\infty \beta(s)\boldsymbol{w}(s)\, ds\Big) = \boldsymbol{\psi}, \\ &\boldsymbol{w} - \frac{1}{\rho}\boldsymbol{v} + D\boldsymbol{w} = \boldsymbol{\theta}, \quad \boldsymbol{w}(0) = 0. \end{aligned} \tag{2.9}$$

From the first and third equations and $\boldsymbol{w}(0) = 0$ we get

$$\boldsymbol{v} = \rho(\boldsymbol{u} - \boldsymbol{\phi}), \quad \boldsymbol{w}(s) = \int_0^s e^{r-s}\big(\frac{1}{\rho}\boldsymbol{v} + \boldsymbol{\theta}(r)\big)\, dr.$$

Substituting these into the second equation of (2.9), we get

$$\rho(\boldsymbol{u} - \boldsymbol{\phi}) + A\Big(\tilde{\gamma}\boldsymbol{u} + \int_0^\infty \beta(s) \int_0^s e^{r-s}\big(\boldsymbol{u} - \boldsymbol{\phi} + \boldsymbol{\theta}(r)\big)\, dr\, ds\Big) = \boldsymbol{\psi},$$

and hence

$$\boldsymbol{u} + \kappa A\boldsymbol{u} = \boldsymbol{\phi} + \frac{1}{\rho}\Big(\boldsymbol{\psi} + \int_0^\infty \beta(s)e^{-s} \int_0^s e^r A\big(\boldsymbol{\phi} - \boldsymbol{\theta}(r)\big)\, dr\, ds\Big), \tag{2.10}$$

where $\kappa = \frac{1}{\rho}\big(1 - \int_0^\infty \beta(s)e^{-s} ds\big) > 0$. Now we need to show that this equation has a solution. The weak form is to find $\boldsymbol{u} \in V$ such that

$$b(\boldsymbol{u}, \boldsymbol{v}) = L(\boldsymbol{v}) \quad \forall \boldsymbol{v} \in V,$$

with the bilinear form

$$b(\boldsymbol{u}, \boldsymbol{v}) = (\boldsymbol{u}, \boldsymbol{v}) + \kappa a(\boldsymbol{u}, \boldsymbol{v}) \quad \text{for } \boldsymbol{u}, \boldsymbol{v} \in V,$$

and the linear form

$$L(\boldsymbol{v}) = (\boldsymbol{\phi}, \boldsymbol{v}) + \frac{1}{\rho}(\boldsymbol{\psi}, \boldsymbol{v}) + \frac{1}{\rho}\int_0^\infty \beta(s)e^{-s} \int_0^s e^r a\big(\boldsymbol{\phi} - \boldsymbol{\theta}(r), \boldsymbol{v}\big)\, dr\, ds.$$

Clearly $b(\cdot,\cdot)$ is bounded and coercive on V, and L is bounded on V. Therefore by the Riesz representation theorem, there is a unique solution, hence $R(I-\mathcal{A})=Z$.

Since Z is a Hilbert space, it follows from [11, Theorem 1.4.6], that $\overline{\mathcal{D}(\mathcal{A})}=Z$. So we have verified all the hypotheses of the Lumer-Philips theorem to complete the first part of the proof.

3. The unique strong solution of (2.6), is given by

$$\boldsymbol{z}(t)=e^{t\mathcal{A}}\boldsymbol{z}^0+\int_0^t e^{(t-s)\mathcal{A}}F(s)\,ds,$$

and $\|e^{t\mathcal{A}}\|_Z\le 1$, since $\mathcal{A}$ generates a C_0 semigroup of contractions. Therefore

$$\|\boldsymbol{z}(t)\|_Z\le\|\boldsymbol{z}^0\|_Z+\int_0^t\|F(s)\|_Z\,ds.$$

Since $\boldsymbol{v}=\rho\dot{\boldsymbol{u}}$, $\boldsymbol{z}^0=(\boldsymbol{u}^0,\,\rho\boldsymbol{v}^0,\,0)$ and $\|F(s)\|_Z=\|\tilde{\boldsymbol{f}}(s)\|=\|\boldsymbol{f}(s)-\xi(s)A\boldsymbol{u}^0\|$, we have

$$\begin{aligned}\Big(\tilde{\gamma}\rho\|\boldsymbol{u}(t)\|_V^2+\rho^2\|\dot{\boldsymbol{u}}(t)\|^2+\rho\int_0^\infty\beta(s)\|\boldsymbol{w}(t)\|_V^2\,ds\Big)^{1/2}&\\ \le\big(\tilde{\gamma}\rho\|\boldsymbol{u}^0\|_V^2+\rho^2\|\boldsymbol{v}^0\|^2\big)^{1/2}+\int_0^t\big(\|\boldsymbol{f}(s)\|+\xi(s)\|A\boldsymbol{u}^0\|\big)\,ds.&\end{aligned}$$

Consequently, we have the estimate (2.8) with $C=C(\tilde{\gamma},\rho,T)$. □

2.2. **Regularity.** In order to prove higher regularity we specialize to the homogeneous Dirichlet boundary condition, that is, $\Gamma_N=\varnothing$, and assume that the polygonal domain Ω is convex. This guarantees that we have the elliptic regularity estimate,

$$\|\boldsymbol{u}\|_{H^2}\le C\|A\boldsymbol{u}\|,\quad \boldsymbol{u}\in H^2(\Omega)^d\cap V. \tag{2.11}$$

We first choose $\boldsymbol{h}(t)=\boldsymbol{u}^0+t\boldsymbol{v}^0$ in (2.3), so that

$$\rho\ddot{\boldsymbol{u}}(t)+\tilde{\gamma}A\boldsymbol{u}(t)+\int_0^\infty\beta(s)A\boldsymbol{w}(t,s)\,ds=\breve{\boldsymbol{f}}(t), \tag{2.12}$$

where

$$\boldsymbol{w}(t,s)=\begin{cases}\boldsymbol{u}(t)-\boldsymbol{u}(t-s), & s\in[0,t],\\ \boldsymbol{u}(t)-\boldsymbol{u}^0-(t-s)\boldsymbol{v}^0, & s\in[t,\infty),\end{cases}$$

and, in view of (1.4),

$$\breve{\boldsymbol{f}}(t)=\boldsymbol{f}(t)-A\boldsymbol{v}^0\int_t^\infty(t-s)\beta(s)\,ds-\xi(t)A\boldsymbol{u}^0. \tag{2.13}$$

Then differentiating the equation (2.12) in time we get

$$\rho\,\ddot{\boldsymbol{u}}(t) + \tilde{\gamma} A\dot{\boldsymbol{u}}(t) + \int_0^\infty \beta(s) A\dot{\boldsymbol{w}}(t,s)\,ds = \dot{\check{\boldsymbol{f}}}(t), \tag{2.14}$$

which, with an underline instead of one time derivative, can be written as

$$\rho\dot{\underline{\boldsymbol{u}}}(t) + \tilde{\gamma} A\underline{\boldsymbol{u}}(t) + \int_0^\infty \beta(s) A\underline{\boldsymbol{w}}(t,s)\,ds = \check{\underline{\boldsymbol{f}}}(t), \tag{2.15}$$

with the initial values

$$\underline{\boldsymbol{u}}(0) = \underline{\boldsymbol{u}}^0 = \boldsymbol{v}^0\,,\ \ \dot{\underline{\boldsymbol{u}}}(0) = \underline{\boldsymbol{v}}^0 = \frac{1}{\rho}\big(\boldsymbol{f}(0) - A\boldsymbol{u}^0\big), \tag{2.16}$$

and

$$\check{\underline{\boldsymbol{f}}}(t) = \dot{\check{\boldsymbol{f}}}(t) = \dot{\boldsymbol{f}}(t) - \xi(t)A\boldsymbol{v}^0 + \beta(t)A\boldsymbol{u}^0, \tag{2.17}$$

and

$$\underline{\boldsymbol{w}}(t,s) = \begin{cases} \dot{\boldsymbol{u}}(t) - \dot{\boldsymbol{u}}(t-s), & s\in[0,t], \\ \dot{\boldsymbol{u}}(t) - \boldsymbol{v}^0, & s\in[t,\infty), \end{cases}$$

so that $\underline{\boldsymbol{w}}(t,0) = 0$. We note that $\underline{\boldsymbol{w}}$ is continuous and $\underline{\boldsymbol{w}}(t,\cdot) \in \mathcal{D}(D)$ for $t \geq 0$.

Then, in the same way as in §2.1 with $\underline{\boldsymbol{v}} = \rho\dot{\underline{\boldsymbol{u}}}$, we can reformulate (2.15)–(2.16) as the abstract Cauchy problem

$$\begin{aligned} \dot{\underline{\boldsymbol{z}}}(t) &= \mathcal{A}\,\underline{\boldsymbol{z}}(t) + \check{\underline{F}}(t), \quad 0<t<T, \\ \underline{\boldsymbol{z}}(0) &= \underline{\boldsymbol{z}}^0, \end{aligned} \tag{2.18}$$

where $\check{\underline{F}}(t) = (0\,,\,\check{\underline{\boldsymbol{f}}}(t)\,,\,0)$ and $\underline{\boldsymbol{z}}^0 = \big(\underline{\boldsymbol{u}}^0,\,\rho\underline{\boldsymbol{v}}^0,\,0\big)$, since

$$\underline{\boldsymbol{w}}(0,s) = \dot{\boldsymbol{u}}(0) - \boldsymbol{v}^0 = \boldsymbol{v}^0 - \boldsymbol{v}^0 = 0. \tag{2.19}$$

Lemma 2. *Let* $\underline{\boldsymbol{z}} = (\underline{\boldsymbol{u}},\underline{\boldsymbol{v}},\underline{\boldsymbol{w}})$ *be a strong solution of* (2.18) *with* $\underline{\boldsymbol{z}}^0 = \big(\underline{\boldsymbol{u}}^0,\,\rho\underline{\boldsymbol{v}}^0,\,0\big)$. *Then* $\boldsymbol{u}(t) = \boldsymbol{u}^0 + \int_0^t \underline{\boldsymbol{u}}(s)\,ds$ *is a solution of* (2.1) *with initial conditions* (2.2).

Proof. Clearly $\boldsymbol{u}(0) = \boldsymbol{u}^0$ and $\dot{\boldsymbol{u}} = \underline{\boldsymbol{u}}$. Hence $\underline{\boldsymbol{z}}^0 = \big(\underline{\boldsymbol{u}}^0,\,\underline{\boldsymbol{v}}^0,\,0\big)$ implies $\dot{\boldsymbol{u}}(0) = \underline{\boldsymbol{u}}(0) = \underline{\boldsymbol{u}}^0 = \boldsymbol{v}^0$, so that (2.2) holds. Then since $\dot{\underline{\boldsymbol{z}}}(t) = \mathcal{A}\,\underline{\boldsymbol{z}}(t) + \check{\underline{F}}(t)$ a.e. on $[0,T]$, we have,

$$\begin{aligned} \dot{\underline{\boldsymbol{u}}}(t) &= \frac{1}{\rho}\underline{\boldsymbol{v}}(t), \quad t\in(0,T), \\ \dot{\underline{\boldsymbol{v}}}(t) &= -A\Big(\tilde{\gamma}\underline{\boldsymbol{u}}(t) + \int_0^\infty \beta(s)\underline{\boldsymbol{w}}(t,s)\,ds\Big) + \check{\underline{\boldsymbol{f}}}(t), \quad t\in(0,T), \\ \dot{\underline{\boldsymbol{w}}}(t,s) &= \frac{1}{\rho}\underline{\boldsymbol{v}}(t) - D\underline{\boldsymbol{w}}(t,s), \quad s\in(0,\infty),\ t\in(0,T). \end{aligned}$$

The first and the third equation with $\underline{\boldsymbol{w}}(t,0)=0$, $\underline{\boldsymbol{w}}(0,s)=0$ has the unique solution $\underline{\boldsymbol{w}}(t,s)=\underline{\boldsymbol{u}}(t)-\underline{\boldsymbol{u}}(t-s)$ that implies, by integration with respect to t, $\boldsymbol{w}(t,s):=\boldsymbol{u}(t)-\boldsymbol{u}(t-s)=\int_0^t \underline{\boldsymbol{w}}(\tau,s)\,d\tau$. By the first equation we have $\ddot{\boldsymbol{u}}=\dot{\underline{\boldsymbol{u}}}=\frac{1}{\rho}\underline{\boldsymbol{v}}$, so that the second equation is (2.14). The proof is completed by backward calculation from (2.14). □

In the next theorem we find the circumstances under which there is a unique solution of (2.1) with more regularity.

Theorem 2. *Assume that* $\Gamma_N=\varnothing$ *and that* Ω *is a convex polygonal domain. There is a unique solution* $\boldsymbol{u}=\boldsymbol{u}(t)$ *of* (2.1)–(2.2) *if* $\boldsymbol{v}^0\in\mathcal{D}(A)$, $\boldsymbol{f}(0)-A\boldsymbol{u}^0\in V$, *and* $\dot{\boldsymbol{f}}:[0,T]\to H$ *is Lipschitz continuous. Moreover, for some* $C=C(\tilde{\gamma},\rho,T)$, *we have the regularity estimate*

(2.20)
$$\|\dot{\boldsymbol{u}}(t)\|_V+\|\ddot{\boldsymbol{u}}(t)\|+\|\boldsymbol{u}(t)\|_{H^2} \le C\Big(\|\boldsymbol{f}(0)-A\boldsymbol{u}^0\|+\|\boldsymbol{v}^0\|_{H^2}+\int_0^t\|\dot{\boldsymbol{f}}\|\,ds\Big),\quad t\in[0,T].$$

Proof. 1. From the assumptions on $\boldsymbol{u}^0$, $\boldsymbol{v}^0$ and $\boldsymbol{f}(0)$ and recalling (2.16) and (2.19), we have $\underline{\boldsymbol{z}}^0=\big(\underline{\boldsymbol{u}}^0,\rho\underline{\boldsymbol{v}}^0,\underline{\boldsymbol{w}}^0(\cdot)\big)\in\mathcal{D}(\mathcal{A})$. We split the load term $\underline{\breve{F}}$ in (2.18) as
$$\underline{\breve{F}}=\underline{\breve{F}}_1+\underline{\breve{F}}_2,$$
where
$$\begin{aligned}\underline{\breve{F}}_1(t)&=(0\,,\,\underline{\breve{\boldsymbol{f}}}_1(t)\,,\,0)=(0\,,\,\dot{\boldsymbol{f}}(t)-\xi(t)A\boldsymbol{v}^0\,,\,0),\\ \underline{\breve{F}}_2(t)&=(0\,,\,\underline{\breve{\boldsymbol{f}}}_2(t)\,,\,0)=(0\,,\,\beta(t)A\boldsymbol{u}^0\,,\,0).\end{aligned}$$
We show that each one of the abstract Cauchy problems, for $i=1,\,2$,

$$\begin{aligned}&\dot{\underline{\boldsymbol{z}}}(t)=\mathcal{A}\,\underline{\boldsymbol{z}}(t)+\underline{\breve{F}}_i(t),\quad 0<t<T,\\ &\underline{\boldsymbol{z}}(0)=\underline{\boldsymbol{z}}^0,\end{aligned}\tag{2.21}$$

has a unique strong solution, and consequently there is a unique strong solution of (2.18). We recall that the linear operator $\mathcal{A}$ is an infinitesimal generator of a C_0-semigroup of contractions $e^{t\mathcal{A}}$ on Z by the proof of Theorem 1.

Considering
$$\dot{\underline{\breve{\boldsymbol{f}}}}_1(t)=\ddot{\boldsymbol{f}}(t)+\beta(t)A\boldsymbol{v}^0,$$
and the assumptions on $\boldsymbol{v}^0$ and $\dot{\boldsymbol{f}}$, $\underline{\breve{F}}_1$ is differentiable a.e. on $[0,T]$ and $\dot{\underline{\breve{F}}}_1\in L_1((0,T);Z)$. By [11, Corollary 4.2.10], there is a unique strong solution of (2.21) for $i=1$. On the other hand $\underline{\breve{F}}_2(t)$ is continuous on $(0,T)$, $\underline{\breve{F}}_2(t)\in\mathcal{D}(\mathcal{A})$, $t\in(0,T)$, and $\mathcal{A}\underline{\breve{F}}_2\in L_1((0,T);Z)$, since $\beta(t)$ is continuous on $(0,T)$

and $A\boldsymbol{u}^0 \in V$. Therefore, by [11, Corollary 4.2.6], there is a unique classical solution of (2.21) with $i = 2$. Since any classical solution is a strong solution, the proof of existence and uniqueness is completed by Lemma 2.

2. We have the unique strong solution of (2.18), i.e.

$$\underline{\boldsymbol{z}}(t) = e^{t\mathcal{A}}\underline{\boldsymbol{z}}^0 + \int_0^t e^{(t-s)\mathcal{A}}\underline{\breve{F}}(s)\,ds,$$

with $\|e^{t\mathcal{A}}\|_Z \leq 1$. Following step 3 of Theorem 1, using (2.11), we get (2.20). □

3. The continuous Galerkin method

Recalling the function spaces $H = L_2(\Omega)^d$, $H_{\Gamma_N} = L_2(\Gamma_N)^d$ and $V = \{\boldsymbol{v} \in H^1(\Omega)^d : \boldsymbol{v}|_{\Gamma_D} = 0\}$ $(d = 2, 3)$, we provide some definitions which will be used in the forthcoming discussions.

Let $0 = t_0 < t_1 < \cdots < t_{n-1} < t_n < \cdots < t_N = T$ be a partition of the time interval $I = [0, T]$. To each time subinterval $I_n = (t_{n-1}, t_n)$ of length $k_n = t_n - t_{n-1}$, we associate a triangulation $\mathcal{T}_n$ of Ω with piecewise constant mesh function h_n defined by

$$h_n(x) = \text{diam}(K), \quad \text{for } x \in K,\ K \in \mathcal{T}_n, \tag{3.1}$$

and the corresponding finite element space V_n of vector-valued continuous piecewise linear polynomials, that vanish on Γ_D (This requires that the mesh is adjusted to fit Γ_D.). We also define the spaces, for $q = 0,\ 1$,

$$W^{(q)} = \Big\{\boldsymbol{w} : \boldsymbol{w}|_{\Omega\times I_n} = \boldsymbol{w}^n \in W_n^{(q)},\ n = 1, \ldots, N\Big\},$$

where,

$$W_n^{(q)} = \Big\{\boldsymbol{w} : \boldsymbol{w}(x,t) = \sum_{i=0}^{q} \boldsymbol{w}_i(x)t^i,\ \boldsymbol{w}_i \in V_n\Big\}.$$

Note that $\boldsymbol{w} \in W^{(q)}$ may be discontinuous at $t = t_n$, and $w \in W^{(0)}$ is piecewise constant in time.

With $\mathbb{P}_q^d$ denoting the set of all vector-valued polynomials of degree at most q, the orthogonal projections $\mathcal{R}_{h,n} : V \to V_n$, $\mathcal{P}_{h,n} : H \to V_n$ and $\mathcal{P}_{k,n} : L_2(I_n)^d \to \mathbb{P}_{q-1}^d(I_n)$ are defined, respectively, by

$$\begin{aligned} &a(\mathcal{R}_{h,n}\boldsymbol{v} - \boldsymbol{v}, \boldsymbol{\chi}) = 0, && \forall \boldsymbol{v} \in V,\ \boldsymbol{\chi} \in V_n,\\ &(\mathcal{P}_{h,n}\boldsymbol{v} - \boldsymbol{v}, \boldsymbol{\chi}) = 0, && \forall \boldsymbol{v} \in H,\ \boldsymbol{\chi} \in V_n,\\ &\int_{I_n} (\mathcal{P}_{k,n}\boldsymbol{v} - \boldsymbol{v}) \cdot \boldsymbol{\psi}\,dt = 0, && \forall \boldsymbol{v} \in L_2(I_n)^d,\ \boldsymbol{\psi} \in \mathbb{P}_{q-1}^d. \end{aligned} \tag{3.2}$$

Correspondingly, we define $\mathcal{R}_h\boldsymbol{v}$, $\mathcal{P}_h\boldsymbol{v}$ and $\mathcal{P}_k\boldsymbol{v}$ by $(\mathcal{R}_h\boldsymbol{v})(t) = \mathcal{R}_{h,n}\boldsymbol{v}(t)$, $(\mathcal{P}_h\boldsymbol{v})(t) = \mathcal{P}_{h,n}\boldsymbol{v}(t)$ for $t \in I_n$, and $\mathcal{P}_k\boldsymbol{v} = \mathcal{P}_{k,n}(\boldsymbol{v}\,|_{I_n})$, $(n = 1, \cdots, N)$.

We also define the orthogonal projections, $R_n : L_2(I_n, V) \to W_n^{(q-1)}$ and $P_n : L_2(I_n, H) \to W_n^{(q-1)}$, such that

$$
\begin{aligned}
&\int_{I_n} a(R_n \boldsymbol{u} - \boldsymbol{u}, \boldsymbol{\psi})\, dt = 0, \quad \forall \boldsymbol{\psi} \in W_n^{(q-1)}, \ \boldsymbol{u} \in L_2(I_n, V), \\
&\int_{I_n} (P_n \boldsymbol{u} - \boldsymbol{u}, \boldsymbol{\psi})\, dt = 0, \qquad \forall \boldsymbol{\psi} \in W_n^{(q-1)}, \ \boldsymbol{u} \in L_2(I_n, H).
\end{aligned} \tag{3.3}
$$

Correspondingly, we define $R : L_2(I, V) \to W^{(0)}$, $P : L_2(I, H) \to W^{(0)}$ in the obvious way.

One can easily show that

$$
R = \mathcal{R}_h \mathcal{P}_k = \mathcal{P}_k \mathcal{R}_h, \quad P = \mathcal{P}_h \mathcal{P}_k = \mathcal{P}_k \mathcal{P}_h, \tag{3.4}
$$

and for $\boldsymbol{u} \in W_n^{(1)}$, $\boldsymbol{v} \in W_n^{(0)}$,

$$
\int_{I_n} (\boldsymbol{u}, \boldsymbol{v})\, dt = \int_{I_n} (\mathcal{P}_{k,n} \boldsymbol{u}, \boldsymbol{v})\, dt, \quad \int_{I_n} a(\boldsymbol{u}, \boldsymbol{v})\, dt = \int_{I_n} a(\mathcal{P}_{k,n} \boldsymbol{u}, \boldsymbol{v})\, dt. \tag{3.5}
$$

We introduce the linear operator $A_{n,r} : V_r \to V_n$ by

$$
a(\boldsymbol{v}_r, \boldsymbol{w}_n) = (A_{n,r} \boldsymbol{v}_r, \boldsymbol{w}_n) \qquad \forall \boldsymbol{v}_r \in V_r,\ \boldsymbol{w}_n \in V_n.
$$

We set $A_n = A_{n,n}$, with discrete norms

$$
|\boldsymbol{v}_n|_{h,l} = \|A_n^{l/2} \boldsymbol{v}_n\| = \sqrt{(\boldsymbol{v}_n, A_n^l \boldsymbol{v}_n)}, \quad \boldsymbol{v}_n \in V_n \text{ and } l \in \mathbb{R},
$$

and A_h so that $A_h \boldsymbol{v}|_{I_n} = A_n \boldsymbol{v}$ for $\boldsymbol{v} \in V_n$. For later use in our error analysis we note that $\mathcal{P}_h A = A_h \mathcal{R}_h$.

We define the bilinear and linear forms $B : \mathcal{W} \times \mathcal{W} \to \mathbb{R}$ and $L : \mathcal{W} \to \mathbb{R}$ by

$$
\begin{aligned}
B(\boldsymbol{u}, \boldsymbol{v}) &= \sum_{n=1}^{N} \int_{I_n} \Big(- a(\boldsymbol{u}_2, \boldsymbol{v}_1) + a(\dot{\boldsymbol{u}}_1, \boldsymbol{v}_1) + \rho(\dot{\boldsymbol{u}}_2, \boldsymbol{v}_2) + a(\boldsymbol{u}_1, \boldsymbol{v}_2) \Big)\, dt \\
&\quad - \sum_{n=1}^{N} \int_{I_n} \int_0^t \beta(t-s) a\big(\boldsymbol{u}_1(s), \boldsymbol{v}_2(t)\big) ds\, dt, \\
L(\boldsymbol{w}) &= \sum_{n=1}^{N} \int_{I_n} (\boldsymbol{f}, \boldsymbol{w}_2) + (\boldsymbol{g}, \boldsymbol{w}_2)_{\Gamma_N}\, dt,
\end{aligned}
$$

where $\mathcal{W}$ is the space of pairs of vector-valued functions $\boldsymbol{u}(t) = (\boldsymbol{u}_1(t),\ \boldsymbol{u}_2(t))$ in V that are piecewise smooth with respect to the temporal mesh. We may note that $(W^{(q)})^2 \subset \mathcal{W}$ for $q \geq 0$.

The continuous Galerkin method of degree $(1,1)$ is based on the variational formulation (1.7) and reads: Find $U = (U_1, U_2) \in \big(W^{(1)}\big)^2$ such that,

$U_{1,0}^{-} = \boldsymbol{u}^0$, $U_{2,0}^{-} = \boldsymbol{v}^0$, and, for $n = 1, \ldots, N$,

$$
\begin{aligned}
&\int_{I_n} a(\dot{U}_1, V_1) - a(U_2, V_1)\, dt = 0,\\
&\int_{I_n} \Big(\rho(\dot{U}_2, V_2) + a(U_1, V_2) - \int_0^t \beta(t-s) a\big(U_1(s), V_2(t)\big)\, ds\Big)\, dt \\
&\qquad = \int_{I_n} (\boldsymbol{f}, V_2)\, dt + \int_{I_n} (\boldsymbol{g}, V_2)_{\Gamma_N}\, dt, \quad \forall (V_1, V_2) \in \big(W_n^{(0)}\big)^2,\\
&U_{1,n-1}^{+} = \mathcal{R}_{h,n} U_{1,n-1}^{-}\,, \quad U_{2,n-1}^{+} = \mathcal{P}_{h,n} U_{2,n-1}^{-}\,,
\end{aligned}
\tag{3.6}
$$

where $U_{i,n}^{\pm} = \lim_{s \to 0^{\pm}} U_i(t_n + s)$, $i = 1, 2$. Hence $U \in \big(W^{(1)}\big)^2$ defined in (3.6) satisfies:

$$
\begin{aligned}
&B(U, \mathcal{P}_k V) = L(\mathcal{P}_k V), \qquad\qquad \forall V \in \big(W^{(1)}\big)^2,\\
&U_{1,n-1}^{+} = \mathcal{R}_{h,n} U_{1,n-1}^{-}\,, \quad U_{2,n-1}^{+} = \mathcal{P}_{h,n} U_{2,n-1}^{-}\,,\\
&U_{1,0}^{-} = \boldsymbol{u}^0, \quad U_{2,0}^{-} = \boldsymbol{v}^0,
\end{aligned}
$$

where $\mathcal{P}_k V = (\mathcal{P}_k V_1, \mathcal{P}_k V_2)$.

Since the variational equation (1.7) can be written as: Find $\boldsymbol{u} \in \mathcal{W}$ such that

$$
B(\boldsymbol{u}, \boldsymbol{v}) = L(\boldsymbol{v}), \quad \forall \boldsymbol{v} \in \mathcal{W},
$$

we may, for later reference, note the Galerkin orthogonality

$$
B(U - \boldsymbol{u}, \mathcal{P}_k V) = 0, \quad \forall V \in \big(W^{(1)}\big)^2. \tag{3.7}
$$

Considering the fact that functions in $W_n^{(0)}$ are constant with respect to time, we can write (3.6) as

$$
\begin{aligned}
&A_n(U_{1,n}^{-} - U_{1,n-1}^{+}) - \frac{k_n}{2} A_n(U_{2,n}^{-} + U_{2,n-1}^{+}) = 0,\\
&A_n\Big((\frac{k_n}{2} - \gamma\omega_{n,n}^{-}) U_{1,n}^{-} + (\frac{k_n}{2} - \gamma\omega_{n,n-1}^{+}) U_{1,n-1}^{+} \Big)\\
&\qquad\qquad\qquad + \rho\, (U_{2,n}^{-} - U_{2,n-1}^{+}) = H_n + b_n,
\end{aligned}
$$

where

$$\begin{aligned}
b_n &= k_n(P_n\boldsymbol{f} + P_n^{\Gamma_N}\boldsymbol{g}),\\
H_n &= \gamma\sum_{r=1}^{n-1} k_r A_{n,r}(\omega_{n,r}^- U_{1,r}^- + \omega_{n,r-1}^+ U_{1,r-1}^+),\\
\omega_{n,r}^- &= \int_{I_n}\int_{t_{r-1}}^{t_r\wedge t}\beta(t-s)\psi_r^-(s)\,ds\,dt\,, \qquad t_r\wedge t = \min(t_r,t),\\
\omega_{n,r-1}^+ &= \int_{I_n}\int_{t_{r-1}}^{t_r\wedge t}\beta(t-s)\psi_{r-1}^+(s)\,ds\,dt,
\end{aligned}$$

and ψ_n^-, ψ_{n-1}^+ are the linear Lagrange basis functions on I_n, so that, for $i = 1, 2$,

$$U_i(x,t)\,|_{\Omega\times I_n} = \psi_{n-1}^+(t)U_{i,n-1}^+(x) + \psi_n^-(t)U_{i,n}^-(x).$$

From now on we assume, for simplicity, that $\mathcal{T}_{n-1}\subset\mathcal{T}_n$, $n = 2,\dots,N$, which means that the spatial mesh is allowed to be refined (or unchanged) at t_{n-1}. Then $V_{n-1}\subset V_n$ $(n = 2,\dots,N)$, $U_{i,n-1}^+ = U_{i,n-1}^-$ $(n = 1,\dots,N,\ i = 1, 2)$, and the initial conditions in (3.6) reduce to $U_1(\cdot,0) = \boldsymbol{u}_h^0 = \mathcal{R}_{h,1}\boldsymbol{u}^0$ and $U_2(\cdot,0) = \boldsymbol{v}_h^0 = \mathcal{P}_{h,1}\boldsymbol{v}^0$. In this case U is continuous with respect to t.

4. Stability estimates

We consider a modified problem by adding an extra load function, say $\boldsymbol{f}_1 = \boldsymbol{f}_1(t)$, to the first equation of (3.6). This kind of problem will occur in our error analysis below. Moreover, in the error equations the term corresponding to the surface load is zero, i.e., $\boldsymbol{g} = 0$. In this section we therefore consider the problem: Find $U\in\big(W^{(1)}\big)^2$ such that, for $n = 1,\dots,N$,

$$\begin{aligned}
&\int_{I_n} a(\dot U_1,V_1) - a(U_2,V_1)\,dt = \int_{I_n} a(\boldsymbol{f}_1,V_1)\,dt,\\
&\int_{I_n}\Big(\rho(\dot U_2,V_2) + a(U_1,V_2) - \int_0^t\beta(t-s)a\big(U_1(s),V_2(t)\big)ds\Big)\,dt\\
&\qquad\qquad = \int_{I_n}(\boldsymbol{f}_2,V_2)\,dt\,, \quad \forall (V_1,V_2)\in\big(W_n^{(0)}\big)^2,\\
&U_1,\,U_2 \ \text{continuous at } t_{n-1},\\
&U_1(\cdot,0) = \boldsymbol{u}_h^0\,, \quad U_2(\cdot,0) = \boldsymbol{v}_h^0\,.
\end{aligned}\tag{4.1}$$

Then U satisfies:

$$\begin{aligned}
&B(U,\mathcal{P}_k V) = \hat L(\mathcal{P}_k V), \quad \forall\, V\in\big(W^{(1)}\big)^2,\\
&U_1,\,U_2 \ \text{continuous at } t_{n-1},\\
&U_1(\cdot,0) = \boldsymbol{u}_h^0\,, \quad U_2(\cdot,0) = \boldsymbol{v}_h^0\,,
\end{aligned}\tag{4.2}$$

where the linear form $\hat{L} : \mathcal{W} \to \mathbb{R}$ is defined by,

$$\hat{L}(\boldsymbol{w}) = \sum_{n=1}^{N} \int_{I_n} \Big(a(\boldsymbol{f}_1, \boldsymbol{w}_1) + (\boldsymbol{f}_2, \boldsymbol{w}_2) \Big)\, dt.$$

In the next theorem we prove an energy identity for problem (4.1) which will be used for proving the error estimates in the next section.

Theorem 3. *Let $U = (U_1, U_2)$ be a solution of (4.1). Then for any $l \in \mathbb{R}$, $T > 0$, we have the equality*

$$\begin{aligned}
&\rho |U_{2,N}|_{h,l}^2 + \tilde{\xi}(T) |U_{1,N}|_{h,l+1}^2 + \int_0^T \beta |U_1|_{h,l+1}^2 \, dt \\
&\quad + \int_0^T \int_0^t \beta(t-s) D_t |W_1(t,s)|_{h,l+1}^2 \, ds\, dt \\
&= \rho |\boldsymbol{v}_h^0|_{h,l}^2 + |\boldsymbol{u}_h^0|_{h,l+1}^2 \\
&\quad + 2 \int_0^T (P\boldsymbol{f}_2, A_h^l U_2)\, dt + 2 \int_0^T \tilde{\xi} a(R\boldsymbol{f}_1, A_h^l U_1)\, dt \\
&\quad + 2 \int_0^T \int_0^t \beta(t-s) a\big(R\boldsymbol{f}_1(t), A_h^l W_1(t,s)\big)\, ds\, dt,
\end{aligned} \tag{4.3}$$

where $W_1(t,s) = U_1(t) - U_1(s)$ and, recalling (1.4),

$$\tilde{\xi}(t) = \xi(t) + 1 - \gamma = 1 - \int_0^t \beta(s)\, ds, \tag{4.4}$$

with $0 < 1 - \gamma < \tilde{\xi}(t) \le 1$. All terms on the left side of (4.3) are non-negative.

Proof. Throughout the proof we note that U_i $(i = 1, 2)$ are continuous, hence piecewise differentiable with respect to t. We organize our proof in 6 steps.

1. Expressing U_2 in terms of U_1. The first equation of (4.1) may be written as

$$\int_{I_n} a(\mathcal{P}_{k,n} U_2, V_1)\, dt = \int_{I_n} a(\dot{U}_1 - R_n \boldsymbol{f}_1, V_1)\, dt, \quad \forall V_1 \in W_n^{(0)}.$$

Therefore, we get

$$\mathcal{P}_k U_2 = \dot{U}_1 - R\boldsymbol{f}_1 \in W^{(0)}. \tag{4.5}$$

2. Using the calculation

$$\begin{aligned} U_1(t) - \int_0^t \beta(t-s) U_1(s)\,ds &= U_1(t) + \int_0^t \beta(t-s)\big(U_1(t) - U_1(s)\big)\,ds \\ &\quad - \int_0^t \beta(s)\,ds\, U_1(t) \\ &= \tilde{\xi}(t) U_1(t) + \int_0^t \beta(t-s) W_1(t,s)\,ds, \end{aligned}$$

and recalling the definitions of the P and P^{Γ_N} in (3.3) and the functions W_1 and $\tilde{\xi}$, we can write the second equation of (4.1) in the form

$$\begin{aligned} \int_0^T \Big(\rho(\dot{U}_2, V_2) + \tilde{\xi} a(U_1, V_2) + \int_0^t \beta(t-s) a\big(W_1(t,s), V_2(t)\big)\,ds \Big)\,dt \\ = \int_0^T (P_n \boldsymbol{f}_2, V_2)\,dt, \quad \forall V_2 \in W^{(0)}. \end{aligned}$$

Choosing $V_2 = A_h^l \mathcal{P}_k U_2$ gives us

$$\begin{aligned} \int_0^T \rho(\dot{U}_2, A_h^l \mathcal{P}_k U_2)\,dt + \int_0^T \tilde{\xi} a(U_1, A_h^l \mathcal{P}_k U_2)\,dt \\ + \int_0^T \int_0^t \beta(t-s) a\big(W_1(t,s), A_h^l \mathcal{P}_k U_2(t)\big)\,ds\,dt \\ = \int_0^T (P \boldsymbol{f}_2, A_h^l U_2)\,dt. \end{aligned} \tag{4.6}$$

We study the three terms in the left side of the above equation.

3. Using $\dot{U}_2 \in W^{(0)}$, by (3.5), we can write the first term of the left side of (4.6) as

$$\begin{aligned} \int_0^T \rho\big(\dot{U}_2, A_h^l \mathcal{P}_k U_2\big)\,dt &= \rho \int_0^T \big(\dot{U}_2, A_h^l U_2\big)\,dt = \frac{\rho}{2} \int_0^T D_t |U_2|_{h,l}^2\,dt \\ &= \frac{\rho}{2} \sum_{n=1}^N \Big(|U_{2,n}|_{h,l}^2 - |U_{2,n-1}|_{h,l}^2 \Big) \\ &= \frac{\rho}{2} \Big(|U_{2,N}|_{h,l}^2 - |\boldsymbol{v}_h^0|_{h,l}^2 \Big), \end{aligned}$$

where in the last equality we used the continuity of U_2.

4. With (4.5), we can write the second term as

$$\int_0^T \tilde{\xi} a\big(U_1, A_h^l \mathcal{P}_k U_2\big)\,dt = \frac{1}{2} \sum_{n=1}^N \int_{I_n} \tilde{\xi} D_t |U_1|_{h,l+1}^2\,dt - \int_0^T \tilde{\xi} a(U_1, A_h^l R \boldsymbol{f}_1)\,dt.$$

Then we integrate by parts in the first term of the right hand side and use the facts that $\dot{\tilde{\xi}}(t) = -\beta(t)$ and $\tilde{\xi}(0) = 1$, to get

$$\begin{aligned}
\int_0^T \tilde{\xi} a\big(U_1, A_h^l \mathcal{P}_k U_2\big)\, dt
&= \frac{1}{2}\sum_{n=1}^{N} \Big(\tilde{\xi}(t_n)|U_{1,n}|_{h,l+1}^2 - \tilde{\xi}(t_{n-1})|U_{1,n-1}|_{h,l+1}^2\Big) \\
&\quad - \frac{1}{2}\sum_{n=1}^{N} \int_{I_n} \dot{\tilde{\xi}} |U_1|_{h,l+1}^2\, dt - \int_0^T \tilde{\xi} a(U_1, A_h^l R\boldsymbol{f}_1)\, dt \\
&= \frac{1}{2}\Big(\tilde{\xi}(T)|U_{1,N}|_{h,l+1}^2 - |\boldsymbol{u}_h^0|_{h,l+1}^2\Big) + \frac{1}{2}\int_0^T \beta |U_1|_{h,l+1}^2\, dt \\
&\quad - \int_0^T \tilde{\xi} a(R\boldsymbol{f}_1, A_h^l U_1)\, dt\,,
\end{aligned}$$

where we used the continuity of U_1.

5. Consider now the third term in the left hand side of (4.6). Using (4.5) and the fact that $\dot{U}_1(t) = D_t W(t,s)$ we have

$$\begin{aligned}
\int_0^T \int_0^t \beta(t-s) a\big(W_1(t,s), A_h^l \mathcal{P}_k U_2\big)\, ds\, dt
&= \frac{1}{2}\int_0^T \int_0^t \beta(t-s) D_t |W_1(t,s)|_{h,l+1}^2 ds\, dt \\
&\quad - \int_0^T \int_0^t \beta(t-s) a\big(A_h^l W_1(t,s), R\boldsymbol{f}_1(t)\big)\, ds\, dt.
\end{aligned}$$

The first term on the right hand side is non-negative. To prove this, take $\epsilon \in (0,T)$. Then

(4.7)
$$\begin{aligned}
\int_\epsilon^T \int_0^{t-\epsilon} & \beta(t-s) D_t |W_1(t,s)|_{h,l+1}^2\, ds\, dt \\
&= \int_0^{T-\epsilon} \int_{s+\epsilon}^T \beta(t-s) D_t |W_1(t,s)|_{h,l+1}^2 dt\, ds \\
&= \int_0^{T-\epsilon} \beta(T-s)|W_1(T,s)|_{h,l+1}^2\, ds - \int_0^{T-\epsilon} \beta(\epsilon)|W_1(s+\epsilon,s)|_{h,l+1}^2\, ds \\
&\quad - \int_0^{T-\epsilon} \int_{s+\epsilon}^T \dot{\beta}(t-s)|W_1(t,s)|_{h,l+1}^2 dt\, ds \\
&\geq -\beta(\epsilon) \int_0^{T-\epsilon} |W_1(s+\epsilon,s)|_{h,l+1}^2\, ds\,,
\end{aligned}$$

where we used the facts that $\dot{\beta}(t) \leq 0$ and $\beta(t) \geq 0$ for the last inequality. Then using $W_1(s+\epsilon, s) = U_1(s+\epsilon) - U_1(s) = \int_s^{s+\epsilon} \dot{U}_1(t)\,dt$ we get

$$|W_1(s+\epsilon, s)|_{h,l+1}^2 \leq \Big(\int_s^{s+\epsilon} |\dot{U}_1(t)|_{h,l+1}\,dt \Big)^2 \leq \epsilon^2 \max_{0\leq t\leq T} |\dot{U}_1(t)|_{h,l+1}^2.$$

So (4.7) yields

$$\int_\epsilon^T \int_0^{t-\epsilon} \beta(t-s) D_t |W_1(t,s)|_{h,l+1}^2 \,ds\,dt \geq -\epsilon^2 \beta(\epsilon) \max_{0\leq t\leq T} |\dot{U}_1(t)|_{h,l+1}^2.$$

Therefore, for a fixed mesh, we let $\epsilon \to 0$ and conclude

$$\int_0^T \int_0^t \beta(t-s) D_t |W_1(t,s)|_{h,l+1}^2 \,ds\,dt \geq 0\,.$$

6. Putting the results from steps 3, 4, and 5 into (4.6) completes the proof. □

Remark 1. In [4] the auxiliary function $\boldsymbol{w}(t,s) = \boldsymbol{u}(t) - \boldsymbol{u}(t-s)$ was used in the same way as in our §2, to obtain stability estimates for the spatially semidiscrete finite element method. This does not work here because $s \mapsto U_1(t) - U_1(t-s)$ does not belong to $W^{(1)}$ if the temporal mesh is non-uniform.

We use (4.3) to obtain a stability estimate to be used in the error analysis. To this end, from (4.3), we have

$$\begin{aligned} \rho |U_{2,N}|_{h,l}^2 + \tilde{\xi}(T) |U_{1,N}|_{h,l+1}^2 \leq \rho |\boldsymbol{v}_h^0|_{h,l}^2 + |\boldsymbol{u}_h^0|_{h,l+1}^2 &+ 2\int_0^T (A_h^l P \boldsymbol{f}_2, U_2)\,dt \\ &+ 2\int_0^T a(A_h^l R \boldsymbol{f}_1, U_1)\,dt \\ &+ 2\int_0^T \int_0^t \beta(t-s) a\big(A_h^l R \boldsymbol{f}_1(t), W_1(t,s)\big)\,ds\,dt. \end{aligned}$$

Therefore using (3.4), $1-\gamma<\tilde{\xi}(t)\le 1$ and $\int_0^t\beta(s)\,ds\le\gamma$, we get

$$
\begin{aligned}
&\rho|U_{2,N}|_{h,l}^2+(1-\gamma)|U_{1,N}|_{h,l+1}^2\\
&\quad\le \rho|\boldsymbol{v}_h^0|_{h,l}^2+|\boldsymbol{u}_h^0|_{h,l+1}^2+2\int_0^T(A_h^{l/2}\mathcal{P}_k\mathcal{P}_h\boldsymbol{f}_2,A_h^{l/2}U_2)\,dt\\
&\qquad+2\int_0^T a(A_h^{l/2}\mathcal{P}_k\mathcal{R}_h\boldsymbol{f}_1,A_h^{l/2}U_1)\,dt\\
&\qquad+2\int_0^T\int_0^t\beta(t-s)a\big(A_h^{l/2}\mathcal{P}_k\mathcal{R}_h\boldsymbol{f}_1(t),A_h^{l/2}W_1(t,s)\big)\,ds\,dt\\
&\quad\le \rho|\boldsymbol{v}_h^0|_{h,l}^2+|\boldsymbol{u}_h^0|_{h,l+1}^2+2\int_0^T|\mathcal{P}_k\mathcal{P}_h\boldsymbol{f}_2|_{h,l}|U_2|_{h,l}\,dt\\
&\qquad+2\int_0^T|\mathcal{P}_k\mathcal{R}_h\boldsymbol{f}_1|_{h,l+1}|U_1|_{h,l+1}\,dt\\
&\qquad+2\gamma\int_0^T|\mathcal{P}_k\mathcal{R}_h\boldsymbol{f}_1(t)|_{h,l+1}\max_{0\le s\le T}|W_1(t,s)|_{h,l+1}\,dt\\
&\quad\le \rho|\boldsymbol{v}_h^0|_{h,l}^2+|\boldsymbol{u}_h^0|_{h,l+1}^2+\frac{1}{2}\rho\max_{[0,T]}|U_2|_{h,l}^2+C\Big(\int_0^T|\mathcal{P}_k\mathcal{P}_h\boldsymbol{f}_2|_{h,l}\,dt\Big)^2\\
&\qquad+\frac{1}{2}(1-\gamma)\max_{[0,T]}|U_1|_{h,l+1}^2+C\Big(\int_0^T|\mathcal{P}_k\mathcal{R}_h\boldsymbol{f}_1|_{h,l+1}\,dt\Big)^2
\end{aligned}
$$

where $C=C(\rho,\gamma)$. Using that, for piecewise linear functions, we have

$$\max_{[0,T]}|U_i|\le\max_{[0,T]}|U_{i,n}|,$$

and

$$\int_0^T|\mathcal{P}_k\boldsymbol{f}|\,dt\le\int_0^T|\boldsymbol{f}|\,dt,$$

and that the above inequality holds for arbitrary N, we conclude in a standard way

$$
\begin{aligned}
&|U_{2,N}|_{h,l}+|U_{1,N}|_{h,l+1}\\
&\quad\le C\bigg(|\boldsymbol{v}_h^0|_{h,l}+|\boldsymbol{u}_h^0|_{h,l+1}+\int_0^T\Big(|\mathcal{R}_h\boldsymbol{f}_1|_{h,l+1}+|\mathcal{P}_h\boldsymbol{f}_2|_{h,l}\Big)\,dt\bigg),
\end{aligned}
\tag{4.8}
$$

with $C=C(\rho,\gamma)$.

5. A priori error estimates

To simplify the notation we denote the Sobolev norms $\|\cdot\|_{H^i(\Omega)}$ by $\|\cdot\|_i$. We define the standard interpolant $I_k\boldsymbol{v}\in W^{(1)}$ by

$$I_k\boldsymbol{v}(t_n)=\boldsymbol{v}(t_n),\quad n=0,1,\cdots,N. \tag{5.1}$$

By standard arguments in approximation theory we see that, for $q = 0,\ 1$,

$$\int_0^T \|I_k \boldsymbol{v} - \boldsymbol{v}\|_i \, dt \leq C k^{q+1} \int_0^T \|D_t^{q+1} \boldsymbol{v}\|_i \, dt, \quad \text{for } i = 0,\ 1,\ 2, \tag{5.2}$$

where $k = \max_{1 \leq n \leq N} k_n$.

We assume the elliptic regularity estimate $\|\boldsymbol{v}\|_2 \leq C\|A\boldsymbol{v}\|,\ \forall\, \boldsymbol{v} \in \mathcal{D}(A)$, so that the following error estimates for the Ritz projection (3.2), hold true

$$\|\mathcal{R}_h v - v\| \leq C h^s \|v\|_s\,, \quad \forall \boldsymbol{v} \in H^s \cap V, \quad s = 1,\ 2. \tag{5.3}$$

Hence, as in §2.2, we must specialize to the pure Dirichlet boundary condition and a convex polygonal domain. We note that the energy norm $\|\cdot\|_V$ is equinalent to $\|\cdot\|_1$ on V.

Theorem 4. *Assume that $\Gamma_N = \varnothing$, Ω is a convex polygonal domain, and $V_{n-1} \subset V_n$, $n = 2, \dots, N$. Let $\boldsymbol{u}$ and U be the solutions of* (1.7) *and* (4.1)*. Then, with $\boldsymbol{e} = U - \boldsymbol{u}$ and $C = C(\rho, \gamma)$, we have*

$$\begin{aligned}
\|\boldsymbol{e}_{2,N}\| &\leq C h^2 \Big(\|\boldsymbol{v}^0\|_2 + \|\boldsymbol{u}_{2,N}\|_2 + \int_0^T \|\dot{\boldsymbol{u}}_2\|_2 \, dt \Big) \\
&\quad + C k^2 \int_0^T \big(\|\ddot{\boldsymbol{u}}_2\|_1 + \|\ddot{\boldsymbol{u}}_1\|_2 \big) \, dt, \\
\|\boldsymbol{e}_{1,N}\|_1 &\leq C h \Big(\|\boldsymbol{u}_{1,N}\|_2 + \|\boldsymbol{v}^0\|_1 + \int_0^T \|\dot{\boldsymbol{u}}_2\|_1 \, dt \Big) \\
&\quad + C k^2 \int_0^T \big(\|\ddot{\boldsymbol{u}}_2\|_1 + \|\ddot{\boldsymbol{u}}_1\|_2 \big) \, dt, \\
\|\boldsymbol{e}_{1,N}\| &\leq C h^2 \Big(\|\boldsymbol{u}_{1,N}\|_2 + \int_0^T \|\boldsymbol{u}_2\|_2 \, dt \Big) + C k^2 \int_0^T \big(\|\ddot{\boldsymbol{u}}_2\| + \|\ddot{\boldsymbol{u}}_1\|_1 \big) \, dt.
\end{aligned}$$

Proof. We set

$$\boldsymbol{e} = \boldsymbol{\theta} + \boldsymbol{\eta} + \boldsymbol{\rho} = (U - \pi \boldsymbol{u}) + (\pi \boldsymbol{u} - J\boldsymbol{u}) + (J\boldsymbol{u} - \boldsymbol{u}), \tag{5.4}$$

for some suitable operators π and J which will be specified in terms of the time interpolant I_k in (5.1) and projectors $\mathcal{R}_h$ and $\mathcal{P}_h$ in (3.2), so that $\pi \boldsymbol{u} \in W^{(1)}$ and $\boldsymbol{\eta}$ and $\boldsymbol{\rho}$ will correspond to the temporal and spatial errors, respectively. Due to (5.2)–(5.3) we just need to estimate $\boldsymbol{\theta}$. To this end, using the Galerkin orthogonality (3.7) and the definition of $\boldsymbol{\theta}$, we get

(5.5)
$$\begin{aligned}
B(\boldsymbol{\theta},\mathcal{P}_kV) &= -B(\boldsymbol{\eta},\mathcal{P}_kV) - B(\boldsymbol{\rho},\mathcal{P}_kV)\\
&= \int_0^T a(\boldsymbol{\eta}_2,\mathcal{P}_kV_1) - a(\dot{\boldsymbol{\eta}}_1,\mathcal{P}_kV_1) - \rho(\dot{\boldsymbol{\eta}}_2,\mathcal{P}_kV_2) - a(\boldsymbol{\eta}_1,\mathcal{P}_kV_2)\,dt\\
&\quad + \int_0^T\int_0^t \beta(t-s)a\big(\boldsymbol{\eta}_1(s),\mathcal{P}_kV_2(t)\big)\,ds\,dt\\
&\quad + \int_0^T a(\boldsymbol{\rho}_2,\mathcal{P}_kV_1) - a(\dot{\boldsymbol{\rho}}_1,\mathcal{P}_kV_1) - \rho(\dot{\boldsymbol{\rho}}_2,\mathcal{P}_kV_2) - a(\boldsymbol{\rho}_1,\mathcal{P}_kV_2)\,dt\\
&\quad + \int_0^T\int_0^t \beta(t-s)a\big(\boldsymbol{\rho}_1(s),\mathcal{P}_kV_2(t)\big)\,ds\,dt\\
&= \sum_{j=1}^{10} E_j, \quad \forall\, V \in \big(W^{(1)}\big)^2.
\end{aligned}$$

We consider two different choices of the operators π and J. In order to prove the first two error estimates we set, for $i = 1,\ 2$,

$$\boldsymbol{\theta}_i = U_i - I_k\mathcal{R}_h\boldsymbol{u}_i, \quad \boldsymbol{\eta}_i = (I_k - I)\mathcal{R}_h\boldsymbol{u}_i, \quad \boldsymbol{\rho}_i = (\mathcal{R}_h - I)\boldsymbol{u}_i.$$

Integrating by parts in E_2 and E_3 with respect to time and using (5.1) we have for both cases

$$E_2 = E_3 = 0. \tag{5.6}$$

Moreover, by the definitions of $\boldsymbol{\eta}$ and $\boldsymbol{\rho}$, we have

$$E_6 = E_7 = E_9 = E_{10} = 0.$$

Therefore,

$$\begin{aligned}
B(\boldsymbol{\theta},\mathcal{P}_kV) &= \int_0^T a(\boldsymbol{\eta}_2,\mathcal{P}_kV_1)\,dt\\
&\quad + \int_0^T \Big(a\big(-\boldsymbol{\eta}_1 + \int_0^t \beta(t-s)\boldsymbol{\eta}_1(s)ds,\mathcal{P}_kV_2\big) - \rho\big(\dot{\boldsymbol{\rho}}_2,\mathcal{P}_kV_2\big)\Big)\,dt\\
&= \hat{L}(\mathcal{P}_kV), \quad \forall\, V \in \big(W^{(1)}\big)^2,
\end{aligned}$$

which is of the form (4.2) with $\boldsymbol{f}_1 = \boldsymbol{\eta}_2$, $\boldsymbol{f}_2 = A_h\big(-\boldsymbol{\eta}_1 + \int_0^t \beta(t-s)\boldsymbol{\eta}_1(s)\,ds\big) - \rho\dot{\boldsymbol{\rho}}_2$.

Applying the stability inequality (4.8) with $l = 0$, and considering the fact that $|\cdot|_{0,h} = \|\cdot\|$ and $|\cdot|_{h,1} = \|\cdot\|_1$, we have

$$
\begin{aligned}
&\|\boldsymbol{\theta}_{2,N}\| + \|\boldsymbol{\theta}_{1,N}\|_1 \\
&\qquad \leq C\Big(\|\boldsymbol{\theta}_2(0)\| + \|\boldsymbol{\theta}_1(0)\|_1\Big) + C\int_0^T \|\mathcal{R}_h\boldsymbol{\eta}_2\|_1\,dt \\
&\qquad\quad + C\int_0^T \Big(\|\mathcal{P}_hA_h\boldsymbol{\eta}_1\| + \Big\|\mathcal{P}_hA_h\int_0^t \beta(t-s)\boldsymbol{\eta}_1(s)\,ds\Big\| + \rho\|\mathcal{P}_h\dot{\boldsymbol{\rho}}_2\|\Big)\,dt,
\end{aligned}
$$

where $\boldsymbol{\theta}_1(0) = 0$, since $U_1(0) = \mathcal{R}_{h,1}\boldsymbol{u}^0$. Since $\|\mathcal{R}_h\boldsymbol{v}\|_1 \leq C\|\boldsymbol{v}\|_1$, $\|\mathcal{P}_h\boldsymbol{v}\| \leq \|\boldsymbol{v}\|$, $\forall \boldsymbol{v} \in V$ and $A_h\mathcal{R}_h = \mathcal{P}_hA$, we have

$$
\begin{aligned}
\|\mathcal{R}_h\boldsymbol{\eta}_2\|_1 &= \|(I_k - I)\mathcal{R}_h\boldsymbol{u}_2\|_1 \leq C\|(I_k - I)\boldsymbol{u}_2\|_1, \\
\|\mathcal{P}_hA_h\boldsymbol{\eta}_1\| &= \|A_h\boldsymbol{\eta}_1\| = \|(I_k - I)A_h\mathcal{R}_h\boldsymbol{u}_1\| = \|(I_k - I)\mathcal{P}_hA\boldsymbol{u}_1\| \\
&\leq \|(I_k - I)A\boldsymbol{u}_1\| \leq C\|(I_k - I)\boldsymbol{u}_1\|_2,
\end{aligned}
$$

and

$$
\begin{aligned}
\int_0^T \Big\|\mathcal{P}_hA_h\int_0^t \beta(t-s)\boldsymbol{\eta}_1(s)\,ds\Big\|\,dt &\leq \int_0^T \Big\|A_h\int_0^t \beta(t-s)\boldsymbol{\eta}_1(s)\,ds\Big\|\,dt \\
&\leq C\int_0^T\int_0^t \beta(t-s)\|(I_k - I)\boldsymbol{u}_1(s)\|_2\,ds\,dt \\
&\leq C\gamma\int_0^T \|(I_k - I)\boldsymbol{u}_1\|_2\,dt.
\end{aligned}
$$

Therefore by $\boldsymbol{\theta} = \boldsymbol{e} - \boldsymbol{\eta} - \boldsymbol{\rho}$, $\boldsymbol{\eta}(t_n) = 0$ and $\boldsymbol{\theta}_1(0) = 0$, we get

$$
\begin{aligned}
\|\boldsymbol{e}_{2,N}\| &\leq \|\boldsymbol{\rho}_{2,N}\| + C\|\boldsymbol{\theta}_2(0)\| \\
&\quad + C\int_0^T \Big(\|(I_k - I)\boldsymbol{u}_2\|_1 + \|(I_k - I)\boldsymbol{u}_1\|_2 + \|(\mathcal{R}_h - I)\dot{\boldsymbol{u}}_2\|\Big)\,dt, \\
\|\boldsymbol{e}_{1,N}\|_1 &\leq \|\boldsymbol{\rho}_{1,N}\|_1 + C\|\boldsymbol{\theta}_2(0)\| \\
&\quad + C\int_0^T \Big(\|(I_k - I)\boldsymbol{u}_2\|_1 + \|(I_k - I)\boldsymbol{u}_1\|_2 + \|(\mathcal{R}_h - I)\dot{\boldsymbol{u}}_2\|\Big)\,dt,
\end{aligned}
$$

which implies the first two estimates by (5.2) and (5.3).

Finally, we choose

$$
\begin{aligned}
\boldsymbol{\theta}_1 &= U_1 - I_k\mathcal{R}_h\boldsymbol{u}_1, \quad \boldsymbol{\eta}_1 = (I_k - I)\mathcal{R}_h\boldsymbol{u}_1, \quad \boldsymbol{\rho}_1 = (\mathcal{R}_h - I)\boldsymbol{u}_1, \\
\boldsymbol{\theta}_2 &= U_2 - I_k\mathcal{P}_h\boldsymbol{u}_2, \quad \boldsymbol{\eta}_2 = (I_k - I)\mathcal{P}_h\boldsymbol{u}_2, \quad \boldsymbol{\rho}_2 = (\mathcal{P}_h - I)\boldsymbol{u}_2.
\end{aligned}
$$

By the definitions of $\mathcal{R}_h$ and $\mathcal{P}_h$ in (3.2), this implies

$$E_7 = E_8 = E_9 = E_{10} = 0,$$

and we still have (5.6). Therefore, (5.5) becomes

$$\begin{aligned} B(\boldsymbol{\theta}, \mathcal{P}_k V) &= \int_0^T a(\boldsymbol{\eta}_2 + \boldsymbol{\rho}_2, \mathcal{P}_k V_1)\, dt \\ &\quad + \int_0^T a\big(-\boldsymbol{\eta}_1 + \int_0^t \beta(t-s)\boldsymbol{\eta}_1(s) ds, \mathcal{P}_k V_2\big)\, dt \\ &= \hat{L}(\mathcal{P}_k V), \quad \forall V \in \big(W^{(1)}\big)^2, \end{aligned}$$

which is of the form (4.2) with $\boldsymbol{f}_1 = \boldsymbol{\eta}_2 + \boldsymbol{\rho}_2$, $\boldsymbol{f}_2 = A_h\big(-\boldsymbol{\eta}_1 + \int_0^t \beta(t - s)\boldsymbol{\eta}_1(s)\, ds\big)$.

Again applying the stability inequality (4.8), this time with $l = -1$, and using $|\cdot|_{h,0} = \|\cdot\|$, we have

$$\begin{aligned} \|\boldsymbol{\theta}_{1,N}\| &\leq C \int_0^T \Big(\|\mathcal{R}_h \boldsymbol{\eta}_2\| + \|\mathcal{R}_h \boldsymbol{\rho}_2\| \Big)\, dt \\ &\quad + C \int_0^T \Big(|\mathcal{P}_h A_h \boldsymbol{\eta}_1|_{h,-1} + |\mathcal{P}_h A_h \int_0^t \beta(t-s)\boldsymbol{\eta}_1(s) ds|_{h,-1} \Big)\, dt, \end{aligned}$$

where we used that $\boldsymbol{\theta}(0) = 0$, since $U_1(0) = \mathcal{R}_{h,1}\boldsymbol{u}^0$ and $U_2(0) = \mathcal{P}_{h,1}\boldsymbol{v}^0$. Then, since

$$\begin{aligned} \|\mathcal{R}_h \boldsymbol{\eta}_2\| &= \|(I_k - I)\mathcal{P}_h \boldsymbol{u}_2\| \leq \|(I_k - I)\boldsymbol{u}_2\|, \\ \|\mathcal{R}_h \boldsymbol{\rho}_2\| &= \|\mathcal{P}_h (I - \mathcal{R}_h)\boldsymbol{u}_2\| \leq \|(\mathcal{R}_h - I)\boldsymbol{u}_2\|, \\ |\mathcal{P}_h A_h \boldsymbol{\eta}_1|_{h,-1} &= |A_h \mathcal{R}_h (I_k - I)\boldsymbol{u}_1|_{h,-1} = |\mathcal{R}_h (I_k - I)\boldsymbol{u}_1|_{h,1} \\ &\leq C\|(I_k - I)\boldsymbol{u}_1\|_1\,, \end{aligned}$$

and

$$\begin{aligned} \int_0^T \Big| \mathcal{P}_h A_h \int_0^t \beta(t-s)\boldsymbol{\eta}_1(s)\, ds \Big|_{h,-1} dt & \\ &\leq C \int_0^T \int_0^t \beta(t-s)\|(I_k - I)\boldsymbol{u}_1(s)\|_1\, ds\, dt \\ &\leq C\,\gamma \int_0^T \|(I_k - I)\boldsymbol{u}_1\|_1\, dt, \end{aligned}$$

we conclude

$$\begin{aligned} \|\boldsymbol{e}_{1,N}\| &\leq \|\boldsymbol{\rho}_{1,N}\| \\ &\quad + C \int_0^T \Big(\|(I_k - I)\boldsymbol{u}_2\| + \|(\mathcal{R}_h - I)\boldsymbol{u}_2\| + \|(I_k - I)\boldsymbol{u}_1\|_1 \Big)\, dt, \end{aligned}$$

which implies the last estimate by (5.2) and (5.3). □

6. Numerical example

In this section we illustrate the numerical method by solving a simple but realistic example for a two dimensional structure, see Figure 1 (a), using piecewise linear polynomials. This shows that the model captures the mechanical behaviour of the material.

We consider the initial conditions: $\boldsymbol{u}(x,0) = 0\,\mathrm{m}$, $\dot{\boldsymbol{u}}(x,0) = 0\,\mathrm{m/s}$, the boundary conditions: $\boldsymbol{u} = 0$ at $x = 0$, $\boldsymbol{g} = (0,-1)\,\mathrm{Pa}$ at $x = 1.5$ and zero on the rest of the boundary. The volume load is assumed to be $\boldsymbol{f} = 0\,\mathrm{N/m^3}$. And the model parameters are: $\gamma = 0.5$, $\tau = 0.25$, $\nu = 0.3$, $E = 5\mathrm{MPa}$ and $\rho = 7000\,\mathrm{kg/m^3}$. The deformed mesh at $t/\tau = 9$ for $\alpha = 1/2$ is displayed in Figure 1 (b), with the displacement magnified by the factor 10^5, and the computed vertical displacement at the point $(1.5, 1.5)$ for different α is shown in Figure 2. We note that for small α there is less damping, that is what we expect, since in the limit $\alpha = 0$ there is no convolution term in the model. While at the other limit $\alpha = 1$ we expect strong damping, since the kernel β with $\alpha = 1$ is an exponential function, see (1.3).

We also verify numerically the temporal rate of convergence $O(k^2)$ for $\|e_{1,N}\|$. In the lack of an explicit solution we compare with a numerical solution with fine mesh sizes h, k. Here we consider $h = 0.0223$, $k_{\min} = 0.0266$, $\alpha = 1/2$, $\tau = 1/4$. The result is displayed in Figure 3.

Acknowledgment. We would like thank the anonymous referees for constructive comments that helped us improve the manuscript.

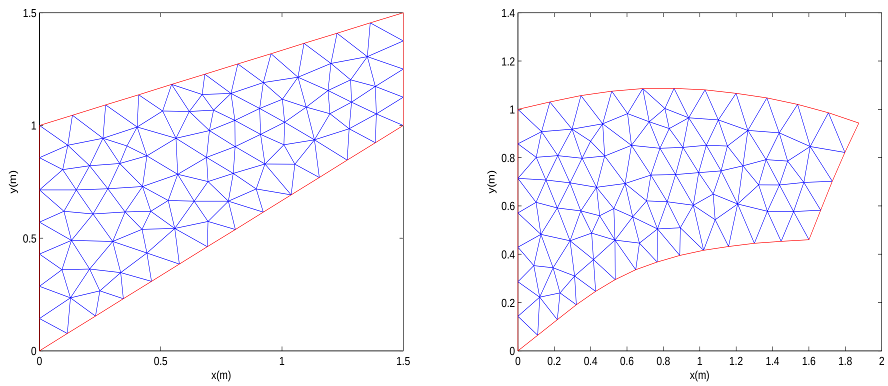


FIGURE 1. (a) Undeformed mesh. (b) Deformed mesh at $t/\tau = 9$ for $\alpha = 1/2$.

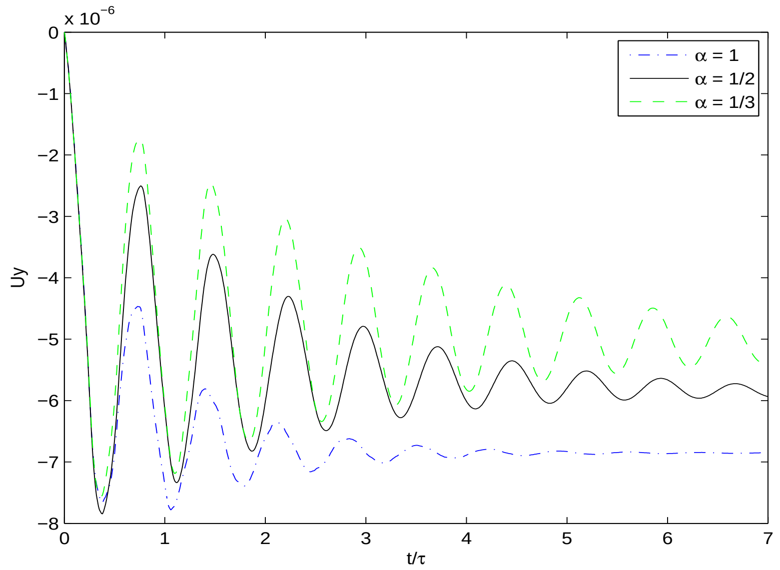


FIGURE 2. Vertical displacement for different α.

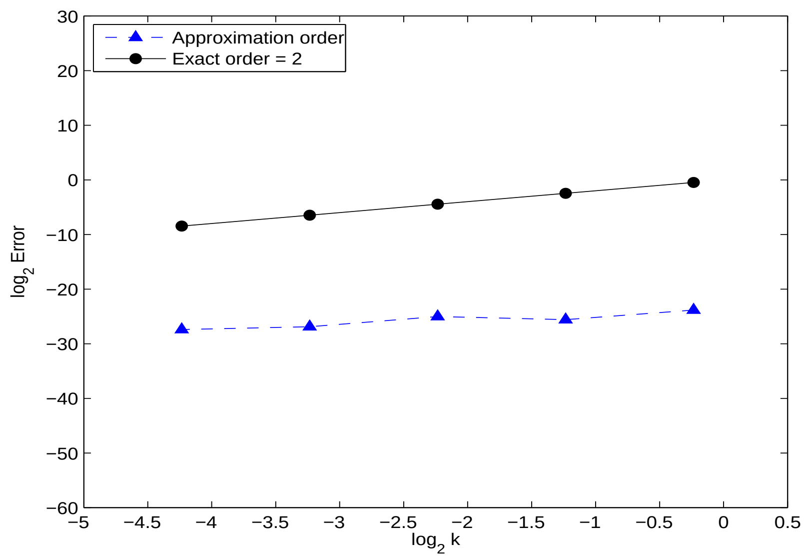


FIGURE 3. Convergence order for time discretization.

References

1. K. Adolfsson, M. Enelund, and S. Larsson, *Adaptive discretization of an integro-differential equation with a weakly singular convolution kernel*, Comput. Methods Appl. Mech. Engrg. **192** (2003), 5285–5304.
2. ———, *Adaptive discretization of fractional order viscoelasticity using sparse time history*, Comput. Methods Appl. Mech. Engrg. **193** (2004), 4567–4590.
3. ———, *Space-time discretization of an integro-differential equation modeling quasi-static fractional order viscoelasticity*, J. Vib. Control **14** (2008), 1631–1649.
4. K. Adolfsson, M. Enelund, S. Larsson, and M. Racheva, *Discretization of integro-differential equations modeling dynamic fractional order viscoelasticity*, LNCS **3743** (2006), 76–83.
5. R. L. Bagley and P. J. Torvik, *Fractional calculus—a different approach to the analysis of viscoelastically damped structures*, AIAA J. **21** (1983), 741–748.
6. R. H. Fabiano and K. Ito, *Semigroup theory and numerical approximation for equations in linear viscoelasticity*, SIAM J. Math. Anal. **21** (1990), 374–393.
7. W. McLean and K. Mustapha, *A second-order accurate numerical method for a fractional wave equation*, Numer. Math. **105** (2007), 481–510.
8. W. McLean and V. Thomée, *Numerical solution of an evolution equation with a positive-type memory term*, J. Austral. Math. Soc. Ser. B **35** (1993), 23–70.
9. W. McLean, V. Thomée, and L. B. Wahlbin, *Discretization with variable time steps of an evolution equation with a positive-type memory term*, J. Comput. Appl. Math. **69** (1996), 49–69.
10. A. K. Pani, V. Thomée, and L. B. Wahlbin, *Numerical methods for hyperbolic and parabolic integro-differential equations*, J. Integral Equations Appl. **4** (1992), 533–584.
11. A. Pazy, *Semigroups of Linear Operators and Applications to Partial Differential Equations*, Springer-Verlag, 1983.
12. B. Rivière, S. Shaw, and J. R. Whiteman, *Discontinuous Galerkin finite element methods for dynamic linear solid viscoelasticity problems*, Numer. Methods Partial Differential Equations **23** (2007), 1149–1166.
13. A. Schädle, M. López-Fernández, and Ch. Lubich, *Fast and oblivious convolution quadrature*, SIAM J. Sci. Comput. **28** (2006), 421–438.
14. ———, *Adaptive, fast, and oblivious convolution in evolution equations with memory*, SIAM J. Sci. Comput. **30** (2008), 1015–1037.
15. S. Shaw and J. R. Whiteman, *A posteriori error estimates for space-time finite element approximation of quasistatic hereditary linear viscoelasticity problems*, Comput. Methods Appl. Mech. Engrg. **193** (2004), 5551–5572.
16. I. H. Sloan and V. Thomée, *Time discretization of an integro-differential equation of parabolic type*, SIAM J. Numer. Anal. **23** (1986), 1052–1061.

Department of Mathematical Sciences, Chalmers University of Technology and University of Gothenburg, SE–412 96 Göteborg, Sweden
E-mail address: `stig@chalmers.se`

Department of Mathematical Sciences, Chalmers University of Technology and University of Gothenburg, SE–412 96 Göteborg, Sweden
E-mail address: `fardin@chalmers.se`

EXISTENCE AND UNIQUENESS OF THE SOLUTION OF AN INTEGRO-DIFFERENTIAL EQUATION WITH WEAKLY SINGULAR KERNEL

FARDIN SAEDPANAH

Abstract. A hyperbolic type integro-differential equation with weakly singular kernel is considered together with mixed homogeneous Dirichlet and non-homogeneous Neumann bounadry conditions. Existence and uniqueness of the solution is proved by means of Galerkin method. Regularity estimates are proved and the limitations of the regularity is discussed.

1. Introduction

We study a model problem, which is a hyperbolic type integro-differential equation with weakly singular kernel. This problem arises as a model for fractional order viscoelasticity. The fractional order viscoelastic model, that is, the linear viscoelastic model with fractional order operators in the constitutive equations, is capable of describing the behavior of many viscoelastic materials by using only a few parameters.

There is an extensive literature regarding well-posedness and numerical treatment for integro-differential equations, see, e.g., [1], [3], [9], [10], [11], [12], [13], [14], [15], [16], and [17]. Existence, uniqueness and regularity of a reformed model has been studied in [12] by means of Fourier series. One may also see [6], where the theory of analytic semigroups is used in terms of interpolation spaces. An abstract Volterra equation, as an abstract model for equations of linear viscoelasticity, has been studied in [4]. In a previous work [9], well-posedness and regularity of the model problem was studied in the framework of the semigroup of linear operators. The drawback of the framework is that this does not admit non-homogeneous Neumann boundary condition. While in practice mixed homogeneous Dirichlet and non-homogeneous Neumann boundary conditions are of special interest. Here

Date: April 14, 2009.

1991 *Mathematics Subject Classification.* 65M60, 45K05.

Key words and phrases. integro-differential equation, fractional order viscoelasticity, Galerkin approximation, weakly singular kernel, regularity, a priori estimate.

we investigate existence, uniqueness and regularity of the solution of the problem (2.5) by means of the Galerkin approximation method.

In the sequel, in §2 we describe the construction of the model and we define a weak (generalized) solution. Then in §3 we study well-posedness, regularity and limitations for higher regularity.

2. The model problem and weak formulation

Let σ_{ij}, ϵ_{ij} and u_i denote, respectively, the usual stress tensor, strain tensor and displacement vector. We recall that the linear strain tensor is defined by

$$\epsilon_{ij} = \frac{1}{2}\Big(\frac{\partial u_i}{\partial x_j} + \frac{\partial u_j}{\partial x_i}\Big).$$

With the decompositions

$$s_{ij} = \sigma_{ij} - \tfrac{1}{3}\sigma_{kk}\delta_{ij}, \quad e_{ij} = \epsilon_{ij} - \tfrac{1}{3}\epsilon_{kk}\delta_{ij},$$

the constitutive equations are formulated in [2] as

$$\begin{aligned} s_{ij}(t) + \tau_1^{\alpha_1} D_t^{\alpha_1} s_{ij}(t) &= 2G_\infty e_{ij}(t) + 2G\tau_1^{\alpha_1} D_t^{\alpha_1} e_{ij}(t), \\ \sigma_{kk}(t) + \tau_2^{\alpha_2} D_t^{\alpha_2} \sigma_{kk}(t) &= 3K_\infty \epsilon_{kk}(t) + 3K\tau_2^{\alpha_2} D_t^{\alpha_2} \epsilon_{kk}(t), \end{aligned} \tag{2.1}$$

with initial conditions

$$s_{ij}(0+) = 2Ge_{ij}(0+), \quad \sigma_{kk}(0+) = 3K\epsilon_{kk}(0+),$$

meaning that the initial response follows Hooke's elastic law. Here G, K are the instantaneous (unrelaxed) moduli, and G_∞, K_∞ are the long-time (relaxed) moduli. Note that we have two relaxation times, $\tau_1, \tau_2 > 0$, and fractional orders of differentiation, α_1, $\alpha_2 \in (0,1)$, where the fractional order derivative is defined by

$$D_t^\alpha f(t) = D_t D_t^{-(1-\alpha)} f(t) = D_t \frac{1}{\Gamma(1-\alpha)} \int_0^t (t-s)^{-\alpha} f(s)\, ds.$$

The contitutive equations (2.1) can be solved for σ by means of Laplace transformation, [7]:

$$\begin{aligned} s_{ij}(t) &= 2G\Big(e_{ij}(t) - \frac{G-G_\infty}{G}\int_0^t \theta_1(t-s) e_{ij}(s)\, ds\Big), \\ \sigma_{kk}(t) &= 3K\Big(\epsilon_{kk}(t) - \frac{K-K_\infty}{K}\int_0^t \theta_2(t-s)\epsilon_{kk}(s)\, ds\Big), \end{aligned}$$

where

$$\theta_i(t) = -\frac{d}{dt}E_{\alpha_i}\Big(-\Big(\frac{t}{\tau_i}\Big)^{\alpha_i}\Big), \quad E_{\alpha_i}(t) = \sum_{n=0}^{\infty} \frac{t^n}{\Gamma(1+n\alpha_i)},$$

and E_{α_i} is the Mittag-Leffler function of order α_i. We make the simplifying assumption (synchronous viscoelasticity):

$$\alpha = \alpha_1 = \alpha_2, \quad \tau = \tau_1 = \tau_2, \quad f = f_1 = f_2.$$

Then we may define a parameter γ, a kernel β, and the Lamé constants μ, λ,

$$\gamma = \frac{G - G_\infty}{G} = \frac{K - K_\infty}{K}, \quad \beta(t) = \gamma\theta(t), \quad \mu = G, \quad \lambda = K - \tfrac{2}{3}G,$$

and the constitutive equations become

$$\begin{aligned}\sigma_{ij}(t) &= \Big(2\mu\epsilon_{ij}(t) + \lambda\epsilon_{kk}(t)\delta_{ij}\Big) - \int_0^t \beta(t-s)\Big(2\mu\epsilon_{ij}(s) + \lambda\epsilon_{kk}(s)\delta_{ij}\Big)\,ds,\\ &= (\sigma_0)_{ij}(t) - \int_0^t \beta(t-s)(\sigma_0)_{ij}(s)\,ds.\end{aligned}$$

Note that the viscoelastic part of the model contains only three parameters:

$$0 < \gamma < 1, \quad 0 < \alpha < 1, \quad \tau > 0.$$

The kernel is weakly singular:

$$\begin{aligned}\beta(t) &= -\gamma\frac{d}{dt}E_\alpha\Big(-\Big(\frac{t}{\tau}\Big)^\alpha\Big) = \gamma\frac{\alpha}{\tau}\Big(\frac{t}{\tau}\Big)^{-1+\alpha}E'_\alpha\Big(-\Big(\frac{t}{\tau}\Big)^\alpha\Big)\\ &\approx Ct^{-1+\alpha}, \quad t \to 0,\end{aligned} \tag{2.2}$$

and we note the properties

$$\begin{aligned}&\beta(t) \geq 0,\\ &\|\beta\|_{L_1(\mathbb{R}^+)} = \int_0^\infty \beta(t)\,dt = \gamma\Big(E_\alpha(0) - E_\alpha(\infty)\Big) = \gamma < 1.\end{aligned} \tag{2.3}$$

The equations of motion now become, (we denote time derivatives with '$\cdot$'):

$$\begin{aligned}\rho\ddot{u}_i - \sigma_{ij,j} &= f_i &&\text{in } \Omega,\\ u_i &= 0 &&\text{on } \Gamma_{\mathrm{D}},\\ \sigma_{ij}n_j &= g_i &&\text{on } \Gamma_{\mathrm{N}},\end{aligned} \tag{2.4}$$

where ρ is the (constant) mass density, and f and g represent, respectively, the volume and the surface loads. We let $\Omega \subset \mathbb{R}^d$, $d = 2, 3$, be a bounded polygonal domain with boundary $\Gamma_{\mathrm{D}} \cup \Gamma_{\mathrm{N}} = \partial\Omega$, $\Gamma_{\mathrm{D}} \cap \Gamma_{\mathrm{N}} = \varnothing$ and $\mathrm{meas}(\Gamma_{\mathrm{D}}) \neq 0$. We set

$$(Au)_i = -\big(2\mu\epsilon_{ij}(u) + \lambda\epsilon_{kk}(u)\delta_{ij}\big)_{,j}\,,$$

that is, $Au = -\nabla \cdot \sigma_0(u)$, and we write the equations of motion (2.4) in the form

$$
\begin{aligned}
&\rho\ddot{u}(x,t) + Au(x,t) \\
&\qquad - \int_0^t \beta(t-s)Au(x,s)\,ds = f(x,t) && \text{in } \Omega\times(0,T), \\
&u(x,t) = 0 && \text{on } \Gamma_{\mathrm{D}}\times(0,T), \\
&\sigma(u;x,t)\cdot n = g(x,t) && \text{on } \Gamma_{\mathrm{N}}\times(0,T), \\
&u(x,0) = u^0(x) && \text{in } \Omega, \\
&\dot{u}(x,0) = v^0(x) && \text{in } \Omega.
\end{aligned} \tag{2.5}
$$

We introduce the function spaces $H = L_2(\Omega)^d$, $H_{\Gamma_{\mathrm{N}}} = L_2(\Gamma_{\mathrm{N}})^d$, and $V = \{v \in H^1(\Omega)^d : v|_{\Gamma_{\mathrm{D}}} = 0\}$. We define the bilinear form (with the usual summation convention)

$$a(u,v) = \int_\Omega \big(2\mu\epsilon_{ij}(u)\epsilon_{ij}(v) + \lambda\epsilon_{ii}(u)\epsilon_{jj}(v)\big)\,dx, \quad \forall u,v\in V,$$

which is coercive on V. We denote the norms in H and $H_{\Gamma_{\mathrm{N}}}$ by $\|\cdot\|$ and $\|\cdot\|_{\Gamma_{\mathrm{N}}}$, respectively, and we equip V with the inner product $a(\cdot,\cdot)$ and norm $\|v\|_V^2 = a(v,v)$.

Now we define a weak solution to be a function $u = u(x,t)$ that satisfies

$$u \in L_2((0,T);V), \quad \dot{u} \in L_2((0,T);H), \quad \ddot{u} \in L_2((0,T);V^*), \tag{2.6}$$

$$
\begin{aligned}
\rho\langle\ddot{u}(t),v\rangle + a(u(t),v) &- \int_0^t \beta(t-s)a(u(s),v)\,ds \\
&= (f(t),v) + (g(t),v)_{\Gamma_{\mathrm{N}}}, \quad \forall v\in V, \text{ a.e. } t\in(0,T),
\end{aligned} \tag{2.7}
$$

$$u(0) = u^0, \quad \dot{u}(0) = v^0. \tag{2.8}$$

Here $(g(t),v)_{\Gamma_{\mathrm{N}}} = \int_{\Gamma_{\mathrm{N}}} g(t)\cdot v\,dS$, and $\langle\cdot,\cdot\rangle$ denotes the pairing of V^* and V. We note that (2.6) implies, by a classical result for Sobolev spaces, that $u \in \mathcal{C}([0,T];H)$, $\dot{u} \in \mathcal{C}([0,T];V^*)$ so that the initial conditions (2.8) make sense for $u^0 \in H$, $v^0 \in V^*$.

3. Existence, uniqueness and regularity

In this section we prove existence and uniqueness as well as regularity of a weak solution of (2.5) using Galerkin method, in a similar way as for hyperbolic PDE's in [8], [5]. To this end, we first introduce the Galerkin approximation of a weak solution of (2.5) in a classical way, and we obtain a priori estimates for approximate solutions. These will be used to construct a weak solution and then we will verify uniqueness as well as regularity.

We recall (2.2), (2.3) and we define the function

$$\xi(t) = \gamma - \int_0^t \beta(s)\,ds = \int_t^\infty \beta(s)\,ds = \gamma E_\alpha(t), \tag{3.1}$$

and it is easy to see that

$$D_t\xi(t) = -\beta(t) < 0, \quad \xi(0) = \gamma, \quad \lim_{t\to\infty} \xi(t) = 0, \quad 0 < \xi(t) \le \gamma. \tag{3.2}$$

Besides, ξ is a completely monotone function, that is,

$$(-1)^j D_t^j \xi(t) \ge 0, \quad t \in (0,\infty),\, j \in \mathbb{N},$$

since the Mittag-Leffler function E_α, $\alpha \in [0,1]$ is completely monotone. Consequently, an important property of ξ, is that it is a positive type kernel, that is, it is continuous and, for any $T \ge 0$, satisfies

$$\int_0^T \int_0^t \xi(t-s)\phi(t)\phi(s)\,ds\,dt \ge 0, \quad \forall \phi \in \mathcal{C}([0,T]). \tag{3.3}$$

3.1. **Galerkin approximations.** Let $\{(\lambda_j,\varphi_j)\}_{j=1}^\infty$ be the eigenpairs of the weak eigenvalue problem

$$a(\varphi, v) = \lambda(\varphi, v), \quad \forall v \in V. \tag{3.4}$$

It is known that $\{\varphi_j\}_{j=1}^\infty$ can be chosen to be an ON-basis in H and an orthogonal basis for V.

Now, for a fixed positive integer $m \in \mathbb{N}$, we seek a function of the form

$$u_m(t) = \sum_{j=1}^m d_j(t)\varphi_j \tag{3.5}$$

to satisfy

$$\begin{aligned} \rho(\ddot{u}_m(t),\varphi_k) + a(u_m(t),\varphi_k) - \int_0^t \beta(t-s)a(u_m(s),\varphi_k)\,ds \\ = (f(t),\varphi_k) + (g(t),\varphi_k)_{\Gamma_N}, \quad k = 1,\ldots,m,\, t \in (0,T), \end{aligned} \tag{3.6}$$

with initial conditions

$$u_m(0) = \sum_{j=1}^m (u^0,\varphi_j)\varphi_j, \quad \dot{u}_m(0) = \sum_{j=1}^m (v^0,\varphi_j)\varphi_j. \tag{3.7}$$

Theorem 1. *For each $m \in \mathbb{N}$, there exists a unique function u_m of the form* (3.5) *satisfying* (3.6)–(3.7)*. Moreover, if $u^0 \in V, v^0 \in H$, $f \in L_2((0,T);H)$, $g \in W_1^1((0,T);H^{\Gamma_N})$, there is a constant $C = C(\Omega,\gamma,\rho,T)$ such that,*

$$\begin{aligned} \|u_m\|_{L_\infty((0,T);V)} + \|\dot{u}_m\|_{L_\infty((0,T);H)} + \|\ddot{u}_m\|_{L_2((0,T);V^*)} \\ \le C\big\{\|u^0\|_V + \|v^0\| + \|g\|_{W_1^1((0,T);H^{\Gamma_N})} + \|f\|_{L_2((0,T);H)}\big\}. \end{aligned} \tag{3.8}$$

Proof. Using (3.5) and the fact that $\{\varphi_j\}_{j=1}^{\infty}$ is an ON-basis for H and a solution of the eigenvalue problem (3.4), we obtain from (3.6) that,

$$\rho\ddot{d}_k(t) + \lambda_k d_k(t) - \lambda_k(\beta * d_k)(t) = f_k(t) + g_k(t), \quad k = 1, \ldots, m, \; t \in (0,T), \tag{3.9}$$

where $*$ denotes the convolution, and $f_k(t) = (f(t), \varphi_k)$, $g_k(t) = (g(t), \varphi_k)_{\Gamma_{\mathrm{N}}}$. This is a system of second order ODE's with the initial conditions

$$d_k(0) = (u^0, \varphi_k), \quad \dot{d}_k(0) = (v^0, \varphi_k), \quad k = 1, \ldots, m. \tag{3.10}$$

The Laplace transform can be used, for example, to find the unique solution of the system. Indeed, the Laplace transform of the Mittag-Leffler function is,

$$\mathcal{L}(E_\alpha(ax^\alpha)) = \frac{s^{\alpha-1}}{s^\alpha - a}, \quad s > |a|^{1/\alpha}.$$

Hence for the kernel β defined in (2.2) we have,

$$\begin{aligned} \mathcal{L}(\beta(t)) &= -\gamma s \mathcal{L}(E_\alpha(-\tau^{-\alpha}t^\alpha)) + \gamma E_\alpha(0) \\ &= -\gamma s \frac{s^{\alpha-1}}{s^\alpha + \tau^{-\alpha}} + \gamma = \gamma - \gamma \frac{s^\alpha}{s^\alpha + \tau^{-\alpha}} = \frac{\gamma}{(\tau s)^\alpha + 1}. \end{aligned} \tag{3.11}$$

Then, taking the Laplace transform of (3.9) we get,

$$\begin{aligned} &\Big(\rho s^2 + \lambda_k - \lambda_k \mathcal{L}(\beta)(s)\Big)\mathcal{L}(d_k)(s) \\ &\qquad = \mathcal{L}(f_k)(s) + \mathcal{L}(g_k)(s) + \rho d_k(0)s + \rho\dot{d}_k(0), \end{aligned} \tag{3.12}$$

where the inverse Laplace transform is computable. Therefore, there is a unique solution for the system (3.9) with the initial conditions (3.10).

Now we prove the a priori estimate (3.8). Since $\beta(t-s) = D_s\xi(t-s)$, by (3.2), we can write (3.6), after partial integration, as

$$\begin{aligned} \rho(\ddot{u}_m(t), \varphi_k) + \tilde{\gamma} a(u_m(t), \varphi_k) + \int_0^t \xi(t-s) a(\dot{u}_m(s), \varphi_k)\, ds \\ = (f(t), \varphi_k) + (g(t), \varphi_k)_{\Gamma_{\mathrm{N}}} \\ - \xi(t) a(u_m(0), \varphi_k), \quad k = 1, \ldots, m, \; t \in (0,T), \end{aligned}$$

where $\tilde{\gamma} = 1 - \gamma$. Then multiplying by $\dot{d}_k(t)$ and summing $k = 1, \ldots, m$, we have,

$$\begin{aligned} \rho(\ddot{u}_m(t), \dot{u}_m(t)) + \tilde{\gamma} a(u_m(t), \dot{u}_m(t)) + \int_0^t \xi(t-s) a(\dot{u}_m(s), \dot{u}_m(t))\, ds \\ = (f(t), \dot{u}_m(t)) + (g(t), \dot{u}_m(t))_{\Gamma_{\mathrm{N}}} \\ - \xi(t) a(u_m(0), \dot{u}_m(t)), \quad t \in (0,T). \end{aligned}$$

Then integrating with respect to t, we have,

$$\begin{aligned}\rho\|\dot{u}_m(t)\|^2+&\tilde{\gamma}\|u_m(t)\|_V^2+2\int_0^t\int_0^r \xi(r-s)a(\dot{u}_m(s),\dot{u}_m(r))\,ds\,dr\\&=\rho\|\dot{u}_m(0)\|^2+\tilde{\gamma}\|u_m(0)\|_V^2\\&\quad+2\int_0^t (f(r),\dot{u}_m(r))\,dr+2\int_0^t (g(r),\dot{u}_m(r))_{\Gamma_{\mathrm{N}}}\,dr\\&\quad-2\int_0^t \xi(r)a(u_m(0),\dot{u}_m(r))\,dr.\end{aligned}$$

Since ξ is a positive type kernel, by (3.3), the third term of the left hand side is non-negative. Then integration by parts in the last two terms at the right side yields

$$\begin{aligned}\rho\|\dot{u}_m(t)\|^2&+\tilde{\gamma}\|u_m(t)\|_V^2\\&\leq \rho\|\dot{u}_m(0)\|^2+\tilde{\gamma}\|u_m(0)\|_V^2+2\int_0^t (f(r),\dot{u}_m(r))\,dr\\&\quad-2\int_0^t (\dot{g}(r),u_m(r))_{\Gamma_{\mathrm{N}}}\,dr+2(g(t),u_m(t))_{\Gamma_{\mathrm{N}}}-2(g(0),u_m(0))_{\Gamma_{\mathrm{N}}}\\&\quad-2\int_0^t \beta(r)a(u_m(0),u_m(r))\,dr\\&\quad-2\xi(t)a(u_m(0),u_m(t))+2\xi(0)a(u_m(0),u_m(0)),\end{aligned}$$

that using the Cauchy-Schwarz inequality, the trace theorem, $\|\beta\|_{L_1(\mathbb{R}^+)}=\gamma$, and $\xi(0)=\gamma$, we obtain

$$\begin{aligned}\rho\|\dot{u}_m(t)\|^2&+\tilde{\gamma}\|u_m(t)\|_V^2\\&\leq \rho\|\dot{u}_m(0)\|^2+\tilde{\gamma}\|u_m(0)\|_V^2\\&\quad+2/C_1\max_{0\leq r\leq t}\|\dot{u}_m(r)\|^2+C_1\Big(\int_0^t\|f(r)\|\,dr\Big)^2\\&\quad+2C_{Trace}/C_2\max_{0\leq r\leq t}\|u_m(r)\|_V^2+2C_{Trace}C_2\Big(\int_0^t\|\dot{g}(r)\|_{\Gamma_{\mathrm{N}}}\,dr\Big)^2\\&\quad+2C_{Trace}/C_3\|u_m(t)\|_V^2+2C_{Trace}C_3\|g(t)\|_{\Gamma_{\mathrm{N}}}^2\\&\quad+2C_{Trace}/C_4\|u_m(0)\|_V^2+2C_{Trace}C_4\|g(0)\|_{\Gamma_{\mathrm{N}}}^2\\&\quad+2/C_5\|u_m(0)\|_V^2+2\gamma^2C_5\max_{0\leq r\leq t}\|u_m(r)\|_V^2\\&\quad+2/C_6\|u_m(0)\|_V^2+2C_6\xi^2(t)\|u_m(t)\|_V^2+2\gamma\|u_m(0)\|_V^2.\end{aligned}$$

Hence, considering the facts that $C_{Trace} = C(\Omega)$, $\|\dot{u}_m(0)\| \leq \|v^0\|$, and $\|u_m(0)\|_V \leq \|u^0\|_V$, in a standard way we get,

$$\begin{aligned}\|\dot{u}_m\|^2_{L_\infty((0,T);H)} &+ \|u_m\|^2_{L_\infty((0,T);V)} \\ &\leq C\big\{\|v^0\|^2 + \|u^0\|^2_V + \|g\|^2_{L_\infty((0,T);H^{\Gamma_{\rm N}})} \\ &\qquad + \|f\|^2_{L_1((0,T);H)} + \|\dot{g}\|^2_{L_1((0,T);H^{\Gamma_{\rm N}})}\big\},\end{aligned}$$

for some constant $C = C(\Omega, \gamma, \rho, T)$. This, and the facts that $\|f\|_{L_1((0,T);H)} \leq C\|f\|_{L_2((0,T);H)}$, and by Sobolev's inequality

$$\|g\|_{L_\infty((0,T);H^{\Gamma_{\rm N}})} \leq C\|g\|_{W_1^1((0,T);H^{\Gamma_{\rm N}})},$$

implies

$$\begin{aligned}\|\dot{u}_m\|^2_{L_\infty((0,T);H)} &+ \|u_m\|^2_{L_\infty((0,T);V)} \\ &\leq C\big\{\|v^0\|^2 + \|u^0\|^2_V + \|g\|^2_{W_1^1((0,T);H^{\Gamma_{\rm N}})} + \|f\|^2_{L_2((0,T);H)}\big\}.\end{aligned} \tag{3.13}$$

Now we need to find a bound for $\ddot{u}_m$. For any fixed $v \in V$ with $\|v\|_V \leq 1$, we write $v = v^1 + v^2$, where $v^1 \in \text{span}\{\varphi_j\}_{j=1}^m$, $v^2 \in \text{span}(\{\varphi_j\}_{j=1}^m)^\perp$. Then from (3.6) we obtain,

$$\begin{aligned}\rho\langle\ddot{u}_m(t), v\rangle = \rho(\ddot{u}_m(t), v^1) &= (f(t), v^1) + (g(t), v^1)_{\Gamma_{\rm N}} - a(u_m(t), v^1) \\ &\qquad + \int_0^t \beta(t-s) a(u_m(s), v^1)\, ds,\end{aligned}$$

that, using the Cuachy-Schwarz inequality and the trace theorem, implies

$$\begin{aligned}\rho\int_0^T \|\ddot{u}_m(t)\|^2_{V^*}\, dt \leq \int_0^T \Big\{&\|f(t)\|\|v^1\| + C_{Trace}\|g(t)\|_{\Gamma_{\rm N}}\|v^1\|_V \\ &+ \|u_m(t)\|_V\|v^1\|_V + \int_0^t \beta(t-s)\|u_m(s)\|_V\|v^1\|_V\, ds\Big\}\, dt.\end{aligned}$$

This, using $\|v^1\|_V \leq 1$ and (3.13), concludes

$$\begin{aligned}\rho\|\ddot{u}_m\|^2_{L_2((0,T);V^*)} &\leq C_\Omega\big\{\|f\|^2_{L_2((0,T);H)} + \|g\|^2_{L_2((0,T);H^{\Gamma_{\rm N}})}\big\} \\ &\quad + C_{\gamma,T}\|u_m\|^2_{L_\infty((0,T);V)} \\ &\leq C\big\{\|f\|^2_{L_2((0,T);H)} + \|g\|^2_{L_2((0,T);H^{\Gamma_{\rm N}})} \\ &\qquad + \|v^0\|^2 + \|u^0\|^2_V + \|g\|^2_{W_1^1((0,T);H^{\Gamma_{\rm N}})}\big\}.\end{aligned}$$

Therefore, for some constant $C = C(\Omega, \gamma, \rho, T)$,

$$\begin{aligned}&\|\ddot{u}_m\|^2_{L_2((0,T);V^*)} \\ &\qquad \leq C\big\{\|v^0\|^2 + \|u^0\|^2_V + \|g\|^2_{W_1^1((0,T);H^{\Gamma_{\rm N}})} + \|f\|^2_{L_2((0,T);H)}\big\}.\end{aligned}$$

This and (3.13) imply the estimate (3.8), and the proof is complete. □

3.2. **Existence and uniqueness of the weak solution.** Now, we use Theorem 1 to prove existence and uniqueness of the weak solution of (2.5), that is, a solution of (2.6)–(2.8).

Theorem 2. *If* $u^0 \in V$, $v^0 \in H$, $g \in W_1^1((0,T);H^{\Gamma_N})$, $f \in L_2((0,T);H)$, *there exists a unique weak solution of* (2.5).

Proof. 1. We note that the estimate (3.8) does not depend on m, so we have

$$\begin{aligned}\|u_m\|_{L_\infty((0,T);V)} + \|\dot{u}_m\|_{L_\infty((0,T);H)} + \|\ddot{u}_m\|_{L_2((0,T);V^*)}\\ \leq K = K(\Omega,\gamma,T,u,v,f,g).\end{aligned}$$

That is,

$$\begin{aligned}&\{u_m\}_{m=1}^\infty \text{ is bounded in } L_\infty((0,T);V) \subset L_2((0,T);V),\\ &\{\dot{u}_m\}_{m=1}^\infty \text{ is bounded in } L_\infty((0,T);H) \subset L_2((0,T);H),\\ &\{\ddot{u}_m\}_{m=1}^\infty \text{ is bounded in } L_2((0,T);V^*).\end{aligned} \tag{3.14}$$

2. First we prove existence. From (3.14) and a classical result in functional analysis, we conclude that the sequences $\{u_m\}_{m=1}^\infty$, $\{\dot{u}_m\}_{m=1}^\infty$, $\{\ddot{u}_m\}_{m=1}^\infty$ are weakly precompact. That is there are subsequences of $\{u_m\}_{m=1}^\infty$, $\{\dot{u}_m\}_{m=1}^\infty$, $\{\ddot{u}_m\}_{m=1}^\infty$, such that,

$$\begin{aligned}u_l &\rightharpoonup u \quad \text{in} \quad L_2((0,T);V),\\ \dot{u}_l &\rightharpoonup \dot{u} \quad \text{in} \quad L_2((0,T);H),\\ \ddot{u}_l &\rightharpoonup \ddot{u} \quad \text{in} \quad L_2((0,T);V^*),\end{aligned} \tag{3.15}$$

where the index l is a replacement of the label of the subsequences and '$\rightharpoonup$' denotes weak convergence. Consequently, (2.6) holds true and we need to verify (2.7) and (2.8). To show (2.7) we fix a positive integer N and we choose $v \in \mathcal{C}([0,T];V)$ of the form

$$v(t) = \sum_{j=1}^N h_j(t)\varphi_j. \tag{3.16}$$

Then we take $l \geq N$ and from (3.6) we have

$$\begin{aligned}\int_0^T \Big(\rho\langle \ddot{u}_l, v\rangle + a(u_l,v) - \int_0^t \beta(t-s)a(u_l(s),v)\,ds\Big)\,dt\\ = \int_0^T \big((f,v) + (g,v)_{\Gamma_\mathrm{N}}\big)\,dt.\end{aligned} \tag{3.17}$$

This, by (3.15), implies in the limit,

$$\int_0^T \Big(\rho\langle \ddot{u}, v\rangle + a(u,v) - \int_0^t \beta(t-s)a(u(s),v)\,ds\Big)\,dt = \int_0^T \big((f,v) + (g,v)_{\Gamma_{\rm N}}\big)\,dt. \tag{3.18}$$

Since functions of the form (3.16) are dense in $L_2((0,T);V)$, this equality then holds for all functions $v \in L_2((0,T);V)$, and further it implies (2.7).

Now, we need to show that u satisfies the initial conditions (2.8). Let $v \in \mathcal{C}^2([0,T];V)$ be any function which is zero in a neighborhood of T (or simply $v(T) = \dot{v}(T) = 0$). Then by partial integration in (3.17) we have

$$\int_0^T \Big(\rho\langle u_l, \ddot{v}\rangle + a(u_l,v) - \int_0^t \beta(t-s)a(u_l(s),v)\,ds\Big)\,dt = \int_0^T \big((f,v) + (g,v)_{\Gamma_{\rm N}}\big)\,dt - \rho(u_l(0),\dot{v}(0)) + \rho(\dot{u}_l(0),v(0)),$$

so that, recalling (3.15) and (3.7), in the limit we conclude,

$$\int_0^T \Big(\rho\langle u, \ddot{v}\rangle + a(u,v) - \int_0^t \beta(t-s)a(u(s),v)\,ds\Big)\,dt = \int_0^T \big((f,v) + (g,v)_{\Gamma_{\rm N}}\big)\,dt - \rho(u^0,\dot{v}(0)) + \rho(v^0,v(0)).$$

On the other hand integration by parts in (3.18) gives,

$$\int_0^T \Big(\rho\langle u, \ddot{v}\rangle + a(u,v) - \int_0^t \beta(t-s)a(u(s),v)\,ds\Big)\,dt = \int_0^T \big((f,v) + (g,v)_{\Gamma_{\rm N}}\big)\,dt - \rho(u(0),\dot{v}(0)) + \rho(v(0),v(0)).$$

Compairing the last two identities we conclude (2.8), since $v(0)$, $\dot{v}(0)$ are arbitrary. Hence u is a weak solution of (2.5).

3. It remains to prove uniqueness. To this end, we show that $u = 0$ is the only solution of (2.6)–(2.8) for $u^0 = v^0 = f = g = 0$. Let us fix $r \in [0,T]$ and define

$$v(t) = \begin{cases} \displaystyle\int_t^r u(\omega)\,d\omega & 0 \le t \le r, \\ 0 & r \le t \le T. \end{cases}$$

We note that

$$v(t) \in V, \quad v(r) = 0, \quad \dot{v}(t) = -u(t). \tag{3.19}$$

Then inserting v in (2.7) and integrating with respect to t, we have

$$\int_0^r \rho\langle\ddot{u}, v\rangle\,dt + \int_0^r a(u,v)\,dt - \int_0^r\int_0^t \beta(t-s)a(u(s),v(t))\,ds\,dt = 0. \tag{3.20}$$

For the last term we obtain

$$\begin{aligned}
-\int_0^r\int_0^t \beta(t-s)a(u(s),v(t))\,ds\,dt &= \int_0^r\int_s^r D_t\xi(t-s)a(u(s),v(t))\,dt\,ds\\
&= \int_0^r \xi(r-s)a(u(s),v(r))\,ds\\
&\quad - \int_0^r \xi(0)a(u(s),v(s))\,ds\\
&\quad - \int_0^r\int_s^r \xi(t-s)a(u(s),\dot{v}(t))\,dt\,ds\\
&= -\gamma\int_0^r a(u(s),v(s))\,ds\\
&\quad + \int_0^r\int_0^t \xi(t-s)a(u(s),u(t))\,ds\,dt,
\end{aligned}$$

where we changed the order of integrals and we used integration by parts, $\xi(0) = \gamma$ from (3.2) and $v(r) = 0$ from (3.19). Therefore integration by parts in the first term of (3.20), recalling $\tilde{\gamma} = 1-\gamma$, yields

$$-\rho\int_0^r (\dot{u},\dot{v})\,dt + \tilde{\gamma}\int_0^r a(u,v)\,dt + \int_0^r\int_0^t \xi(t-s)a(u(s),u(t))\,ds\,dt = 0.$$

This, using $\dot{v} = -u$ from (3.19), implies

$$\begin{aligned}
\rho\|u(r)\|^2 - \rho\|u(0)\|^2 - \tilde{\gamma}\|v(r)\|_V^2 + \tilde{\gamma}\|v(0)\|_V^2&\\
+ 2\int_0^r\int_0^t \xi(t-s)a(u(s),u(t))\,ds\,dt &= 0.
\end{aligned}$$

Consequently, recalling (3.3), $v(r) = 0$, and $u(0) = 0$, we have $u = 0$ a.e., and this completes the proof. □

3.3. **Regularity.** Here we study the regularity of the unique weak solution of (2.5), that is, a solution of (2.6)–(2.8).

Corollary 1. *If* $u^0 \in V, v^0 \in H, g \in W_1^1((0,T);H^{\Gamma_N})$, *and* $f \in L_2((0,T);H)$, *then for the unique solution* u *of* (2.6)–(2.8) *we have*

$$u \in L_\infty((0,T);V), \quad \dot{u} \in L_\infty((0,T);H), \quad \ddot{u} \in L_2((0,T);V^*). \tag{3.21}$$

Moreover we have the estimate

(3.22)
$$\|u\|_{L_\infty((0,T);V)} + \|\dot u\|_{L_\infty((0,T);H)} + \|\ddot u\|_{L_2((0,T);V^*)} \le C\big\{\|u^0\|_V + \|v^0\| + \|g\|_{W_1^1((0,T);H^{\Gamma_N})} + \|f\|_{L_2((0,T);H)}\big\}.$$

Proof. It is known that if $u_m \rightharpoonup u$, then

$$\|u\| \le \lim_{m\to\infty} \inf \|u_m\|. \tag{3.23}$$

Then, recalling (3.15) and the a priori estimates (3.8), we conclude (3.21) and (3.22). □

It is known from the theory of the elliptic operators, that global higher spatial regularity can not be obtained with mixed boundary conditions. Therefore we specialize to the homogeneous Dirichlet boundary condition, that is $\Gamma_{\mathrm{N}} = \varnothing$, and assume that the polygonal domain Ω is convex. We recall the usual Sobolev space $H^2 = H^2(\Omega)$ and we note that here $V = H_0^1(\Omega)$. We then use the extension of the operator $Au = -\nabla\cdot\sigma_0(u)$ to an abstract operator A with $\mathcal{D}(A) = H^2(\Omega)^d \cap V$ such that $a(u,v) = (Au,v)$ for sufficiently smooth u, v. We note that, the elliptic regularity holds, that is,

$$\|u\|_{H^2} \le C\|Au\|, \quad u \in H^2(\Omega)^d \cap V. \tag{3.24}$$

Theorem 3. *We assume that $\Gamma_N = \varnothing$. If $u^0 \in H^2$, $v^0 \in V$, and $\dot f \in L_2((0,T);H)$, then for the unique solution u of* (2.6)–(2.8) *we have*

$$\begin{aligned} &u \in L_\infty((0,T);H^2), \quad \dot u \in L_\infty((0,T);V),\\ &\ddot u \in L_\infty((0,T);H), \quad \dddot u \in L_2((0,T);V^*). \end{aligned} \tag{3.25}$$

Moreover we have the estimate

$$\begin{aligned} &\|u\|_{L_\infty((0,T);H^2)} + \|\dot u\|_{L_\infty((0,T);V)} + \|\ddot u\|_{L_\infty((0,T);H)} + \|\dddot u\|_{L_2((0,T);V^*)}\\ &\qquad \le C\big\{\|u^0\|_{H^2} + \|v^0\|_V + \|f\|_{H^1((0,T);H)}\big\}. \end{aligned} \tag{3.26}$$

Proof. Writing

$$\int_0^t \beta(t-s)a(u_m(s),\varphi_k)\,ds = \int_0^t \beta(s)a(u_m(t-s),\varphi_k)\,ds,$$

differentiating (3.6) with respect to time, and writing $\underline{v} = \dot v$, we have

(3.27)
$$\rho(\underline{\ddot u}_m(t),\varphi_k) + a(\underline{u}_m(t),\varphi_k) - \int_0^t \beta(t-s)a(\underline{u}_m(s),\varphi_k)\,ds = (\dot f(t),\varphi_k) + \beta(t)a(u_m(0),\varphi_k), \quad k = 1, \ldots, m,\ t \in (0,T),$$

with the initial conditions

$$
\begin{aligned}
&\underline{u}_m(0) = \dot{u}_m(0) = \sum_{j=1}^{m} (v^0, \varphi_j)\varphi_j, \\
&\dot{\underline{u}}_m(0) = \ddot{u}_m(0) = \sum_{j=1}^{m} \big(f(0) - Au_m(0), \varphi_j\big)\varphi_j.
\end{aligned} \tag{3.28}
$$

Then, using $\beta(t-s) = D_s\xi(t-s)$ from (3.2) and partial integration, we have

$$
\begin{aligned}
\rho(\ddot{\underline{u}}_m(t), \varphi_k) + \tilde{\gamma} a(\underline{u}_m(t), \varphi_k) + \int_0^t \xi(t-s) a(\dot{\underline{u}}_m(s), \varphi_k)\, ds& \\
= (\dot{f}(t), \varphi_k) + \beta(t) a(u_m(0), \varphi_k)& \\
- \xi(t) a(\underline{u}_m(0), \varphi_k), \quad k = 1, \ldots, m,\ t \in (0, T)&.
\end{aligned}
$$

Now, multiplying by $\ddot{d}_k(t)$ and summing $k = 1, \ldots, m$, we have

$$
\begin{aligned}
\rho(\ddot{\underline{u}}_m(t), \dot{\underline{u}}_m(t)) + \tilde{\gamma} a(\underline{u}_m(t), \dot{\underline{u}}_m(t)) + \int_0^t \xi(t-s) a(\dot{\underline{u}}_m(s), \dot{\underline{u}}_m(t))\, ds& \\
= (\dot{f}(t), \dot{\underline{u}}_m(t)) + \beta(t) a(u_m(0), \dot{\underline{u}}_m(t))& \\
- \xi(t) a(\underline{u}_m(0), \dot{\underline{u}}_m(t)), \quad t \in (0, T)&.
\end{aligned}
$$

Integrating over $[0, t]$ and partial integration in the last term, we have

$$
\begin{aligned}
\rho\|\dot{\underline{u}}_m(t)\|^2 &+ \tilde{\gamma}\|\underline{u}_m(t)\|_V^2 \\
&\le \rho\|\dot{\underline{u}}_m(0)\|^2 + \tilde{\gamma}\|\underline{u}_m(0)\|_V^2 \\
&\quad + 2\int_0^t (\dot{f}(r), \dot{\underline{u}}_m(r))\, dr + 2\int_0^t \beta(r) a(u_m(0), \dot{\underline{u}}_m(r))\, dr \\
&\quad - 2\int_0^t \beta(r) a(\underline{u}_m(0), \underline{u}_m(r))\, dr \\
&\quad - 2\xi(t) a(\underline{u}_m(0), \underline{u}_m(t)) + 2\xi(0) a(\underline{u}_m(0), \underline{u}_m(0)),
\end{aligned}
$$

that using the Cauchy-Schwarz inequality, the trace theorem, $\|\beta\|_{L_1(\mathbb{R}^+)} = \gamma$, and $\xi(0) = \gamma$, we obtain

$$\begin{aligned}
\rho\|\dot{\underline{u}}_m(t)\|^2 &+ \tilde{\gamma}\|\underline{u}_m(t)\|_V^2 \\
&\leq \rho\|\dot{\underline{u}}_m(0)\|^2 + (1+\gamma)\|\underline{u}_m(0)\|_V^2 \\
&\quad + 2/C_1 \max_{0\leq r\leq t}\|\dot{\underline{u}}_m(r)\|^2 + C_1\Big(\int_0^t \|\dot{f}(r)\|\,dr\Big)^2 \\
&\quad + 2/C_2\|u_m(0)\|_{H^2}^2 + 2\gamma^2 C_2 \max_{0\leq r\leq t}\|\dot{\underline{u}}_m(r)\|^2 \\
&\quad + 2/C_3\|\underline{u}_m(0)\|_V^2 + 2\gamma^2 C_3 \max_{0\leq r\leq t}\|\underline{u}_m(r)\|_V^2 \\
&\quad + 2/C_4\|\underline{u}_m(0)\|_V^2 + 2\gamma^2 C_4\|\underline{u}_m(t)\|_V^2.
\end{aligned}$$

This implies, for some constant C,

$$\begin{aligned}
\|\dot{\underline{u}}_m\|^2_{L_\infty((0,T);H)} &+ \|\underline{u}_m\|^2_{L_\infty((0,T);V)} \\
&\leq C\big\{\|\dot{\underline{u}}_m(0)\|^2 + \|\underline{u}_m(0)\|_V^2 + \|u_m(0)\|_{H^2}^2 + \|\dot{f}\|^2_{L_1((0,T);H)}\big\}.
\end{aligned}$$

Then recalling $\underline{u} = \dot{u}$, the initial data from (3.28), and using

$$\|u_m(0)\|_{H^2} \leq \|u^0\|_{H^2}, \quad \|\dot{u}_m(0)\|_V \leq \|v^0\|_V,$$

we have

$$\begin{aligned}
\|\ddot{u}_m\|^2_{L_\infty((0,T);H)} &+ \|\dot{u}_m\|^2_{L_\infty((0,T);V)} \\
&\leq C\big\{\|u^0\|_{H^2}^2 + \|v^0\|_V^2 + \|f(0)\|^2 + \|\dot{f}\|^2_{L_1((0,T);H)}\big\}.
\end{aligned} \tag{3.29}$$

We now find a bound for $\|u_m(t)\|_{H^2}$. We recall the eigenvalue problem (3.4) with eigenpairs $\{(\lambda_j, \varphi_j)\}_{j=1}^\infty$, $a(u,v) = (Au, v)$. Then we multiply (3.6) by $\lambda_k d_k(t)$ and add for $k = 1, \ldots, m$ to obtain

$$a(u_m, Au_m) = (f - \rho\ddot{u}_m, Au_m) + \int_0^t \beta(t-s)a(u_m(s), Au_m(t))\,ds.$$

This implies

$$\begin{aligned}
\|Au_m(t)\|^2 \leq \frac{C}{\epsilon}&\Big(\|f(t)\|^2 + \rho\|\ddot{u}_m(t)\|^2\Big) + \epsilon\|Au_m(t)\|^2 \\
&+ \gamma \max_{0\leq r\leq t}\|Au_m(s)\|^2,
\end{aligned}$$

that gives us, by elliptic regularity (3.24),

$$\|u_m\|^2_{L_\infty((0,T);H^2)} \leq C\Big(\|f\|^2_{L_\infty((0,T);H)} + \|\ddot{u}_m\|^2_{L_\infty((0,T);H)}\Big).$$

From this and (3.29) we conclude

$$\begin{aligned}\|\ddot{u}_m\|^2_{L_\infty((0,T);H)} + \|\dot{u}_m\|^2_{L_\infty((0,T);V)} + \|u_m\|^2_{L_\infty((0,T);H^2)} \\ \leq C\big\{\|u^0\|^2_{H^2} + \|v^0\|^2_V + \|f\|^2_{L_\infty((0,T);H)} + \|\dot{f}\|^2_{L_1((0,T);H)}\big\},\end{aligned}$$

that using $\|f\|_{L_\infty((0,T);H)} \leq C\|f\|_{W^1_1((0,T);H)}$, by Sobolev inequality, we have

$$\begin{aligned}\|\ddot{u}_m\|^2_{L_\infty((0,T);H)} &+ \|\dot{u}_m\|^2_{L_\infty((0,T);V)} + \|u_m\|^2_{L_\infty((0,T);H^2)} \\ &\leq C\big\{\|u^0\|^2_{H^2} + \|v^0\|^2_V + \|f\|^2_{W^1_1((0,T);H)}\big\} \\ &\leq C\big\{\|u^0\|^2_{H^2} + \|v^0\|^2_V + \|f\|^2_{H^1((0,T);H)}\big\}.\end{aligned}$$

Finally from (3.27) , similar to the proof of Theorem 1, we obtain

$$\|\dddot{u}_m\|^2_{L_2((0,T);V^*)} \leq C\big\{\|u^0\|^2_{H^2} + \|v^0\|^2_V + \|f\|^2_{H^1((0,T);H)}\big\}.$$

The last two estimates then, in the limit, imply the desired estimate (3.26), and the proof is now complete. ☐

Remark 1. If we continue differentiation in time to investigate more regularity, we obtain from (3.27)

$$\begin{aligned}\rho(\dddot{\underline{u}}_m(t), \varphi_k) + a(\dot{\underline{u}}_m(t), \varphi_k) - \int_0^t \beta(t-s) a(\dot{\underline{u}}_m(s), \varphi_k)\, ds \\ = (\ddot{f}(t), \varphi_k) + \dot{\beta}(t) a(u_m(0), \varphi_k) \\ + \beta(t) a(\underline{u}_m(0), \varphi_k), \quad k = 1, \ldots, m,\, t \in (0,T),\end{aligned}$$

Further the term $\dot{\beta}(t)a(u_m(0), \varphi_k)$ leads to

$$\dot{\beta}(t) a(u_m(0), \ddot{\underline{u}}_m(t)).$$

But $\dot{\beta}$ is not integrable. Besides, after integration in time, we can not use partial integration to transfer one time derivative from $\dot{\beta}$ to $\ddot{\underline{u}}_m(t)$, since there is not enough regularity to handle $\dddot{\underline{u}}_m$. This means that we can not get more regularity with weakly singular kernel β. This also indicates that with smoother kernel we can get higher regularity in case of homogeneous Dirichlet boundary condition under the appropriate assumption on the data, that is, more regularity and compatibility.

References

1. K. Adolfsson, M. Enelund, and S. Larsson, *Space-time discretization of an integro-differential equation modeling quasi-static fractional-order viscoelasticity*, J. Vib. Control **14** (2008), 1631–1649.
2. R. L. Bagley and P. J. Torvik, *Fractional calculus–a different approach to the analysis of viscoelastically damped structures*, AIAA J. **21** (1983), 741–748.

3. H. Brunner and P. J. van der Houwen, *The Numerical Solution of Volterra Equations*, CWI Monographs, vol. 3, North-Holland Publishing Co., 1986.
4. C. M. Dafermos, *An abstract Volterra equation with applications to linear viscoelasticity*, J. Differential Equations **7** (1970), 554–569.
5. R. Dautray and J. L. Lions, *Mathematical Analysis and Numerical Methods for Science and Technology. Vol. 5*, Springer-Verlag, 1992.
6. W. Desch and E. Fašanga, *Stress obtained by interpolation methods for a boundary value problem in linear viscoelasticity*, J. Differential Equations **217** (2005), 282–304.
7. M. Enelund and P. Olsson, *Damping described by fading memory-analysis and application to fractional derivative models*, International J. Solids Structures **36** (1999), 939–970.
8. L. C. Evans, *Partial Differential Equations*, American Mathematical Society, Providence, Rhode Island, 1998.
9. S. Larsson and F. Saedpanah, *The continuous Galerkin method for an integro-differential equation modeling dynamic fractional order viscoelasticity*, IMA J. Numer. Anal., to appear.
10. Ch. Lubich, I. H. Sloan, and V. Thomée, *Nonsmooth data error estimates for approximations of an evolution equation with a positive-type memory term*, Math. Comp. **65** (1996), 1–17.
11. W. McLean, I. H. Sloan, and V. Thomée, *Time discretization via Laplace transformation of an integro-differential equation of parabolic type*, Numer. Math. **102** (2006), 497–522.
12. W. McLean and V. Thomée, *Numerical solution of an evolution equation with a positive-type memory term*, J. Austral. Math. Soc. Ser. B **35** (1993), 23–70.
13. A. K. Pani, V. Thomée, and L. B. Wahlbin, *Numerical methods for hyperbolic and parabolic integro-differential equations*, J. Integral Equations Appl. **4** (1992), 533–584.
14. M. Renardy, W. J. Hrusa, and J.—A. Nohel, *Mathematical Problems in Viscoelasticity*, Pitman Monographs and Surveys in Pure and Applied Mathematics, vol. 35, Longman Scientific & Technical, 1987.
15. B. Riviére, S. Shaw, and J. R. Whiteman, *Discontinuous Galerkin finite element methods for dynamic linear solid viscoelasticity problems*, Numer. Methods Partial Differential Equations **23** (2007), 1149–1166.
16. S. Shaw and J. R. Whiteman, *A posteriori error estimates for space-time finite element approximation of quasistatic hereditary linear viscoelasticity problems*, Comput. Methods Appl. Mech. Engrg. **193** (2004), 5551–5572.
17. O. J. Staffans, *Well-posedness and stabilizability of a viscoelastic equation in energy space*, Trans. Amer. Math. Soc. **345** (1994), 527–575.

Department of Mathematical Sciences, Chalmers University of Technology and University of Gothenburg, SE–412 96 Göteborg, Sweden
E-mail address: `fardin@chalmers.se`

A CONTINUOUS SPACE-TIME FINITE ELEMENT METHOD FOR AN INTEGRO-DIFFERENTIAL EQUATION MODELING DYNAMIC FRACTIONAL ORDER VISCOELASTICITY

FARDIN SAEDPANAH

ABSTRACT. A fractional order integro-differential equation with a weakly singular kernel is considered. A weak form is formulated, and stability of the primal and the dual problem is studied. The continuous Galerkin method of degree one is formulated, and optimal a priori error estimate is obtained by duality argument. A posteriori error representation based on space-time cells is presented such that it can be used for adaptive strategies based on dual weighted residual methods. Some global a posteriori error estimates are also proved.

1. INTRODUCTION

We study an initial-boundary value problem, modeling dynamic fractional order viscoelasticity of the form (we use '$\cdot$' to denote the time derivative),

(1.1)

$$\begin{aligned}
&\rho\ddot{u}(x,t) - \nabla\cdot\sigma_0(u;x,t) \\
&\qquad + \int_0^t \beta(t-s)\nabla\cdot\sigma_0(u;x,s)\,ds = f(x,t) && \text{in } \Omega\times(0,T), \\
&u(x,t) = 0 && \text{on } \Gamma_{\mathrm{D}}\times(0,T), \\
&\sigma(u;x,t)\cdot n = g(x,t) && \text{on } \Gamma_{\mathrm{N}}\times(0,T), \\
&u(x,0) = u^0(x), \quad \dot{u}(x,0) = v^0(x) && \text{in } \Omega.
\end{aligned}$$

Here u is the displacement vector, ρ is the (constant) mass density, f and g represent, respectively, the volume and surface loads. The stress $\sigma = \sigma(u;x,t)$ is determined by

$$\sigma(t) = \sigma_0(t) - \int_0^t \beta(t-s)\sigma_0(s)\,ds,$$

Date: April 12, 2009.

1991 *Mathematics Subject Classification.* 65M60, 45K05.

Key words and phrases. fractional order viscoelasticity, Galerkin approximation, weakly singular kernel, stability, a priori estimate, a posteriori estimate.

with

$$\sigma_0(t) = 2\mu_0\epsilon(t) + \lambda_0 \operatorname{Tr}(\epsilon(t))\mathrm{I},$$

where I is the identity operator, ϵ is the strain which is defined by $\epsilon = \frac{1}{2}(\nabla u + (\nabla u)^T)$, and $\mu_0, \lambda_0 > 0$ are elastic constants of Lamé type. The kernel is defined by:

$$\begin{aligned}\beta(t) &= -\gamma \frac{d}{dt} E_\alpha\Big(- \Big(\frac{t}{\tau}\Big)^\alpha\Big) = \gamma \frac{\alpha}{\tau}\Big(\frac{t}{\tau}\Big)^{-1+\alpha} E_\alpha'\Big(- \Big(\frac{t}{\tau}\Big)^\alpha\Big) \\ &\approx Ct^{-1+\alpha}, \quad t \to 0,\end{aligned} \tag{1.2}$$

which means that it is weakly singular with the properties

$$\begin{aligned}&\beta(t) \geq 0, \\ &\|\beta\|_{L_1(\mathbb{R}^+)} = \int_0^\infty \beta(t)\,dt = \gamma\Big(E_\alpha(0) - E_\alpha(\infty)\Big) = \gamma < 1,\end{aligned} \tag{1.3}$$

where $\gamma \in (0,1)$ is a constant and $\tau > 0$ is the relaxation time. Here E_α is the Mittag-Leffler function of order $\alpha \in (0,1)$ and defined by,

$$E_\alpha(z) = \sum_{n=0}^{\infty} \frac{z^n}{\Gamma(1+n\alpha)}.$$

There is an extensive literature on theoretical and numerical analysis of integro-differential equations modeling linear and fractional order viscoelasticity, see, e.g., [1], [7], [8], [9], [10], [12], and [13]. Existence, uniqueness and regularity of solution of a problem in the form of (1.1) has been studied in [7], [11]. A posteriori analysis of temporal finite element approximation of a parabolic type problem and discontinuous Galerkin finite element approximation of a quasi-static ($\rho\ddot{u} \approx 0$) linear viscoelasticity problem has been studied, respectively, in [1] and [13]. The present work extends previous works, e.g., [1], [2], and [13].

The outline of this paper is as follows. In §2 we define a weak form of (1.1) and the corresponding dual (adjoint) problem, and we study the stability. In §3 we formulate a continuous Galerkin method of degree one and we obtain stability estimates. Then in §4 we obtain an optimal a priori error estimate. We present an a posteriori error representation based on the dual weighted residual method in §5, and we prove some global a posteriori error estimates.

2. Weak formulation and stability

We let $\Omega \subset \mathbb{R}^d$, $d = 2, 3$, be a bounded polygonal domain with boundary $\Gamma = \Gamma_{\mathrm{D}} \cup \Gamma_{\mathrm{N}}$ where Γ_{D} and Γ_{N} are disjoint and $\operatorname{meas}(\Gamma_{\mathrm{D}}) \neq 0$. We introduce the function spaces $H = L_2(\Omega)^d$, $H_{\Gamma_{\mathrm{N}}} = L_2(\Gamma_{\mathrm{N}})^d$, and $V = \{v \in H^1(\Omega)^d :$

$v|_{\Gamma_{\rm D}}=0\}$. We denote the norms in H and $H_{\Gamma_{\rm N}}$ by $\|\cdot\|$ and $\|\cdot\|_{\Gamma_{\rm N}}$, respectively. We also define a bilinear form (with the usual summation convention)

$$a(v,w)=\int_\Omega\big(2\mu_0\epsilon_{ij}(v)\epsilon_{ij}(w)+\lambda_0\epsilon_{ii}(v)\epsilon_{jj}(w)\big)\,dx,\quad v,w\in V\,, \tag{2.1}$$

which is coercive on V, and we equip V with the inner product $a(\cdot,\cdot)$ and norm $\|v\|_V^2=a(v,v)$. We define $Au=-\nabla\cdot\sigma_0(u)$, which is a selfadjoint, positive definite, unbounded linear operator, with $\mathcal{D}(A)=H^2(\Omega)^d\cap V$, and we use the norms $\|v\|_s=\|A^{s/2}v\|$.

We use a "velocity-displacement" formulation of (1.1) which is obtained by introducing a new velocity variable. Henceforth we use the new variables $u_1=u$, $u_2=\dot u$ and $u=(u_1,u_2)$ the pair of vector valued functions. Now we define the bilinear and linear forms $\mathcal{A}:\mathcal{V}\times\mathcal{W}\to\mathbb{R}$, $\mathcal{A}^*_\tau:\mathcal{W}^*\times\mathcal{V}^*\to\mathbb{R}$, $F:\mathcal{W}\to\mathbb{R}$, $J_\tau:\mathcal{W}^*\to\mathbb{R}$, for $\tau\in\mathbb{R}^{\geq 0}$, by

$$\begin{aligned}
\mathcal{A}(u,w)&=\int_0^T\Big\{(\dot u_1,w_1)-(u_2,w_1)+\rho(\dot u_2,w_2)+a(u_1,w_2)\\
&\qquad-\int_0^t\beta(t-s)a\big(u_1(s),w_2\big)\,ds\Big\}\,dt\\
&\quad+\big(u_1(0),w_1(0)\big)+\rho\big(u_2(0),w_2(0)\big),\\
\mathcal{A}^*_\tau(w,z)&=\int_\tau^T\Big\{-(w_1,\dot z_1)+a(w_1,z_2)-\int_t^T\beta(s-t)a\big(w_1,z_2(s)\big)\,ds\\
&\qquad-\rho(w_2,\dot z_2)-(w_2,z_1)\Big\}\,dt+\big(w_1(T),z_1(T)\big)+\rho\big(w_2(T),z_2(T)\big),\\
F(w)&=\int_0^T\Big\{(f,w_2)+(g,w_2)_{\Gamma_{\rm N}}\Big\}\,dt+\big(u^0,w_1(0)\big)+\rho\big(v^0,w_2(0)\big),\\
J_\tau(w)&=\int_\tau^T\Big\{(w_1,j_1)+(w_2,j_2)\Big\}\,dt+\big(w_1(T),z_1^T\big)+\rho\big(w_2(T),z_2^T\big),
\end{aligned}$$

where j_1, j_2 and z_1^T, z_2^T represent, respectively, the load terms and the initial data of the dual (adjoint) problem. In case of $\tau=0$, we use the notation $\mathcal{A}^*$, J for short. Here

$$\begin{aligned}
\mathcal{V}&=H^1\big((0,T);V\big)\times H^1\big((0,T);H\big),\\
\mathcal{V}^*&=H^1\big((0,T);H\big)\times H^1\big((0,T);V\big),\\
\mathcal{W}&=\big\{w=(w_1,w_2):w\in L_2\big((0,T);H\big)\times L_2\big((0,T);V\big),\\
&\qquad\qquad w_i\text{ are right continuous in time}\big\},\\
\mathcal{W}^*&=\big\{w=(w_1,w_2):w\in L_2\big((0,T);V\big)\times L_2\big((0,T);H\big),\\
&\qquad\qquad w_i\text{ are left continuous in time}\big\},
\end{aligned} \tag{2.2}$$

and we note that $\mathcal{V} \subset \mathcal{W}^*$, $\mathcal{V}^* \subset \mathcal{W}$.

The variational formulation analogue to (1.1) is then to find $u \in \mathcal{V}$ such that,

$$\mathcal{A}(u, w) = F(w), \quad \forall w \in \mathcal{W}. \tag{2.3}$$

Here the definition of the velocity $u_2 = \dot{u}_1$ is enforced in the L_2 sense, and the initial data are placed in the bilinear form in a weak sense. A variant is used in [7] where the velocity has been enforced in the H^1 sense, without placing the initial data in the bilinear form. We also note that the initial data are retained by the choice of the function space $\mathcal{W}$, that consists of right continuous functions with respect to time.

To obtain the dual (adjoint) problem we note that $\mathcal{A}^*$ is the adjoint form of $\mathcal{A}$. Indeed, integrating by parts with respect to time in $\mathcal{A}$, then changing the order of integrals in the convolution term as well as changing the role of the variables s, t, we have,

$$\mathcal{A}(v, w) = \mathcal{A}^*(v, w), \quad \forall v \in \mathcal{V},\, w \in \mathcal{V}^*. \tag{2.4}$$

Then the variational formulation of the dual problem is to find $z \in \mathcal{V}^*$ such that,

$$\mathcal{A}^*(w, z) = J(w), \quad \forall w \in \mathcal{W}^*, \tag{2.5}$$

that is a weak formulation of

$$\rho \ddot{z}_2 + A z_2 - \int_t^T \beta(s-t) A z_2(s)\, ds = j_1 - \frac{\partial}{\partial t} j_2,$$

with initial data z_1^T, z_2^T, and function $j = (j_1, j_2)$ that is defined by $J(w) = \int_0^T (w, j)\, dt$.

In the analysis below we use a positive type kernel ξ. Indeed, we recall (1.2), (1.3) and we define the function

$$\xi(t) = \gamma - \int_0^t \beta(s)\, ds = \int_t^\infty \beta(s)\, ds = \gamma E_\alpha(t), \tag{2.6}$$

and it is easy to see that

$$D_t \xi(t) = -\beta(t) < 0, \quad \xi(0) = \gamma, \quad \lim_{t \to \infty} \xi(t) = 0, \quad 0 < \xi(t) \leq \gamma. \tag{2.7}$$

Besides, ξ is a completely monotone function, that is,

$$(-1)^j D_t^j \xi(t) \geq 0, \quad t \in (0, \infty),\, j \in \mathbb{N},$$

since the Mittag-Leffler function E_α, $\alpha \in [0, 1]$ is completely monotone, see, e.g., [5]. Consequently, an important property of ξ is that, it is a positive type kernel, that is, it is continuous and, for any $T \geq 0$, satisfies

$$\int_0^T \int_0^t \xi(t-s)\phi(t)\phi(s)\,ds\,dt \geq 0, \quad \forall \phi \in \mathcal{C}([0,T]). \tag{2.8}$$

Theorem 1. *Let u be the solution of (2.3) with sufficiently smooth data u^0, v^0, f, g. Then for $l \in \mathbb{R}, T > 0$, we have the identity,*

$$\begin{aligned} \rho\|u_2(T)\|_l^2 + \tilde{\gamma}\|u_1(T)\|_{l+1}^2 + 2\int_0^T \int_0^t \xi(t-s)a\big(\dot{u}(s), A^l\dot{u}(t)\big)\,ds\,dt \\ = \rho\|v^0\|_l^2 + (1+\gamma)\|u^0\|_{l+1}^2 + 2\int_0^T \big(f(t), A^l\dot{u}_1(t)\big)\,dt \\ + 2\int_0^T \big(g(t), A^l\dot{u}_1(t)\big)_{\Gamma_N}\,dt \\ - 2\int_0^T \beta(t)a\big(u^0, A^l u_1(t)\big)\,dt - 2\xi(T)a\big(u^0, A^l u_1(T)\big), \end{aligned} \tag{2.9}$$

where $\tilde{\gamma} = 1-\gamma$. Moreover, with $\Gamma_N = \varnothing$ or $\Gamma_N \neq \varnothing$, $g = 0$, we have,

$$\|u_2(T)\|_l + \|u_1(T)\|_{l+1} \leq C\Big\{\|v^0\|_l + \|u^0\|_{l+1} + \int_0^T \|f(t)\|_l\,dt\Big\}, \tag{2.10}$$

for some $C = C(\rho,\gamma,T)$. And with $\Gamma_N \neq \varnothing$, $g \neq 0$, $l = 0$, we have the estimate,

$$\begin{aligned} \|u_2(T)\| + \|u_1(T)\|_1 \\ \leq C\big\{\|v^0\| + \|u^0\|_1 + \|f\|_{L_1((0,t);H)} + \|g\|_{W_1^1((0,t);H^{\Gamma_N})}\big\}, \end{aligned} \tag{2.11}$$

for some $C = C(\Omega,\rho,\gamma,T)$.

Proof. Since u is a solution of (2.3), we obviously have $u_2 = \dot{u}_1$. We recall $\tilde{\gamma} = 1-\gamma$, and from (2.7) we have that $\beta(t-s) = D_s\xi(t-s)$, $\xi(0) = \gamma$. These and partial integration in the convolution term in $\mathcal{A}$, yield

$$\begin{aligned} \mathcal{A}(u,w) = \int_0^T \Big\{\rho(\dot{u}_2, w_2) + \tilde{\gamma}a(u_1, w_2) + \int_0^t \xi(t-s)a\big(\dot{u}_1(s), w_2\big)\,ds \\ + \xi(t)a(u^0, w_2)\Big\}\,dt + \big(u_1(0), w_1(0)\big) + \rho\big(u_2(0), w_2(0)\big). \end{aligned} \tag{2.12}$$

Putting this with $w=(w_1,w_2)=(A^{l+1}u_1,A^lu_2)=(A^{l+1}u_1,A^l\dot{u}_1)$ in (2.3) we obtain,

$$\begin{aligned}
&\rho\|u_2(T)\|_l^2+\tilde{\gamma}\|u_1(T)\|_{l+1}^2+2\int_0^T\int_0^t\xi(t-s)a\big(\dot{u}_1(s),A^l\dot{u}_1(t)\big)\,ds\,dt\\
&\quad=\rho\|v^0\|_l^2+\tilde{\gamma}\|u^0\|_{l+1}^2+2\int_0^T(f,A^l\dot{u}_1)\,dt+2\int_0^T(g,A^l\dot{u}_1)_{\Gamma_{\mathrm{N}}}\,dt\\
&\qquad-2\int_0^T\xi(t)a\big(u^0,A^l\dot{u}_1(t)\big)\,dt.\\
&\quad=\rho\|v^0\|_l^2+\tilde{\gamma}\|u^0\|_{l+1}^2+2\int_0^T(f,A^l\dot{u}_1)\,dt+2\int_0^T(g,A^l\dot{u}_1)_{\Gamma_{\mathrm{N}}}\,dt\\
&\qquad-2\int_0^T\beta(t)a\big(u^0,A^lu_1(t)\big)\,dt-2\xi(T)a\big(u^0,A^lu_1(T)\big)+2\gamma\|u^0\|_{l+1}^2,
\end{aligned}$$

where for the last equality we used partial integration and $D_t\xi(t)=-\beta(t)$. This, having $\tilde{\gamma}=1-\gamma$, implies the identity (2.9).

Now we prove the stimates (2.10) and (2.11). First, in (2.9), we use (2.8), integration by parts in the term with the surface load g, and the Cauchy-Schwarz inequality and conclude,

$$\begin{aligned}
&\rho\|u_2(T)\|_l^2+\tilde{\gamma}\|u_1(T)\|_{l+1}^2\\
&\quad\le\rho\|v^0\|_l^2+(1+\gamma)\|u^0\|_{l+1}^2\\
&\qquad+2/C_1\max_{0\le t\le T}\|\dot{u}_1(t)\|_l^2+2C_1\Big(\int_0^T\|f(t)\|_l\,dt\Big)^2\\
&\qquad-2\int_0^T\big(\dot{g}(t),A^lu_1(t)\big)_{\Gamma_{\mathrm{N}}}\,dt+2\big(g(T),A^lu_1(T)\big)_{\Gamma_{\mathrm{N}}}-2\big(g(0),A^lu^0\big)_{\Gamma_{\mathrm{N}}}\\
&\qquad+2\|u^0\|_{l+1}\max_{0\le t\le T}\|u_1(t)\|_{l+1}\int_0^T\beta(t)\,dt\\
&\qquad+2\gamma/C_2\|u^0\|_{l+1}^2+2\gamma C_2\|u_1(T)\|_{l+1}^2.
\end{aligned}$$

This with $\Gamma_{\mathrm{N}}=\varnothing$ or $\Gamma_{\mathrm{N}}\neq\varnothing$, $g=0$, considering $\|\beta\|_{L_1(\mathbb{R}^+)}=\gamma$, implies (2.10). But for the case that the surface load $g\neq0$ ($\Gamma_{\mathrm{N}}\neq\varnothing$), we need to restrict to $l=0$, due to the trace theorem. That is, in this case with $l=0$ and using $\|u_1(t)\|_{\Gamma_{\mathrm{N}}}\le C(\Omega)\|u_1(t)\|_1$ by the trace theorem, in a standard way, we have the estimate

$$\begin{aligned}
\|u_2(T)\|+\|u_1(T)\|_1\le C\big\{&\|v^0\|+\|u^0\|_1+\|f\|_{L_1((0,t);H)}\\
&+\|\dot{g}\|_{L_1((0,T);H^{\Gamma_{\mathrm{N}}})}+\|g\|_{L_\infty((0,T);H^{\Gamma_{\mathrm{N}}})}\big\}.
\end{aligned}$$

This and the fact that $\|g\|_{L_\infty((0,T);H^{\Gamma_{\mathrm{N}}})}\le C\|g\|_{W_1^1((0,T);H^{\Gamma_{\mathrm{N}}})}$, by Sobolev inequality, imply (2.11). □

Remark 1. An identity, slightly different from (2.9) has been presented in [2], by using a function $w(t,s)=u(t)-u(t-s)$ that must belong to a weighted L_2-space, introduced in [6], see also [7] where the abstract framework by [6] has been, specifically, applied to this problem. The proof presented here avoids using function w and seems to be simple and straightforward.

Remark 2. An important property, used here is that the kernel is of positive type. This means that the technique presented here can be applied to the problems with positive type kernels. For example, with $g=0$, we could consider problems with positive type kernels and of the form

$$\rho\ddot{u}+Au-\int_0^t \beta(t-s)Bu(s)\,ds=f,$$

where B is a selfadjoint, positive definite linear operator such that, for some suitable constant C,

$$(Bv,w)\le C(Av,w),\quad \forall v,w\in\mathcal{D}(A).$$

For example, with kernel β defined in (1.2), we should have $1-\gamma C>0$. Then a similar argument can be applied with $l=0$, and also with $l\neq 0$ provided B and A be comutative.

Theorem 2. *Let z be the solution of the dual problem* (2.5) *with sufficiently smooth data z_1^T, z_2^T, j_1, j_2. Then for $l\in\mathbb{R}$, $0<t<T$, we have the identity,*

(2.13)

$$\begin{aligned}\|z_1(t)\|_l^2&+\rho\tilde{\gamma}\|z_2(t)\|_{l+1}^2+2\rho\int_t^T\int_r^T \xi(s-r)a(A^l\dot{z}_2(r),\dot{z}_2(s))\,ds\,dr\\&=\|z_1^T\|_l^2+\rho(1+\gamma)\|z_2^T\|_{l+1}^2+2\int_t^T\Big\{(A^lz_1,j_1)+(A^{l+1}z_2,j_2)\Big\}\,dr\\&\quad-2\rho\int_t^T\beta(T-r)a(A^lz_2(r),z_2^T)\,dr\\&\quad-2\rho\xi(T-t)a(A^lz_2(t),z_2^T).\end{aligned}$$

where $\tilde{\gamma}=1-\gamma$. Moreover, for some constant $C=C(\rho,\gamma,T)$, we have stability estimates

(2.14)

$$\|z_1(t)\|_l+\|z_2(t)\|_{l+1}\le C\Big\{\|z_1^T\|_l+\|z_2^T\|_{l+1}+\int_t^T\Big(\|j_1\|_l+\|j_2\|_{l+1}\Big)\,dr\Big\}.$$

Proof. Since z is a solution of (2.5), we obviously have $z_1=-\rho\dot{z}_2$ and, for $t\in[0,T)$, z satisfies

$$\mathcal{A}_t^*(w,z)=J_t(w),\quad \forall v\in\mathcal{W}^*. \tag{2.15}$$

We recall $\tilde{\gamma} = 1-\gamma$, and from (2.7) we have $\beta(s-t) = -D_s\xi(s-t)$, $\xi(0) = \gamma$. Then, by partial integration with respect to time in the convolution term in $\mathcal{A}_t^*$, we obtain

(2.16)

$$\begin{aligned}\mathcal{A}_t^*(w,z) = \int_t^T \Big\{ &-(w_1,\dot{z}_1) + \tilde{\gamma}a(w_1,z_2) - \int_r^T \xi(s-r)a\big(w_1,\dot{z}_2(s)\big)\,ds \\ &+ \xi(T-r)a\big(w_1,z_2(T)\big)\Big\}\,dr \\ &+ \big(w_1(T),z_1(T)\big) + \rho\big(w_2(T),z_2(T)\big).\end{aligned}$$

Putting this with $w = (w_1,w_2) = (A^l z_1, A^{l+1}z_2) = (-\rho A^l\dot{z}_2, A^{l+1}z_2)$ in (2.15) we have

$$\begin{aligned}\int_t^T \Big\{ -\frac{1}{2}D_t\|z_1(r)\|_l^2 - \frac{1}{2}\rho\tilde{\gamma}D_t\|z_2(r)\|_{l+1}^2 + \rho\int_r^T \xi(s-r)a\big(A^l\dot{z}_2(r),\dot{z}_2(s)\big)\,ds \\ - \rho\xi(T-r)a\big(A^l\dot{z}_2(r),z_2^T\big)\Big\}\,dr = \int_t^T \Big\{(A^l z_1, j_1) + (A^{l+1}z_2, j_2)\Big\}\,dr,\end{aligned}$$

that implies

$$\begin{aligned}\|z_1(t)\|_l^2 + \rho\tilde{\gamma}\|z_2(t)\|_{l+1}^2 + 2\rho\int_t^T\int_r^T \xi(s-r)a\big(A^l\dot{z}_2(r),\dot{z}_2(s)\big)\,ds\,dr \\ = \|z_1^T\|_l^2 + \rho\tilde{\gamma}\|z_2^T\|_{l+1}^2 + 2\int_t^T \Big\{(A^l z_1, j_1) + (A^{l+1}z_2, j_2)\Big\}\,dr \\ + 2\rho\int_t^T \xi(T-r)a\big(A^l\dot{z}_2(r),z_2^T\big)\,dr.\end{aligned}$$

Now integration by parts in the last term, having $D_r\xi(T-r) = \beta(T-r)$ by (2.7), gives the identity (2.13). Then, using (2.8) in (2.13), similar to the proof of (2.10), in a standard way, we conclude the inequality (2.14), and this completes the proof. □

3. The continuous Galerkin method

Let $0 = t_0 < t_1 < \cdots < t_{n-1} < t_n < \cdots < t_N = T$ be a partition of the time interval $[0,T]$. To each discrete time level t_n we associate a triangulation $\mathcal{T}_h^n$ of the polygonal domain Ω with the mesh function,

(3.1) $$h_n(x) = h_K = \operatorname{diam}(K), \quad x\in K,\ K\in\mathcal{T}_h^n,$$

and a finite element space V_h^n consisting of continuous piecewise linear polynomials. For each time subinterval $I_n = (t_{n-1},t_n)$ of length $k_n = t_n - t_{n-1}$, we define intermediate triangulaion $\bar{\mathcal{T}}_h^n$ which is composed of mutually finest meshes of the neighboring meshes $\mathcal{T}_h^n$, $\mathcal{T}_h^{n-1}$ defined at discrete time levels t_n, t_{n-1}, respectively. The mesh function $\bar{h}_n$ is then defined by

$$\bar{h}_n(x) = \bar{h}_K = \text{diam}(K), \quad x \in K,\, K \in \bar{\mathcal{T}}_h^n. \tag{3.2}$$

Correspondingly, we define the finite element spaces $\bar{V}_h^n$ consisting of continuous piecewise linear polynomials. This construction is used in order to allow continuity in time of the trial functions when the meshes change with time. Hence we obtain a decomposition of each time slab $\Omega^n = \Omega \times I_n$ into space-time cells $K^n = K \times I_n$, $K \in \bar{\mathcal{T}}_h^n$ (prisms, for example, in case of $\Omega \subset \mathbb{R}^2$). We note the difference between the mesh functions h_n and $\bar{h}_n$, and this is important in our a posteriori error analysis. The trial and test function spaces for the discrete form are, respectively:

$$\begin{aligned}
\mathcal{V}_{hk} = \Big\{ U = (U_1, U_2) : \; & U \text{ continuous in } \Omega \times [0,T], \\
& U(x,t)|_{I_n} \text{ linear in } t, \\
& U(\cdot,t_n) \in (V_h^n)^2, U(\cdot,t)|_{I_n} \in (\bar{V}_h^n)^2 \Big\}, \\
\mathcal{W}_{hk} = \Big\{ V = (V_1, V_2) : \; & V(\cdot,t) \text{ continuous in } \Omega, \\
& V(\cdot,t)|_{I_n} \in (V_h^n)^2, \\
& V(x,t)|_{I_n} \text{ piecewise constant in } t \Big\}.
\end{aligned} \tag{3.3}$$

We note that global continuity of the trail functions in $\mathcal{V}_{hk}$ requires the use of '*hanging nodes*' if the spatial mesh changes across a time level t_n. We allow one hanging node per edge or face.

Remark 3. If we do not change the spatial mesh or just refine the spatial mesh from one time level to the next one, i.e.,

$$V_h^{n-1} \subset V_h^n, \quad n = 1, \dots, N, \tag{3.4}$$

then we have $\bar{V}_h^n = V_h^n$.

In the construction of $\mathcal{V}_{hk}$ we have associated the triangulation $\mathcal{T}_h^n$ with discrete time levels instead of the time slabs Ω^n, and in the interior of time slabs we let U be from the union of the finite element spaces defined on the triangulations at the two adjacent time levels. This construction is necessary to allow for trial functions that are continuous also at the discrete time leveles even if grids change between time steps. Associating triangulation with time slabs instead of time levels would yield a variant scheme which includes jump terms due to discontinuity at discrete time leveles, when coarsening happens. This means that there are extra degrees of freedom that one might use suitable projections for transfering solution at the time levels t_n, see [7].

The continuous Galerkin method, based on the variational fomulation (2.3), is to find $U \in \mathcal{V}_{hk}$ such that,

$$\mathcal{A}(U, W) = F(W), \quad \forall\, W \in \mathcal{W}_{hk}. \tag{3.5}$$

The Galerkin orthogonality, with $u = (u_1, u_2)$ being the exact solution of (2.3), is then,

$$\mathcal{A}(U - u, W) = 0, \quad \forall\, W \in \mathcal{W}_{hk}. \tag{3.6}$$

Similarly the continuous Galerkin method, based on the dual variational formulation (2.5), is to find $Z \in \mathcal{V}_{hk}$ such that,

$$\mathcal{A}^*(W, Z) = J(W), \quad \forall\, W \in \mathcal{W}_{hk}. \tag{3.7}$$

Then, Z also satisfies, for $n = 0, 1, \ldots, N-1$,

$$\mathcal{A}^*_{t_n}(W, Z) = J_{t_n}(W), \quad \forall\, W \in \mathcal{W}_{hk}. \tag{3.8}$$

We notice that, rather than using the dual formulation of the discrete problem (3.5), we formulated the same finite element method for the continuous dual problem (2.5).

From (3.5) we can recover the time stepping scheme,

$$\begin{aligned}
&\int_{I_n} \Big\{ (\dot{U}_1, W_1) - (U_2, W_1) \Big\}\, dt = 0,\\
&\int_{I_n} \Big\{ \rho(\dot{U}_2, W_2) + a(U_1, W_2) - \int_0^t \beta(t-s) a\big(U_1(s), W_2(t)\big)\, ds \Big\}\, dt\\
&\qquad\qquad = \int_{I_n} \big\{ (f, W_2)\, dt + (g, W_2)_{\Gamma_N} \big\}\, dt, \quad \forall\, W_1, W_2 \in \mathcal{W}_{hk},\\
&U_1(0) = u_h^0, \quad U_2(0) = v_h^0,
\end{aligned} \tag{3.9}$$

for suitable choice of $u_h^0, v_h^0 \in V_h^0$ as approximations of the initial data u^0, v^0. Here, as a natural choice, we have

$$u_h^0 = \mathcal{P}_h u^0, \quad v_h^0 = \mathcal{P}_h v^0. \tag{3.10}$$

Typical functions $U = (U_1, U_2) \in \mathcal{V}_{hk}$, $W = (W_1, W_2) \in \mathcal{W}_{hk}$ are as follows:

$$\begin{aligned}
U_i(x, t_n) &= U_i^n(x) = \sum_{j=1}^{m_n} U_{i,j}^n \varphi_j^n(x),\\
U_i(x, t)|_{I_n} &= \psi_{n-1}(t) U_i^{n-1}(x) + \psi_n(t) U_i^n(x),\\
W_i(x, t)|_{I_n} &= \sum_{j=1}^{m_n} W_{i,j}^n \varphi_j^n(x),
\end{aligned} \tag{3.11}$$

where m_n is the number of degrees of freedom in $\mathcal{T}_h^n$, $\{\varphi_j^n(x)\}_{j=1}^{m_n}$ are the nodal basis functions for V_h^n defined on triangulation $\mathcal{T}_h^n$, and $\psi_n(t)$ is the nodal basis function defined at time level t_n. Hence (3.9) yields

$$
\begin{aligned}
&(U_1^n - U_1^{n-1}, W_1) - \frac{k_n}{2}\big(U_2^n + U_2^{n-1}, W_1) = 0,\\
&\rho(U_2^n - U_2^{n-1}, W_2) + \frac{k_n}{2}(U_1^n + U_1^{n-1}, W_2)\\
&\qquad\qquad - \sum_{l=1}^{n} a(U_1^{l-1}, W_2) \int_{I_n}\int_{t_{l-1}}^{t\wedge t_l} \beta(t-s)\psi_{l-1}(s)\,ds\,dt\\
&\qquad\qquad - \sum_{l=1}^{n} a(U_1^{l}, W_2) \int_{I_n}\int_{t_{l-1}}^{t\wedge t_l} \beta(t-s)\psi_{l}(s)\,ds\,dt\\
&\qquad\qquad\quad = \int_{I_n} \big\{(f, W_2)\,dt + (g, W_2)_{\Gamma_N}\big\}\,dt, \quad \forall\, W_1, W_2 \in V_h^n,\\
&U_1^0 = u_h^0, \quad U_2^0 = v_h^0.
\end{aligned}
$$

This implies the discrete linear system,

$$
\begin{aligned}
&M^n\tilde{U}_1^n - \frac{k_n}{2}M^n\tilde{U}_2^n = M^{n-1,n}\tilde{U}_1^{n-1} + \frac{k_n}{2}M^{n-1,n}\tilde{U}_2^{n-1},\\
&\rho M^n\tilde{U}_2^n + (\frac{k_n}{2} - \omega_{n,n}^-)S^n\tilde{U}_1^n = \rho M^{n-1,n}\tilde{U}_2^{n-1} + (-\frac{k_n}{2} + \omega_{n,n-1})S^{n-1,n}\tilde{U}_1^{n-1}\\
&\qquad\qquad + S^{0,n}\tilde{U}_1^0\omega_{n,0}^+ + \sum_{l=1}^{n-2}\omega_{n,l}S^{l,n}\tilde{U}_1^l + B^n,\\
&\tilde{U}_1^0 = u_h^0, \quad \tilde{U}_2^0 = v_h^0,
\end{aligned}
$$

where

$$
\begin{aligned}
\omega_{n,0}^+ &= \int_{I_n}\int_0^{t\wedge t_1} \beta(t-s)\psi_l(s)\,ds\,dt, \quad \omega_{n,n}^- = \int_{I_n}\int_{t_{n-1}}^{t} \beta(t-s)\psi_l(s)\,ds\,dt,\\
\omega_{n,l} &= \int_{I_n}\int_{t_{l-1}}^{t\wedge t_{l+1}} \beta(t-s)\psi_l(s)\,ds\,dt,\\
B^n &= (B_j^n)_j = \Big(\int_{I_n}\big\{(f,\varphi_j) + (g,\varphi_j)_{\Gamma_N}\big\}\,dt\Big)_j,\\
M^n &= (M_{ij}^n)_{ij} = \big((\varphi_i^n, \varphi_j^n)\big)_{ij}, \quad M^{n-1,n} = (M_{ij}^{n-1,n})_{ij} = \big((\varphi_i^{n-1}, \varphi_j^n)\big)_{ij},\\
S^{l,n} &= (S_{ij}^{l,n})_{ij} = \big(a(\varphi_i^l, \varphi_j^n)\big)_{ij},
\end{aligned}
$$

and $\tilde{U}_i^n = (U_{i,j}^n)_{j=1}^{m_n}$ with $U_{i,j}^n$ introduced in (3.11).

We define the orthogonal projections $\mathcal{R}_{h,n}: V \to V_h^n$, $\mathcal{P}_{h,n}: H \to V_h^n$ and $\mathcal{P}_{k,n}: L_2(I_n)^d \to \mathbb{P}_0^d(I_n)$, respectively, by

$$
\begin{aligned}
a(\mathcal{R}_{h,n}v - v, \chi) = 0, \quad \forall v \in V,\, \chi \in V_h^n,\\
(\mathcal{P}_{h,n}v - v, \chi) = 0, \quad \forall v \in H,\, \chi \in V_h^n,\\
\int_{I_n} (\mathcal{P}_{k,n}v - v) \cdot \psi \, dt = 0, \quad \forall v \in L_2(I_n)^d,\, \psi \in \mathbb{P}_0^d(I_n)\,,
\end{aligned}
\tag{3.12}
$$

with $\mathbb{P}_0^d$ denoting the set of all vector-valued constant polynomials. Correspondingly, we define $\mathcal{R}_h v$, $\mathcal{P}_h v$ and $\mathcal{P}_k v$ for $t \in I_n\,(n = 1, \cdots, N)$, by $(\mathcal{R}_h v)(t) = \mathcal{R}_{h,n}v(t)$, $(\mathcal{P}_h v)(t) = \mathcal{P}_{h,n}v(t)$, and $\mathcal{P}_k v = \mathcal{P}_{k,n}(v|_{I_n})$.

Remark 4. In the case of assumption (3.4), by Remark 3 and the definition of the L_2-projection $\mathcal{P}_k$, we have $\dot{V}$, $\mathcal{P}_k V \in \mathcal{W}_{hk}$, for any $V \in \mathcal{V}_{hk}$.

We introduce the linear operator $A_{n,r}: V_h^r \to V_h^n$ by

$$
a(v_r, w_n) = (A_{n,r}v_r, w_n), \quad \forall v_r \in V_h^r,\, w_n \in V_h^n.
$$

We set $A_n = A_{n,n}$, with discrete norms

$$
\|v_n\|_{h,l} = \|A_n^{l/2} v_n\| = \sqrt{(v_n, A_n^l v_n)}, \quad v_n \in V_h^n \text{ and } l \in \mathbb{R}\,,
$$

and A_h so that $A_h v = A_n v$ for $v \in V_h^n$. We use $\bar{A}_h$ when it acts on $\bar{V}_h^n$. For later use in our error analysis we note that $\mathcal{P}_h A = A_h \mathcal{R}_h$.

Theorem 3. *Let Z be the solution of* (3.7) *with sufficiently smooth data z_1^T, z_2^T, j_1, j_2. Further, we assume* (3.4)*. Then for $l \in \mathbb{R}$, we have the identity,*

$$
\begin{aligned}
&\|Z_1(t_n)\|_{h,l}^2 + \rho\tilde{\gamma}\|Z_2(t_n)\|_{h,l+1}^2 + 2\rho \int_{t_n}^T \int_t^T \xi(s-t) a\big(A_h^l \dot{Z}_2(t), \dot{Z}_2(s)\big)\, ds\, dt\\
&\quad = \|Z_1(T)\|_{h,l}^2 + \rho(1+\gamma)\|Z_2(T)\|_{h,l+1}^2\\
&\qquad + 2\int_{t_n}^T (A_h^l Z_1, \mathcal{P}_k\mathcal{P}_h j_1)\, dt + 2\int_{t_n}^T (A_h^{l+1} Z_2, \mathcal{P}_k\mathcal{P}_h j_2)\, dt\\
&\qquad - 2\int_{t_n}^T \int_t^T \beta(s-t) a\big(A_h^l \mathcal{P}_k\mathcal{P}_h j_2, Z_2(s)\big)\, ds\, dt\\
&\qquad - 2\rho \int_{t_n}^T \beta(T-t) a\big(A_h^l Z_2(t), Z_2(T)\big)\, dt\\
&\qquad - 2\rho\xi(T-t_n) a\big(A_h^l Z_2(t_n), Z_2(T)\big).
\end{aligned}
\tag{3.13}
$$

where $\tilde{\gamma} = 1 - \gamma$. Moreover, for some constant $C = C(\rho, \gamma, T)$, we have stability estimate

(3.14)
$$\|Z_1(t_n)\|_{h,l} + \|Z_2(t_n)\|_{h,l+1} \leq C\Big\{ \|\mathcal{P}_h z_1^T\|_{h,l} + \|\mathcal{P}_h z_2^T\|_{h,l+1} + \int_{t_n}^T \Big(\|\mathcal{P}_h j_1\|_{h,l} + \|\mathcal{P}_h j_2\|_{h,l+1} \Big)\, dt \Big\}.$$

Proof. The solution Z of (3.7) also satisfies (3.8), for $n = N-1, \dots, 1, 0$. Then recalling Remark 4 for the assumption (3.4), we obviously have,

$$\mathcal{P}_k Z_1 = -\rho \dot{Z}_2 - \mathcal{P}_k \mathcal{P}_h j_2. \tag{3.15}$$

Using this in (3.8) and recalling the initial data $Z_i(T) = \mathcal{P}_h z_i^T,\ i = 1, 2$, we obtain

$$\int_{t_n}^T \Big\{ -(W_1, \dot{Z}_1) + a(W_1, Z_2) - \int_t^T \beta(s-t) a\big(W_1, Z_2(s)\big)\, ds \Big\}\, dt + \big(W_1(T), \mathcal{P}_h z_1^T\big) + \rho\big(W_2(T), \mathcal{P}_h z_2^T\big) = \int_{t_n}^T (W_1, j_1)\, dt + \big(W_1(T), z_1^T\big) + \rho\big(W_2(T), z_2^T\big).$$

The terms concerning the initial data are canceled by the definition of the orthogonal projection $\mathcal{P}_h$. Besides, for the convolution term we recall $\beta(s - t) = -D_s \xi(s-t)$ from (2.7) and then partial integration yields,

$$-\int_{t_n}^T \int_t^T \beta(s-t) a\big(W_1, Z_2(s)\big)\, ds\, dt = -\int_{t_n}^T \int_t^T \xi(s-t) a\big(W_1, \dot{Z}_2(s)\big)\, ds\, dt + \int_{t_n}^T \xi(T-t) a\big(W_1, Z_2(T)\big)\, dt - \gamma \int_{t_n}^T a\big(W_1, Z_2(t)\big)\, dt.$$

These and $\tilde{\gamma} = 1 - \gamma$ imply that the solution Z satisfies,

$$\int_{t_n}^T \Big\{ -(W_1, \dot{Z}_1) + \tilde{\gamma} a(W_1, Z_2) - \int_t^T \xi(s-t) a\big(W_1, \dot{Z}_2(s)\big)\, ds + \xi(T-t) a(W_1, Z_2(T)) \Big\}\, dt = \int_{t_n}^T (W_1, \mathcal{P}_h j_1)\, dt.$$

Now we set $W_1 = A_h^l \mathcal{P}_k Z_1$, and we have

$$\begin{aligned}\int_{t_n}^T \Big\{ &- (A_h^l \mathcal{P}_k Z_1, \dot{Z}_1) + \tilde{\gamma} a(A_h^l \mathcal{P}_k Z_1, Z_2) \\ &- \int_t^T \xi(s-t) a\big(A_h^l \mathcal{P}_k Z_1, \dot{Z}_2(s)\big)\, ds \\ &+ \xi(T-t) a\big(A_h^l \mathcal{P}_k Z_1, Z_2(T)\big) \Big\}\, dt = \int_{t_n}^T (A_h^l \mathcal{P}_k Z_1, \mathcal{P}_h j_1)\, dt.\end{aligned} \tag{3.16}$$

We study the four terms at the left side of the above equation. For the first term we have

$$\begin{aligned}\int_{t_n}^T -(A_h^l \mathcal{P}_k Z_1, \dot{Z}_1)\, dt &= -\frac{1}{2} \int_{t_n}^T D_t \|Z_1(t)\|_{h,l}^2\, dt \\ &= -\frac{1}{2} \|Z_1(T)\|_{h,l}^2 + \frac{1}{2} \|Z_1(t_n)\|_{h,l}^2.\end{aligned} \tag{3.17}$$

With (3.15) we can write the second term as

$$\begin{aligned}\tilde{\gamma} \int_{t_n}^T a(A_h^l \mathcal{P}_k Z_1, Z_2)\, dt &= -\rho\tilde{\gamma} \int_{t_n}^T a(A_h^l \dot{Z}_2, Z_2)\, dt \\ &\quad - \tilde{\gamma} \int_{t_n}^T a(A_h^l \mathcal{P}_k \mathcal{P}_h j_2, Z_2)\, dt \\ &= -\frac{\rho\tilde{\gamma}}{2} \|Z_2(T)\|_{h,l+1}^2 + \frac{\rho\tilde{\gamma}}{2} \|Z_2(t_n)\|_{h,l+1}^2 \\ &\quad - \tilde{\gamma} \int_{t_n}^T a(A_h^l \mathcal{P}_k \mathcal{P}_h j_2, Z_2)\, dt.\end{aligned} \tag{3.18}$$

For the third term in (3.16), by virtue of (3.15) and integration by parts, we obtain

$$
\begin{aligned}
-\int_{t_n}^{T}\int_{t}^{T} \xi(s-t)a\big(A_h^l\mathcal{P}_k Z_1, \dot{Z}_2(s)\big)\,ds\,dt \\
&= \rho\int_{t_n}^{T}\int_{t}^{T} \xi(s-t)a\big(A_h^l\dot{Z}_2(t), \dot{Z}_2(s)\big)\,ds\,dt \\
&\quad + \int_{t_n}^{T}\int_{t}^{T} \xi(s-t)a\big(A_h^l\mathcal{P}_k\mathcal{P}_h j_2, \dot{Z}_2(s)\big)\,ds\,dt \\
&= \rho\int_{t_n}^{T}\int_{t}^{T} \xi(s-t)a\big(A_h^l\dot{Z}_2(t), \dot{Z}_2(s)\big)\,ds\,dt \\
&\quad + \int_{t_n}^{T}\int_{t}^{T} \beta(s-t)a\big(A_h^l\mathcal{P}_k\mathcal{P}_h j_2, Z_2(s)\big)\,ds\,dt \\
&\quad + \int_{t_n}^{T} \xi(T-t)a(A_h^l\mathcal{P}_k\mathcal{P}_h j_2, Z_2(T))\,dt \\
&\quad - \gamma\int_{t_n}^{T} a\big(A_h^l\mathcal{P}_k\mathcal{P}_h j_2, Z_2(t)\big)\,dt.
\end{aligned} \tag{3.19}
$$

Finally, for the last term at the left side of (3.16), we use (3.15) and integration by parts to have

$$
\begin{aligned}
\int_{t_n}^{T} \xi(T-t)a\big(A_h^l\mathcal{P}_k Z_1, Z_2(T)\big)\,dt &= -\rho\int_{t_n}^{T} \xi(T-t)a\big(A_h^l\dot{Z}_2(t), Z_2(T)\big)\,dt \\
&\quad - \int_{t_n}^{T} \xi(T-t)a\big(A_h^l\mathcal{P}_k\mathcal{P}_h j_2, Z_2(T)\big)\,dt \\
&= \rho\int_{t_n}^{T} \beta(T-t)a\big(A_h^l Z_2(t), Z_2(T)\big)\,dt \\
&\quad - \rho\gamma\|Z_2(T)\|_{h,l+1}^2 \\
&\quad + \rho\xi(T-t_n)a\big(A_h^l Z_2(t_n), Z_2(T)\big) \\
&\quad - \int_{t_n}^{T} \xi(T-t)a\big(A_h^l\mathcal{P}_k\mathcal{P}_h j_2, Z_2(T)\big)\,dt.
\end{aligned} \tag{3.20}
$$

Putting (3.17)–(3.20) in (3.16) we conclude the identity (3.13).

Now we prove the estimate (3.14). We recall, from (2.8), that ξ is a positive type kernel. Then, using the Cauchy-Schwarz inequality in (3.13), and $\|\beta\|_{L_1(\mathbb{R}^+)} = \gamma,\ \xi(t) \leq \gamma$, we get

$$\begin{aligned}
&\|Z_1(t_n)\|_{h,l}^2 + \rho\tilde{\gamma}\|Z_2(t_n)\|_{h,l+1}^2 \\
&\qquad \leq \|Z_1(T)\|_{h,l}^2 + \rho(1+\gamma)\|Z_2(T)\|_{h,l+1}^2 \\
&\qquad\quad + 2\int_{t_n}^T (A_h^l Z_1, \mathcal{P}_k\mathcal{P}_h j_1)\,dt + 2\int_{t_n}^T (A_h^{l+1} Z_2, \mathcal{P}_k\mathcal{P}_h j_2)\,dt \\
&\qquad\quad - 2\int_{t_n}^T\int_t^T \beta(s-t)a\big(A_h^l\mathcal{P}_k\mathcal{P}_h j_2, Z_2(s)\big)\,ds\,dt \\
&\qquad\quad - 2\rho\int_{t_n}^T \beta(T-t)a\big(A_h^l Z_2(t), Z_2(T)\big)\,dt \\
&\qquad\quad - 2\rho\xi(T-t_n)a\big(A_h^l Z_2(t_n), Z_2(T)\big) \\
&\qquad \leq \|Z_1(T)\|_{h,l}^2 + \rho(1+\gamma)\|Z_2(T)\|_{h,l+1}^2 \\
&\qquad\quad + C_1 \max_{t_n\leq t\leq T}\|Z_1\|_{h,l}^2 + 1/C_1\Big(\int_{t_n}^T \|\mathcal{P}_k\mathcal{P}_h j_1\|_{h,l}\,dt\Big)^2 \\
&\qquad\quad + C_2 \max_{t_n\leq t\leq T}\|Z_2\|_{h,l+1}^2 + 1/C_2\Big(\int_{t_n}^T \|\mathcal{P}_k\mathcal{P}_h j_2\|_{h,l+1}\,dt\Big)^2 \\
&\qquad\quad + C_3\|Z_2(T)\|_{h,l+1}^2 + 1/C_3 \max_{t_n\leq t\leq T}\|Z_2\|_{h,l+1}^2 \\
&\qquad\quad + C_4\|Z_2(T)\|_{h,l+1}^2 + 1/C_4\|Z_2(t_n)\|_{h,l+1}^2.
\end{aligned}$$

Using that, for piecewise linear functions, we have

$$\max_{[0,T]} |U_i| \leq \max_{0\leq n\leq N} |U_i(t_n)|, \tag{3.21}$$

and

$$\int_0^T |\mathcal{P}_k f|\,dt \leq \int_0^T |f|\,dt, \tag{3.22}$$

and that the above inequality holds for arbitrary N, in a standard way, we conclude the estimate inequality (3.14). Now the proof is complete. □

4. A priori error estimates

We define the standard interpolant I_k with $I_k v$ belong to the space of continuous piecewise linear polynomials, and

$$I_k v(t_n) = v(t_n), \quad n = 0,\, 1,\, \cdots,\, N. \tag{4.1}$$

By standard arguments in approximation theory we see that, for $q = 0,\, 1$,

$$\int_0^T \|I_k v - v\|_i\,dt \leq Ck^{q+1}\int_0^T \|D_t^{q+1} v\|_i\,dt, \quad \text{for } i = 0,\, 1, \tag{4.2}$$

where $k = \max_{1\le n\le N} k_n$.

We assume the elliptic regularity estimate $\|v\|_2 \le C\|Av\|,\ \forall\, v \in \mathcal{D}(A)$, so that the following error estimates for the Ritz projection (3.12), hold true

$$\|\mathcal{R}_h v - v\| \le Ch^s\|v\|_s\,, \quad \forall v \in H^s \cap V, \quad s = 1,\, 2. \tag{4.3}$$

Hence we must specialize to the pure Dirichlet boundary condition and a convex polygonal domain. We note that the energy norm $\|\cdot\|_V$ is equinalent to $\|\cdot\|_1$ on V.

Theorem 4. *Assume that $\Gamma_N = \varnothing$, Ω is a convex polygonal domain, and* (3.4)*. Let u and U be the solutions of* (2.3) *and* (3.5)*. Then, with $e = U - u$ and $C = C(\rho, \gamma, T)$, we have*

$$\begin{aligned}\|e_1(T)\| \le Ch^2\Big(\|u^0\|_2 + \|u_1(T)\|_2 + \int_0^T \|\dot u_1\|_2\, dt\Big)\\ + Ck^2\int_0^T \big(\|\ddot u_2\| + \|\ddot u_1\|_1\big)\, dt.\end{aligned}$$

Proof. We recall Remark 4 for the assumption (3.4). We set $e = U - u = \theta + \eta + \omega$ with

$$\begin{aligned}\theta_1 &= U_1 - I_k\mathcal{R}_h u_1, \quad \eta_1 = (I_k - I)\mathcal{R}_h u_1, \quad \omega_1 = (\mathcal{R}_h - I)u_1,\\ \theta_2 &= U_2 - I_k\mathcal{P}_h u_2, \quad \eta_2 = (I_k - I)\mathcal{P}_h u_2, \quad \omega_2 = (\mathcal{P}_h - I)u_2.\end{aligned}$$

Now, putting $W = \mathcal{P}_k\theta$ in (3.7) we have

$$J(\mathcal{P}_k\theta) = \mathcal{A}^*(\mathcal{P}_k\theta, Z),$$

where by definition

$$\begin{aligned}J(\mathcal{P}_k\theta) = \int_0^T \big\{(\mathcal{P}_k\theta_1, j_1) + (\mathcal{P}_k\theta_2, j_2)\big\}\, dt\\ + \big((\mathcal{P}_k\theta_1)(T), z_1^T\big) + \rho\big((\mathcal{P}_k\theta_2)(T), z_2^T\big),\end{aligned}$$

and by partial integration

$$\begin{aligned}\mathcal{A}^*(\mathcal{P}_k\theta, Z) = \mathcal{A}(\theta, \mathcal{P}_k Z) + \big((\mathcal{P}_k\theta_1)(T), Z_1(T)\big) - \big(\theta_1(T), Z_1(T)\big)\\ + \rho\big((\mathcal{P}_k\theta_2)(T), Z_2(T)\big) - \rho\big(\theta_2(T), Z_2(T)\big).\end{aligned}$$

We set $j_1 = j_2 = 0$ and $z_2^T = 0,\ \mathcal{P}_h z_1^T = \theta_1(T)$, and we recall that $Z_i(T) = \mathcal{P}_h z_i^T,\ i = 1, 2$. Hence using the definition of the orthogonal projection $\mathcal{P}_h$ we have

$$\|\theta_1(T)\|^2 = \mathcal{A}(\theta, \mathcal{P}_k Z),$$

that, using $\theta = e - \eta - \omega$ and the Galerkin orthogonality (3.6), implies,

$$\begin{aligned}
\|\theta_1(T)\|^2 &= -\mathcal{A}(\eta, \mathcal{P}_k Z) - \mathcal{A}(\omega, \mathcal{P}_k Z)\\
&= \int_0^T \Big\{ -(\dot\eta_1, \mathcal{P}_k Z_1) + (\eta_2, \mathcal{P}_k Z_1) + \rho(\dot\eta_2, \mathcal{P}_k Z_2) - a(\eta_1, \mathcal{P}_k Z_2)\\
&\qquad + \int_0^t \beta(t-s) a\big(\eta_1(s), \mathcal{P}_k Z_2\big)\, ds \Big\}\, dt\\
&\quad - \big(\eta_1(0), \mathcal{P}_k Z_1(0)\big) - \rho\big(\eta_2(0), \mathcal{P}_k Z_2(0)\big)\\
&\quad + \int_0^T \Big\{ -(\dot\omega_1, \mathcal{P}_k Z_1) + (\omega_2, \mathcal{P}_k Z_1) + \rho(\dot\omega_2, \mathcal{P}_k Z_2) - a(\omega_1, \mathcal{P}_k Z_2)\\
&\qquad + \int_0^t \beta(t-s) a\big(\omega_1(s), \mathcal{P}_k Z_2\big)\, ds \Big\}\, dt\\
&\quad - \big(\omega_1(0), \mathcal{P}_k Z_1(0)\big) - \rho\big(\omega_2(0), \mathcal{P}_k Z_2(0)\big).
\end{aligned}$$

By the definition of η, that indicates the interpolation error, terms including $\dot\eta_i,\ \eta_i(0)$ vanish. We also use the definition of ω, that indicates the projection error, and we conclude

$$\begin{aligned}
\|\theta_1(T)\|^2 &= \int_0^T \Big\{ (\eta_2, \mathcal{P}_k Z_1) - a(\eta_1, \mathcal{P}_k Z_2) + \int_0^t \beta(t-s) a\big(\eta_1(s), \mathcal{P}_k Z_2\big)\, ds \Big\}\, dt\\
&\quad - \int_0^T (\dot\omega_1, \mathcal{P}_k Z_1)\, dt - \big(\omega_1(0), \mathcal{P}_k Z_1(0)\big),
\end{aligned}$$

that by the Cauchy-Schwarz inequality implies

$$\begin{aligned}
\|\theta_1(T)\|^2 \le C_1 \max_{0\le t\le T} \|\mathcal{P}_k Z_1\|^2 &+ 1/C_1 \Big(\int_0^T \|\eta_2\|\, dt \Big)^2\\
+ C_2 \max_{0\le t\le T} \|\mathcal{P}_k Z_2\|_1^2 &+ 1/C_2 \Big(\int_0^T \|\eta_1\|_1\, dt \Big)^2\\
+ C_3 \max_{0\le t\le T} \|\mathcal{P}_k Z_2\|_1^2 &+ 1/C_3 \Big(\int_0^T \big(\beta * \|\eta_1\|_1\big)(t)\, dt \Big)^2\\
+ C_4 \max_{0\le t\le T} \|\mathcal{P}_k Z_1\|^2 &+ 1/C_4 \Big(\int_0^T \|\dot\omega_1\|\, dt \Big)^2\\
+ C_5 \|\mathcal{P}_k Z_1(0)\|^2 &+ 1/C_5 \|\omega_1(0)\|^2.
\end{aligned}$$

Using (3.21), $\|\beta\|_{L_1(\mathbb{R}^+)} = \gamma$, and the stability estimate (3.14) with $l = 0$, in a standard way, we have

$$\|\theta_1(T)\| \le C\Big\{ \|\omega_1(0)\| + \int_0^T \Big(\|\eta_2\| + \|\eta_1\|_1 + \|\dot\omega_1\| \Big)\, dt \Big\}.$$

Recalling $\theta_1(T) = e_1(T) - \omega_1(T)$, and stability of the projections $\mathcal{P}_h$, $\mathcal{R}_h$ with respect to $\|\cdot\|, \|\cdot\|_1$, respectively, we have

$$\|\theta_1(T)\| \leq C\Big\{\|(\mathcal{R}_h - I)u^0\| + \|(\mathcal{R}_h - I)u_1(T)\| \\ + \int_0^T \Big(\|(I_k - I)u_2\| + \|(I_k - I)u_1\|_1 + \|(\mathcal{R}_h - I)\dot{u}_1\|\Big)\,dt\Big\}.$$

This completes the proof by (4.2), (4.3). □

5. A posteriori error estimates

Having certain regularity on the data, i.e., initial data u^0, v^0 and the force terms f, g, there are still two types of limitation for higher global regularity of a weak solution of (1.1). One is due to the mixed Dirichlet-Neumann boundary condition. This type of boundary condition are natural in practice, and a pure Dirichlet boundary condition can not be realistic in applications. Other limitation is the singularity of the convolution kernel β. This means that even with the pure Dirichlet boundary condition, higher regularity of a weak solution is limited, see [7], [11], though with smoother kernels we can get higher regularity. Besides, the stability and a priori error estimates presented in Theorem 3 and Theorem 4 do not admit adaptive meshes. These, and other general motivations such as no practical use of a priori error estimates, call for adaptive meshes based on a posteriori error analysis.

Here a space-time cellwise error representation is given. The main framework is adapted from [3], and a general linear goal functional $J(\cdot)$ is used. This error representation can be used for goal-oriented adaptive strategies based on dual weighted residual method. For more details on dual weighted residual method and its practical aspects for differential equations, see [3] and references therein.

Theorem 5. *Let u and U be the solutions of* (2.3) *and* (3.5)*, and $J(\cdot)$ the linear functional defined in* §2*. Then, with $e = U - u$, we have the error representation*

$$J(e) = \sum_{K \in \mathcal{T}_h^0} \Theta_{0,K} + \sum_{n=1}^{N} \sum_{K \in \bar{\mathcal{T}}_h^n} \sum_{i=1}^{6} \Theta_{i,K}^n, \tag{5.1}$$

where, with $z_{hk} \in \mathcal{W}_{hk}$ *being an approximation of the dual solution* z *and* $E_{hk}z = z_{hk} - z$ *being the error operator,*

$$
\begin{aligned}
&\Theta_{0,K} = \big(U_1(0) - u^0, E_{hk}z_1(0)\big)_K + \rho\big(U_2(0) - v^0, E_{hk}z_2(0)\big)_K,\\
&\Theta^n_{1,K} = (\dot U_1 - U_2, E_{hk}z_1)_{K^n}, \quad \Theta^n_{2,K} = (\rho\dot U_2 - f, E_{hk}z_2)_{K^n},\\
&\Theta^n_{3,K} = (r_h, E_{hk}z_2)_{\partial K^n}, \qquad\quad \Theta^n_{4,K} = (g_h - g, E_{hk}z_2)_{\partial K^n},\\
&\Theta^n_{5,K} = -\Big(r_h, \int_s^T \beta(t-s)E_{hk}z_2(t)\,dt\Big)_{\partial K^n},\\
&\Theta^n_{6,K} = -\Big(g_h, \int_s^T \beta(t-s)E_{hk}z_2(t)\,dt\Big)_{\partial K^n}.
\end{aligned}
\tag{5.2}
$$

Here $K^n = K \times I_n$ *and* $\partial K^n = \partial K \times I_n$ *are the space-time cells, and* r_h, g_h *are defined below,* (5.6), (5.7), *respectively.*

It should be noticed that ∂K^n is not the boundary of K^n.

Proof. Using the identity (2.4) and the Galerkin orthogonality (3.6) we have,

$$
\begin{aligned}
J(e) &= \mathcal{A}^*(e,z) = \mathcal{A}(e,z) = \mathcal{A}(e, E_{hk}z)\\
&= \mathcal{A}(U, E_{hk}z) - F(E_{hk}z) = \mathrm{R}(U; E_{hk}z),
\end{aligned}
\tag{5.3}
$$

where $\mathrm{R}(U;\cdot)$ is the residual of the Galerkin approximation U as a functional on the solution space $\mathcal{V}^*$. Then by the definition of $\mathcal{A},\ F$ we have

$$
\begin{aligned}
J(e) &= \big(U_1(0) - u^0, E_{hk}z_1(0)\big) + \rho\big(U_2(0) - v^0, E_{hk}z_2(0)\big)\\
&\quad + \int_0^T \Big\{(\dot U_1, E_{hk}z_1) - (U_2, E_{hk}z_1) + \rho(\dot U_2, E_{hk}z_2) + a(U_1, E_{hk}z_2)\\
&\qquad - \int_0^t \beta(t-s)a\big(U_1(s), E_{hk}z_2(t)\big)\,ds\Big\}\,dt\\
&\quad - \int_0^T \Big\{(f, E_{hk}z_2) + (g, E_{hk}z_2)_{\Gamma_{\mathrm{N}}}\Big\}\,dt.
\end{aligned}
\tag{5.4}
$$

Now, by partial integration with respect to the space variable, we obtain

$$
\begin{aligned}
\int_0^T a(U_1, z_2)\, dt &= \sum_{n=1}^{N} \int_{I_n} \sum_{K\in\bar{\mathcal{T}}_h^n} a(U_1, z_2)_K\, dt \\
&= \sum_{n=1}^{N} \int_{I_n} \sum_{K\in\bar{\mathcal{T}}_h^n} \big(\sigma_0(U_1)\cdot n, z_2\big)_{\partial K}\, dt \\
&= \sum_{n=1}^{N} \int_{I_n} \Big\{ \sum_{E\in\mathcal{E}_I^n} \big(-[\sigma_0(U_1)\cdot n], z_2\big)_E \\
&\qquad + \sum_{E\in\mathcal{E}_{\Gamma_\mathrm{N}}^n} \big(\sigma_0(U_1)\cdot n, z_2\big)_E \Big\}\, dt \\
&= \sum_{n=1}^{N} \int_{I_n} \sum_{K\in\bar{\mathcal{T}}_h^n} \big\{(r_h, z_2)_{\partial K} + (g_h, z_2)_{\partial K}\big\}\, dt \\
&= \sum_{n=1}^{N} \sum_{K\in\bar{\mathcal{T}}_h^n} \big\{(r_h, z_2)_{\partial K^n} + (g_h, z_2)_{\partial K^n}\big\}
\end{aligned} \tag{5.5}
$$

where $\mathcal{E}_I^n$, $\mathcal{E}_{\Gamma_\mathrm{N}}^n$ are, respectively, the sets of the interior edges and the edges on the Neumann boundary, corresponding to the triangulation $\bar{\mathcal{T}}_h^n$. Here r_h are the residuals representing the jumps of the normal derivatives $\sigma_0(U_1)\cdot n$, and determined by,

$$
r_h|_\Gamma = \begin{cases} -\frac{1}{2}[\sigma_0(U_1)\cdot n] & \text{if } \Gamma\subset \partial K\setminus\partial\Omega, \\ 0 & \text{if } \Gamma\subset\partial\Omega, \end{cases} \tag{5.6}
$$

and g_h is the contribution from the Neumann boundary defined as

$$
g_h|_\Gamma = \begin{cases} \sigma_0(U_1)\cdot n & \text{if } \Gamma\subset \partial K\cap\Gamma_\mathrm{N}, \\ 0 & \text{if } \Gamma\subset\partial\Omega. \end{cases} \tag{5.7}
$$

For the convolution term in (5.4) we first change the order of the time integrals, then similar to (5.5), we have,

$$
\begin{aligned}
&\int_0^T\int_0^t \beta(t-s)a\big(U_1(s),z_2(t)\big)\,ds\,dt\\
&\qquad\qquad = \int_0^T\int_s^T \beta(t-s)a\big(U_1(s),z_2(t)\big)\,dt\,ds\\
&\qquad\qquad = \sum_{n=1}^N\int_{I_n}\sum_{K\in\bar{\mathcal{T}}_h^n} a\Big(U_1(s),\int_s^T\beta(t-s)z_2(t)\,dt\Big)_K\,ds\\
&\qquad\qquad = \sum_{n=1}^N\sum_{K\in\bar{\mathcal{T}}_h^n}\Big\{\Big(r_h,\int_s^T\beta(t-s)z_2(t)\,dt\Big)_{\partial K^n}\\
&\qquad\qquad\qquad + \Big(g_h,\int_s^T\beta(t-s)z_2(t)\,dt\Big)_{\partial K^n}\Big\}.
\end{aligned}
\tag{5.8}
$$

Now, using (5.5), (5.8) and space-time cellwise representation of the other terms in (5.4) we conclude the error representation (5.1). □

The error representation (5.1) leads us to the weighted a posteriori estimate,

$$
J(e)\le \sum_{K\in\mathcal{T}_h^0}\mathrm{R}_{0,K}\omega_{0,K}+\sum_{n=1}^N\sum_{K\in\bar{\mathcal{T}}_h^n}\sum_{i=1}^3\mathrm{R}_{i,K}^n\omega_{i,K}^n,
\tag{5.9}
$$

with the residuals and weights defined as

$$
\begin{aligned}
\mathrm{R}_{0,K} &= \big(\|U_1(0)-u^0\|_K^2+\|U_2(0)-v^0\|_K^2\big)^{1/2},\\
\omega_{0,K} &= \big(\|E_{hk}z_1(0)\|_K^2+\|E_{hk}z_2(0)\|_K^2\big)^{1/2},\\
\mathrm{R}_{1,K}^n &= \|\dot U_1-U_2\|_{K^n},\quad \omega_{1,K}^n=\|E_{hk}z_1\|_{K^n},\\
\mathrm{R}_{2,K}^n &= \Big(\|\rho\dot U_2-f\|_{K^n}^2+2\bar h_K^{-1}\|r_h\|_{\partial K^n}^2\Big)^{1/2},\\
\omega_{2,K}^n &= \bigg(\|E_{hk}z_2\|_{K^n}^2+\bar h_K\|E_{hk}z_2\|_{\partial K^n}^2\\
&\qquad +\bar h_K\Big\|\int_s^T\beta(t-s)E_{hk}z_2(t)\,dt\Big\|_{\partial K^n}^2\bigg)^{1/2},\\
\mathrm{R}_{3,K}^n &= \Big(\bar h_K^{-1}\|g_h-g\|_{\partial K^n}^2+\bar h_K^{-1}\|g_h\|_{\partial K^n}^2\Big)^{1/2},\\
\omega_{3,K}^n &= \Big(\bar h_K\|E_{hk}z_2\|_{\partial K^n}^2+\bar h_K\Big\|\int_s^T\beta(t-s)E_{hk}z_2(t)\,dt\Big\|_{\partial K^n}^2\Big)^{1/2}.
\end{aligned}
$$

In order to evaluate the a posteriori error representation (5.1) or the a posteriori estimate (5.9), we need information about the continuous dual solution z. Such information has to be obtained either through a priori analysis in form of bounds for z in certain Sobolev norms or through computation by solving the dual problem numerically. In this context we provide information through a priori analysis and we leave the investigation on the second case to a latter work.

In the following, the target functional $J(\cdot)$ will be the global L_2-norm of the approximation displacement $u_1 = u_1(x,t)$. We first present a weighted global a posteriori error estimate, using global L_2-projections $\mathcal{P}_k, \mathcal{P}_h$ defined in (3.12), and error estimates of $\mathcal{P}_h$ in a weighted L_2-norm.

We recall the weighted global error estimates of the L_2-projection $\mathcal{P}_h$ (3.12), see [4]. First we recall some notation. Let $\mathcal{T}$ be a given triangulation with mesh function h, and for any simplex $K \in \mathcal{T}$, ρ_K denote the radius of the largest ball contained in the closure of K, that is $\bar{K}$. A family $\mathcal{F}$ of triangulations $\mathcal{T}$ is called non-degenerate, if there exsist a constant c_0 such that we have

$$c_0 = \max_{\mathcal{T}\in\mathcal{F}} \max_{K\in\mathcal{T}} \frac{h_K}{\rho_K}.$$

Let $S_K = \{K' \in \mathcal{T} : \bar{K}' \cap \bar{K} \neq \varnothing\}$ and $\delta_{\mathcal{T}}$ be a measure for the given triangulation $\mathcal{T}$ defined by

$$\delta_{\mathcal{T}} = \max_{K\in\mathcal{T}} \max_{K'\in S_K} |1 - h_{K'}^2/h_K^2|.$$

We define a measure, $\delta_{\mathcal{F}}$, for a given family $\mathcal{F}$, by

$$\delta_{\mathcal{F}} = \max_{\mathcal{T}\in\mathcal{F}} \delta_{\mathcal{T}}. \tag{5.10}$$

We define the error operators E_{hk}, E_h, and E_k by

$$E_{hk}v = (\mathcal{P}_k\mathcal{P}_h - I)v, \quad E_h v = (\mathcal{P}_h - I)v, \quad E_k v = (\mathcal{P}_k - I)v, \tag{5.11}$$

and we note that

$$E_{hk} = E_h + E_k\mathcal{P}_h. \tag{5.12}$$

Lemma 1. *Assume that the family $\mathcal{F}$ of traingulations $\mathcal{T}$ be non-degenerate. Then for sufficiently small $\delta_{\mathcal{F}}$, there exists a constant C such that for any triangulation $\mathcal{T} \in \mathcal{F}$ we have, for all $v \in H^2$,*

$$\|h^{-s}E_h v\| \le C\|\nabla^s v\|, \quad s = 1,\, 2,\ \forall v \in H^s, \tag{5.13}$$

$$\|h^{-1}\nabla E_h v\| \le C\|\nabla^2 v\|, \quad \forall v \in H^2, \tag{5.14}$$

where '∇' denotes the usual gradient.

For more details on the practical aspects of $\delta_{\mathcal{F}}$, see [4].

For the next theorem we recall the mesh functions h_n, $\bar{h}_n$ from (3.1), (3.2), and we define the notations

$$\bar{h}_{min,n} = \min_{K\in\bar{\mathcal{T}}_h^n} \bar{h}_K, \quad h_{max,n} = \max_{K\in\mathcal{T}_h^n} h_K.$$

Theorem 6. *Let u be the solutions of (2.3), and U be the solution of (3.5) with a non-degenerate family $\mathcal{F}_h$ of triangulations $\mathcal{T}_h^n$, $n = 0, 1, \ldots, N$, with sufficiently small $\delta_{\mathcal{F}_h}$, such that the weighted global error estimates (5.13) and (5.14) hold. Then, with $e = U - u$, we have the weighted a posteriori error estimate*

(5.15)

$$\begin{aligned}
\|e_1(T)\| \leq C\Big\{ &\|h_0\big(U_1(0) - u^0\big)\| + \|h_0^2\big(U_2(0) - v^0\big)\| \\
&+ \sum_{n=1}^{N} \int_{I_n} \Big\{ \|h_n(\dot{U}_1 - U_2)\| + \|h_n^2(\rho\dot{U}_2 - f)\| \\
&\quad + (\zeta_n + \zeta_{n,N})\Big(\sum_{K\in\bar{\mathcal{T}}_h^n} \bar{h}_K^3 \|r_h\|_{\partial K}^2\Big)^{1/2} \\
&\quad + \zeta_n\Big(\sum_{K\in\bar{\mathcal{T}}_h^n} \bar{h}_K^3 \|g_h - g\|_{\partial K}^2\Big)^{1/2} + \zeta_{n,N}\Big(\sum_{K\in\bar{\mathcal{T}}_h^n} \bar{h}_K^3 \|g_h\|_{\partial K}^2\Big)^{1/2} \\
&\quad + k_n\|\dot{U}_1 - U_2\| + k_n\|E_k\bar{A}_hU_1\| \\
&\quad + k_n\Big\|E_k \int_0^t \beta(t-s)\bar{A}_hU_1(s)\,ds\Big\| \\
&\quad + k_n\|E_kf\| + k_n\Big(\sum_{K\in\mathcal{T}_h^n} h_K^{-1}\|E_kg\|_{\partial K}^2\Big)^{1/2}\Big\}\,dt\Big\},
\end{aligned}$$

where

$$\zeta_n = \bar{h}_{min,n}^{-2} h_{max,n}^2, \quad \zeta_{n,N} = \bar{h}_{min,n}^{-2} \max_{n\leq j\leq N} h_{max,j}^2. \tag{5.16}$$

Proof. Let $z \in \mathcal{V}^*$ be the solution of the dual problem (2.5). From the definition of the L_2 projections $\mathcal{P}_k, \mathcal{P}_h$ in (3.12) and the test space $\mathcal{W}_{hk}$ in (3.3) we have $\mathcal{P}_k\mathcal{P}_hz \in \mathcal{W}_{hk}$. Therefore, using (5.3) and the error operators (5.11) we have,

$$J(e) = \mathrm{R}(U; E_{hk}z) = \mathrm{R}(U; E_hz) + \mathrm{R}(U; E_k\mathcal{P}_hz), \tag{5.17}$$

where we used (5.12). We study the two terms at the right side of this equation.

For the first term we can write,

$$\begin{aligned}
\mathrm{R}(U;E_h z) &= \big(U_1(0)-u^0, E_h z_1(0)\big) + \rho\big(U_2(0)-v^0, E_h z_2(0)\big) \\
&\quad + \sum_{n=1}^{N} \int_{I_n} \Big\{ \big(\dot{U}_1 - U_2, E_h z_1\big) + \big(\rho \dot{U}_2 - f, E_h z_2\big) \\
&\quad + a(U_1, E_h z_2) - (g, E_h z_2)_{\Gamma_{\mathrm{N}}} \\
&\quad - \int_0^t \beta(t-s) a\big(U_1(s), E_h z_2\big)\, ds \Big\}\, dt.
\end{aligned}$$

Now by partial integration in space, similar to (5.5), (5.8), we have

(5.18)

$$\begin{aligned}
\mathrm{R}(U;E_h z) &= \sum_{K\in\mathcal{T}_h^0} \big\{ \big(U_1(0)-u^0, E_h z_1(0)\big)_K + \rho\big(U_2(0)-v^0, E_h z_2(0)\big)_K \big\} \\
&\quad + \sum_{n=1}^{N} \int_{I_n} \sum_{K\in\bar{\mathcal{T}}_h^n} \Big\{ \big(\dot{U}_1 - U_2, E_h z_1\big)_K + (\rho \dot{U}_2 - f, E_h z_2)_K \\
&\quad + (r_h, E_h z_2)_{\partial K} + (g_h - g, E_h z_2)_{\partial K} \\
&\quad - \Big(r_h, \int_t^T \beta(s-t) E_h z_2(s)\, ds\Big)_{\partial K} \\
&\quad - \Big(g_h, \int_t^T \beta(s-t) E_h z_2(s)\, ds\Big)_{\partial K} \Big\}\, dt \\
&= \sum_{i=1}^{2} \mathcal{I}_i + \sum_{i=1}^{6} \mathcal{II}_i.
\end{aligned}$$

We then, for each term, use the Cauchy-Schwarz inequality twice. First on the local elements $K, \partial K$, to obtain local L_2-norms, and then on the sum over the elements to obtain global norms such that the weighted global error estimates (5.13), (5.14) can be used. For $\mathcal{I}_1$ we have

$$\mathcal{I}_1 \le \|h_0(U_1(0)-u^0)\|\|h_0^{-1} E_h z_1(0)\| \le C\|h_0(U_1(0)-u^0)\|\|\nabla z_1(0)\|, \tag{5.19}$$

and in a similar way we have,

$$\mathcal{I}_2 \le C\|h_0^2(U_2(0)-v^0)\|\|\nabla^2 z_2(0)\|. \tag{5.20}$$

For the next term, using the error estimate (5.13), we have,

$$
\begin{aligned}
\mathcal{I}\mathcal{I}_1 &= \sum_{n=1}^{N} \int_{I_n} \sum_{K\in\mathcal{T}_h^n} \big(\dot{U}_1 - U_2, E_h z_1\big)_K \, dt \\
&\le \sum_{n=1}^{N} \int_{I_n} \|h_n(\dot{U}_1 - U_2)\| \|h_n^{-1} E_h z_1\| \, dt \qquad (5.21)\\
&\le C \max_{[0,T]} \|\nabla z_1(t)\| \sum_{n=1}^{N} \int_{I_n} \|h_n(\dot{U}_1 - U_2)\| \, dt,
\end{aligned}
$$

and similarly we obtain,

$$
\mathcal{I}\mathcal{I}_2 \le C \max_{[0,T]} \|\nabla^2 z_2(t)\| \sum_{n=1}^{N} \int_{I_n} \|h_n^2(\rho\dot{U}_2 - f)\| \, dt. \tag{5.22}
$$

For $\mathcal{I}\mathcal{I}_3$, we first have,

$$
\mathcal{I}\mathcal{I}_3 \le \sum_{n=1}^{N} \int_{I_n} \Big(\sum_{K\in\bar{\mathcal{T}}_h^n} \bar{h}_K^3 \|r_h(t)\|_{\partial K}^2 \Big)^{1/2} \Big(\sum_{K\in\bar{\mathcal{T}}_h^n} \bar{h}_K^{-3} \|E_h z_2(t)\|_{\partial K}^2 \Big)^{1/2} \, dt.
$$

Then by a scaled trace inequality and the weighted global error estimates (5.13), (5.14), we obtain

$$
\begin{aligned}
\sum_{K\in\bar{\mathcal{T}}_h^n} \bar{h}_K^{-3} \|E_h z_2\|_{\partial K}^2 &\le C \sum_{K\in\bar{\mathcal{T}}_h^n} \big\{ \bar{h}_K^{-4} \|E_h z_2\|_K^2 + \bar{h}_K^{-2} \|\nabla E_h z_2\|_K^2 \big\} \\
&\le C \big\{ \bar{h}_{min,n}^{-4} h_{max,n}^4 \|h_n^{-2} E_h z_2\|^2 \\
&\qquad + \bar{h}_{min,n}^{-2} h_{max,n}^2 \|h_n^{-1} \nabla E_h z_2\|^2 \big\} \\
&\le C \bar{h}_{min,n}^{-4} h_{max,n}^4 \|\nabla^2 z_2\|^2.
\end{aligned}
$$

These imply the estimate

$$
\mathcal{I}\mathcal{I}_3 \le C \max_{[0,T]} \|\nabla^2 z_2(t)\| \sum_{n=1}^{N} \int_{I_n} \bar{h}_{min,n}^{-2} h_{max,n}^2 \Big(\sum_{K\in\bar{\mathcal{T}}_h^n} \bar{h}_K^3 \|r_h\|_{\partial K}^2 \Big)^{1/2} \, dt. \tag{5.23}
$$

In a similar way, we have

$$
\mathcal{I}\mathcal{I}_4 \le C \max_{[0,T]} \|\nabla^2 z_2(t)\| \sum_{n=1}^{N} \int_{I_n} \bar{h}_{min,n}^{-2} h_{max,n}^2 \Big(\sum_{K\in\bar{\mathcal{T}}_h^n} \bar{h}_K^3 \|g_h - g\|_{\partial K}^2 \Big)^{1/2} \, dt. \tag{5.24}
$$

Finally we study $\mathcal{II}_5$ and a similar result will hold for $\mathcal{II}_6$. To this end, first we note that,

$$\mathcal{II}_5 \leq \sum_{n=1}^{N} \int_{I_n} \Big(\sum_{K \in \bar{\mathcal{T}}_h^n} \bar{h}_K^3 \|r_h(t)\|_{\partial K}^2 \Big)^{1/2} \times \Big(\sum_{K \in \bar{\mathcal{T}}_h^n} \bar{h}_K^{-3} \Big\| \int_t^T \beta(s-t) E_h z_2(s)\, ds \Big\|_{\partial K}^2 \Big)^{1/2} dt.$$

Then, using Minkowski's inequality, the Cuachy-Schwarz inequality, and the fact that $\|\beta\|_{L_1(\mathbb{R}^+)} = \gamma$, we have

$$\begin{aligned}
\sum_{K \in \bar{\mathcal{T}}_h^n} \bar{h}_K^{-3} \Big\| \int_t^T \beta(s-t) E_h z_2(s)\, ds \Big\|_{\partial K}^2
&\leq \sum_{K \in \bar{\mathcal{T}}_h^n} \bar{h}_K^{-3} \Big(\int_t^T \beta(s-t) \|E_h z_2(s)\|_{\partial K} \Big)^2 ds \\
&\leq \sum_{K \in \bar{\mathcal{T}}_h^n} \bar{h}_K^{-3} \int_t^T \beta(s-t)\, ds \int_t^T \beta(s-t) \|E_h z_2(s)\|_{\partial K}^2\, ds \\
&\leq \gamma \int_t^T \beta(s-t) \sum_{K \in \bar{\mathcal{T}}_h^n} \bar{h}_K^{-3} \|E_h z_2(s)\|_{\partial K}^2\, ds,
\end{aligned}$$

that using a scaled trace inequality and the error estimates (5.13), (5.14), we have

$$\begin{aligned}
\sum_{K \in \bar{\mathcal{T}}_h^n} \bar{h}_K^{-3} \Big\| \int_t^T \beta(s-t) E_h z_2(s)\, ds \Big\|_{\partial K}^2
&\leq C \int_t^T \beta(s-t) \sum_{K \in \bar{\mathcal{T}}_h^n} \big\{ \bar{h}_K^{-4} \|E_h z_2(s)\|_K^2 + \bar{h}_K^{-2} \|\nabla E_h z_2(s)\|_K^2 \big\}\, ds \\
&\leq C \sum_{j=n}^{N} \int_{t \vee t_{j-1}}^{t_j} \beta(s-t) \big\{ \bar{h}_{min,n}^{-4} h_{max,j}^4 \|h_j^{-2} E_h z_2(s)\|^2 \\
&\qquad + \bar{h}_{min,n}^{-2} h_{max,j}^2 \|h_j^{-1} \nabla E_h z_2(s)\|^2 \big\}\, ds \\
&\leq C \sum_{j=n}^{N} \int_{t \vee t_{j-1}}^{t_j} \beta(s-t) \bar{h}_{min,n}^{-4} h_{max,j}^4 \|\nabla^2 z_2(s)\|^2\, ds \\
&\leq C\gamma \max_{[t,T]} \|\nabla^2 z_2(s)\|^2 \bar{h}_{min,n}^{-4} \max_{n \leq j \leq N} h_{max,j}^4.
\end{aligned}$$

Hence we obtain,

$$
\begin{aligned}
\mathcal{II}_5 \leq C \max_{[0,T]} \|\nabla^2 z_2(t)\| \\
\times \sum_{n=1}^{N} \int_{I_n} \bar{h}_{min,n}^{-2} \max_{n\leq j\leq N} h_{max,j}^2 \Big(\sum_{K\in \bar{\mathcal{T}}_h^n} \bar{h}_K^3 \|r_h\|_{\partial K}^2 \Big)^{1/2} dt.
\end{aligned} \tag{5.25}
$$

A similar estimate for $\mathcal{II}_6$ holds, that is,

$$
\begin{aligned}
\mathcal{II}_6 \leq C \max_{[0,T]} \|\nabla^2 z_2(t)\| \\
\times \sum_{n=1}^{N} \int_{I_n} \bar{h}_{min,n}^{-2} \max_{n\leq j\leq N} h_{max,j}^2 \Big(\sum_{K\in \bar{\mathcal{T}}_h^n} \bar{h}_K^3 \|g_h\|_{\partial K}^2 \Big)^{1/2} dt.
\end{aligned} \tag{5.26}
$$

Putting (5.19)–(5.26) in (5.18) we conclude,

$$
\begin{aligned}
\mathrm{R}(U; E_h z) \leq C \max\big(\max_{[0,T]} \|\nabla^2 z_2(t)\|, \max_{[0,T]} \|\nabla z_1(t)\|\big) \\
\times \Big\{ \|h_0\big(U_1(0) - u^0\big)\| + \|h_0^2\big(U_2(0) - v^0\big)\| \\
+ \sum_{n=1}^{N} \int_{I_n} \Big\{ \|h_n(\dot{U}_1 - U_2)\| + \|h_n^2(\rho \dot{U}_2 - f)\| \\
+ \bar{h}_{min,n}^{-2} h_{max,n}^2 \Big(\sum_{K\in \bar{\mathcal{T}}_h^n} \bar{h}_K^3 \|r_h\|_{\partial K}^2 \Big)^{1/2} \\
+ \bar{h}_{min,n}^{-2} h_{max,n}^2 \Big(\sum_{K\in \bar{\mathcal{T}}_h^n} \bar{h}_K^3 \|g_h - g\|_{\partial K}^2 \Big)^{1/2} \\
+ \bar{h}_{min,n}^{-2} \max_{n\leq j\leq N} h_{max,j}^2 \Big(\sum_{K\in \bar{\mathcal{T}}_h^n} \bar{h}_K^3 \|r_h\|_{\partial K}^2 \Big)^{1/2} \\
+ \bar{h}_{min,n}^{-2} \max_{n\leq j\leq N} h_{max,j}^2 \Big(\sum_{K\in \bar{\mathcal{T}}_h^n} \bar{h}_K^3 \|g_h\|_{\partial K}^2 \Big)^{1/2} \Big\} dt \Big\}.
\end{aligned} \tag{5.27}
$$

Now we study the second term in (5.17), that is,

$$\begin{aligned}
\mathrm{R}(U;E_k\mathcal{P}_hz) &= \mathrm{R}(U;E_k\mathcal{P}_hz) \pm \int_0^T \big\{(\mathcal{P}_kf, E_k\mathcal{P}_hz_2) + (\mathcal{P}_kg, E_k\mathcal{P}_hz_2)_{\Gamma_\mathrm{N}}\big\}\,dt\\
&= \big(U_1(0)-u^0, E_k\mathcal{P}_hz_1(0)\big) + \rho\big(U_2(0)-v^0, E_k\mathcal{P}_hz_2(0)\big)\\
&\quad + \sum_{n=1}^N \int_{I_n} \Big\{ \big(\dot U_1 - U_2, E_k\mathcal{P}_hz_1\big)\\
&\qquad + a(U_1, E_k\mathcal{P}_hz_2) - \int_0^t \beta(t-s) a\big(U_1(s), E_k\mathcal{P}_hz_2\big)\,ds\\
&\qquad + \big(\rho\dot U_2 - \mathcal{P}_kf, E_k\mathcal{P}_hz_2\big) - (\mathcal{P}_kg, E_k\mathcal{P}_hz_2)_{\Gamma_\mathrm{N}}\\
&\qquad + (E_kf, E_k\mathcal{P}_hz_2) + (E_kg, E_k\mathcal{P}_hz_2)_{\Gamma_\mathrm{N}} \Big\}\,dt.
\end{aligned}$$

Recalling the initial condition $U_i(0) = \mathcal{P}_hu_i(0)$, $i = 1,2$, the first two terms on the right side vanish. Besides, from the second equation of (3.9) we have, for $W \in V_h^n$,

$$\begin{aligned}
&\int_{I_n} \big\{\rho(\dot U_2, W) - (\mathcal{P}_kf, W) - (\mathcal{P}_kg, W)\big\}\,dt\\
&\qquad = -\int_{I_n} \Big\{ a(\mathcal{P}_kU_1, W) - a\Big(\mathcal{P}_k \int_0^t \beta(t-s)U_1(s)\,ds, W\Big)\Big\}\,dt.
\end{aligned}$$

Hence, we conclude

(5.28)

$$\begin{aligned}
\mathrm{R}(U;E_k\mathcal{P}_hz) = \sum_{n=1}^N \int_{I_n} \Big\{ &\big(\dot U_1 - U_2, E_k\mathcal{P}_hz_1\big)\\
&- a(E_kU_1, E_k\mathcal{P}_hz_2) + a\Big(E_k \int_0^t \beta(t-s)U_1(s)\,ds, E_k\mathcal{P}_hz_2\Big)\\
&+ (E_kf, E_k\mathcal{P}_hz_2) + (E_kg, E_k\mathcal{P}_hz_2)_{\Gamma_\mathrm{N}} \Big\}\,dt.
\end{aligned}$$

For the last term we have,

$$\begin{aligned}
&\sum_{n=1}^N \int_{I_n} (E_kg, E_k\mathcal{P}_hz_2)_{\Gamma_\mathrm{N}}\,dt\\
&\qquad \le \sum_{n=1}^N \int_{I_n} \Big(\sum_{K\in\mathcal{T}_h^n} h_K^{-1} \|E_kg\|_{\partial K}^2 \Big)^{1/2} \Big(\sum_{K\in\mathcal{T}_h^n} h_K \|E_k\mathcal{P}_hz_2\|_{\partial K}^2 \Big)^{1/2}.
\end{aligned}$$

By a scaled trace inequality and local inverse inequality we have,

$$\begin{aligned}\sum_{K\in\bar{\mathcal{T}}_h^n} h_K\|E_k\mathcal{P}_hz_2\|_{\partial K}^2 &\le C\sum_{K\in\mathcal{T}_h^n}\big\{\|E_k\mathcal{P}_hz_2\|_K^2+h_K^2\|\nabla E_k\mathcal{P}_hz_2\|_K^2\big\}\\ &\le C\sum_{K\in\mathcal{T}_h^n}\big\{\|E_k\mathcal{P}_hz_2\|_K^2+\|E_k\mathcal{P}_hz_2\|_K^2\big\}\\ &= C\|E_k\mathcal{P}_hz_2\|^2.\end{aligned}$$

Hence,

$$\sum_{n=1}^N\int_{I_n}(E_kg,E_k\mathcal{P}_hz_2)_{\Gamma_{\mathrm{N}}}\,dt\le C\Big(\sum_{K\in\mathcal{T}_h^n}h_K^{-1}\|E_kg\|_{\partial K}^2\Big)^{1/2}\|E_k\mathcal{P}_hz_2\|.$$

Considering this in (5.28) and using the Cauchy-Schwarz inequality we have

$$\begin{aligned}\mathrm{R}(U;E_k\mathcal{P}_hz)\le C\sum_{n=1}^N\int_{I_n}\Big\{&\|\dot{U}_1-U_2\|\|E_k\mathcal{P}_hz_1\|+\|E_k\bar{A}_hU_1\|\|E_k\mathcal{P}_hz_2\|\\ &+\Big\|E_k\int_0^t\beta(t-s)\bar{A}_hU_1(s)\,ds\Big\|\|E_k\mathcal{P}_hz_2\|\\ &+\|E_kf\|\|E_k\mathcal{P}_hz_2\|+\Big(\sum_{K\in\mathcal{T}_h^n}h_K^{-1}\|E_kg\|_{\partial K}^2\Big)^{1/2}\|E_k\mathcal{P}_hz_2\|\Big\}.\end{aligned}$$

This, L_2-stability of the L_2-projection $\mathcal{P}_h$, and a standard error estimation of the error operator E_k, conclude

$$\begin{aligned}\mathrm{R}(U;E_k\mathcal{P}_hz)\le C\max\big(&\max_{[0,T]}\|\dot{z}_1(t)\|,\max_{[0,T]}\|\dot{z}_2(t)\|\big)\\ &\times\sum_{n=1}^N\int_{I_n}\Big\{k_n\|\dot{U}_1-U_2\|+k_n\|E_k\bar{A}_hU_1\|\\ &+k_n\Big\|E_k\int_0^t\beta(t-s)\bar{A}_hU_1(s)\,ds\Big\|+k_n\|E_kf\|\\ &+k_n\Big(\sum_{K\in\mathcal{T}_h^n}h_K^{-1}\|E_kg\|_{\partial K}^2\Big)^{1/2}\Big\}\,dt.\end{aligned}\tag{5.29}$$

We now set $j_1=j_2=z_2^T=0$ and $z_1^T=A^{-1/2}e_1(T)$. Then, putting (5.27) and (5.29) in (5.17), using the stability estimates (2.14) and a standard argument, we conclude the a posteriori error estimate (5.15), and this completes the proof. □

Remark 5. We note that for the error estimate (5.15) there are two types of restriction on the triangulations; One by ζ_n, $\zeta_{n,N}$, that measures the quasiuniformity of the family of triangulation, and the other by $\delta_{\mathcal{F}_h}$, that measure the regularity of the family of triangulations in a slightly different sense. Although maybe not explicitly, but ζ_n, $\zeta_{n,N}$ and $\delta_{\mathcal{F}_h}$ can be related. In practice we use finitely many triangulations, that means quasiuniformity holds, though possibly with big ζ_n, $\zeta_{n,N}$. This means that we still can use the a posteriori error estimate (5.15). But when $\delta_{\mathcal{F}_h}$ is not sufficiently small, the error estimate (5.15) does not hold. This calls for using local interpolants instead of global L_2-projection $\mathcal{P}_h$. In the next theorem we present an a posteriori error estimate using interpolation, linear in space and constant in time on each space-time cell. We also note that a possible, but not neccesserily optimal, way of ignoring these limitations could be using global error estimates of $\mathcal{P}_h$ with global mesh size h_{max}. That is a posteriori error estimate in the form, with $k(t) = k_n$ for $t \in I_n$,

$$
\begin{aligned}
\|e_1(T)\| &\le C\big\{h_{max,0}\|U_1(0) - u^0\| + h^2_{max,0}\|U_2(0) - v^0\|\big\} \\
&\quad + Ch_{max}\int_0^T \|\dot{U}_1 - U_2\|\,dt \\
&\quad + Ch^2_{max}\int_0^T \Big\{\|\rho\dot{U}_2 - f\| + \|\tilde{r}_h\| + \|\tilde{g}_h - \tilde{g}\|_{\Gamma_{\rm N}} + \|\tilde{g}_h\|_{\Gamma_{\rm N}}\Big\}\,dt \\
&\quad + \int_0^T k\Big\{\|\dot{U}_1 - U_2\| + \|E_k\bar{A}_h U_1\| \\
&\qquad + \Big\|E_k\int_0^t \beta(t-s)\bar{A}_h U_1(s)\,ds\Big\| + \|E_k f\| + \|E_k g\|_{\Gamma_{\rm N}}\Big\}\,dt,
\end{aligned}
$$

where $\tilde{r}_h|_K = h_K^{-1}\max_{\partial K}|r_h|$, $\tilde{g}_h|_K = h_K^{-1/2}|g_h|$, and $\tilde{g}|_K = h_K^{-1/2}|g|$.

We recall the decomposition of the space-time slab $\Omega^n = \Omega \times I_n$ into cells $K^n = K \times I_n$, $K \in \bar{\mathcal{T}}_h^n$. Let I_{hk} be the standard interpolant, such that $I_{hk}v|_{K^n}$ be linear in space and constant in time. We define the error operator E_{hk} by $E_{hk}v = (I_{hk} - I)v$. A variant of the Bramble-Hilbert lemma then implies the error estimates,

$$\|E_{hk}v\|_{K^n} \le C\big(h_K^r\|\nabla^r v\|_{\tilde{K}^n} + k_n\|\dot{v}\|_{\tilde{K}^n}\big), \quad r = 1,\, 2, \tag{5.30}$$

$$\|\nabla E_{hk}v\|_{K^n} \le C\big(h_K^2\|\nabla^2 v\|_{\tilde{K}^n} + k_n\|\nabla\dot{v}\|_{\tilde{K}^n}\big), \tag{5.31}$$

where $\tilde{K}$ is a patch of space cells suitably chosen around K.

We recall the mesh function $\bar{h}_n$ from (3.2), and we define the notation

$$\bar{h}_{max,n} = \max_{K\in\bar{\mathcal{T}}_h^n} \bar{h}_K.$$

We will also use the fact that

$$\|v\|^2_{\Omega^n} = \int_{I_n}\|v(t)\|^2\,dt \le k_n\max_{I_n}\|v(t)\|. \tag{5.32}$$

Theorem 7. *Let u and U be the solutions of* (2.3) *and* (3.5)*. Then, with $e = U - u$, we have the weighted a posteriori error estimate*

(5.33)
$$
\begin{aligned}
&\|e_1(T)\| \\
&\leq C\Big\{\|h_0\big(U_1(0)-u^0\big)\| + \|h_0^2\big(U_2(0)-v^0\big)\|\Big\} \\
&\quad + C\sum_{n=1}^{N} k_n^{1/2}\bigg\{\|\bar{h}_n(\dot{U}_1 - U_2)\|_{\Omega^n} + k_n\|\dot{U}_1 - U_2\|_{\Omega^n} \\
&\quad + \|\bar{h}_n^2(\rho\dot{U}_2 - f)\|_{\Omega^n} + k_n\|\rho\dot{U}_2 - f\|_{\Omega^n} \\
&\quad + \Big(\sum_{K\in\bar{\mathcal{T}}_h^n} \bar{h}_K^3\|r_h\|_{\partial K^n}^2\Big)^{1/2} + \Big(\sum_{K\in\bar{\mathcal{T}}_h^n} \bar{h}_K^3\|g_h - g\|_{\partial K^n}^2\Big)^{1/2} \\
&\quad + k_n\Big(\sum_{K\in\bar{\mathcal{T}}_h^n} \bar{h}_K^{-1}\|r_h\|_{\partial K^n}^2\Big)^{1/2} + k_n\Big(\sum_{K\in\bar{\mathcal{T}}_h^n} \bar{h}_K^{-1}\|g_h - g\|_{\partial K^n}^2\Big)^{1/2} \\
&\quad + k_n\Big(\sum_{K\in\bar{\mathcal{T}}_h^n} \bar{h}_K\|r_h\|_{\partial K^n}^2\Big)^{1/2} + k_n\Big(\sum_{K\in\bar{\mathcal{T}}_h^n} \bar{h}_K\|g_h - g\|_{\partial K^n}^2\Big)^{1/2} \\
&\quad + \big(\bar{h}_{max,n}^2 + k_n\big)\bigg(\int_{I_n}\bigg(\sum_{j=1}^{n}\Big(\sum_{K\in\bar{\mathcal{T}}_h^j} \bar{h}_K^{-1}\|(\beta * r_h)_j(t)\|_{\partial K}^2\Big)^{1/2}\bigg)^2 dt\bigg)^{1/2} \\
&\quad + \big(\bar{h}_{max,n}^2 + k_n\big)\bigg(\int_{I_n}\bigg(\sum_{j=1}^{n}\Big(\sum_{K\in\bar{\mathcal{T}}_h^j} \bar{h}_K^{-1}\|(\beta * g_h)_j(t)\|_{\partial K}^2\Big)^{1/2}\bigg)^2 dt\bigg)^{1/2} \\
&\quad + \big(\bar{h}_{max,n} + k_n\big) \\
&\qquad\quad \times\bigg(\int_{I_n}\bigg(\sum_{j=1}^{n}\bar{h}_{max,j}\Big(\sum_{K\in\bar{\mathcal{T}}_h^j} \bar{h}_K^{-1}\|(\beta * r_h)_j(t)\|_{\partial K}^2\Big)^{1/2}\bigg)^2 dt\bigg)^{1/2} \\
&\quad + \big(\bar{h}_{max,n} + k_n\big) \\
&\qquad\quad \times\bigg(\int_{I_n}\bigg(\sum_{j=1}^{n}\bar{h}_{max,j}\Big(\sum_{K\in\bar{\mathcal{T}}_h^j} \bar{h}_K^{-1}\|(\beta * g_h)_j(t)\|_{\partial K}^2\Big)^{1/2}\bigg)^2 dt\bigg)^{1/2}\bigg\}.
\end{aligned}
$$

Proof. We write the error representation (5.1) as

$$
J(e) = \sum_{K\in\mathcal{T}_h^0}\Theta_{0,K} + \sum_{i=1}^{6}\sum_{n=1}^{N}\sum_{K\in\bar{\mathcal{T}}_h^n}\Theta_{i,K}^n = I_0 + \sum_{i=1}^{6} I_i. \tag{5.34}
$$

First we estimate I_0. To this end, recalling $\Theta_{0,K}$ from (5.2), we use the Cauchy-Schwarz inequality and the interpolation error estimate (5.30) to obtain,

$$\begin{aligned}\sum_{K\in\mathcal{T}_h^0}\big(U_1(0)-u^0,E_{hk}z_1(0)\big)_K &\le \sum_{K\in\mathcal{T}_h^0}\|U_1(0)-u^0\|_K\|E_{hk}z_1(0)\|_K\\ &\le C\sum_{K\in\mathcal{T}_h^0}\|U_1(0)-u^0\|_K h_K\|\nabla z_1(0)\|_K\\ &\le C\|\nabla z_1(0)\|\Big(\sum_{K\in\mathcal{T}_h^0}h_K^2\|U_1(0)-u^0\|_K^2\Big)^{1/2}\\ &= C\|\nabla z_1(0)\|\|h_0\big(U_1(0)-u^0\big)\|.\end{aligned}$$

Similarly we have

$$\sum_{K\in\mathcal{T}_h^0}\rho\big(U_2(0)-v^0,E_{hk}z_2(0)\big)_K\le C\|\nabla^2 z_2(0)\|\|h_0^2\big(U_2(0)-v^0\big)\|.$$

From these two estimates we conclude

$$\begin{aligned}I_0 &\le C\max(\|\nabla z_1(0)\|,\|\nabla^2 z_2(0)\|)\\ &\quad\times\big\{\|h_0\big(U_1(0)-u^0\big)\|+\|h_0^2\big(U_2(0)-v^0\big)\|\big\}.\end{aligned}\tag{5.35}$$

For the next term, using the Cauchy-Schwarz inequality and the error estimate (5.30), we have

$$\begin{aligned}I_1 &\le \sum_{n=1}^N\sum_{K\in\bar{\mathcal{T}}_h^n}\|\dot U_1-U_2\|_{K^n}\|E_{hk}z_1\|_{K^n}\\ &\le C\sum_{n=1}^N\sum_{K\in\bar{\mathcal{T}}_h^n}\|\dot U_1-U_2\|_{K^n}\big(\bar h_K\|\nabla z_1\|_{\tilde K^n}+k_n\|\dot z_1\|_{\tilde K^n}\big)\\ &\le C\sum_{n=1}^N\Big(\sum_{K\in\bar{\mathcal{T}}_h^n}\bar h_K^2\|\dot U_1-U_2\|_{K^n}^2\Big)^{1/2}\Big(\sum_{K\in\bar{\mathcal{T}}_h^n}\|\nabla z_1\|_{\tilde K^n}^2\Big)^{1/2}\\ &\quad+C\sum_{n=1}^N k_n\Big(\sum_{K\in\bar{\mathcal{T}}_h^n}\|\dot U_1-U_2\|_{K^n}^2\Big)^{1/2}\Big(\sum_{K\in\bar{\mathcal{T}}_h^n}\|\dot z_1\|_{\tilde K^n}^2\Big)^{1/2}\\ &= C\sum_{n=1}^N\Big\{\|\bar h_n(\dot U_1-U_2)\|_{\Omega^n}\|\nabla z_1\|_{\Omega^n}+k_n\|\dot U_1-U_2\|_{\Omega^n}\|\dot z_1\|_{\Omega^n}\Big\},\end{aligned}$$

that using (5.32) we have

$$
\begin{aligned}
I_1 &\leq C \max\Big(\max_{[0,T]} \|\nabla z_1(t)\|, \max_{[0,T]} \|\dot{z}_1(t)\|\Big) \\
&\quad \times \sum_{n=1}^{N} k_n^{1/2} \big\{ \|\bar{h}_n(\dot{U}_1 - U_2)\|_{\Omega^n} + k_n \|\dot{U}_1 - U_2\|_{\Omega^n} \big\}.
\end{aligned} \tag{5.36}
$$

In the same way we obtain

$$
\begin{aligned}
I_2 &\leq C \max\Big(\max_{[0,T]} \|\nabla^2 z_2(t)\|, \max_{[0,T]} \|\dot{z}_2(t)\|\Big) \\
&\quad \times \sum_{n=1}^{N} k_n^{1/2} \big\{ \|\bar{h}_n^2(\rho\dot{U}_2 - f)\|_{\Omega^n} + k_n \|\rho\dot{U}_2 - f\|_{\Omega^n} \big\}.
\end{aligned} \tag{5.37}
$$

Now for I_3, we use the Cauchy-Schwarz inequality, a trace inequality, and the error estimates (5.30), (5.31) to obtain,

$$
\begin{aligned}
I_3 &\leq \sum_{n=1}^{N} \sum_{K \in \bar{\mathcal{T}}_h^n} \bar{h}_K^{-1/2} \|r_h\|_{\partial K^n} \bar{h}_K^{1/2} \|E_{hk} z_2\|_{\partial K^n} \\
&\leq C \sum_{n=1}^{N} \sum_{K \in \bar{\mathcal{T}}_h^n} \bar{h}_K^{-1/2} \|r_h\|_{\partial K^n} \big\{ \|E_{hk} z_2\|_{K^n} + \bar{h}_K \|\nabla E_{hk} z_1\|_{K^n} \big\} \\
&\leq C \sum_{n=1}^{N} \sum_{K \in \bar{\mathcal{T}}_h^n} \bar{h}_K^{-1/2} \|r_h\|_{\partial K^n} \\
&\quad \times \big\{ 2\bar{h}_K^2 \|\nabla^2 z_2\|_{\tilde{K}^n} + k_n \|\dot{z}_2\|_{\tilde{K}^n} + \bar{h}_K k_n \|\nabla \dot{z}_2\|_{\tilde{K}^n} \big\} \\
&\leq C \sum_{n=1}^{N} \bigg\{ \Big(\sum_{K \in \bar{\mathcal{T}}_h^n} \bar{h}_K^3 \|r_h\|_{\partial K^n}^2 \Big)^{1/2} \|\nabla^2 z_2\|_{\Omega^n} \\
&\qquad + k_n \Big(\sum_{K \in \bar{\mathcal{T}}_h^n} \bar{h}_K^{-1} \|r_h\|_{\partial K^n}^2 \Big)^{1/2} \|\dot{z}_2\|_{\Omega^n} \\
&\qquad + k_n \Big(\sum_{K \in \bar{\mathcal{T}}_h^n} \bar{h}_K \|r_h\|_{\partial K^n}^2 \Big)^{1/2} \|\nabla \dot{z}_2\|_{\Omega^n} \bigg\},
\end{aligned}
$$

that using (5.32) we have

(5.38)
$$
\begin{aligned}
I_3 \leq C \max\Big(\max_{[0,T]} \|\nabla^2 z_2(t)\|, \max_{[0,T]} \|\dot{z}_2(t)\|, \max_{[0,T]} \|\nabla \dot{z}_2(t)\|\Big) \\
\times \sum_{n=1}^{N} k_n^{1/2} \Big\{ \Big(\sum_{K\in \bar{\mathcal{T}}_h^n} \bar{h}_K^3 \|r_h\|^2_{\partial K^n}\Big)^{1/2} + k_n \Big(\sum_{K\in \bar{\mathcal{T}}_h^n} \bar{h}_K^{-1} \|r_h\|^2_{\partial K^n}\Big)^{1/2} \\
+ k_n \Big(\sum_{K\in \bar{\mathcal{T}}_h^n} \bar{h}_K \|r_h\|^2_{\partial K^n}\Big)^{1/2} \Big\}.
\end{aligned}
$$

And similarly

$$
\begin{aligned}
I_4 \leq C \max\Big(\max_{[0,T]} \|\nabla^2 z_2(t)\|, \max_{[0,T]} \|\dot{z}_2(t)\|, \max_{[0,T]} \|\nabla \dot{z}_2(t)\|\Big) \\
\times \sum_{n=1}^{N} k_n^{1/2} \Big\{ \Big(\sum_{K\in \bar{\mathcal{T}}_h^n} \bar{h}_K^3 \|g_h - g\|^2_{\partial K^n}\Big)^{1/2} \\
+ k_n \Big(\sum_{K\in \bar{\mathcal{T}}_h^n} \bar{h}_K^{-1} \|g_h - g\|^2_{\partial K^n}\Big)^{1/2} \\
+ k_n \Big(\sum_{K\in \bar{\mathcal{T}}_h^n} \bar{h}_K \|g_h - g\|^2_{\partial K^n}\Big)^{1/2} \Big\}.
\end{aligned}
\tag{5.39}
$$

Finally we study I_5, I_6 which include the convolution terms. We find an estimate for I_5 and a similar argument holds for I_6. First, recalling the definition of $\Theta^n_{5,K}$ from (5.2), we can write I_5 as,

$$
I_5 = \sum_{n=1}^{N} \sum_{K\in \bar{\mathcal{T}}_h^n} \Theta^n_{5,K} = \sum_{n=1}^{N} \int_{I_n} \int_s^T \beta(t-s) \sum_{K\in \bar{\mathcal{T}}_h^n} \big(r_h(s), E_{hk} z_2(t)\big)_{\partial K} \, dt\, ds.
$$

Then we change the order of the time integrals and we obtain,

$$
\begin{aligned}
I_5 &= \sum_{n=1}^{N} \int_{I_n} \sum_{j=1}^{n} \int_{t_{j-1}}^{t\wedge t_j} \beta(t-s) \sum_{K\in \bar{\mathcal{T}}_h^j} \big(r_h(s), E_{hk} z_2(t)\big)_{\partial K} \, ds\, dt \\
&= \sum_{n=1}^{N} \int_{I_n} \sum_{j=1}^{n} \sum_{K\in \bar{\mathcal{T}}_h^j} \big((\beta * r_h)_j(t), E_{hk} z_2(t)\big)_{\partial K} \, dt,
\end{aligned}
$$

where

$$
(\beta * v)_j(t) = \int_{t_{j-1}}^{t\wedge t_j} \beta(t-s) v(s)\, ds.
$$

Now by a trace inequality and then Cauchy-Schwarz inequality in the sum over the triangles, and further Cuachy-Schwarz inequality in the integral over I_n, we have,

$$
\begin{aligned}
I_5 &\leq C \sum_{n=1}^{N} \int_{I_n} \sum_{j=1}^{n} \Big(\sum_{K \in \bar{\mathcal{T}}_h^j} \bar{h}_K^{-1} \|(\beta * r_h)_j(t)\|_{\partial K}^2 \Big)^{1/2} \\
&\qquad \times \big\{ \|E_{hk} z_2(t)\| + \bar{h}_{max,j} \|\nabla E_{hk} z_2(t)\| \big\} \, dt \\
&= C \sum_{n=1}^{N} \int_{I_n} \|E_{hk} z_2(t)\| \sum_{j=1}^{n} \Big(\sum_{K \in \bar{\mathcal{T}}_h^j} \bar{h}_K^{-1} \|(\beta * r_h)_j(t)\|_{\partial K}^2 \Big)^{1/2} dt \\
&\quad + C \sum_{n=1}^{N} \int_{I_n} \|\nabla E_{hk} z_2(t)\| \sum_{j=1}^{n} \bar{h}_{max,j} \Big(\sum_{K \in \bar{\mathcal{T}}_h^j} \bar{h}_K^{-1} \|(\beta * r_h)_j(t)\|_{\partial K}^2 \Big)^{1/2} dt \\
&\leq C \sum_{n=1}^{N} \Big(\int_{I_n} \|E_{hk} z_2(t)\|^2 \, dt \Big)^{1/2} \\
&\qquad \times \bigg(\int_{I_n} \Big(\sum_{j=1}^{n} \Big(\sum_{K \in \bar{\mathcal{T}}_h^j} \bar{h}_K^{-1} \|(\beta * r_h)_j(t)\|_{\partial K}^2 \Big)^{1/2} \Big)^2 dt \bigg)^{1/2} \\
&\quad + C \sum_{n=1}^{N} \Big(\int_{I_n} \|\nabla E_{hk} z_2(t)\|^2 \, dt \Big)^{1/2} \\
&\qquad \times \bigg(\int_{I_n} \Big(\sum_{j=1}^{n} \bar{h}_{max,j} \Big(\sum_{K \in \bar{\mathcal{T}}_h^j} \bar{h}_K^{-1} \|(\beta * r_h)_j(t)\|_{\partial K}^2 \Big)^{1/2} \Big)^2 dt \bigg)^{1/2}.
\end{aligned}
$$

Since by the error estimate (5.30) we have

$$
\begin{aligned}
\int_{I_n} \|E_{hk} z_2(t)\|^2 \, dt &= \int_{I_n} \sum_{K \in \bar{\mathcal{T}}_h^n} \|E_{hk} z_2(t)\|_K^2 \, dt = \sum_{K \in \bar{\mathcal{T}}_h^n} \|E_{hk} z_2(t)\|_{K^n}^2 \\
&\leq C \sum_{K \in \bar{\mathcal{T}}_h^n} \big(\bar{h}_K^4 \|\nabla^2 z_2\|_{\tilde{K}^n}^2 + k_n^2 \|\dot{z}_2\|_{\tilde{K}^n}^2 \big) \\
&\leq C \big(\bar{h}_{max,n}^4 \|\nabla^2 z_2\|_{\Omega^n}^2 + k_n^2 \|\dot{z}_2\|_{\Omega^n}^2 \big),
\end{aligned}
$$

and similarly by (5.31)

$$\begin{aligned}\int_{I_n} \|\nabla E_{hk} z_2(t)\|^2\,dt &= \int_{I_n} \sum_{K\in\bar{\mathcal{T}}_h^n} \|\nabla E_{hk} z_2(t)\|_K^2\,dt = \sum_{K\in\bar{\mathcal{T}}_h^n} \|\nabla E_{hk} z_2(t)\|_{K^n}^2\\ &\le C \sum_{K\in\bar{\mathcal{T}}_h^n} \big(\bar{h}_K^2 \|\nabla^2 z_2\|_{\tilde{K}^n}^2 + k_n^2 \|\nabla \dot{z}_2\|_{\tilde{K}^n}^2\big)\\ &\le C\big(\bar{h}_{max,n}^2 \|\nabla^2 z_2\|_{\Omega^n}^2 + k_n^2 \|\nabla \dot{z}_2\|_{\Omega^n}^2\big),\end{aligned}$$

then, recalling (5.32), we conclude the estimate

(5.40)

$$\begin{aligned} I_5 \le C \max\Big(&\max_{[0,T]} \|\nabla^2 z_2(t)\|, \max_{[0,T]} \|\dot{z}_2(t)\|, \max_{[0,T]} \|\nabla \dot{z}_2(t)\|\Big)\\ &\times \Bigg\{ \sum_{n=1}^N k_n^{1/2} \big(\bar{h}_{max,n}^2 + k_n\big)\\ &\quad \times \Bigg(\int_{I_n} \Big(\sum_{j=1}^n \Big(\sum_{K\in\bar{\mathcal{T}}_h^j} \bar{h}_K^{-1} \|(\beta * r_h)_j(t)\|_{\partial K}^2 \Big)^{1/2}\Big)^2 dt\Bigg)^{1/2}\\ &+ C \sum_{n=1}^N k_n^{1/2} \big(\bar{h}_{max,n} + k_n\big)\\ &\quad \times \Bigg(\int_{I_n} \Big(\sum_{j=1}^n \bar{h}_{max,j} \Big(\sum_{K\in\bar{\mathcal{T}}_h^j} \bar{h}_K^{-1} \|(\beta * r_h)_j(t)\|_{\partial K}^2 \Big)^{1/2}\Big)^2 dt\Bigg)^{1/2}\Bigg\}. \end{aligned}$$

The same estimate holds for I_6 with r_h be replaced by g_h.

We now set $j_1 = j_2 = z_2^T = 0$ and $z_1^T = A^{-1/2} e_1(T)$. Then, putting (5.35)-(5.40), and the counterpart of (5.40) for I_6, in (5.1), using the stability estimates (2.14) and a standard argument, we conclude the a posteriori error estimate (5.33), and this completes the proof. $\square$

Remark 6. We can compute

$$\begin{aligned}\|r_h\|_{\partial K^n} &= \Big(\int_{I_n} \|\psi_{n-1}(t) r_h(t_{n-1}) + \psi_n(t) r_h(t_n)\|_{\partial K}^2\,dt\Big)^{1/2}\\ &\le \frac{k_n^{1/2}}{\sqrt{3}} \Big(\|r_h(t_{n-1})\|_{\partial K}^2 + \|r_h(t_n)\|_{\partial K}^2\Big)^{1/2}\\ &\le \sqrt{\frac{2}{3}} k_n^{1/2} \big(\|r_h(t_{n-1})\|_{\partial K} + \|r_h(t_n)\|_{\partial K}\big).\end{aligned}$$

Remark 7. We note that the last a posteriori error estimate presented in (5.33), does not have the restrictions that were mentioned in Remark 5.

References

1. K. Adolfsson, M. Enelund, and S. Larsson, *Space-time discretization of an integro-differential equation modeling quasi-static fractional-order viscoelasticity*, J. Vib. Control **14** (2008), 1631–1649.
2. K. Adolfsson, M. Enelund, S. Larsson, and M. Racheva, *Discretization of integro-differential equations modeling dynamic fractional order viscoelasticity*, LNCS **3743** (2006), 76–83.
3. W. Bangerth and R. Rannacher, *Adaptive Finite Element Methods for Differential Equations*, Lectures in Mathematics ETH Zürich, Birkhäuser Verlag, 2003.
4. M. Boman, *Estimates for the L_2-projection onto continuous finite element spaces in a weighted L_p-norm*, BIT **46** (2006), 249–260.
5. A. Erdélyi, W. Magnus, F. Oberhettinger, and F. G. Tricomi, *Higher Transcendental Functions. Vol. III*, Robert E. Krieger Publishing Co. Inc., 1981, Based on notes left by Harry Bateman, Reprint of the 1955 original.
6. R. H. Fabiano and K. Ito, *Semigroup theory and numerical approximation for equations in linear viscoelasticity*, SIAM J. Math. Anal. **21** (1990), 374–393.
7. S. Larsson and F. Saedpanah, *The continuous Galerkin method for an integro-differential equation modeling dynamic fractional order viscoelasticity*, IMA J. Numer. Anal., to appear.
8. W. McLean, V. Thomée, and L. B. Wahlbin, *Discretization with variable time steps of an evolution equation with a positive-type memory term*, J. of Comput. Appl. Math. **69** (1996), 49–69.
9. A. K. Pani, V. Thomée, and L. B. Wahlbin, *Numerical methods for hyperbolic and parabolic integro-differential equations*, J. Integral Equations Appl. **4** (1992), 533–584.
10. B. Riviére, S. Shaw, and J. R. Whiteman, *Discontinuous Galerkin finite element methods for dynamic linear solid viscoelasticity problems*, Numer. Methods Partial Differential Equations **23** (2007), 1149–1166.
11. F. Saedpanah, *Existence and uniqueness of the solution of an integro-differential equation with weakly singular kernel*, Preprint 2009:16, Dept. of Mathematical Sciences, Chalmers University of Technology and University of Gothenburg.
12. S. Shaw and J. R. Whiteman, *Applications and numerical analysis of partial differential Volterra equations: a brief survey*, Comput. Methods Appl. Mech. Engrg. **150** (1997), 397–409.
13. ———, *A posteriori error estimates for space-time finite element approximation of quasistatic hereditary linear viscoelasticity problems*, Comput. Methods Appl. Mech. Engrg. **193** (2004), 5551–5572.

Department of Mathematical Sciences, Chalmers University of Technology and University of Gothenburg, SE–412 96 Göteborg, Sweden

E-mail address: fardin@chalmers.se

FINITE ELEMENT APPROXIMATION OF THE LINEAR STOCHASTIC WAVE EQUATION WITH ADDITIVE NOISE

MIHÁLY KOVÁCS[1], STIG LARSSON[1,2], AND FARDIN SAEDPANAH

ABSTRACT. Semidiscrete finite element approximation of the linear stochastic wave equation with additive noise is studied in a semigroup framework. Optimal error estimates for the deterministic problem are obtained under minimal regularity assumptions. These are used to prove strong convergence estimates for the stochastic problem. The theory presented here applies to multi-dimensional domains and spatially correlated noise. Numerical examples illustrate the theory.

1. INTRODUCTION

We study the finite element approximation of the linear stochastic wave equation driven by additive noise,

$$\begin{aligned} &\mathrm{d}\dot{u} - \Delta u \,\mathrm{d}t = \mathrm{d}W && \text{in } \mathcal{D}\times(0,\infty),\\ &u = 0 && \text{in } \partial\mathcal{D}\times(0,\infty),\\ &u(\cdot,0)=u_0,\quad \dot{u}(\cdot,0)=v_0 && \text{in } \mathcal{D}, \end{aligned} \tag{1.1}$$

where $\mathcal{D}\subset\mathbb{R}^d$, $d=1,2,3$, is a bounded convex polygonal domain with boundary $\partial\mathcal{D}$, and $\{W(t)\}_{t\geq 0}$ is a $L_2(\mathcal{D})$-valued Wiener process on a filtered probability space $(\Omega,\mathcal{F},\mathbf{P},\{\mathcal{F}_t\}_{t\geq0})$ with respect to the normal filtration $\{\mathcal{F}_t\}_{t\geq0}$. We let u_0, v_0 be $\mathcal{F}_0$-measurable random variables.

For introduction to the stochastic wave equation and its applications we refer to [1], [6], [14], [16], [23] and the references therein.

The stochastic heat equation and its numerical approximation has been extensively researched in the literature, see, for example, [6], [11], [12], [13], [23], [25], [26], and the references therein. The numerical analysis of the stochastic wave equation is less studied, see [15], [18], [20], [24] for existing

Date: April 12, 2009.

1991 *Mathematics Subject Classification.* 65M60, 60H15, 60H35, 65C30.

Key words and phrases. finite element method, stochastic wave equation, additive noise, Wiener process, stability, a priori error estimate, strong convergence.

[1]Supported by the Swedish Research Council (VR).

[2]Supported by the Swedish Foundation for Strategic Research (SSF) through GMMC, the Gothenburg Mathematical Modelling Centre.

results. In particular, these works do not deal with multiple dimensions or correlated noise. This is the purpose of the present work.

We use the semigroup framework of [16] in which the weak solution of (1.1) is represented as a stochastic convolution

$$u(t) = \int_0^t \Lambda^{-1/2} \sin((t-s)\Lambda^{1/2})\, \mathrm{d}W(s),$$

where, for simplicity, we have set the initial values $u_0 = v_0 = 0$. Here $\Lambda = -\Delta$ with $D(\Lambda) = H^2(\mathcal{D}) \cap H_0^1(\mathcal{D})$, and $v(t) = \Lambda^{-1/2}\sin(t\Lambda^{1/2})f$ is the solution of

$$\begin{aligned} &\ddot{v} + \Lambda v = 0, \quad t > 0, \\ &v(0) = 0, \quad \dot{v}(0) = f. \end{aligned} \tag{1.2}$$

We show that, if Q denotes the covariance operator of W, and if

$$\|\Lambda^{(\beta-1)/2} Q^{1/2}\|_{\mathrm{HS}} < \infty, \tag{1.3}$$

for some $\beta \geq 0$, then we have spatial regularity of order β,

$$\Big(\mathbf{E}\big(\|u(t)\|^2_{\dot{H}^\beta}\big)\Big)^{1/2} \leq Ct^{1/2}\|\Lambda^{(\beta-1)/2}Q^{1/2}\|_{\mathrm{HS}},$$

where $\dot{H}^\beta = \mathcal{D}(\Lambda^{\beta/2})$. In particular, if $\mathrm{Tr}(Q) = \|Q^{1/2}\|^2_{\mathrm{HS}} < \infty$ (spatially correlated noise), then we may take $\beta = 1$. On the other hand if $Q = I$ (uncorrelated noise), then $\beta < 1 - d/2$, that is, $\beta < 1/2$, $d = 1$. See Section 3 for details.

We discretize (1.1) in the spatial variables with a standard piecewise linear finite element method, and we show strong convergence estimates in various norms. For example,

$$\Big(\mathbf{E}\big(\|u_h(t) - u(t)\|^2\big)\Big)^{1/2} \leq C(t)h^{\frac{2}{3}\beta}\|\Lambda^{(\beta-1)/2}Q^{1/2}\|_{\mathrm{HS}}, \quad \beta \in [0,3], \tag{1.4}$$

where again $u_0 = v_0 = 0$ and $u_h(t)$ is the approximate solution with maximal meshsize h, see Theorem 5.1.

As a comparison, we recall from [25] that for the stochastic heat equation we have

$$\begin{aligned} \Big(\mathbf{E}\big(\|u(t)\|^2_{\dot{H}^\beta}\big)\Big)^{1/2} &\leq C\|\Lambda^{(\beta-1)/2}Q^{1/2}\|_{\mathrm{HS}}, \quad \beta \geq 0, \\ \Big(\mathbf{E}\big(\|u_h(t) - u(t)\|^2\big)\Big)^{1/2} &\leq Ch^{\beta}\|\Lambda^{(\beta-1)/2}Q^{1/2}\|_{\mathrm{HS}}, \quad \beta \in [0,2]. \end{aligned}$$

Here the order of regularity coincides with the order of convergence.

The main tools for the proof of (1.4) are the Itô-isometry (2.5) and error estimates for the deterministic problem (1.2) with minimal regularity assumptions,

$$\|v_h(t) - v(t)\| \leq C(t)h^2\|f\|_{\dot{H}^2}, \tag{1.5}$$

$$\|v_h(t) - v(t)\| \leq C\|f\|_{\dot{H}^{-1}},$$

and, hence by interpolation, see Corollary 4.3,

$$\|v_h(t) - v(t)\| \leq C(t)h^{\frac{2}{3}\beta}\|f\|_{\dot{H}^\beta}, \quad \beta \in [0,3].$$

As mentioned above, when we specialize to $Q = I,\ d = 1$, we have $\beta < 1/2$ and thus the order of strong convergence is $O(h^\alpha),\ \alpha < 1/3$. This is the same order as in [18], where spatial semi-discretization of the nonlinear stochastic wave equation with a standard difference scheme of uniform meshsize h is considered for $d = 1$ and with space-time white noise ($Q = I$). We note that the order of convergence is less than the order of regularity, which is $\beta < 1/2$. However, it is known that in (1.5), $\|f\|_{\dot{H}^2}$ can not be replaced by $\|f\|_{\dot{H}^{2-\epsilon}}$ for any $\epsilon > 0$, see [19] and Remark 4.4 below. Therefore, $O(h^\alpha)$, $\alpha < 1/3$, is the best that one can expect. This explains the discrepancy in the convergence behavior between the heat and wave equations.

In [24] the leap-frog scheme is applied to the nonlinear stochastic wave equation in the unbounded domain $\mathcal{D} = \mathbb{R}$, and a strong convergence rate $O(h^{1/2})$ is proved. The proofs in both [18] and [24] are based on representation of the exact and approximate solutions by means of Green's functions. The difference in convergence rate between the two is explained by the fact that in $\mathbb{R}$ the Green's functions for the wave equation and the leap-frog scheme coincide at mesh points, see Remark 5.2 for more details.

In summary we may say that we extend the results of [18] to the finite element method in multiple dimensions and correlated noise. But we only consider the linear equation with additive noise. We also explain the discrepancy between [18] and [24]. We plan to address the nonlinear equation $\mathrm{d}\dot{u} - \Delta u\,\mathrm{d}t = f(u)\,\mathrm{d}t + g(u)\,\mathrm{d}W$ in future work.

The paper is organized as follows. In Section 2 some preliminaries are provided and a rigorous meaning to the infinite dimensional Wiener process $\{W(t)\}_{t\geq 0}$ and the stochastic integral are given together with the definition of a weak solution of (1.1). Existence, uniqueness, and regularity of weak solutions are discussed in Section 3. In Section 4 the finite element method for the deterministic problem is formulated and analyzed. The results obtained here are used in Section 5 to derive strong convergence estimates for finite element approximation of the stochastic equation (1.1). Finally, numerical experiments are presented in Section 6 in order to illustrate the theory.

2. Preliminaries

Throughout the paper we use '$\cdot$' to denote the time derivative '$\frac{\partial}{\partial t}$', and C to denote a generic positive constant, not necessarily the same at different occurrences. We refer to [16] and [17] for more details on stochastic integration and for some concepts that we cannot explain here.

Let $\big(U, (\cdot,\cdot)_U\big)$ and $\big(H, (\cdot,\cdot)_H\big)$ be separable Hilbert spaces with corresponding norms $\|\cdot\|_U$ and $\|\cdot\|_H$. We suppress the subscripts when it causes no confusion. Let $\mathcal{L}(U,H)$ denote the space of bounded linear operators from U to H, and $\mathcal{L}_2(U,H)$ the space of Hilbert-Schmidt operators, endowed with norm $\|\cdot\|_{\mathcal{L}_2(U,H)}$. That is, $T \in \mathcal{L}_2(U,H)$ if $T \in \mathcal{L}(U,H)$ and

$$\|T\|^2_{\mathcal{L}_2(U,H)} := \sum_{k=1}^{\infty} \|Te_k\|_H^2 < \infty,$$

where $\{e_k\}_{k=1}^{\infty}$ is an arbitrary ON-basis in U. If $H = U$ we write $\mathcal{L}(U) = \mathcal{L}(U,U)$ and $\mathrm{HS} = \mathcal{L}_2(U,U)$. It is well known that if $S \in \mathcal{L}(U)$ and $T \in \mathcal{L}_2(U,H)$, then $TS \in \mathcal{L}_2(U,H)$ and we have the norm inequality

$$\|TS\|_{\mathcal{L}_2(U,H)} \le \|T\|_{\mathcal{L}_2(U,H)} \|S\|_{\mathcal{L}(U)}. \tag{2.1}$$

Let $(\Omega, \mathcal{F}, \mathbf{P})$ be a probability space. We define $L_2(\Omega, H)$ to be the space of H-valued square integrable random variables with norm

$$\|v\|_{L_2(\Omega,H)} = \mathbf{E}(\|v\|_H^2)^{1/2} = \Big(\int_\Omega \|v(\omega)\|_H^2 \,\mathrm{d}\mathbf{P}(\omega)\Big)^{1/2},$$

where $\mathbf{E}$ stands for expected value. Let $Q \in \mathcal{L}(U)$ be a selfadjoint, positive semidefinite operator, with $\mathrm{Tr}(Q) < \infty$, where $\mathrm{Tr}(Q)$ denotes the trace of Q. We say that $\{W(t)\}_{t\ge0}$ is a *U-valued Q-Wiener process with respect to* $\{\mathcal{F}_t\}_{t\ge0}$ if

(i) $W(0) = 0$,
(ii) W has continuous trajectories (almost surely),
(iii) W has independent increments,
(iv) $W(t) - W(s)$, $0 \le s \le t$, is a U-valued Gaussian random variable with zero mean and covariance operator $(t-s)Q$,

and

(v) $\{W(t)\}_{t\ge0}$ is adapted to $\{\mathcal{F}_t\}_{t\ge0}$; that is, $W(t)$ is $\mathcal{F}_t$ measurable for all $t \ge 0$;
(vi) the random variable $W(t) - W(s)$ is independent of $\mathcal{F}_s$ for all fixed $s \in [0,t]$.

It is known, see, e.g., [17, Section 2.1], that for a given Q-Wiener process satisfying (i)–(iv) one can always find a normal filtration $\{\mathcal{F}_t\}_{t\ge0}$ so that (v)–(vi) holds. Furthermore, $W(t)$ has the orthogonal expansion

$$W(t) = \sum_{j=1}^{\infty} \gamma_j^{1/2} \beta_j(t) e_j, \tag{2.2}$$

where $\{(\gamma_j, e_j)\}_{j=1}^{\infty}$ are the eigenpairs of Q with orthonormal eigenvectors, and $\{\beta_j\}_{j=1}^{\infty}$ is a sequence of real-valued mutually independent standard Brownian motions. We note that the series in (2.2) converges in $L_2(\Omega, U)$, since for $t \geq 0$, we have

$$\begin{aligned} \|W(t)\|_{L_2(\Omega,U)}^2 &= \mathbf{E}\Big(\Big\| \sum_{j=1}^{\infty} \gamma_j^{1/2} e_j \beta_j(t) \Big\|_U^2\Big) = \sum_{j=1}^{\infty} \gamma_j \mathbf{E}\big(\beta_j(t)\big)^2 \\ &= t \sum_{j=1}^{\infty} \gamma_j = t \operatorname{Tr}(Q) < \infty. \end{aligned} \tag{2.3}$$

We need only a special case of the Itô integral where the integrand is deterministic. If a function $\Phi : [0, \infty) \to L(U, H)$ is strongly measurable and

$$\int_0^t \|\Phi(s) Q^{1/2}\|_{\mathrm{HS}}^2 \, \mathrm{d}s < \infty, \tag{2.4}$$

then the stochastic integral $\int_0^t \Phi(s) \, \mathrm{d}W(s)$ is well defined and Itô's isometry,

$$\mathbf{E}\Big\| \int_0^t \Phi(s) \, \mathrm{d}W(s) \Big\|_{L_2(\Omega,H)}^2 = \int_0^t \big\|\Phi(s) Q^{1/2}\big\|_{\mathcal{L}_2(U,H)}^2 \, \mathrm{d}s, \tag{2.5}$$

holds.

More generally, if $Q \in \mathcal{L}(U)$ is a selfadjoint, positive semidefinite operator with eigenpairs $\{(\gamma_j, e_j)\}_{j=1}^{\infty}$, but not trace class, that is, $\operatorname{Tr}(Q) = \infty$, then the series (2.2) does not converge in $L_2(\Omega, U)$. However, it converges in a suitably chosen (usually larger) Hilbert space and the stochastic integral $\int_0^t \Phi(s) \, \mathrm{d}W(s)$ can still be defined and the isometry (2.5) holds, as long as (2.4) is satisfied. In this case W is called a cylindrical Wiener process. In particular, we may have $Q = I$ (the identity operator).

Next we consider the abstract stochastic differential equation

$$\mathrm{d}X(t) = AX(t) \, \mathrm{d}t + B \, \mathrm{d}W(t), \quad t > 0; \quad X(0) = X_0, \tag{2.6}$$

and assume that

(a1) $A : D(A) \subset H \to H$ is the generator of a strongly continuous semigroup (C_0-semigroup) of bounded linear operators $\{E(t)\}_{t \geq 0}$ on H,
(a2) $B \in \mathcal{L}(U, H)$,
(a3) X_0 is an $\mathcal{F}_0$-measurable H-valued random variable.

An H-valued predictable process $\{X(t)\}_{t\geq 0}$ is called a *weak solution* of (2.6), if the trajectories of X are $\mathbf{P}$-a.s. Bochner integrable and, for all $\eta \in D(A^*)$ and all $t \geq 0$,

(2.7)
$$\big(X(t),\eta\big) = (X_0,\eta) + \int_0^t \big(X(s),A^*\eta\big)\,\mathrm{d}s + \int_0^t \big(B\,\mathrm{d}W(s),\eta\big), \quad \mathbf{P}\text{-a.s.}$$

3. Abstract framework and regularity

As in the introduction, let $\Lambda = -\Delta$ be the Laplace operator with $D(\Lambda) = H^2(\mathcal{D}) \cap H_0^1(\mathcal{D})$ and let $U = L_2(\mathcal{D})$ with the usual inner product $(\cdot,\cdot)$ and norm $\|\cdot\|$. In order to describe the spatial regularity of functions we introduce the following spaces and norms. Let

$$\dot{H}^\alpha = D(\Lambda^{\alpha/2}), \quad \|v\|_\alpha = \|\Lambda^{\alpha/2}v\| = \Big(\sum_{j=1}^{\infty} \lambda_j^\alpha (v,\phi_j)^2\Big)^{1/2}, \quad \alpha \in \mathbb{R},\ v \in \dot{H}^\alpha,$$

where $\{(\lambda_j,\phi_j)\}_{j=1}^\infty$ are the eigenpairs of Λ with orthonormal eigenvectors. Then $\dot{H}^\alpha \subset \dot{H}^\beta$ for $\alpha \geq \beta$. It is known that $\dot{H}^0 = U$, $\dot{H}^1 = H_0^1(\mathcal{D})$, $\dot{H}^2 = H^2(\mathcal{D}) \cap H_0^1(\mathcal{D})$ with equivalent norms and that $\dot{H}^{-\beta}$ can be identified with the dual space $(\dot{H}^\beta)^*$ for $\beta > 0$, see [22]. We note that the inner product in $\dot{H}^1$ is $(\cdot,\cdot)_1 = (\nabla\cdot,\nabla\cdot)$. We also introduce

$$H^\alpha := \dot{H}^\alpha \times \dot{H}^{\alpha-1}, \quad |||v|||_\alpha^2 := \|v_1\|_\alpha^2 + \|v_2\|_{\alpha-1}^2, \quad \alpha \in \mathbb{R}, \tag{3.1}$$

and set $H = H^0 = \dot{H}^0 \times \dot{H}^{-1}$ with corresponding norm $|||\cdot||| = |||\cdot|||_0$.

Next we write (1.1) as an abstract stochastic differential equation (2.6). To this end, we put $u_1 = u$, $u_2 = \dot{u}$ and note that (1.1) is formally

$$\mathrm{d}\begin{bmatrix} u_1 \\ u_2 \end{bmatrix} = \begin{bmatrix} 0 & I \\ -\Lambda & 0 \end{bmatrix} \begin{bmatrix} u_1 \\ u_2 \end{bmatrix} \mathrm{d}t + \begin{bmatrix} 0 \\ I \end{bmatrix} \mathrm{d}W.$$

We therefore define

$$\begin{aligned} &A := \begin{bmatrix} 0 & I \\ -\Lambda & 0 \end{bmatrix}, \quad B := \begin{bmatrix} 0 \\ I \end{bmatrix}, \quad X := \begin{bmatrix} u_1 \\ u_2 \end{bmatrix}, \quad X_0 := \begin{bmatrix} u_0 \\ v_0 \end{bmatrix}, \\ &H := H^0 = \dot{H}^0 \times \dot{H}^{-1}, \quad U := \dot{H}^0, \end{aligned}$$

with

$$D(A) = \Big\{ x \in H : Ax = \begin{bmatrix} x_2 \\ -\Lambda x_1 \end{bmatrix} \in H = \dot{H}^0 \times \dot{H}^{-1} \Big\} = H^1 = \dot{H}^1 \times \dot{H}^0.$$

Here Λ is regarded as an operator $\dot{H}^1 \to \dot{H}^{-1}$. The operator A is the generator of a strongly continuous semigroup (C_0-semigroup) $E(t) = \mathrm{e}^{tA}$ on

H and

$$E(t) = \mathrm{e}^{tA} = \begin{bmatrix} C(t) & \Lambda^{-1/2}S(t) \\ -\Lambda^{1/2}S(t) & C(t) \end{bmatrix}, \tag{3.2}$$

where $C(t) = \cos(t\Lambda^{1/2})$ and $S(t) = \sin(t\Lambda^{1/2})$ are the so-called cosine and sine operators. For example, using $\{(\lambda_j, \phi_j)\}_{j=1}^{\infty}$, the eigenpairs of Λ, we have

$$\Lambda^{-1/2}S(t)v = \Lambda^{-1/2}\sin(t\Lambda^{1/2})v = \sum_{j=1}^{\infty} \lambda_j^{-1/2}\sin(t\lambda_j^{1/2})(v, \phi_j)\phi_j.$$

We also note that $B \in \mathcal{L}(U, H)$ and we let X_0 be an $\mathcal{F}_0$-measurable H-valued random variable to fulfill the assumptions (a1)–(a3). We assume that W is a Q-Wiener process or a cylindrical Wiener process on U. Now (1.1) is set in the form (2.6), which is given a rigorous meaning by the weak formulation (2.7). Next we consider the existence, uniqueness, and regularity of the weak solution. Recall that we write $\mathrm{HS} = \mathcal{L}_2(U, U)$ for the Hilbert-Schmidt operators on U.

Theorem 3.1. *With the above definitions and if* $\|\Lambda^{(\beta-1)/2}Q^{1/2}\|_{\mathrm{HS}} < \infty$ *for some* $\beta \geq 0$*, then* (2.6) *has a unique weak solution, which is given by the variation of constants formula,*

$$X(t) = E(t)X_0 + \int_0^t E(t-s)B\,\mathrm{d}W(s), \quad t \geq 0. \tag{3.3}$$

Moreover,

$$\|X(t)\|_{L_2(\Omega, H^\beta)} \leq C\Big(\|X_0\|_{L_2(\Omega, H^\beta)} + t^{1/2}\|\Lambda^{(\beta-1)/2}Q^{1/2}\|_{\mathrm{HS}}\Big), \quad t \geq 0. \tag{3.4}$$

Proof. To prove that (3.3) is the unique weak solution it is enough to show that, for fixed t,

$$\int_0^t \|E(s)BQ^{1/2}\|^2_{\mathcal{L}_2(U,H)}\,\mathrm{d}s < \infty, \tag{3.5}$$

see [16, Theorem 5.4]. Indeed, with $\{e_k\}_{k=1}^{\infty}$, an arbitrary ON-basis in U, and for any $\beta \geq 0$, we have

$$\begin{aligned}
\int_0^t \|E(s)BQ^{1/2}\|^2_{\mathcal{L}_2(U,H^\beta)}\,\mathrm{d}s &= \int_0^t \sum_{k=1}^{\infty} \Big|\Big|\Big|E(s)BQ^{1/2}e_k\Big|\Big|\Big|^2_\beta\,\mathrm{d}s \\
&= \int_0^t \sum_{k=1}^{\infty} \{\|\Lambda^{-1/2}S(s)Q^{1/2}e_k\|^2_\beta + \|C(s)Q^{1/2}e_k\|^2_{\beta-1}\}\,\mathrm{d}s \\
&= \int_0^t \{\|\Lambda^{(\beta-1)/2}S(s)Q^{1/2}\|^2_{\mathrm{HS}} + \|\Lambda^{(\beta-1)/2}C(s)Q^{1/2}\|^2_{\mathrm{HS}}\}\,\mathrm{d}s \\
&\leq 2t\|\Lambda^{(\beta-1)/2}Q^{1/2}\|_{\mathrm{HS}},
\end{aligned} \tag{3.6}$$

where, for the last inequality, we used the fact that the Λ commutes with $C(s)$, $S(s)$ and (2.1) together with the boundedness of the cosine and the sine operators in U. With $\beta = 0$, this implies (3.5), and therefore it implies existence and uniqueness of the weak solution. Finally, (3.4) follows from (3.3), the boundedness of $E(t)$ in H^β, the Itô isometry (2.5), and (3.6):

$$\begin{aligned}
&\|X(t)\|^2_{L_2(\Omega,H^\beta)} \\
&\quad\leq 2\Big(\|E(t)X_0\|^2_{L_2(\Omega,H^\beta)} + \Big\|\int_0^t E(t-s)B\,\mathrm{d}W(s)\Big\|^2_{L_2(\Omega,H^\beta)}\Big) \\
&\quad\leq 2\Big(\|X_0\|^2_{L_2(\Omega,H^\beta)} + \int_0^t \|E(s)BQ^{1/2}\|^2_{\mathcal{L}_2(U,H^\beta)}\,\mathrm{d}s\Big).
\end{aligned}$$

□

Remark 3.2. The parameter β in the condition $\|\Lambda^{(\beta-1)/2}Q^{1/2}\|_{\mathrm{HS}} < \infty$ quantifies the spatial correlation of the noise. We highlight three special cases.

- If Q is of trace class, then $\beta = 1$, because $\|Q^{1/2}\|^2_{\mathrm{HS}} = \mathrm{Tr}(Q) < \infty$.
- If $Q = I$, which corresponds to space-time white noise, then we have $\|\Lambda^{(\beta-1)/2}\|_{\mathrm{HS}} < \infty$ if and only if $d = 1$ and $\beta < 1/2$. Indeed, the eigenvalues of Λ behave asymptotically like $\lambda_j \approx j^{2/d}$, so that
$$\|\Lambda^{(\beta-1)/2}\|^2_{\mathrm{HS}} = \sum_{j=1}^{\infty} \lambda_j^{\beta-1} \approx \sum_{j=1}^{\infty} j^{2(\beta-1)/d},$$
and the series converges if and only if $\beta < 1 - d/2$, that is, $d = 1$, $\beta < 1/2$.
- Similarly, if $Q = \Lambda^{-s}$, $s > 0$, then $\beta < 1 + s - d/2$.

Thus, in order to have a positive order of regularity in multiple dimensions ($d > 1$) we need correlated noise.

4. The finite element method for the deterministic problem

In this section we first study the spatially semidiscrete finite element method for the deterministic linear wave equation,

$$\begin{aligned} &\ddot{u} - \Delta u = f && \text{in } \mathcal{D}\times(0,\infty), \\ &u = 0 && \text{on } \partial\mathcal{D}\times(0,\infty) \\ &u(\cdot,0) = u_0, \quad \dot{u}(\cdot,0) = v_0 && \text{in } \mathcal{D}, \end{aligned} \tag{4.1}$$

where $\mathcal{D} \in \mathbb{R}^d$, $d = 1,2,3$, is a bounded convex polygonal domain with boundary $\partial\mathcal{D}$. Then we specialize to the homogeneous equation and derive error estimates which will used to prove strong convergence of the finite element approximation of the stochastic equation.

4.1. **Error estimates for the non-homogeneous problem.** Let $\{\mathcal{T}_h\}$ be a regular family of triangulations of $\mathcal{D}$ with $h_K = \operatorname{diam}(K)$, $h = \max_{K\in\mathcal{T}_h} h_K$, and denote by V_h the space of piecewise linear continuous functions with respect to $\mathcal{T}_h$ which vanish on $\partial\mathcal{D}$. Hence, $V_h \subset H_0^1(\mathcal{D}) = \dot{H}^1$.

The assumption that $\mathcal{D}$ is convex and polygonal guarantees that the triangulations can be exacly fitted to $\partial\mathcal{D}$ and that we have the elliptic regularity $\|v\|_{H^2(\mathcal{D})} \le C\|\Lambda v\|$ for $v \in D(\Lambda)$. We can now quote basic results from the theory of finite elements. We use the norms $\|\cdot\|_s = \|\cdot\|_{\dot{H}^s}$.

For the orthogonal projectors $\mathcal{P}_h : \dot{H}^0 \to V_h$, $\mathcal{R}_h : \dot{H}^1 \to V_h$ defined by

$$(\mathcal{P}_h v, \chi) = (v,\chi), \quad (\nabla\mathcal{R}_h v, \nabla\chi) = (\nabla v, \nabla\chi), \quad \forall\chi\in V_h,$$

we have the following error estimates:

$$\|(\mathcal{R}_h - I)v\|_r \le Ch^{s-r}\|v\|_s, \quad r = 0,1,\ s = 1,2, \quad v \in \dot{H}^s, \tag{4.2}$$

$$\|(\mathcal{P}_h - I)v\|_r \le Ch^{s-r}\|v\|_s, \quad r = -1,0,\ s = 1,2, \quad v \in \dot{H}^s. \tag{4.3}$$

If $\{\mathcal{T}_h\}$ is a quasi-uniform family, then $\mathcal{P}_h$ is bounded in $\dot{H}^1$,

$$\|\mathcal{P}_h v\|_1 \le C\|v\|_1, \quad v \in \dot{H}^1. \tag{4.4}$$

Then we have also

$$\|(\mathcal{P}_h - I)v\|_1 \le Ch^{s-1}\|v\|_s, \quad s = 1,2, \quad v \in \dot{H}^s. \tag{4.5}$$

Remark 4.1. We note that the assumption of quasi-uniformity for the validity of (4.4) can be relaxed, see [4], [5], and [7].

We define a discrete variant of the norm $\|\cdot\|_\alpha$:

$$\|v_h\|_{h,\alpha} = \|\Lambda_h^{\alpha/2} v_h\|, \quad v_h \in V_h,\ \alpha\in\mathbb{R},$$

where $\Lambda_h : V_h \to V_h$ is the discrete Laplace operator defined by

$$(\Lambda_h v_h, \chi) = (\nabla v_h, \nabla\chi), \quad \forall\chi\in V_h.$$

It is clear that $\|v_h\|_{1,h} = \|\nabla v_h\| = \|v_h\|_1$ and

$$\|\mathcal{P}_h f\|_{-1,h} \le \|f\|_{-1}, \quad f \in \dot{H}^{-1} \tag{4.6}$$

follows from the calculation

$$\begin{aligned}
\|\Lambda_h^{-\frac12}\mathcal{P}_h f\| &= \sup_{v_h\in V_h} \frac{|(\Lambda_h^{-\frac12}\mathcal{P}_h f, v_h)|}{\|v_h\|} = \sup_{v_h\in V_h} \frac{|(f, \Lambda_h^{-\frac12} v_h)|}{\|v_h\|} \\
&= \sup_{w_h\in V_h} \frac{|(f, w_h)|}{\|\Lambda_h^{\frac12} w_h\|} = \sup_{w_h\in V_h} \frac{|(f, w_h)|}{\|w_h\|_1} \\
&\le \sup_{w\in \dot{H}^1} \frac{|(f, w)|}{\|w\|_1} = \|f\|_{-1}.
\end{aligned}$$

With $u_1 = u$, $u_2 = \dot{u}$, the weak form of (4.1) reads: find $u_1(t)$, $u_2(t) \in \dot{H}^1$, such that

$$\begin{aligned}
&(\nabla \dot{u}_1(t), \nabla v_1) - (\nabla u_2(t), \nabla v_1) = 0, \\
&(\dot{u}_2(t), v_2) + (\nabla u_1(t), \nabla v_2) = (f(t), v_2), \qquad \forall v_1, v_2 \in \dot{H}^1,\ t > 0, \\
&u_1(0) = u_0,\ u_2(0) = v_0.
\end{aligned} \tag{4.7}$$

The semidiscrete analogue of (4.7) is then to find $u_{h,1}(t)$, $u_{h,2}(t) \in V_h$ such that

$$\begin{aligned}
&(\nabla \dot{u}_{h,1}(t), \nabla \chi_1) - (\nabla u_{h,2}(t), \nabla \chi_1) = 0, \\
&(\dot{u}_{h,2}(t), \chi_2) + (\nabla u_{h,1}(t), \nabla \chi_2) = (f(t), \chi_2), \qquad \forall \chi_1, \chi_2 \in V_h,\ t > 0, \\
&u_{h,1}(0) = u_{h,0}, \quad u_{h,1}(0) = v_{h,0},
\end{aligned} \tag{4.8}$$

with initial values $u_{h,0}, v_{h,0} \in V_h$.

In our error analysis we will use the stability of the slightly more general problem of finding $u_{h,1}(t)$, $u_{h,2}(t) \in V_h$ such that

$$\begin{aligned}
&(\nabla \dot{u}_{h,1}(t), \nabla \chi_1) - (\nabla u_{h,2}(t), \nabla \chi_1) = (\nabla f_1(t), \nabla \chi_1), \\
&(\dot{u}_{h,2}(t), \chi_2) + (\nabla u_{h,1}(t), \nabla \chi_2) = (f_2(t), \chi_2), \qquad \forall \chi_1, \chi_2 \in V_h,\ t > 0, \\
&u_{h,1}(0) = u_{h,0}, \quad u_{h,1}(0) = v_{h,0},
\end{aligned} \tag{4.9}$$

We set $\chi_i = \Lambda_h^{\alpha} u_{h,i}$, $i = 1, 2$, $\alpha \in \mathbb{R}$, in (4.9) and conclude in a standard way that

$$\begin{aligned}
\|u_{h,1}(t)\|_{h,\alpha+1} + \|u_{h,2}(t)\|_{h,\alpha} \le C\Big\{ &\|u_{h,0}\|_{h,\alpha+1} + \|v_{h,0}\|_{h,\alpha} \\
&+ \int_0^t \|\mathcal{R}_h f_1(s)\|_{h,\alpha+1}\,\mathrm{d}s + \int_0^t \|\mathcal{P}_h f_2(s)\|_{h,\alpha}\,\mathrm{d}s \Big\}.
\end{aligned} \tag{4.10}$$

Next, we obtain optimal order error estimates in $L_\infty([0,\infty),\dot H^s)$ with $s=0,1$ for $u_{h,1}$ and $s=0$ for $u_{h,2}$. The regularity requirement is minimal, see Remark 4.6.

Theorem 4.2. *Let u_1,u_2 and $u_{h,1},u_{h,2}$ be the solutions of* (4.7) *and* (4.8), *respectively, and set $e_i:=u_{h,i}-u_i$, $i=1,2$. Then, for $t\ge 0$, we have*

$$\begin{aligned}\|e_1(t)\|_1 &\le C\big\{\|u_{h,0}-\mathcal{R}_h u_0\|_1+\|v_{h,0}-\mathcal{R}_h v_0\|\big\}\\ &\quad + Ch\Big\{\|u_1(t)\|_2+\int_0^t \|\dot u_2(s)\|_1\,\mathrm{d}s\Big\},\end{aligned}\tag{4.11}$$

$$\begin{aligned}\|e_2(t)\| &\le C\big\{\|u_{h,0}-\mathcal{R}_h u_0\|_1+\|v_{h,0}-\mathcal{R}_h v_0\|\big\}\\ &\quad + Ch^2\Big\{\|u_2(t)\|_2+\int_0^t \|\dot u_2(s)\|_2\,\mathrm{d}s\Big\},\end{aligned}\tag{4.12}$$

$$\begin{aligned}\|e_1(t)\| &\le C\big\{\|u_{h,0}-\mathcal{R}_h u_0\|+\|v_{h,0}-\mathcal{P}_h v_0\|_{-1}\big\}\\ &\quad + Ch^2\Big\{\|u_1(t)\|_2+\int_0^t \|u_2(s)\|_2\,\mathrm{d}s\Big\}.\end{aligned}\tag{4.13}$$

Proof. We set

$$e_i=\theta_i+\rho_i=(u_{h,i}-\pi_i u)+(\pi_i u_i-u_i),\quad i=1,2,\tag{4.14}$$

where π_i will be chosen as $\mathcal{R}_h$ or $\mathcal{P}_h$. By subtraction of (4.7) and (4.8), recalling $V_h\subset\dot H^1$, we obtain

$$\begin{aligned}(\nabla\dot e_1(t),\nabla\chi_1)-(\nabla e_2(t),\nabla\chi_1)&=0,\\ (\dot e_2(t),\chi_2)+(\nabla e_1(t),\nabla\chi_2)&=0,\end{aligned}\qquad \forall\chi_1,\chi_2\in V_h,\ t>0.$$

Hence,

$$\begin{aligned}(\nabla\dot\theta_1,\nabla\chi_1)-(\nabla\theta_2,\nabla\chi_1)&=-(\nabla\dot\rho_1,\nabla\chi_1)+(\nabla\rho_2,\nabla\chi_1),\\ (\dot\theta_2,\chi_2)+(\nabla\theta_1,\nabla\chi_2)&=-(\dot\rho_2,\chi_2)-(\nabla\rho_1,\nabla\chi_2),\end{aligned}\quad \forall\chi_1,\chi_2\in V_h,t>0.$$

First, in order to prove the error estimates (4.11) and (4.12), we set

$$\theta_i=u_{h,i}-\mathcal{R}_h u_i,\quad \rho_i=(\mathcal{R}_h-I)u_i,\quad i=1,2.$$

By the definitions of the operators $\mathcal{R}_h,\mathcal{P}_h$, we have

$$\begin{aligned}(\nabla\dot\theta_1,\nabla\chi_1)-(\nabla\theta_2,\nabla\chi_1)&=0,\\ (\dot\theta_2,\chi_2)+(\nabla\theta_1,\nabla\chi_2)&=-(\dot\rho_2,\chi_2),\end{aligned}\qquad \forall\chi_1,\chi_2\in V_h,\ t>0,$$

that is, θ_1,θ_2 satisfy (4.9) with $f_1=0$, $f_2=-\dot\rho_2$. Therefore, by the stability inequality (4.10) with $\alpha=0$, we obtain

$$\|\theta_1(t)\|_{h,1}+\|\theta_2(t)\|_{h,0}\le C\Big\{\|\theta_1(0)\|_{h,1}+\|\theta_2(0)\|_{h,0}+\int_0^t\|\mathcal{P}_h\dot\rho_2(s)\|_{h,0}\,\mathrm{d}s\Big\},$$

Recalling (4.14) and that $\|v_h\|_{h,0} = \|v_h\|$ and $\|v_h\|_{h,1} = \|v_h\|_1$, $v_h \in V_h$, we have

$$\begin{aligned}\|e_1(t)\|_1 \le C\Big\{&\|u_{h,0} - \mathcal{R}_h u_0\|_1 + \|v_{h,0} - \mathcal{R}_h v_0\| \\ &+ \int_0^t \|(\mathcal{R}_h - I)\dot{u}_2(s)\|\,\mathrm{d}s + \|(\mathcal{R}_h - I)u_1(t)\|_1\Big\},\end{aligned}$$

$$\begin{aligned}\|e_2(t)\| \le C\Big\{&\|u_{h,0} - \mathcal{R}_h u_0\|_1 + \|v_{h,0} - \mathcal{R}_h v_0\| \\ &+ \int_0^t \|(\mathcal{R}_h - I)\dot{u}_2(s)\|\,\mathrm{d}s + \|(\mathcal{R}_h - I)u_2(t)\|\Big\}.\end{aligned}$$

Using (4.2) we conclude (4.11) and (4.12).

Finally, to prove the error estimates (4.13) we alter the choice of π_i in (4.14) and set

$$\begin{aligned}\theta_1 &= u_{h,1} - \mathcal{R}_h u_1, \quad \rho_1 = (\mathcal{R}_h - I)u_1, \\ \theta_2 &= u_{h,2} - \mathcal{P}_h u_2, \quad \rho_2 = (\mathcal{P}_h - I)u_2.\end{aligned} \tag{4.15}$$

Then, similarly to the previous case,

$$\begin{aligned}(\nabla\dot{\theta}_1, \nabla\chi_1) - (\nabla\theta_2, \nabla\chi_1) &= (\nabla\rho_2, \nabla\chi_1), \\ (\dot{\theta}_2, \chi_2) + (\nabla\theta_1, \nabla\chi_2) &= 0,\end{aligned} \qquad \forall \chi_1, \chi_2 \in V_h,\ t > 0,$$

that is, θ_1, θ_2 satisfy (4.9) with $f_1 = \rho_2$, $f_2 = 0$. Therefore, by the stability inequality (4.10) with $\alpha = -1$, we obtain

$$\begin{aligned}&\|\theta_1(t)\|_{h,0} + \|\theta_2(t)\|_{h,-1} \\ &\qquad \le C\Big\{\|\theta_1(0)\|_{h,0} + \|\theta_2(0)\|_{h,-1} + \int_0^t \|\mathcal{R}_h\rho_2(s)\|_{h,0}\,\mathrm{d}s\Big\},\end{aligned}$$

Using (4.6), (4.14), and

$$\|\mathcal{R}_h\rho_2\| = \|\mathcal{P}_h(I - \mathcal{R}_h)u_2\| \le \|(\mathcal{R}_h - I)u_2\|,$$

we have

$$\begin{aligned}\|e_1(t)\| \le C\Big\{&\|u_{h,0} - \mathcal{R}_h u_0\| + \|v_{h,0} - \mathcal{P}_h v_0\|_{-1} \\ &+ \int_0^t \|(\mathcal{R}_h - I)u_2(s)\|\,\mathrm{d}s + \|(\mathcal{R}_h - I)u_1(t)\|\Big\}.\end{aligned}$$

This proves (4.13). □

4.2. **Error estimates for the homogeneous problem.** Here we specialize to the homogeneous problem

$$\begin{aligned} &\ddot{u}(t) + \Lambda u(t) = 0, \qquad t > 0, \\ &u(0) = u_0, \ \dot{u}(0) = v_0, \end{aligned} \tag{4.16}$$

and express the error estimates in terms of the initial values. Differentiating the equation with respect to t, we obtain in a standard way

$$\|\mathrm{D}_t^r \dot{u}(t)\|_\alpha^2 + \|\mathrm{D}_t^r u(t)\|_{\alpha+1}^2 = \|v_0^r\|_\alpha^2 + \|u_0^r\|_{\alpha+1}^2. \tag{4.17}$$

Here, for $k = 0, 1, \ldots,$

$$\begin{aligned} u_0^r &= \Lambda^k u_0, \quad v_0^r = \Lambda^k v_0, \qquad r = 2k, \\ u_0^r &= \Lambda^k v_0, \quad v_0^r = \Lambda^{k+1} u_0, \quad r = 2k+1. \end{aligned}$$

We use the notation from Section 3 and we write (4.16) as

$$\begin{aligned} &\dot{X}(t) = AX(t), \quad t > 0, \\ &X(0) = X_0, \end{aligned} \tag{4.18}$$

and we recall that the linear operator A is the generator of a C_0-semigroup $E(t) = \mathrm{e}^{tA}$ given by (3.2). Therefore the solution is $X(t) = E(t)X_0$. The finite element problem is then to find $X_h(t) \in V_h \times V_h$ such that

$$\begin{aligned} &\dot{X}_h(t) = A_h X_h(t), \quad t > 0, \\ &X_h(0) = X_{h,0}, \end{aligned} \tag{4.19}$$

where

$$A_h = \begin{bmatrix} 0 & I \\ -\Lambda_h & 0 \end{bmatrix}, \quad X_h = \begin{bmatrix} u_{h,1} \\ u_{h,2} \end{bmatrix}, \quad X_{h,0} = \begin{bmatrix} u_{h,0} \\ v_{h,0} \end{bmatrix}. \tag{4.20}$$

Similarly to (3.2), it can be shown that A_h generates a C_0-semigroup $E_h(t)$ given by

$$E_h(t) = \mathrm{e}^{tA_h} = \begin{bmatrix} C_h(t) & \Lambda_h^{-1/2} S_h(t) \\ -\Lambda_h^{1/2} S_h(t) & C_h(t) \end{bmatrix} \tag{4.21}$$

with

$$C_h(t) = \cos(t\Lambda_h^{1/2}), \quad S_h(t) = \sin(t\Lambda_h^{1/2}).$$

For example, similarly to the infinite dimensional case, using the eigenpairs $\{(\lambda_{h,j}, \phi_{h,j})\}_{j=1}^{N_h}$ of the discrete Laplacian Λ_h, with $N_h = \dim(V_h)$, we have

$$\Lambda_h^{-1/2} \sin(t\Lambda_h^{1/2}) v_h = \sum_{j=1}^{N_h} \lambda_{h,j}^{-1/2} \sin(t\lambda_{h,j}^{1/2}) (v_h, \phi_{h,j}) \phi_{h,j}, \quad v_h \in V_h.$$

We may now formulate a consequence of Theorem 4.2, which will be used to prove the strong convergence of the finite element approximation of the stochastic wave equation. Recall $|||v|||_\alpha^2 = \|v_1\|_\alpha^2 + \|v_2\|_{\alpha-1}^2$ from (3.1).

Corollary 4.3. *Denote* $X_0 = [u_0, v_0]^T$ *and*

$$F_h(t)X_0 = \big(C_h(t)\mathcal{P}_h - C(t)\big)u_0 + \big(\Lambda_h^{-1/2}S_h(t)\mathcal{P}_h - \Lambda^{-1/2}S(t)\big)v_0, \tag{4.22}$$

$$G_h(t)X_0 = \big(C_h(t)\mathcal{R}_h - C(t)\big)u_0 + \big(\Lambda_h^{-1/2}S_h(t)\mathcal{P}_h - \Lambda^{-1/2}S(t)\big)v_0, \tag{4.23}$$

$$\dot G_h(t)X_0 = -\big(\Lambda_h^{1/2}S_h(t)\mathcal{R}_h - \Lambda^{1/2}S(t)\big)u_0 + \big(C_h(t)\mathcal{P}_h - C(t)\big)v_0. \tag{4.24}$$

Then we have

$$\|F_h(t)X_0\| \le C(1+t)h^{\frac{2}{3}\beta}|||X_0|||_\beta, \quad t\ge 0, \quad \beta\in[0,3], \tag{4.25}$$

$$\|G_h(t)X_0\|_1 \le C(1+t)h^{\frac{1}{2}(\beta-1)}|||X_0|||_\beta, \quad t\ge 0, \quad \beta\in[1,3], \tag{4.26}$$

$$\|\dot G_h(t)X_0\| \le C(1+t)h^{\frac{2}{3}(\beta-1)}|||X_0|||_\beta, \quad t\ge 0, \quad \beta\in[1,4]. \tag{4.27}$$

Note that F_h and G_h differ only in the choice of initial value: $u_{0,h} = \mathcal{P}_h u_0$ and $u_{0,h} = \mathcal{R}_h u_0$. This is necessary in order to accomodate the lowest order of initial regularity used ($\beta = 0$ and $\beta = 1$).

Proof. We begin with the case $\beta = 0$ of (4.25). By the stability (4.10) with $\alpha = -1$ and its the analogue for the continuous equation, and (4.6), we have

$$\begin{aligned}\|F_h(t)X_0\| &\le \|u_{h,1}(t)\| + \|u_1(t)\|\\ &\le C\big\{\|\mathcal{P}_h u_0\| + \|\mathcal{P}_h v_0\|_{-1,h} + \|u_0\| + \|v_0\|_{-1}\big\}\\ &\le C\big(\|u_0\| + \|v_0\|_{-1}\big) = C|||X_0|||_0.\end{aligned}$$

For the case $\beta = 3$ we use (4.13) with $u_{0,h} = \mathcal{P}_h u_0$ and $v_{0,h} = \mathcal{P}_h v_0$, and (4.17),

$$\begin{aligned}\|F_h(t)X_0\| &= \|e_1(t)\|\\ &\le C\big\{\|\mathcal{P}_h(I-\mathcal{R}_h)u_0\|\big\}\\ &\quad + Ch^2\Big\{\|u_1(t)\|_2 + \int_0^t \|u_2(s)\|_2\,\mathrm{d}s\Big\}\\ &\le Ch^2\big\{\|u_0\|_2 + \|v_0\|_1 + t(\|u_0\|_3 + \|v_0\|_2)\big\}\\ &\le C(1+t)h^2|||X_0|||_3.\end{aligned}$$

The proof is then completed by interpolation between these cases.

For (4.26) we first use (4.10) with $\alpha = 0$,

$$\begin{aligned}\|G_h(t)X_0\|_1 &\le \|u_{h,1}(t)\|_1 + \|u_1(t)\|_1\\ &\le C\big\{\|\mathcal{R}_h u_0\|_1 + \|\mathcal{P}_h v_0\| + \|u_0\|_1 + \|v_0\|\big\}\\ &\le C\big(\|u_0\|_1 + \|v_0\|\big) = C|||X_0|||_1.\end{aligned}$$

Then we use (4.11) with $u_{0,h} = \mathcal{R}_h u_0$ and $v_{0,h} = \mathcal{P}_h v_0$,

$$\begin{aligned}\|G_h(t)X_0\|_1 &= \|e_1(t)\|_1\\ &\le C\big\{\|\mathcal{P}_h(I-\mathcal{R}_h)v_0\|\big\} + Ch\Big\{\|u_1(t)\|_2 + \int_0^t \|\dot{u}_2(s)\|_1\,\mathrm{d}s\Big\}\\ &\le Ch\big\{\|u_0\|_2 + \|v_0\|_1 + t(\|u_0\|_3 + \|v_0\|_2)\big\}\\ &\le C(1+t)h^2|||X_0|||_3.\end{aligned}$$

For (4.27) we apply (4.10) with $\alpha = 0$,

$$\begin{aligned}\|\dot{G}_h(t)X_0\| &\le \|u_{h,2}(t)\| + \|u_2(t)\| \le C\big\{\|\mathcal{R}_h u_0\|_1 + \|\mathcal{P}_h v_0\| + \|u_0\|_1 + \|v_0\|\big\}\\ &\le C\big(\|u_0\|_1 + \|v_0\|\big) = C|||X_0|||_1.\end{aligned}$$

Then we use (4.12) with $u_{0,h} = \mathcal{R}_h u_0$ and $v_{0,h} = \mathcal{P}_h v_0$,

$$\begin{aligned}\|\dot{G}_h(t)X_0\| &= \|e_2(t)\| \le Ch^2\Big\{\|u_2(t)\|_2 + \int_0^t \|\dot{u}_2(s)\|_2\,\mathrm{d}s\Big\}\\ &\le Ch\big\{\|u_0\|_3 + \|v_0\|_2 + t(\|u_0\|_4 + \|v_0\|_3)\big\} \le C(1+t)h^2|||X_0|||_4.\end{aligned}$$

□

Remark 4.4. The regularity assumption on X_0 in Corollary 4.3 cannot be relaxed. This means that $|||X_0|||_\beta$ can not be replaced by $|||X_0|||_{\beta-\epsilon}$ for any $\epsilon > 0$. This is shown in the lemma below for the periodic problem

$$\begin{aligned} &\ddot{u}(x,t) - u_{xx}(x,t) = 0, && (x,t)\in\mathbb{R}\times(0,\infty),\\ &u(x+2\pi,t) = u(x,t), && (x,t)\in\mathbb{R}\times(0,\infty),\\ &u(x,0) = u_0(x),\ \dot{u}(x,0) = v_0(x), && x\in\mathbb{R}.\end{aligned} \tag{4.28}$$

Lemma 4.5. *Let u be the solution of* (4.28) *and u_h its finite element approximation. Assume that, for some $\beta \ge 0$, there is a constant C such that for all $u_0 \in \dot{H}^\alpha_{per}$, $v_0 \in \dot{H}^{\alpha-1}_{per}$ and $h > 0$,*

$$\|u(t) - u_h(t)\| \le Ch^{\frac{2}{3}\beta}\big(\|u_0\|_{\dot{H}^\alpha_{per}} + \|v_0\|_{\dot{H}^{\alpha-1}_{per}}\big), \quad t \ge 0.$$

Then $\alpha \ge \beta$.

Here $\dot{H}^\alpha_{\mathrm{per}}$ stands for the subspace of $\dot{H}^\alpha$ consisting of 2π-periodic functions.

Proof. The proof is adapted from [19]. We omit the details. □

Remark 4.6. Optimal order $L_\infty([0,\infty), \dot{H}^0)$ estimates for the finite element approximation of displacement $u = u_1$ and velocity $\dot{u} = u_2$ were first obtained by [10]. However, the regularity requirement for the initial displacement is not minimal in [10]. This was improved in [3], and in [19] it was shown that the resulting regularity requirement is optimal, see Lemma 4.5

above. The error estimates (4.12) and (4.13) are in agreement with the corresponding ones in [3] and [19]. Furthermore, the proof presented here seems to be more straightforward.

5. The finite element method for the stochastic problem

We now consider the approximation of the stochastic wave equation. The spatially discrete analogue of (2.6) is to find $X_h(t) = (u_{h,1}(t), u_{h,2}(t)) \in V_h \times V_h$ such that

$$
\begin{aligned}
&\mathrm{d}X_h(t) = A_h X_h(t)\,\mathrm{d}t + \mathcal{P}_h B\,\mathrm{d}W(t), \quad t > 0,\\
&X_h(0) = X_{0,h},
\end{aligned} \tag{5.1}
$$

where A_h is defined in (4.20). Recall that A_h generates the C_0-semigroup $E_h(t) = \mathrm{e}^{tA_h}$ on V_h given by (4.21), and therefore the unique mild solution of (5.1) is given by

$$
X_h(t) = E_h(t)X_{0,h} + \int_0^t E_h(t-s)\mathcal{P}_h B\,\mathrm{d}W(s), \quad t \geq 0. \tag{5.2}
$$

Recall $|||v|||_\alpha^2 = \|v_1\|_\alpha^2 + \|v_2\|_{\alpha-1}^2$ from (3.1).

Theorem 5.1. *Let $X_0 = [u_0, v_0]^T$ and let $X = [u_1, u_2]^T$ and $X_h = [u_{h,1}, u_{h,2}]^T$ be given by* (3.3) *and* (5.2)*, respectively. Then, the following estimates hold for $t \geq 0$, where $C(t)$ is an increasing function.*
If $u_{0,h} = \mathcal{P}_h u_0$, $v_{0,h} = \mathcal{P}_h v_0$, and $\beta \in [0,3]$, then

$$
\|u_{h,1}(t) - u_1(t)\|_{L_2(\Omega,\dot{H}^0)} \leq C(t)h^{\frac{2}{3}\beta}\big\{\|X_0\|_{L_2(\Omega,H^\beta)} + \|\Lambda^{\frac{1}{2}(\beta-1)}Q^{1/2}\|_{\mathrm{HS}}\big\}. \tag{5.3}
$$

If $u_{0,h} = \mathcal{R}_h u_0$, $v_{0,h} = \mathcal{P}_h v_0$, and $\beta \in [1,3]$, then

$$
\|u_{h,1}(t) - u_1(t)\|_{L_2(\Omega,\dot{H}^1)} \leq C(t)h^{\frac{1}{2}(\beta-1)}\big\{\|X_0\|_{L_2(\Omega,H^\beta)} + \|\Lambda^{\frac{1}{2}(\beta-1)}Q^{1/2}\|_{\mathrm{HS}}\big\}. \tag{5.4}
$$

If $u_{0,h} = \mathcal{R}_h u_0$, $v_{0,h} = \mathcal{P}_h v_0$, and $\beta \in [1,4]$, then

$$
\|u_{h,2}(t) - u_2(t)\|_{L_2(\Omega,\dot{H}^0)} \leq C(t)h^{\frac{2}{3}(\beta-1)}\big\{\|X_0\|_{L_2(\Omega,H^\beta)} + \|\Lambda^{\frac{1}{2}(\beta-1)}Q^{1/2}\|_{\mathrm{HS}}\big\}. \tag{5.5}
$$

The discrete initial values ($u_{0,h} = \mathcal{R}_h u_0$, or $u_{0,h} = \mathcal{R}_h u_0$, and $v_{0,h} = \mathcal{P}_h v_0$) and the regularity of the initial values ($X_0 \in H^\beta$) are chosen so that the corresponding rates of convergence match those of the stochastic convolution terms. Of course, other choices are possible with different convergence rates that can be derived from Theorem 4.2.

Proof. We prove (5.3); the proofs of the other estimates are similar.

In addition to F_h defined in (4.22) we introduce

$$K_h(t)f = \big(\Lambda_h^{-1/2} S_h(t)\mathcal{P}_h - \Lambda^{-1/2} S(t)\big) f \tag{5.6}$$

and deduce from (4.25) with $u_0 = 0$ that

$$\|K_h(t)f\| \le C(1+t)h^{\frac{2}{3}\beta}\|f\|_{\beta-1}. \tag{5.7}$$

Then we have

$$u_{h,1}(t) - u_1(t) = F_h(t)X_0 + \int_0^t K_h(t-s)\,\mathrm{d}W(s).$$

By Itô's isometry (2.5),

$$\begin{aligned}\|u_{h,1}(t) - u_1(t)\|_{L_2(\Omega,U)} &\le \|F_h(t)X_0\|_{L_2(\Omega,U)} + \Big\| \int_0^t K_h(t-s)\,\mathrm{d}W(s)\Big\|_{L_2(\Omega,U)} \\ &= \|F_h(t)X_0\|_{L_2(\Omega,U)} + \Big(\int_0^t \|K_h(s)Q^{1/2}\|_{\mathrm{HS}}^2\,\mathrm{d}s\Big)^{1/2} \\ &= I + II.\end{aligned}$$

From (4.25) it follows that

$$I^2 = \mathbf{E}\big(\|F_h(t)X_0\|^2\big) \le C(t)h^{\frac{4}{3}\beta}\mathbf{E}\big(|||X_0|||_\beta^2\big).$$

Recalling the definition of the Hilbert-Schmidt norm from Section 2, using an orthonormal basis $\{e_k\}_{k=1}^\infty$ in $U = \dot{H}^0$, we obtain

$$II^2 = \sum_{k=1}^\infty \int_0^t \|K_h(s)Q^{1/2}e_k\|^2\,\mathrm{d}s.$$

Finally, by setting $f = Q^{1/2}e_k$ in (5.7), we conclude that

$$II^2 \le C(t)th^{\frac{4}{3}\beta}\sum_{k=1}^\infty \|Q^{1/2}e_k\|_{\beta-1}^2 = C(t)h^{\frac{4}{3}\beta}\|\Lambda^{(\beta-1)/2}Q^{1/2}\|_{\mathrm{HS}}^2,$$

which completes the proof of (5.3). □

Remark 5.2. Let consider the one dimensional case with space-time white noise, that is, when $d = 1$, $Q = I$. Then $\beta < 1/2$ (see Remark 3.2) and the convergence rate in (5.3) is $O(h^\alpha)$, $\alpha < 1/3$, which is in agreement with [18], while $O(h^{1/2})$ was shown for the leap-frog scheme in [24]. The reason why a higher rate of convergence is obtained in [24] is that the Green's functions of the continuous and the discrete equations coincide at the mesh points.

Another example of a numerical scheme where this happens is Galerkin's method with

$$V_h = \mathrm{span}\{\mathrm{e}^{inx} : |n| \le 1/h\},$$

see [19, Remark 2]. Then instead of (4.25) we would have

$$\|F_h(t)X_0\| \le Ch^\beta |||X_0|||_\beta, \quad t \ge 0,$$

and, under the assumptions of (5.3),

$$\|u_{h,1}(t) - u_1(t)\|_{L_2(\Omega,U)} \le Ch^\beta \big\{ \|X_0\|_{L_2(\Omega,H^\beta)} + \|\Lambda^{(\beta-1)/2} Q^{1/2}\|_{\mathrm{HS}} \big\}.$$

This yields the optimal order $O(h^\alpha)$, $\alpha < 1/2$, for $Q = I$.

The error estimates in Theorem 4.2, and therefore in Corollary 4.3 and Theorem 5.1, can be extended to higher order finite element methods. The reason is that the error estimates for the elliptic and the orthogonal projections in (4.2) and (4.3), respectively, as well as the stability inequality (4.10) hold for higher order finite element spaces V_h consisting of continuous piecewise polynomials of order at most $k \ge 1$. This means that in case of highly correlated noise, one might expect higher order of strong convergence when using a higher order finite element method. In this case the counterpart of Theorem 5.1 reads as follows.

Theorem 5.3. *Let $X_0 = [u_0, v_0]^T$ and let $X = [u_1, u_2]^T$ and $X_h = [u_{h,1}, u_{h,2}]^T$ be given by (3.3) and (5.2), respectively, where the finite element spaces V_h consist of continuous piecewise polynomials of order at most $k \ge 1$. Then, the following estimates hold for $t \ge 0$, where $C(t)$ is an increasing function. If $u_{0,h} = \mathcal{P}_h u_0$, $v_{0,h} = \mathcal{P}_h v_0$, and $\beta \in [0, k+2]$, then*

$$\|u_{h,1}(t) - u_1(t)\|_{L_2(\Omega,\dot H^0)} \le C(t) h^{\frac{k+1}{k+2}\beta} \big\{ \|X_0\|_{L_2(\Omega,H^\beta)} + \|\Lambda^{\frac12(\beta-1)} Q^{1/2}\|_{\mathrm{HS}} \big\}.$$

If $u_{0,h} = \mathcal{R}_h u_0$, $v_{0,h} = \mathcal{P}_h v_0$, and $\beta \in [1, k+2]$, then

$$\|u_{h,1}(t) - u_1(t)\|_{L_2(\Omega,\dot H^1)} \le C(t) h^{\frac{k}{k+1}(\beta-1)} \big\{ \|X_0\|_{L_2(\Omega,H^\beta)} + \|\Lambda^{\frac12(\beta-1)} Q^{1/2}\|_{\mathrm{HS}} \big\}.$$

If $u_{0,h} = \mathcal{R}_h u_0$, $v_{0,h} = \mathcal{P}_h v_0$, and $\beta \in [1, k+3]$, then

$$\|u_{h,2}(t) - u_2(t)\|_{L_2(\Omega,\dot H^0)} \le C(t) h^{\frac{k+1}{k+2}(\beta-1)} \big\{ \|X_0\|_{L_2(\Omega,H^\beta)} + \|\Lambda^{\frac12(\beta-1)} Q^{1/2}\|_{\mathrm{HS}} \big\}.$$

6. Numerical experiments

In this section we demonstrate the order of strong convergence of the finite element method for the linear stochastic wave equation LSWE (1.1) by numerical examples. To this end, the backward Euler method is used for time stepping and some computational analysis on the approximation of the stochastic convolution is reviewed, see [26].

6.1. **Computational analysis.** First recall the matrix form of (5.1),

$$\begin{bmatrix} \mathrm{d}u_{h,1}(t) \\ \mathrm{d}u_{h,2}(t) \end{bmatrix} = \begin{bmatrix} 0 & I \\ -\Lambda_h & 0 \end{bmatrix} \begin{bmatrix} u_{h,1}(t) \\ u_{h,2}(t) \end{bmatrix} \mathrm{d}t + \begin{bmatrix} 0 \\ \mathcal{P}_h \,\mathrm{d}W(t) \end{bmatrix}, \tag{6.1}$$

Let $0 = t_0 < t_1 < \cdots < t_{N_t} = T_N$, be a uniform partition of the time interval $[0, T_N]$ with time step $k = 1/N_t$ and time subintervals $I_n = (t_{n-1}, t_n)$, $n = 1, 2, \cdots, N_t$. Then the backward Euler method is formulated as, for $n = 1, 2, \cdots, N_t$,

$$\begin{bmatrix} U_1^n \\ U_2^n \end{bmatrix} - \begin{bmatrix} U_1^{n-1} \\ U_2^{n-1} \end{bmatrix} = \begin{bmatrix} 0 & kI \\ -k\Lambda_h & 0 \end{bmatrix} \begin{bmatrix} U_1^n \\ U_2^n \end{bmatrix} + \begin{bmatrix} 0 \\ \mathcal{P}_h \,\Delta W^n \end{bmatrix}. \tag{6.2}$$

Here $U_i^n \in V_h$ is an approximation of $u_i(\cdot, t_n)$, $i = 1, 2$, and $[U_1^0, U_2^0]^T = \xi_h$. We multiply (6.2) by

$$\begin{bmatrix} \Lambda_h & 0 \\ 0 & I \end{bmatrix}$$

to take advantage of the resulting skew-symmetric structure, see Subsection 6.3, and rearrange, to obtain, for $n = 1, 2, \cdots, N_t$,

$$\begin{bmatrix} \Lambda_h & -k\Lambda_h \\ k\Lambda_h & I \end{bmatrix} \begin{bmatrix} U_1^n \\ U_2^n \end{bmatrix} = \begin{bmatrix} \Lambda_h & 0 \\ 0 & I \end{bmatrix} \begin{bmatrix} U_1^{n-1} \\ U_2^{n-1} \end{bmatrix} + \begin{bmatrix} 0 \\ \mathcal{P}_h \,\Delta W^n \end{bmatrix}. \tag{6.3}$$

For some other ways of approximating the noise and the stochastic integrals we refer to, for example, [2] and [8].

Recalling the Fourier expansion (2.2) of W, we have, for all $\chi \in V_h$,

$$\Big(\mathcal{P}_h \,\Delta W^n, \chi\Big) = \sum_{j=1}^{\infty} \gamma_j^{1/2} \Delta\beta_j^n (e_j, \chi) \approx \sum_{j=1}^{J} \gamma_j^{1/2} \Delta\beta_j^n (e_j, \chi), \tag{6.4}$$

where we truncated the sum to J terms. Recall that $\{\beta_j(t)\}_{j=1}^J$ are mutually independent standard real-valued Brownian motions, and that the increments in (6.4) are

$$\Delta\beta_j^n = \beta_j(t_n) - \beta_j(t_{n-1}) \sim \sqrt{k}\mathcal{N}(0,1), \tag{6.5}$$

that is, real-valued Gaussian random variables with 0 mean and variance k. We also note that $\gamma_j = 1$ for the white noise.

Recalling the semidiscrete solution u_h from (5.2), we denote by u_h^J the semidiscrete solution obtained by using the truncated noise; that is,

$$u_h^J(t) = E_h(t) X_{0,h} + \sum_{j=1}^{J} \gamma_j^{1/2} \int_0^t E_h(t-s) \mathcal{P}_h B e_j \,\mathrm{d}\beta_j(s). \tag{6.6}$$

The following lemma shows, that under some assumptions on the triangulation and the covariance operator Q, it is enough to take $J \geq N_h$ with $N_h = \dim(V_h)$ in order to preserve the order of the FEM.

Lemma 6.1. *Let u_h^J and u_h be defined by (6.6) and (5.2), respectively. Assume that Λ and Q have a common orthonormal basis of eigenfunctions $\{e_j\}_{j=1}^{\infty}$ and that V_h, with dimension N_h, is defined on a family of quasi-uniform triangulations $\{\mathcal{T}_h\}$ of $\mathcal{D}$. Then for $J \geq N_h$ the following estimates hold, where $C(t)$ is an increasing function.*
If $\|\Lambda^{(\beta-1)/2}Q^{1/2}\|_{\mathrm{HS}} < \infty$ for some $\beta \in [0,3]$, then,

$$\|u_{h,1}^J(t) - u_{h,1}(t)\|_{L_2(\Omega,\dot{H}^0)} \leq C(t)h^{\frac{2}{3}\beta}\|\Lambda^{(\beta-1)/2}Q^{1/2}\|_{\mathrm{HS}}.$$

If $\|\Lambda^{(\beta-1)/2}Q^{1/2}\|_{\mathrm{HS}} < \infty$ for some $\beta \in [1,3]$, then,

$$\|u_{h,1}^J(t) - u_{h,1}(t)\|_{L_2(\Omega,\dot{H}^1)} \leq C(t)h^{\frac{1}{2}(\beta-1)}\|\Lambda^{(\beta-1)/2}Q^{1/2}\|_{\mathrm{HS}}.$$

If $\|\Lambda^{(\beta-1)/2}Q^{1/2}\|_{\mathrm{HS}} < \infty$ for some $\beta \in [1,4]$, then,

$$\|u_{h,2}^J(t) - u_{h,2}(t)\|_{L_2(\Omega,\dot{H}^0)} \leq C(t)h^{\frac{2}{3}(\beta-1)}\|\Lambda^{(\beta-1)/2}Q^{1/2}\|_{\mathrm{HS}}.$$

Proof. We prove the second estimate; the others are proved similarly. From (5.2) and (6.6) it follows that

$$u_{h,1}^J(t) - u_{h,1}(t) = \sum_{j=J+1}^{\infty} \gamma_j^{1/2} \int_0^t \Lambda_h^{-1/2} S_h(t-s)\mathcal{P}_h e_j \,\mathrm{d}\beta_j(s).$$

By Itô's isometry (2.5), the independence of β_j's and recalling the error operator from (5.6), we have

$$\begin{aligned}
\|u_{h,1}^J(t) - u_{h,1}(t)\|_{L_2(\Omega,\dot{H}^1)}^2 &= \sum_{j=J+1}^{\infty} \gamma_j \int_0^t \|\Lambda_h^{-1/2} S_h(s)\mathcal{P}_h e_j\|_1^2 \,\mathrm{d}s \\
&\leq 2 \sum_{j=J+1}^{\infty} \gamma_j \int_0^t \|\Lambda^{-1/2} S(s) e_j\|_1^2 \,\mathrm{d}s \\
&\quad + 2 \sum_{j=J+1}^{\infty} \gamma_j \int_0^t \|K_h(s) e_j\|_1^2 \,\mathrm{d}s \\
&= I + II.
\end{aligned}$$

Let λ_j denote the eigenvalues of Λ corresponding to e_j. Then

$$\|\Lambda^{-1/2}\sin(s\Lambda^{1/2})e_j\|_1^2 = \sin^2(s\lambda_j^{1/2}).$$

Thus,

$$
\begin{aligned}
I &= 2\sum_{j=J+1}^{\infty}\gamma_j\int_0^t \|\Lambda^{-1/2}\sin(s\Lambda^{1/2})e_j\|_1^2\,\mathrm{d}s\\
&= 2\sum_{j=J+1}^{\infty}\gamma_j\int_0^t \sin^2(s\lambda_j^{1/2})\,\mathrm{d}s\\
&\le 2t\sum_{j=J+1}^{\infty}\gamma_j \le 2t\sum_{j=J+1}^{\infty}\lambda_j^{-(\beta-1)}(\lambda_j^{\beta-1}\gamma_j)\\
&\le 2t\lambda_{J+1}^{-(\beta-1)}\sum_{j=J+1}^{\infty}\lambda_j^{\beta-1}\gamma_j \le 2t\lambda_{J+1}^{-(\beta-1)}\|\Lambda^{(\beta-1)/2}Q^{1/2}\|_{\mathrm{HS}}^2.
\end{aligned}
$$

For II, by (4.26) with $u_0=0$, $v_0=e_j$, we have

$$
\begin{aligned}
II &\le C(t)h^{\beta-1}\sum_{j=J+1}^{\infty}\gamma_j\int_0^t\|e_j\|_{\beta-1}^2\,\mathrm{d}s\\
&= C(t)h^{\beta-1}\sum_{j=J+1}^{\infty}\gamma_j\|e_j\|_{\beta-1}^2 \le C(t)h^{\beta-1}\|\Lambda^{(\beta-1)/2}Q^{1/2}\|_{\mathrm{HS}}^2.
\end{aligned}
$$

Hence the proof is completed by the fact that, for a quasi-uniform family of triangulations, we have $N_h \approx h^{-d}$ and therefore,

$$
\lambda_{J+1}^{-1} \le CJ^{-2/d} \le CN_h^{-2/d} \le Ch^2.
$$

□

Remark 6.2. In practice Q and Λ do not have a common orthonormal basis of eigenfunctions and the eigenfunctions of Q are not known explicitly. In this case, one has to solve the eigenvalue problem $Qu=\lambda u$ on S_h in order to represent $\mathcal{P}_hW$. Computationally this could be very expensive if Q is given by an integral operator. However, if the kernel is smooth then this can be done more efficiently, see [21]. Furthermore, similarly to the parabolic case [13], it is enough to keep $J<N_h$ terms, for suitable J depending on the kernel, in the expansion of $\mathcal{P}_hW$.

6.2. **Numerical example.** For the numerical experiments, we consider the LSWE in one spatial dimension,

$$
\begin{aligned}
&\mathrm{d}\dot{u}-\Delta u\mathrm{d}t=\mathrm{d}W, && (x,t)\in(0,1)\times(0,1),\\
&u(0,t)=u(1,t)=0, && t\in(0,1),\\
&u(x,0)=\cos(\pi(x-1/2)),\ u_t(x,0)=0, && x\in(0,1).
\end{aligned}
\tag{6.7}
$$

Clearly, there is no exact solution available from a numerical viewpoint as even the solution of the deterministic problem is given as an infinite Fourier series expansion (see, e.g., [9]). Therefore we take the exact solution to be a finite element approximation on a very fine mesh with mesh size h_{exact} to approximate $u = u(x, 1)$, using the backward Euler method (6.3) for time stepping with a small fixed time step k. We note that we chose the time step k according to $k \leq h^2$, since the rate of convergence of the fully discrete (6.3) for the deterministic problem is $O(k + h^2)$.

Applying the time stepping (6.3) to (6.7) we obtain the discrete system

$$\Sigma X^n = \Xi X^{n-1} + b, \tag{6.8}$$

where $b = [0, b_2]^T$ and b_2 is computed using (6.4). We note that for the deterministic problem $b = 0$, the expected rate of convergence in the L_2-norm for both the displacement $u = u_1$ and the velocity $\dot{u} = u_2$ is 2 by (4.13) and (4.12), respectively, see Figure 1.

If $\{\lambda_j\}_{j=1}^{\infty}$ are the eigenvalues of Λ, and we set $Q = \Lambda^{-s}$, $s \in \mathbb{R}$, then

$$\|\Lambda^{(\beta-1)/2} Q^{1/2}\|_{\mathrm{HS}}^2 = \|\Lambda^{(\beta-s-1)/2}\|_{\mathrm{HS}}^2 = \sum_{j=1}^{\infty} \lambda_j^{\beta-s-1} \approx \sum_{j=1}^{\infty} j^{\frac{2}{d}(\beta-s-1)},$$

which is finite if and only if $\beta < 1 + s - d/2$ with d being the dimension of the domain $\mathcal{D}$. In our example (6.7), where $d = 1$, we consider two different choices for the noise. First, we consider space-time white noise corresponding to $s = 0$ and hence $\beta < 1/2$ and then a correlated noise corresponding $s = -1$ and hence $\beta < 3/2$. We note that since the eigenfunctions of Λ are given as $e_j = \sqrt{2}\sin(j\pi x)$, $j \geq 1$, (e_j, χ) can be computed exactly for $\chi = \varphi_i$, $i = 1, \cdots, N_h$, with $\{\varphi_i\}_{i=1}^{N_h}$ being a basis in V_h. Thus, in the case of space-time white noise, we do not expect convergence for the finite element approximation of velocity $u_{h,2}$ by (5.5), but we expect the rate of convergence to be 1/3 for displacement $u_{h,1}$ by (5.3). These are confirmed by Figure 2. In the second case, the expected rate of strong convergence is 1 and 1/3 for displacement and velocity by (5.3) and (5.5), respectively, as Figure 3 also confirms. We note that we have used a uniform spatial mesh and therefore with $Q = \Lambda^s$, the assumptions of Lemma 6.1 are fulfilled.

6.3. **Comments on numerical linear algebra.** On each time level the linear system (6.8) has to be solved. This can simply be done by the backslash operator "\" in Matlab, but it can be performed faster if instead we perform a minimum degree permutation of the coefficeint matrix Σ and then use the "LU" factorization of the permuted Σ. The coefficient matrix Σ in (6.8) is skew-symmetric, which implies that, in particular, $\Sigma_{ij} \neq 0$ if $\Sigma_{ji} \neq 0$.

This means that the command "symamd" in Matlab can be used. The algorithm for solving the linear system (6.8) is performed in the following steps, with obvious notations,

$$\begin{aligned}(P_{\mathrm{s}}\Sigma P_{\mathrm{s}}^T)P_{\mathrm{s}}X^n &= P_{\mathrm{s}}(\Xi X^{n-1}+b)\\ \hat{\Sigma}\hat{X}^n &= \hat{b}\\ L_{\mathrm{lu}}U_{\mathrm{lu}}\hat{X}^n &= P_{\mathrm{lu}}\hat{b}\\ \hat{X}^n &= U_{\mathrm{lu}}\backslash\big(L_{\mathrm{lu}}\backslash(P_{\mathrm{lu}}\hat{b})\big)\\ X^n &= P_{\mathrm{s}}^{-1}\hat{X}^n,\end{aligned}$$

where $P_{\mathrm{s}} = \mathrm{symamd}(\Sigma)$, and $[P_{\mathrm{lu}}, L_{\mathrm{lu}}, U_{lu}] = \mathrm{lu}\,(\hat{X}^n)$ in Matlab. With $h_{\mathrm{exact}} = 2^{-7}$ and $k = h_{\mathrm{exact}}^2$, the computation time for each realization, that is, the computation time of generating the Brownian motion, computing the exact solution and the approximated solutions with mesh sizes $h = 2^{-1}$ to $h = 2^{-5}$, takes approximatly 40 seconds with "\" while it takes 4 seconds with minimum degree permutation. The reason for this can be seen in Figure 4 and Figure 5, where the structure and the number of nonzero entries in the "LU" factorization of Σ and $\hat{\Sigma}$ are shown. An AMD Opteron computer with 15 Gigabytes RAM memory and 2.2 GHz CPU has been used for these experiments.

Remark 6.3. One might consider two ways to compute the vector b in (6.8). Either using matrix-matrix multiplication, that is, we need to generate the increments (6.5) at once in a big $N_t \times N_h$ matrix, or using vector-matrix multiplications that means we need to generate the increments (6.5) in a loop and each time in a vector $1 \times N_h$. We used the first idea since it is faster and there was enough memory for the computations. However, the size of the matrix of the increments, and hence the memory usage, grows considerably when refining the mesh and taking smaller time steps. For example, with $N_h = 2^7$ and $N_t = 2^{14}$, in our experiments 256 Mbytes RAM was needed for storing the increment matrix, while for $N_h = 2^8$ and $N_t = 2^{16}$, we needed almost 2 Gbytes. In the latter case we used the second approach, that is vector matrix multiplications, and the computation time for each realization took about 6 seconds.

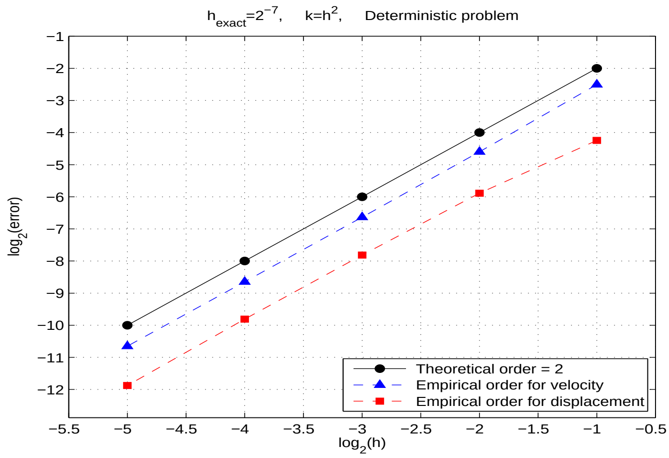


FIGURE 1. Deterministic problem: the order of strong convergence in the L_2-norm is 2 for both the displacement u (dashed-square) and the velocity $\dot{u}$ (dashed-triangle).

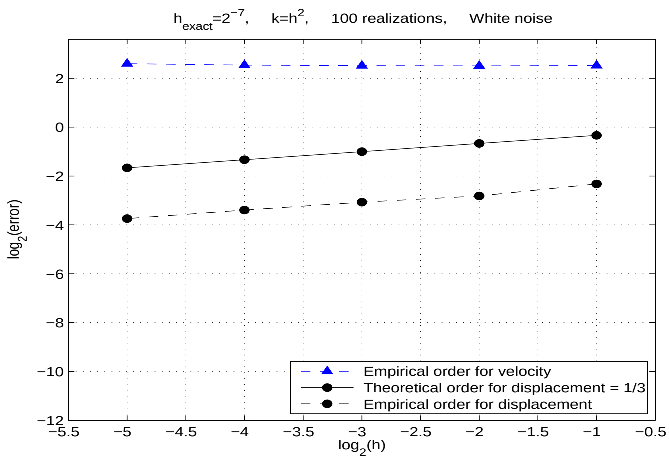


FIGURE 2. LSWE with white noise: the order of strong convergence in the L_2-norm is 1/3 for the displacement u (dashed-circle); but there is no convergence for the velocity $\dot{u}$ (dashed-triangle).

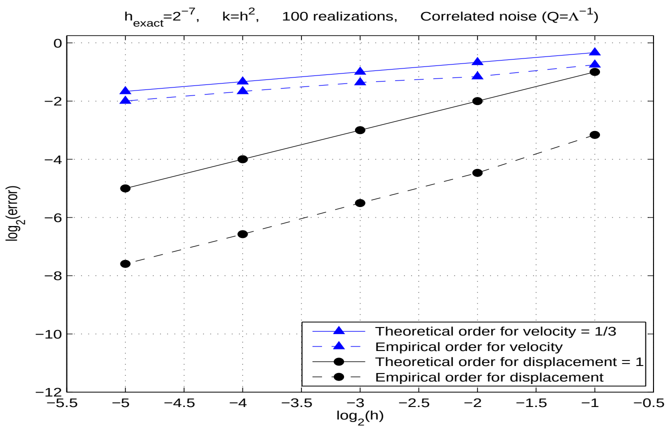


FIGURE 3. LSWE with correlated noise $Q = \Lambda^{-1}$: the order of strong convergence in the L_2-norm is 1 for the displacement u (dashed-circle), and 1/3 for the velocity $\dot{u}$ (dashed-triangle).

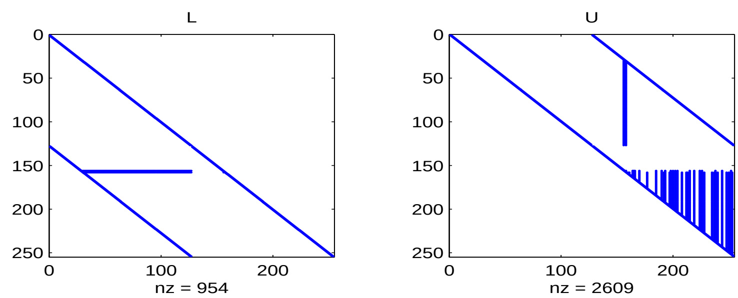


FIGURE 4. Structure and number of nonzero elements of LU(Σ)

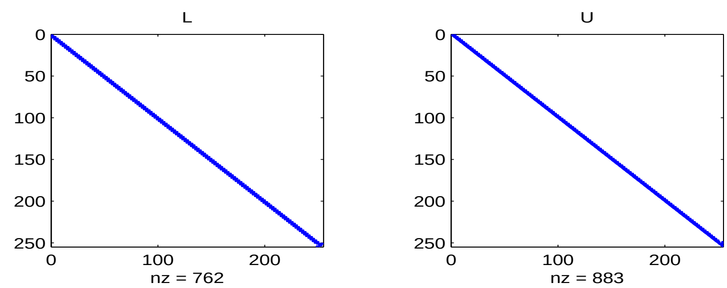


FIGURE 5. Structure and number of nonzero elements of LU($\hat{\Sigma}$).

References

1. E. Allen, *Modeling with Itô Stochastic Differential Equations*, Mathematical Modelling: Theory and Applications, vol. 22, Springer, Dordrecht, 2007.
2. E.J. Allen, S.J. Novosel, and Z. Zhang, *Finite element and difference approximation of some linear stochastic partial differential equations*, Stochastics Stochastics Rep. **64** (1998), 117–142.
3. G.A. Baker, *Error estimates for finite element methods for second order hyperbolic equations*, SIAM J. Numer. Anal. **13** (1976), 564–576.
4. C. Carstensen, *Merging the Bramble-Pasciak-Steinbach and the Crouzeix-Thomée criterion for H^1-stability of the L^2-projection onto finite element spaces*, Math. Comp. **71** (2002), no. 237, 157–163 (electronic).
5. ———, *An adaptive mesh-refining algorithm allowing for an H^1 stable L^2 projection onto Courant finite element spaces*, Constr. Approx. **20** (2004), no. 4, 549–564.
6. P. Chow, *Stochastic Partial Differential Equations*, Chapman & Hall/CRC, Boca Raton, FL, 2007.
7. M. Crouzeix and V. Thomée, *The stability in L_p and W_p^1 of the L_2-projection onto finite element function spaces*, Math. Comp. **48** (1987), no. 178, 521–532.
8. Q. Du and T. Zhang, *Numerical approximation of some linear stochastic partial differential equations driven by special additive noises*, SIAM J. Numer. Anal. **40** (2002), 1421–1445 (electronic).
9. D.G. Duffy, *Green's Functions with Applications*, Chapman & Hall/CRC, Boca Raton, FL, 2001.
10. T. Dupont, *L^2-estimates for Galerkin methods for second order hyperbolic equations*, SIAM J. Numer. Anal. **10** (1973), 880–889.
11. M. Geissert, M. Kovács, and S. Larsson, *Rate of weak convergence of the finite element method for the stochastic heat equation with additive noise*, BIT, to appear.
12. I. Gyöngy, *Approximations of stochastic partial differential equations*, Stochastic partial differential equations and applications (Trento, 2002), Lecture Notes in Pure and Appl. Math., vol. 227, Dekker, New York, 2002, pp. 287–307.
13. M. Kovács, S. Larsson, and F. Lindgren, *Strong convergence of the finite element method with truncated noise for semilinear parabolic stochastic equations with additive noise*, Numer. Algorithms (2009), to appear.
14. A. Martin, *Hyperbolic stochastic partial differential equations: Small balls and simulation; propagation of singularities.*, PhD Thesis, Aachen, Shaker Verlag. Neuherberg: Institut f. Biomathematik u. Biometrie, GSF Forschungszentrum f. Umwelt u. Gesundheit, 2002.
15. A. Martin, S.M. Prigarin, and G. Winkler, *Exact and fast numerical algorithms for the stochastic wave equation*, Int. J. Comput. Math. **80** (2003), 1535–1541.
16. G. Da Prato and J. Zabczyk, *Stochastic Equations in Infinite Dimensions*, Cambridge University Press, Cambridge, 1992.
17. C. Prévôt and M. Röckner, *A Concise Course on Stochastic Partial Differential Equations*, Lecture Notes in Mathematics, vol. 1905, Springer, Berlin, 2007.
18. L. Quer-Sardanyons and M. Sanz-Solé, *Space semi-discretisations for a stochastic wave equation*, Potential Anal. **24** (2006), 303–332.
19. J. Rauch, *On convergence of the finite element method for the wave equation*, SIAM J. Numer. Anal. **22** (1985), 245–249.
20. C. Roth, *Weak approximations of solutions of a first order hyperbolic stochastic partial differential equation*, Monte Carlo Methods Appl. **13** (2007), 117–133.

21. Ch. Schwab and R. A. Todor, *Karhunen-Loève approximation of random fields by generalized fast multipole methods*, J. Comput. Phys. **217** (2006), 100–122.
22. V. Thomée, *Galerkin Finite Element Methods for Parabolic Problems*, second ed., Springer-Verlag, Berlin, 2006.
23. J.B. Walsh, *An introduction to stochastic partial differential equations*, École d'été de probabilités de Saint-Flour, XIV—1984, Lecture Notes in Math., vol. 1180, Springer, Berlin, 1986, pp. 265–439.
24. ______, *On numerical solutions of the stochastic wave equation*, Illinois J. Math. **50** (2006), 991–1018 (electronic).
25. Y. Yan, *Semidiscrete Galerkin approximation for a linear stochastic parabolic partial differential equation driven by an additive noise*, BIT **44** (2004), 829–847.
26. ______, *Galerkin finite element methods for stochastic parabolic partial differential equations*, SIAM J. Numer. Anal. **43** (2005), 1363–1384 (electronic).

Department of Mathematics and Statistics, University of Otago, P. O. Box 56, Dunedin, New Zealand
E-mail address: mihaly@chalmers.se

Department of Mathematical Sciences, Chalmers University of Technology and University of Gothenburg, SE–412 96 Göteborg, Sweden
E-mail address: stig@chalmers.se

Department of Mathematical Sciences, Chalmers University of Technology and University of Gothenburg, SE–412 96 Göteborg, Sweden
E-mail address: fardin@chalmers.se