# Journal of Mechanics of Materials and Structures

Special issue on

## Coupled Field Problems and Multiphase Materials

Volume 12, No. 1 January 2017

msp

# JOURNAL OF MECHANICS OF MATERIALS AND STRUCTURES

msp.org/jomms

Founded by Charles R. Steele and Marie-Louise Steele

Cover photo: Ev Shafrir

See msp.org/jomms for submission guidelines.

JoMMS (ISSN 1559-3959) at Mathematical Sciences Publishers, 798 Evans Hall #6840, c/o University of California, Berkeley, CA 94720-3840, is published in 10 issues a year. The subscription price for 2017 is US $615/year for the electronic version, and $775/year (+$60, if shipping outside the US) for print and electronic. Subscriptions, requests for back issues, and changes of address should be sent to MSP.

JoMMS peer-review and production is managed by EditFlow® from Mathematical Sciences Publishers.

PUBLISHED BY
mathematical sciences publishers
nonprofit scientific publishing
http://msp.org/

JOURNAL OF MECHANICS OF MATERIALS AND STRUCTURES
Vol. 12, No. 1, 2017
dx.doi.org/10.2140/jomms.2017.12.1

# PREFACE

This special issue presents full versions of selected talks given at the minisymposium “Theoretical, Computational and Experimental Mechanics for Coupled Field Problems and Multiphase Materials”, held during the joint scientific meeting that took place in Gdansk, Poland on September 8–11, 2015, uniting the Third Polish Congress on Mechanics (PCM) and the Twenty-First International Conference on Computer Methods in Mechanics (CMM).

The investigation of coupled field problems is of pressing interest in many areas of science. Coupling effects influence experimental measurements, which in turn creates the need for sophisticated physical models and presents challenges for their numerical treatment, bringing to bear a vast array of scientific techniques. Modeling tools designed to incorporate size effects in time and space, such as fractional derivatives, enrich the underlying concepts. This drives new research in many subfields of mechanics.

The papers contained in this special issue address topics related to the constitutive modeling and numerical treatment of coupled field problems and multiphase materials. Emphasis is placed on relating theory to experimental observations, on nonconventional mathematical methods, and on new concepts and developments in the computational algorithms for the solution of the governing equations. The general topics addressed include model adaptation to experimental data, model identification and validation, fractional calculus and its applications in mechanics, nonlocal (scale) effects, and new concepts in the computational treatment for mechanical and thermomechanical problems.

The Guest Editors express their thanks to the *Journal of Mechanics of Materials and Structures* for the opportunity to edit this special issue.

CORINA S. DRAPACA
Pennsylvania State University

STEFAN HARTMANN
Clausthal University of Technology

JACEK LESZCZYŃSKI
AGH University of Science and Technology, Kraków

SIVABAL SIVALOGANATHAN
University of Waterloo

WOJCIECH SUMELKA
Poznań University of Technology

*Guest Editors*

JOURNAL OF MECHANICS OF MATERIALS AND STRUCTURES
Vol. 12, No. 1, 2017
dx.doi.org/10.2140/jomms.2017.12.3

# VARIATIONAL METHODS FOR THE SOLUTION OF FRACTIONAL DISCRETE/CONTINUOUS STURM–LIOUVILLE PROBLEMS

RICARDO ALMEIDA, AGNIESZKA B. MALINOWSKA,
M. LUÍSA MORGADO AND TATIANA ODZIJEWICZ

The fractional Sturm–Liouville eigenvalue problem appears in many situations, e.g., while solving anomalous diffusion equations coming from physical and engineering applications. Therefore, obtaining solutions or approximations of solutions to this problem is of great importance. Here, we describe how the fractional Sturm–Liouville eigenvalue problem can be formulated as a constrained fractional variational principle and show how such formulation can be used in order to approximate the solutions. Numerical examples are given to illustrate the method.

## 1. Introduction

Fractional calculus is a mathematical approach dealing with integral and differential terms of noninteger order. The concept of fractional calculus appeared shortly after calculus itself, but the development of practical applications proceeded very slowly. Only during the last few decades, fractional problems have increasingly attracted the attention of many researchers. Applications of fractional operators include chaotic dynamics [Zaslavsky 2005], material sciences [Mainardi 2010], mechanics of fractal and complex media [Carpinteri and Mainardi 1997; Li and Ostoja-Starzewski 2011], quantum mechanics [Hilfer 2000], physical kinetics [Zaslavsky and Edelman 2004] and many others (see, e.g., [Domek and Pworak 2016; Tarasov 2010]). Fractional derivatives are nonlocal operators and therefore successfully applied in the study of nonlocal or time-dependent processes [Podlubny 1999]. The well-established application of fractional calculus in physics is in the framework of anomalous diffusion behavior [Blaszczyk and Ciesielski 2014; Chen et al. 2012; D'Ovidio 2012; Leonenko et al. 2013; Meerschaert 2012; Metzler and Klafter 2000]: large jumps in space are modeled by space-fractional derivatives of order between 1 and 2, while long waiting times are modeled by the time derivatives of order between 0 and 1. These partial fractional differential equations can be solved by the method of separating variables, which leads to the Sturm–Liouville and the Cauchy equations. It means that, if we are able to solve the fractional Sturm–Liouville problem and the Cauchy problem, then we can find a solution to the fractional diffusion equation. In this paper, we consider two basic approaches to the fractional Sturm–Liouville problem: discrete and continuous. In both cases, we note that the problem can be formulated as a constrained fractional variational principle. A fractional variational problem consists of finding the extremizer of a functional that depends on fractional derivatives (differences) subject to boundary conditions and possibly some extra constraints. It is worth pointing out that the fractional calculus of variations has itself remarkable applications in classical mechanics. Riewe [1996; 1997] showed that a Lagrangian involving

---

*Keywords:* fractional Sturm–Liouville problem, fractional calculus of variations, discrete fractional calculus, continuous fractional calculus.

fractional time derivatives leads to an equation of motion with nonconservative forces such as friction. For more about the fractional calculus of variations, we refer the reader to [Almeida et al. 2015; Klimek 2009; Malinowska and Torres 2012; Malinowska et al. 2015] while for various approaches to fractional Sturm–Liouville problems we refer to [Al-Mdallal 2009; 2010; Klimek et al. 2014; Klimek 2015; 2016; Zayernouri and Karniadakis 2013].

The paper is divided into two main parts dedicated to discrete (Section 2) and continuous (Section 3) fractional problems. In the first part, we give a constructive proof of the existence of orthogonal solutions to the discrete fractional Sturm–Liouville eigenvalue problem (Theorem 2.4) and show that the smallest and largest eigenvalues can be characterized as the optimal values of certain functionals (Theorems 2.5 and 2.7). Our results are illustrated by an example. In the second part, we recall the fractional variational principle and the spectral theorem for the continuous fractional Sturm–Liouville problem. Since for most problems involving fractional derivatives (equations or variational problems) one cannot provide methods to compute the exact solutions analytically, numerical methods should be used for solving such problems. Discretizing both the fractional Sturm–Liouville equation and the related isoperimetric variational problem, we show, by an example, how the variational method can be used for solving the fractional Sturm–Liouville problem.

## 2. Discrete fractional calculus

In this section, we explain a relationship between the fractional Sturm–Liouville difference problem and a constrained discrete fractional variational principle. Namely, it is possible to look for solutions of Sturm–Liouville fractional difference equations by solving finite-dimensional constrained optimization problems. We shall start with necessary preliminaries. There are various versions of fractional differences; we mention here those of [Díaz and Osler 1974; Miller and Ross 1989; Atıcı and Eloe 2009a; 2009b] and the Caputo difference [Abdeljawad 2011]. In this paper, we use the notion of Grünwald and Letnikov [Kaczorek 2011; Podlubny 1999].

Let us define the mesh points $x_j = a + jh$, $j = 0, 1, \dots, N$, where $h$ denotes the uniform space step, and set $D = \{x_0, \dots, x_N\}$. In what follows, $\alpha \in \mathbb{R}$ and $0 < \alpha \le 1$. Moreover, we set

$$a_i^{(\alpha)} := \begin{cases} 1 & \text{if } i = 0, \\ (-1)^i(\alpha(\alpha-1)\cdots(\alpha-i+1))/i! & \text{if } i = 1, 2, \dots. \end{cases} \tag{2-1}$$

**Definition 2.1.** The backward fractional difference of order $\alpha$, where $0 < \alpha \le 1$, of function $f : D \to \mathbb{R}$ is defined by

$${}_0\Delta_k^\alpha f(x_k) := \frac{1}{h^\alpha} \sum_{i=0}^{k} (-1)^i \frac{\alpha(\alpha-1)\cdots(\alpha-i+1)}{i!} f(x_{k-i}) \tag{2-2}$$

while

$${}_k\Delta_N^\alpha f(x_k) := \frac{1}{h^\alpha} \sum_{i=0}^{N-k} (-1)^i \frac{\alpha(\alpha-1)\cdots(\alpha-i+1)}{i!} f(x_{k+i}) \tag{2-3}$$

is the forward fractional difference of function $f$.

Fractional backward and forward differences are linear operators.

**Theorem 2.2** [Ostalczyk 2008]. *Let $f$ and $g$ be two real functions defined on $D$ and $\beta, \gamma \in \mathbb{R}$. Then*

$$
\begin{aligned}
{}_0\Delta_k^\alpha[\gamma f(x_k) + \beta g(x_k)] &= \gamma\, {}_0\Delta_k^\alpha f(x_k) + \beta\, {}_0\Delta_k^\alpha g(x_k),\\
{}_k\Delta_N^\alpha[\gamma f(x_k) + \beta g(x_k)] &= \gamma\, {}_k\Delta_N^\alpha f(x_k) + \beta\, {}_k\Delta_N^\alpha g(x_k)
\end{aligned}
$$

*for all $k$.*

The following formula for summation by parts for fractional operators will be essential for proving results concerning variational problems.

**Lemma 2.3** [Bourdin et al. 2013]. *Let $f$ and $g$ be two real functions defined on $D$. Then*

$$\sum_{k=0}^{N} g(x_k)\, {}_0\Delta_k^\alpha f(x_k) = \sum_{k=0}^{N} f(x_k)\, {}_k\Delta_N^\alpha g(x_k).$$

*If $f(x_0) = f(x_N) = 0$ or $g(x_0) = g(x_N) = 0$, then*

$$\sum_{k=1}^{N} g(x_k)\, {}_0\Delta_k^\alpha f(x_k) = \sum_{k=0}^{N-1} f(x_k)\, {}_k\Delta_N^\alpha g(x_k). \tag{2-4}$$

**2A.** ***The Sturm–Liouville problem.*** The topic of this subsection is the Sturm–Liouville fractional difference equation

$${}_k\Delta_N^\alpha(p(x_k)\, {}_0\Delta_k^\alpha y(x_k)) + q(x_k)y(x_k) = \lambda r(x_k)y(x_k), \quad k = 1, \dots, N-1, \tag{2-5}$$

with boundary conditions

$$y(x_0) = 0, \quad y(x_N) = 0. \tag{2-6}$$

We assume that $p(x_i) > 0$, $r(x_i) > 0$, $q(x_i)$ is defined and real-valued for all $x_i$, $i = 0, \dots, N$, and $\lambda$ is a parameter. It is required to find the eigenfunctions and the eigenvalues of the given boundary value problem, i.e., the nontrivial solutions of (2-5)–(2-6) and the corresponding values of the parameter $\lambda$. The theorem below gives an answer to this question.

**Theorem 2.4.** *The Sturm–Liouville problem* (2-5)–(2-6) *has $N-1$ real eigenvalues, which we denote by*

$$\lambda_1 \le \lambda_2 \le \cdots \le \lambda_{N-1}.$$

*The corresponding eigenfunctions,*

$$y^1, y^2, \dots, y^{N-1} : \{x_1, \dots, x_{N-1}\} \to \mathbb{R},$$

*are mutually orthogonal: if $i \neq j$, then*

$$\langle y^i, y^j\rangle_r := \sum_{k=1}^{N-1} r(x_k) y^i(x_k) y^j(x_k) = 0.$$

*Furthermore, they span $\mathbb{R}^{N-1}$: any vector $\varphi = (\varphi(x_k))_{k=1}^{N-1} \in \mathbb{R}^{N-1}$ has a unique expansion*

$$\varphi(x_k) = \sum_{i=1}^{N-1} c_i y^i(x_k), \quad 1 \le k \le N-1.$$

*The coefficients* $c_i$ *are given by*

$$c_i = \frac{\langle \varphi, y^i \rangle_r}{\langle y^i, y^i \rangle_r}.$$

*Proof.* Observe that equations (2-5)–(2-6) can be considered as a system of $N-1$ linear equations with $N-1$ real unknowns $y(x_1), \dots, y(x_{N-1})$. The corresponding matrix form is

$$Ay^T = \lambda R y^T, \tag{2-7}$$

where the entries $A_{ij}$ of $A$ are

$$A_{ij}^{(\alpha)} = \begin{cases} (1/h^{2\alpha})\big[q(x_i) + \sum_{k=0}^{N-i} (a_k^{(\alpha)})^2 p(x_{i+k})\big] & \text{if } i = j, \\ (1/h^{2\alpha})\big[\sum_{k=0}^{N-i} a_k^{(\alpha)} p(x_{i+k}) \sum_{m=0}^{k+i} a_m^{(\alpha)}\big] \text{ and } k-m+i=j & \text{if } i \neq j \end{cases}$$

and $R = \operatorname{diag}\{r(x_1), \dots, r(x_{N-1})\}$. Writing (2-7) as

$$R^{-1} A y^T = \lambda y^T, \tag{2-8}$$

we get an eigenvalue problem with the symmetric matrix $R^{-1}A$. Because of the equivalence of problem (2-5)–(2-6) to problem (2-8), it follows from matrix theory that the Sturm–Liouville problem (2-5)–(2-6) has $N-1$ linearly pairwise orthogonal real independent eigenfunctions with all eigenvalues real. Now we would like to find constants $c_1, \dots, c_{N-1}$ such that $\varphi(x_k) = \sum_{i=1}^{N-1} c_i y^i(x_k)$, $1 \le k \le N-1$. Note that

$$\langle \varphi, y^j \rangle_r = \left\langle \sum_{i=1}^{N-1} c_i y^i, y^j \right\rangle_r = \sum_{i=1}^{N-1} c_i \langle y^i, y^j \rangle_r = c_j \langle y^j, y^j \rangle_r$$

because of orthogonality. Therefore, $c_i = \langle \varphi, y^i \rangle_r / \langle y^i, y^i \rangle_r$, $1 \le i \le N-1$. □

**2B.** ***Isoperimetric variational problems.*** In this section, we prove two theorems connecting the Sturm–Liouville problem (2-5)–(2-6) to isoperimetric problems of discrete fractional calculus of variations.

**Theorem 2.5.** *Let* $y^1$ *denote the first eigenfunction, normalized to satisfy the isoperimetric constraint*

$$I[y] = \sum_{k=1}^{N} r(x_k)(y(x_k))^2 = 1 \tag{2-9}$$

*associated with the first eigenvalue* $\lambda_1$ *of problem* (2-5)–(2-6). *Then* $y^1$ *is a minimizer of functional*

$$J[y] = \sum_{k=1}^{N} \big[p(x_k)({}_0\Delta_k^{\alpha}\, y(x_k))^2 + q(x_k)(y(x_k))^2\big] \tag{2-10}$$

*subject to boundary conditions* $y(x_0) = 0$ *and* $y(x_N) = 0$ *and isoperimetric constraint* (2-9). *Moreover*, $J[y^1] = \lambda_1$.

*Proof.* Suppose that $y$ is a minimizer of $J$. Then by [Malinowska and Odzijewicz 2016, Theorem 5], there exists a real constant $\lambda$ such that $y$ satisfies

$${}_k\Delta_N^{\alpha}(p(x_k)\, {}_0\Delta_k^{\alpha}\, y(x_k)) + q(x_k)y(x_k) - \lambda r(x_k) y(x_k) = 0, \quad k = 1, \dots, N-1, \tag{2-11}$$

together with $y(x_0)=0$ and $y(x_N)=0$ and isoperimetric constraint (2-9). Let us multiply (2-11) by $y(x_k)$ and sum up from $k=1$ to $N-1$; then

$$\sum_{k=1}^{N-1}\left[y(x_k)\,{}_k\Delta_N^\alpha(p(x_k)\,{}_0\Delta_k^\alpha\, y(x_k))+q(x_k)(y(x_k))^2\right]=\sum_{k=1}^{N-1}\lambda r(x_k)(y(x_k))^2.$$

By summation by parts (2-4),

$$\sum_{k=1}^{N-1} y(x_k)\,{}_k\Delta_N^\alpha(p(x_k)\,{}_0\Delta_k^\alpha\, y(x_k))=\sum_{k=1}^{N} p(x_k)({}_0\Delta_k^\alpha\, y(x_k))^2.$$

As (2-9) holds and $y(x_N)=0$, we obtain

$$J[y]=\lambda.$$

Any solution to problem (2-9)–(2-10) that satisfies (2-11) must be nontrivial since (2-9) holds, so $\lambda$ must be an eigenvalue. According to Theorem 2.4, eigenvalue $\lambda_1$ is the smallest element of the spectrum and has corresponding eigenfunction $y^{(1)}$ normalized to meet the isoperimetric condition. Therefore, $J[y^{(1)}]=\lambda_1$. □

**Definition 2.6.** We will call functional $R$ defined by

$$R[y]=\frac{J[y]}{I[y]},$$

where $J[y]$ is given by (2-10) and $I[y]$ by (2-9), the Rayleigh quotient for the fractional discrete Sturm–Liouville problem (2-5)–(2-6).

**Theorem 2.7.** *Assume that $y$ satisfies boundary conditions $y(x_0)=y(x_N)=0$ and is nontrivial.*

(i) *If $y$ is a minimizer of Rayleigh quotient $R$ for the Sturm–Liouville problem* (2-5)–(2-6), *then the value of $R$ in $y$ is equal to the smallest eigenvalue $\lambda_1$, i.e.,* $R[y]=\lambda_1$.

(ii) *If $y$ is a maximizer of Rayleigh quotient $R$ for the Sturm–Liouville problem* (2-5)–(2-6), *then the value of $R$ in $y$ is equal to the largest eigenvalue $\lambda_{N-1}$, i.e.,* $R[y]=\lambda_{N-1}$.

*Proof.* We give the proof only for (i) as the second case can be proved similarly. Suppose that $y$ satisfying boundary conditions $y(x_0)=y(x_N)=0$ and being nontrivial is a minimizer of Rayleigh quotient $R$ and that value of $R$ in $y$ is equal to $\lambda$. Consider the functions

$$\begin{aligned}
\phi:[-\varepsilon,\varepsilon]&\to\mathbb{R},\\
h&\mapsto I[y+h\eta]=\sum_{k=1}^{N} r(x_k)(y(x_k)+h\eta(x_k))^2,\\
\psi:[-\varepsilon,\varepsilon]&\to\mathbb{R},\\
h&\mapsto J[y+h\eta]=\sum_{k=1}^{N}\left[p(x_k)({}_0\Delta_k^\alpha(y(x_k)+h\eta(x_k)))^2+q(x_k)(y(x_k)+h\eta(x_k))^2\right],\\
\zeta:[-\varepsilon,\varepsilon]&\to\mathbb{R},\\
h&\mapsto R[y+h\eta]=\frac{J[y+h\eta]}{I[y+h\eta]},
\end{aligned}$$

where $\eta : D \to \mathbb{R}$ with $\eta(x_0) = \eta(x_N) = 0$ and $\eta \neq 0$. Since $\zeta$ is of class $C^1$ on $[-\varepsilon, \varepsilon]$ and

$$\zeta(0) \le \zeta(h), \quad |h| \le \varepsilon,$$

we deduce that

$$\zeta'(0) = \frac{d}{dh} R[y + h\eta]\bigg|_{h=0} = 0.$$

Moreover, notice that

$$\zeta'(h) = \frac{1}{\phi(h)}\left(\psi'(h) - \frac{\psi(h)}{\phi(h)}\phi'(h)\right)$$

and

$$\psi'(0) = \frac{d}{dh} J[y + h\eta]\bigg|_{h=0} = 2\sum_{k=1}^{N}\bigl[p(x_k)\, {}_0\Delta_k^\alpha\, y(x_k)\, {}_0\Delta_k^\alpha\, \eta(x_k) + q(x_k) y(x_k)\eta(x_k)\bigr],$$

$$\phi'(0) = \frac{d}{dh} I[y + h\eta]\bigg|_{h=0} = 2\sum_{k=1}^{N} r(x_k) y(x_k)\eta(x_k).$$

Therefore,

$$\begin{aligned}\zeta'(0) &= \frac{d}{dh} R[y + h\eta]\bigg|_{h=0} \\ &= \frac{2}{I[y]}\left[\sum_{k=1}^{N}\bigl[p(x_k)\, {}_0\Delta_k^\alpha\, y(x_k)\, {}_0\Delta_k^\alpha\, \eta(x_k) + q(x_k)y(x_k)\eta(x_k)\bigr] - \frac{J[y]}{I[y]}\sum_{k=1}^{N} r(x_k)y(x_k)\eta(x_k)\right] = 0.\end{aligned}$$

Having in mind that $J[y]/I[y] = \lambda$ and $\eta(x_0) = \eta(x_N) = 0$ and using summation by parts (2-4), we obtain

$$\sum_{k=1}^{N-1}\bigl[{}_k\Delta_N^\alpha(p(x_k)\, {}_0\Delta_k^\alpha\, y(x_k)) + q(x_k)y(x_k) - \lambda r(x_k)y(x_k)\bigr]\eta(x_k) = 0.$$

Since $\eta$ is arbitrary, we have

$${}_k\Delta_N^\alpha(p(x_k)\, {}_0\Delta_k^\alpha\, y(x_k)) + q(x_k)y(x_k) - \lambda r(x_k)y(x_k) = 0, \quad k = 1, \dots, N-1. \tag{2-12}$$

As $y \neq 0$, we have that $\lambda$ is an eigenvalue of (2-12). On the other hand, let $\lambda_i$ be an eigenvalue and $y^i$ the corresponding eigenfunction; then

$${}_k\Delta_N^\alpha(p(x_k)\, {}_0\Delta_k^\alpha\, y_i(x_k)) + q(x_k)y_i(x_k) = \lambda_i r(x_k)y^i(x_k). \tag{2-13}$$

Similarly to the proof of Theorem 2.5, we can obtain

$$\frac{\sum_{k=1}^{N}\bigl[p(x_k)({}_0\Delta_k^\alpha\, y^i(x_k))^2 + q(x_k)(y^i(x_k))^2\bigr]}{\sum_{k=1}^{N} r(x_k)(y^i(x_k))^2} = \lambda_i$$

for any $1 \le i \le N-1$. That is, $R[y^i] = J[y^i]/I[y^i] = \lambda_i$. Finally, since the minimum value of $R$ at $y$ is equal to $\lambda$, i.e.,

$$\lambda \le R[y^i] = \lambda_i \quad \text{for all } i \in \{1, \dots, N-1\},$$

we have $\lambda = \lambda_1$. □

| $\alpha$ | $\lambda_1$ | $\lambda_2$ | $\lambda_3$ |
|---|---|---|---|
| 0.25 | 0.7102065750 | 1.148567387 | 1.349294886 |
| 0.50 | 0.6004483933 | 1.353660384 | 1.831047473 |
| 0.75 | 0.5779798778 | 1.632135974 | 2.496488153 |
| 1 | 0.5857864376 | 2.0 | 3.414213562 |

**Table 1.** Eigenvalues of (2-16) for different values of $\alpha$.

**Example 2.8.** Let us consider the following problem: minimize

$$J[y]=\sum_{k=1}^{N}\left({}_0\Delta_k^\alpha\, y(x_k)\right)^2 \tag{2-14}$$

subject to

$$I[y]=\sum_{k=1}^{N}(y(x_k))^2=1 \tag{2-15}$$

and $y(x_0)=y(x_N)=0$, where $N$ is fixed. In this case, the Euler–Lagrange equation takes the form

$${}_k\Delta_N^\alpha\, {}_0\Delta_k^\alpha\, y(x_k)=\lambda y(x_k), \quad k=1,\dots,N-1. \tag{2-16}$$

Together with boundary conditions $y(x_0)=y(x_N)=0$, it is the Sturm–Liouville eigenvalue problem where $p(x_i)=1$, $r(x_i)=1$ and $q(x_i)=0$ for $k=1,\dots,N-1$. Let us choose $N=4$ and $h=1$. Eigenvalues of (2-16) for different values of $\alpha$ are presented in Table 1. Those results are obtained by solving the matrix eigenvalue problem of the form (2-8).

Observe that problem (2-14)–(2-15) can be treated as a finite-dimensional constrained optimization problem. Namely, the problem is to minimize function $J$ of $N-1$ variables $y_1=y(x_1),\dots,y_{N-1}=y(x_{N-1})$ on the $(N-1)$-dimensional sphere with equation $\sum_{k=1}^{N-1}y_k^2=1$. Table 2 and Figure 1 present the solution to problem (2-14)–(2-15) for $N=4$, $h=1$ and different values of $\alpha$. By Theorem 2.5, the first eigenvalue $\lambda_1$ of (2-16) is the minimum value of $J$ on $\sum_{k=1}^{N-1}y_k^2=1$ and the first eigenfunction of (2-16) is the minimizer of this problem. Other eigenfunctions and eigenvalues of (2-16) we can find by using the first-order necessary optimality conditions (Karush–Kuhn–Tucker conditions), that is, by solving the system of equations

$$\begin{cases}\dfrac{\partial J}{\partial y_k}=\lambda\dfrac{\partial I}{\partial y_k}, \\ \sum_{k=1}^{N-1}y_k^2=1,\end{cases} \qquad k=1,\dots,N-1. \tag{2-17}$$

| $\alpha$ | $y(x_1)$ | $y(x_2)$ | $y(x_3)$ | $\lambda_1$ |
|---|---|---|---|---|
| 0.25 | 0.52042378274 | 0.65949734450 | 0.54242265711 | 0.7102065749 |
| 0.50 | 0.50954825567 | 0.67778735991 | 0.53006119446 | 0.6004483933 |
| 0.75 | 0.50509466979 | 0.69443334582 | 0.51248580736 | 0.5779798777 |
| 1 | 0.49999999999 | 0.70710678118 | 0.5 | 0.5857864376 |

**Table 2.** The solution to problem (2-14)–(2-15) for different values of $\alpha$.

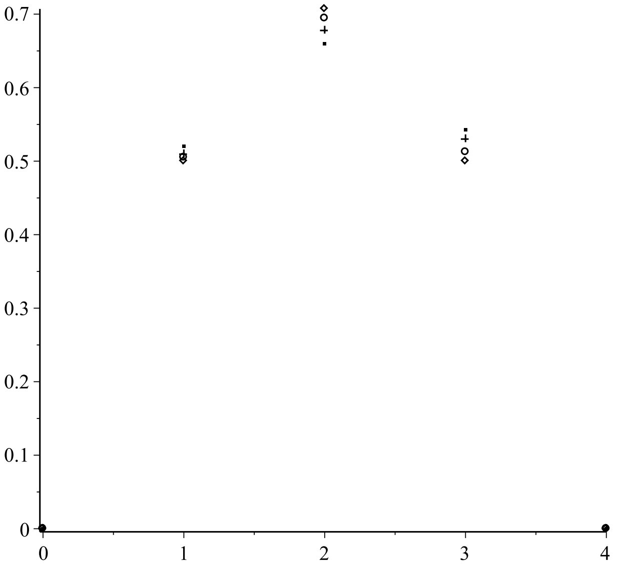

**Figure 1.** The solution to problem (2-14)–(2-15) for $\alpha = \frac{1}{4}$ $(\cdot)$, $\frac{1}{2}$ $(+)$, $\frac{3}{4}$ $(\circ)$ and 1 $(\diamond)$.

## 3. Continuous fractional calculus

This section is devoted to the continuous fractional Sturm–Liouville problem and its formulation as a constrained fractional variational principle. Namely, we shall show that this formulation can be used to approximate the solutions. As in the discrete case, there are several different definitions for fractional derivatives [Kilbas et al. 2006]; the most well-known are the Grünwald–Letnikov, the Riemann–Liouville and the Caputo fractional derivatives.

**Definition 3.1.** Let $f : [a, b] \to \mathbb{R}$ be a function and $\alpha$ a positive real number such that $0 < \alpha < 1$. We define the left and right Riemann–Liouville fractional derivatives of order $\alpha$ by

$$
{}_aD_x^\alpha f(x) := \frac{1}{\Gamma(1-\alpha)} \frac{d}{dx} \int_a^x (x-t)^{-\alpha} f(t)\, dt,
$$

$$
{}_xD_b^\alpha f(x) := \frac{-1}{\Gamma(1-\alpha)} \frac{d}{dx} \int_x^b (t-x)^{-\alpha} f(t)\, dt
$$

and the left and right Caputo fractional derivatives of order $\alpha$ by

$$
{}_a^CD_x^\alpha f(x) := \frac{1}{\Gamma(1-\alpha)} \int_a^x (x-t)^{-\alpha} f'(t)\, dt,
$$

$$
{}_x^CD_b^\alpha f(x) := \frac{-1}{\Gamma(1-\alpha)} \int_x^b (t-x)^{-\alpha} f'(t)\, dt.
$$

The Caputo derivative seems more suitable in applications. Let us recall that the Caputo derivative of a constant is zero, whereas the Riemann–Liouville is not. Moreover, the Laplace transform, which is used for solving fractional differential equations, of the Riemann–Liouville derivative contains the limit values of the Riemann–Liouville fractional derivatives (of order $\alpha - 1$) at the lower terminal $x = a$.

Mathematically such problems can be solved, but there is no physical interpretation for such conditions. On the other hand, the Laplace transform of the Caputo derivative imposes boundary conditions involving the value of the function at the lower point $x = a$, which usually are acceptable physical conditions.

The Grünwald–Letnikov definition is a generalization of the ordinary discretization formulas for integer-order derivatives.

**Definition 3.2.** Let $0 < \alpha < 1$ be real. The left and right Grünwald–Letnikov fractional derivatives of order $\alpha$ of a function $f$ are defined as

$$
\begin{aligned}
{}^{\mathrm{GL}}_{\ a}D^{\alpha}_{x} f(x) &:= \lim_{h\to 0^+} \frac{1}{h^{\alpha}} \sum_{k=0}^{\infty} (-1)^k \binom{\alpha}{k} f(x-kh),\\
{}^{\mathrm{GL}}_{\ x}D^{\alpha}_{b} f(x) &:= \lim_{h\to 0^+} \frac{1}{h^{\alpha}} \sum_{k=0}^{\infty} (-1)^k \binom{\alpha}{k} f(x+kh).
\end{aligned}
$$

Here $\binom{\alpha}{k}$ stands for the generalization of binomial coefficients to real numbers (see (2-1)). However, in this section for historical reasons, we denote

$$
(w^{\alpha}_{k}) := (-1)^k \binom{\alpha}{k}
$$

rather than $a^{(\alpha)}_{i}$.

Relations between those three types of derivatives are given below.

**Proposition 3.3** [Podlubny 1999]. *Let us assume the function $f$ is integrable in $[a, b]$. Then the Riemann–Liouville fractional derivatives exist and coincide with Grünwald–Letnikov fractional derivatives.*

**Proposition 3.4** [Kilbas et al. 2006]. *Let us assume that $f$ is a function for which the Caputo fractional derivatives exist together with the Riemann–Liouville fractional derivatives in $[a, b]$. Then, if $0 < \alpha < 1$,*

$$
\begin{aligned}
{}^{C}_{a}D^{\alpha}_{x} f(x) &= {}_{a}D^{\alpha}_{x} f(x) - \frac{f(a)}{\Gamma(1-\alpha)}(x-a)^{-\alpha},\\
{}^{C}_{x}D^{\alpha}_{b} f(x) &= {}_{x}D^{\alpha}_{b} f(x) - \frac{f(b)}{\Gamma(1-\alpha)}(b-x)^{-\alpha}.
\end{aligned}
\tag{3-1}
$$

*If $f(a) = 0$ or $f(b) = 0$, then ${}^{C}_{a}D^{\alpha}_{x} f(x) = {}_{a}D^{\alpha}_{x} f(x)$ or ${}^{C}_{x}D^{\alpha}_{b} f(x) = {}_{x}D^{\alpha}_{b} f(x)$, respectively.*

It is well-known that we can approximate the Riemann–Liouville fractional derivative using the Grünwald–Letnikov fractional derivative. Given the interval $[a, b]$ and a partition of the interval $x_j = a + jh$, for $j = 0, 1, \ldots, N$ and some $h > 0$ such that $x_N = b$, we have

$$
\begin{aligned}
{}_{a}D^{\alpha}_{x_j} f(x_j) &= \frac{1}{h^{\alpha}} \sum_{k=0}^{j} (w^{\alpha}_{k}) f(x_{j-k}) + O(h),\\
{}_{x_j}D^{\alpha}_{b} f(x_j) &= \frac{1}{h^{\alpha}} \sum_{k=0}^{N-j} (w^{\alpha}_{k}) f(x_{j+k}) + O(h);
\end{aligned}
$$

that is, the truncated Grünwald–Letnikov fractional derivatives are first-order approximations of the Riemann–Liouville fractional derivatives. Using relations (3-1), we deduce a decomposition sum for the Caputo fractional derivatives:

$$ {}_a^C D_{x_j}^{\alpha} f(x_j) \approx \frac{1}{h^{\alpha}} \sum_{k=0}^{j} (w_k^{\alpha}) f(x_{j-k}) - \frac{f(a)}{\Gamma(1-\alpha)} (x_j - a)^{-\alpha} =: {}_a^C \widetilde{D}_{x_j}^{\alpha} f(x_j), \tag{3-2} $$

$$ {}_{x_j}^{C} D_b^{\alpha} f(x_j) \approx \frac{1}{h^{\alpha}} \sum_{k=0}^{N-j} (w_k^{\alpha}) f(x_{j+k}) - \frac{f(b)}{\Gamma(1-\alpha)} (b - x_j)^{-\alpha} =: {}_{x_j}^{C} \widetilde{D}_b^{\alpha} f(x_j). \tag{3-3} $$

**3A.** ***Variational problem.*** Consider the following variational problem: minimize the functional

$$ I[y] = \int_a^b L(x, y(x), {}_a^C D_x^{\alpha} y(x))\, dx, \tag{3-4} $$

subject to the boundary conditions

$$ y(a) = y_a, \quad y(b) = y_b, \quad y_a, y_b \in \mathbb{R}, \tag{3-5} $$

where $0 < \alpha < 1$ and the Lagrange function $L : [a, b] \times \mathbb{R}^2 \to \mathbb{R}$ is differentiable with respect to the second and third arguments.

**Theorem 3.5** [Agrawal 2006]. *If $\bar{y}$ is a solution to* (3-4)–(3-5), *then $\bar{y}$ satisfies the fractional differential equation*

$$ \frac{\partial L}{\partial y}(x, y(x), {}_a^C D_x^{\alpha} y(x)) + {}_x D_b^{\alpha} \frac{\partial L}{\partial\, {}_a^C D_x^{\alpha} y}(x, y(x), {}_a^C D_x^{\alpha} y(x)) = 0, \quad t \in [a, b]. \tag{3-6} $$

Relations like (3-6) are known in the literature as the Euler–Lagrange equation and provide a necessary condition that every solution of the variational problem must satisfy. Adding to problem (3-4)–(3-5) an integral constraint

$$ \int_a^b g(x, y(x), {}_a^C D_x^{\alpha} y(x))\, dx = K, \tag{3-7} $$

where $K$ is a fixed constant and $g : [a, b] \times \mathbb{R}^2 \to \mathbb{R}$ is a differentiable function with respect to the second and third arguments, we get an isoperimetric variational problem. In order to obtain a necessary condition for a minimizer, we define the new function

$$ F := \lambda_0 L(x, y(x), {}_a^C D_x^{\alpha} y(x)) - \lambda g(x, y(x), {}_a^C D_x^{\alpha} y(x)), \tag{3-8} $$

where $\lambda_0$ and $\lambda$ are Lagrange multipliers. Then every solution $\bar{y}$ of the fractional isoperimetric problem given by (3-4)–(3-5) and (3-7) is also a solution to the fractional differential equation [Almeida and Torres 2011]

$$ \frac{\partial F}{\partial y}(x, y(x), {}_a^C D_x^{\alpha} y(x)) + {}_x D_b^{\alpha} \frac{\partial F}{\partial\, {}_a^C D_x^{\alpha} y}(x, y(x), {}_a^C D_x^{\alpha} y(x)) = 0, \quad t \in [a, b]. \tag{3-9} $$

Moreover, if $\bar{y}$ is not a solution to

$$ \frac{\partial g}{\partial y}(x, y(x), {}_a^C D_x^{\alpha} y(x)) + {}_x D_b^{\alpha} \frac{\partial g}{\partial\, {}_a^C D_x^{\alpha} y}(x, y(x), {}_a^C D_x^{\alpha} y(x)) = 0, \quad t \in [a, b], \tag{3-10} $$

then we can put $\lambda_0 = 1$ in (3-8).

**Discretization Method 1.** Using the approximation formula for the Caputo fractional derivative given by (3-2), we can discretize functional (3-4) in the following way. Let $N \in \mathbb{N}$, $h = (b-a)/N$ and the grid $x_j = a + jh$, $j = 0, 1, \dots, N$. Then

$$\begin{aligned} I[y] &= \sum_{k=1}^{N} \int_{x_{k-1}}^{x_k} L(x, y(x), {}_a^C D_x^\alpha\, y(x))\, dx \\ &\approx \sum_{k=1}^{N} hL(x_k, y(x_k), {}_a^C D_{x_k}^\alpha\, y(x_k)) \\ &\approx \sum_{k=1}^{N} hL(x_k, y(x_k), {}_a^C \widetilde{D}_{x_k}^\alpha\, y(x_k)). \end{aligned} \tag{3-11}$$

This is the direct way to solve the problem, using discretization techniques.

**Discretization Method 2.** By the previous discussion, the initial problem of minimization of the functional (3-4), subject to boundary conditions (3-5), can be numerically replaced by the finite-dimensional optimization problem

$$\Phi(y_1, \dots, y_{N-1}) := \sum_{k=1}^{N} hL(x_k, y(x_k), {}_a^C \widetilde{D}_{x_k}^\alpha\, y(x_k)) \to \min,$$

subject to

$$y_0 = y_a, \quad y_N = y_b,$$

where $y_k := y(x_k)$.

Using the first-order necessary optimality conditions given by the system of $N-1$ equations

$$\frac{\partial \Phi}{\partial y_j} = 0 \quad \text{for all } j = 1, \dots, N-1,$$

we get

$$\frac{\partial L}{\partial y}(x_j, y(x_j), {}_a^C \widetilde{D}_{x_j}^\alpha\, y(x_j)) + \sum_{k=0}^{N-j} \frac{(w_k^\alpha)}{h^\alpha} \frac{\partial L}{\partial\, {}_a^C D_x^\alpha\, y}(x_{j+k}, y(x_{j+k}), {}_a^C \widetilde{D}_{x_{j+k}}^\alpha\, y(x_{j+k})) = 0 \tag{3-12}$$

with $j = 1, \dots, N-1$. As $N \to \infty$, that is, as $h \to 0$, the solutions of system (3-12) converge to the solutions of the fractional Euler–Lagrange equation associated with the variational problem [Pooseh et al. 2013, Theorem 4.1]. The constrained variational problem given by (3-4)–(3-5) and (3-7) can be solved similarly. More precisely, in this case, we have to replace the Lagrange function $L$ by the augmented function $F = \lambda_0 L - \lambda g$ and proceed with similar calculations.

### 3B. *Sturm–Liouville problem.*

Consider the fractional differential equation

$$\left[{}^C D_b^\alpha\, p(x)\, {}^C D_a^\alpha + q(x)\right] y(x) = \lambda r_\alpha(x) y(x), \tag{3-13}$$

subject to the boundary conditions

$$y(a) = y(b) = 0. \tag{3-14}$$

Equation (3-13) together with condition (3-14) is called the fractional Sturm–Liouville problem. As in the discrete case, it is required to find the eigenfunctions and the eigenvalues of the given boundary value problem, i.e., the nontrivial solutions of (3-13)–(3-14) and the corresponding values of the parameter $\lambda$.

In what follows, we assume:

**Assumption A.** Let $\frac{1}{2} < \alpha < 1$ and $p$, $q$ and $r_\alpha$ be given functions such that $p \in C^1[a,b]$ and $p(x) > 0$ for all $x \in [a,b]$, $q, r_\alpha \in C[a,b]$, $r_\alpha(x) > 0$ for all $x \in [a,b]$ and $(\sqrt{r_\alpha})'$ is Hölderian, of order $\beta \le \alpha - \frac{1}{2}$, on $[a,b]$.

**Theorem 3.6** [Klimek et al. 2014]. *Under Assumption A, the fractional Sturm–Liouville problem* (3-13)–(3-14) *has an infinite increasing sequence of eigenvalues* $\lambda_1, \lambda_2, \ldots,$ *and to each eigenvalue* $\lambda_k$, *there is a corresponding continuous eigenfunction* $y_k$ *that is unique up to a constant factor.*

The fractional Sturm–Liouville problem can be remodeled as a fractional isoperimetric variational problem.

**Theorem 3.7** [Klimek et al. 2014]. *Let Assumption A hold and* $y^1$ *be the eigenfunction, normalized to satisfy the isoperimetric constraint*

$$I[y] = \int_a^b r_\alpha(x) y^2(x)\, dx = 1, \tag{3-15}$$

*associated with the first eigenvalue* $\lambda_1$ *of problem* (3-13)–(3-14), *and assume that function* $D_b^\alpha(p\, {}^C D_a^\alpha\, y^1)$ *is continuous. Then* $y^1$ *is a minimizer of the variational functional*

$$J[y] = \int_a^b \left[ p(x)({}^C D_a^\alpha\, y(x))^2 + q(x) y^2(x) \right] dx, \tag{3-16}$$

*in the class of* $C[a,b]$ *functions with* ${}^C D_a^\alpha\, y$ *and* $D_b^\alpha(p\, {}^C D_a^\alpha\, y)$ *continuous in* $[a,b]$, *subject to the boundary conditions*

$$y(a) = y(b) = 0 \tag{3-17}$$

*and isoperimetric constraint* (3-15). *Moreover,*

$$J[y^1] = \lambda_1.$$

**Discretization Method 3.** Using the approximation formula for the Caputo fractional derivatives given by (3-2)–(3-3), we can discretize (3-13) in the following way. Let $N \in \mathbb{N}$, $h = (b-a)/N$ and the grid $x_j = a + jh$, $j = 0, 1, \ldots, N$. Then at $x = x_i$, (3-13) may be discretized as

$$\frac{h^{-2\alpha}}{r_\alpha(x_i)} \sum_{k=0}^{N-i} (w_k^\alpha) p(x_{i+k}) \sum_{l=0}^{i+k} (w_l^\alpha) y_{i+k-l} + \frac{q(x_i)}{r_\alpha(x_i)} y_i = \lambda y_i, \quad i = 1, \ldots, N-1,$$

which in matrix form may be written as

$$AY = \lambda Y, \tag{3-18}$$

| $N$ | 5 | 10 | 15 |
|---|---|---|---|
| $\lambda_1$ | 4.603751971 | 4.491185175 | 4.426964914 |

**Table 3.** Values of $\lambda_1$ for $N = 5$, 10, 15.

where $Y = [y_1, y_2, \dots, y_{N-1}]$, $y_i = y(x_i)$, and $A = (c_{ik})$, $i = 1, 2, \dots, N-1$, $k = 1, 2, \dots, N-1$, with

$$c_{ik} = \begin{cases} (h^{-2\alpha}/r_\alpha(x_i)) \sum_{j=0}^{N-i} (w_j^\alpha)^2 p(x_{j+i}) + q(x_i)/r_\alpha(x_i) & \text{if } i = k, \\ (h^{-2\alpha}/r_\alpha(x_i)) \sum_{j=0}^{N-i} (w_j^\alpha)(w_{j+i-k}^\alpha) p(x_{j+i}) & \text{if } i > k, \\ (h^{-2\alpha}/r_\alpha(x_i)) \sum_{j=k-i}^{N-i} (w_j^\alpha)(w_{j+i-k}^\alpha) p(x_{j+i}) & \text{if } i < k, \end{cases}$$

reducing in this way the Sturm–Liouville problem to an algebraic eigenvalue problem.

**Example 3.8.** Let us consider the following problem: minimize the functional

$$\int_0^1 ({}_0^C D_x^\alpha\, y(x))^2\, dx, \tag{3-19}$$

under the restrictions

$$\int_0^1 y^2(x)\, dx = 1, \qquad y(0) = y(1) = 0, \tag{3-20}$$

where $\alpha = \frac{3}{4}$. Since $y(0) = 0$, we have ${}_0^C D_x^\alpha\, y(x) = {}_0 D_x^\alpha\, y(x)$. Using Method 1, we obtain a finite-dimensional constrained optimization problem

$$\sum_{k=1}^{N} N^{2\alpha-1} \left( \sum_{i=0}^{k} (w_i^\alpha) y_{k-i} \right)^2 \to \min, \tag{3-21}$$

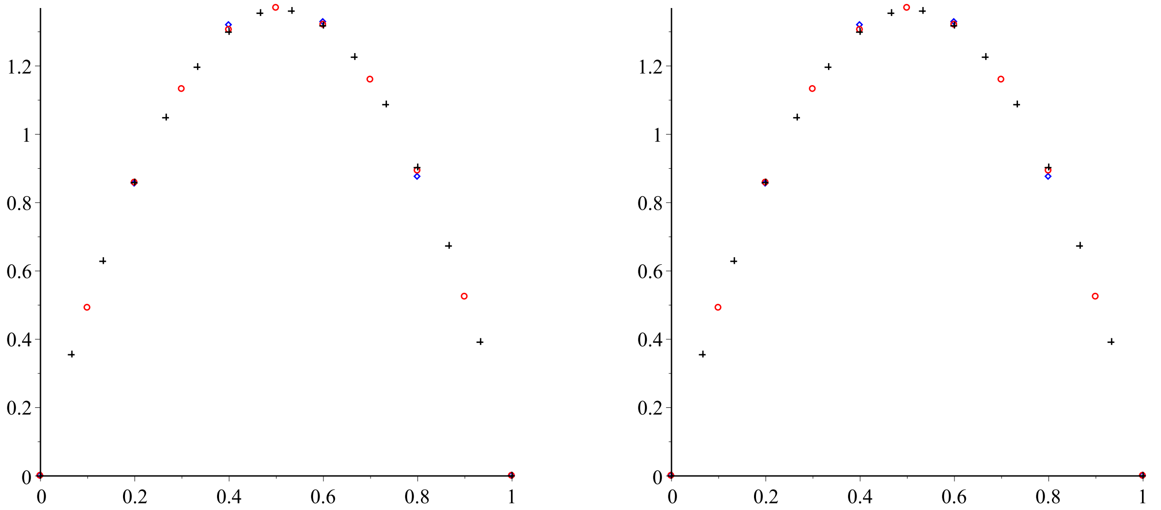

**Figure 2.** Approximations of solutions to problem (3-19)–(3-20) using Methods 1 (left) and 2 (right) with $N = 5$ ($\diamond$), 10 ($\circ$) and 15 ($+$).

| $N$ | 5 | 10 | 15 |
|---|---|---|---|
| $\lambda_1$ | 4.603751969 | 4.491185168 | 4.426964909 |
| $\lambda_2$ | 13.67144835 | 14.31569449 | 14.33350940 |
| $\lambda_3$ | 22.69092491 | 26.35335634 | 26.90113751 |
| $\lambda_4$ | 29.24531071 | 39.48118456 | 41.37391615 |
| $\lambda_5$ | | 52.54234156 | 56.93534748 |
| $\lambda_6$ | | 64.64953668 | 73.03700902 |
| $\lambda_7$ | | 74.96494602 | 89.07875858 |
| $\lambda_8$ | | 82.83813371 | 104.5749014 |
| $\lambda_9$ | | 87.76536891 | 119.0339408 |
| $\lambda_{10}$ | | | 132.0436041 |
| $\lambda_{11}$ | | | 143.2212682 |
| $\lambda_{12}$ | | | 152.2566950 |
| $\lambda_{13}$ | | | 158.8942685 |
| $\lambda_{14}$ | | | 162.9518168 |

**Table 4.** Values of $\lambda_i$ for $N = 5, 10, 15$.

subject to

$$\sum_{k=1}^{N} \frac{y_k^2}{N} = 1, \quad y_0 = y_N = 0. \tag{3-22}$$

Using the Maple package Optimization, we get approximations of the optimal solutions to (3-19)–(3-20) for different values of $N$. Table 3 shows values of $\lambda_1$ for $N = 5, 10, 15$. Note that $\lambda_1$ is the value of (3-21), where $\bar{y} = [0, y_1, \dots, y_{N-1}, 0]$ is the optimal solution to (3-21)–(3-22). In other words, $\lambda_1$ is an approximation of the minimum value of functional (3-19) and the first eigenvalue of the Sturm–Liouville problem (which is the Euler–Lagrange equation for the considered variational problem). Figure 2, left, presents minimizers $\bar{y}$ for $N = 5, 10, 15$.

Observe that the unique solution to the Euler–Lagrange equation (see (3-10)) associated with the integral constraint is $\bar{y}(x) = 0$. As $\bar{y}(x) = 0$ is not a solution to (3-19)–(3-20) (condition $\int_0^1 y^2(x)\,dx = 1$ fails), we can consider $\lambda_0 = 1$ in (3-8). Therefore, the auxiliary function is

$$F := ({}_0^C D_x^\alpha\, y(x))^2 - \lambda y^2(x).$$

Thus,

$$\Phi(y_1, \dots, y_{N-1}) := \sum_{k=1}^{N} h(({}_0^C D_{x_k}^\alpha\, y_k)^2 - \lambda y_k^2),$$

and the computation of $\partial\Phi/\partial y_j$ leads to

$$-\lambda y_j + N^{2\alpha} \sum_{k=0}^{N-j} (w_k^\alpha) \sum_{l=0}^{j+k} (w_l^\alpha) y_{j+k-l} = 0, \quad j = 1, \dots, N-1. \tag{3-23}$$

Solving system of equations (3-23) together with (3-22), we obtain not only an approximation of the optimal solution to problem (3-19)–(3-20) but also other solutions to the Euler–Lagrange equation (3-9)

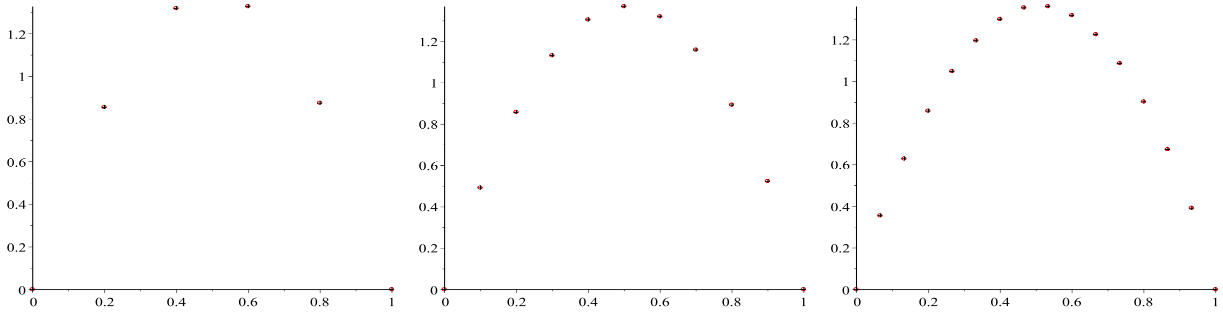

**Figure 3.** Approximations of solutions to problem (3-19)–(3-20) using Methods 1 (∘) and 2 (+) with $N = 5$ (left), 10 (center) and 15 (right).

as $N \to \infty$. In other words, we get some approximations of the eigenvalues and eigenfunctions of the Sturm–Liouville problem. Table 4 presents approximations of the eigenvalues obtained by this procedure for $N = 5$, 10, 15. In Figure 2, right, we present the eigenvectors, for $N = 5$, 10, 15, associated with the eigenvalues $\lambda_1$.

In Figure 3, we compare the approximation of the optimal solutions to (3-19)–(3-20), obtained by solving (3-21)–(3-22) (Method 1) and (3-23)–(3-22) (Method 2), for $N = 5$, 10, 15.

Now let us consider the Sturm–Liouville problem

$$ {}^{C}D_1^{3/4}\,{}^{C}D_0^{3/4}\,y(x) = \lambda y(x), \tag{3-24}$$

subject to the boundary conditions

$$y(0) = y(1) = 0. \tag{3-25}$$

Under the conditions of Theorem 3.7, (3-24) is the Euler–Lagrange equation for isoperimetric problem

| $N$ | 5 | 10 | 20 | 40 | 80 | 160 |
|---|---|---|---|---|---|---|
| $\lambda_1$ | 4.603751972 | 4.491185175 | 4.387575384 | 4.314056432 | 4.264767769 | 4.231946921 |
| $\lambda_2$ | 13.67144835 | 14.31569450 | 14.29943076 | 14.18275912 | 14.08194289 | 14.01015799 |
| $\lambda_3$ | 22.69092491 | 26.35335634 | 27.02784640 | 27.01132309 | 26.88877847 | 26.78184511 |
| $\lambda_4$ | 29.24531071 | 39.48118456 | 41.95747334 | 42.33045300 | 42.25841874 | 42.13429128 |
| $\lambda_5$ | | 52.54234157 | 58.40981791 | 59.60122278 | 59.68496380 | 59.56753673 |
| $\lambda_6$ | | 64.64953668 | 75.99486098 | 78.61012095 | 79.00578911 | 78.93437596 |
| $\lambda_7$ | | 74.96494602 | 94.25189512 | 99.05856280 | 99.96479334 | 99.98764503 |
| $\lambda_8$ | | 82.83813372 | 112.8375161 | 120.7904806 | 122.4632696 | 122.6454529 |
| $\lambda_9$ | | 87.76536891 | 131.3694072 | 143.5891552 | 146.3337902 | 146.7516690 |
| $\lambda_{10}$ | | | 149.5318910 | 167.3194776 | 171.5033012 | 172.2513810 |
| $\lambda_{11}$ | | | 166.9946039 | 191.8029693 | 197.8472756 | 199.0332700 |
| $\lambda_{12}$ | | | 183.4744810 | 216.9142113 | 225.3053712 | 227.0563318 |
| $\lambda_{13}$ | | | 198.6917619 | 242.4962397 | 253.7773704 | 256.2351380 |
| $\lambda_{14}$ | | | 212.4063351 | 268.4292940 | 283.2100111 | 286.5368179 |

**Table 5.** Approximation of the eigenvalues using Method 3.

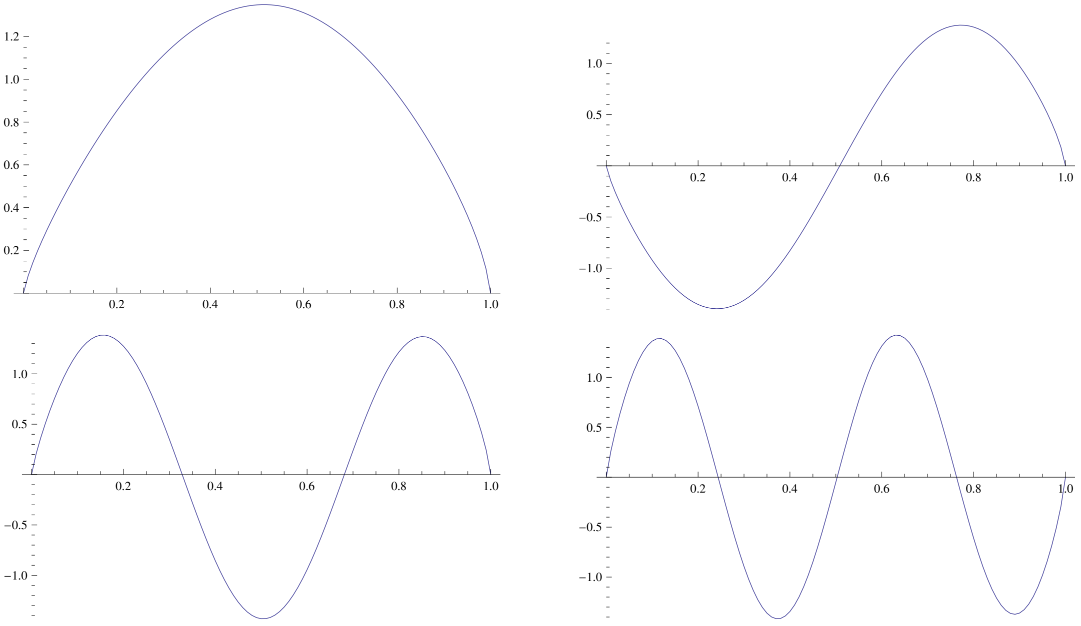

**Figure 4.** Normalized eigenfunctions obtained with $N = 100$, corresponding to the eigenvalues $\lambda_1$ (top left), $\lambda_2$ (top right), $\lambda_3$ (bottom left) and $\lambda_4$ (bottom right).

(3-19)–(3-20). Table 5 presents approximations of the eigenvalues of (3-24) obtained by Method 3 for $N = 5, 10, 20, 40, 80, 160$ (for $N = 20, 40, 80, 160$ only the first 14 eigenvalues are listed). Figure 4 shows normalized eigenfunctions, obtained for $N = 100$, corresponding to the eigenvalues $\lambda_1$, $\lambda_2$, $\lambda_3$, $\lambda_4$.

## 4. Conclusions

Since the seminal works [Nigmatullin 1986; Wyss 1986] were published, fractional differential equations have become a popular way to model anomalous diffusion. As stated in [Meerschaert 2012], this type of approach is the most reasonable: the fractional derivative in space represents large particle jumps (that lead to anomalous superdiffusion) while the time-fractional derivative models time delays between particle motion. Fractional diffusion equations have been used, e.g., to model pollution in ground water [Benson et al. 2001] and flow in porous media [He 1998]. Many other examples can be found in [Meerschaert 2012; Meerschaert and Sikorskii 2012].

It was proved in [Klimek et al. 2016] that, under appropriate assumptions, the space-time fractional diffusion equation

$$ {}^{C}D^{\beta}_{0+,t}\, u(t,x) = -\frac{1}{r_\alpha(x)}\big[{}^{C}D^{\alpha}_{b-,x}\, p(x)\, {}^{C}D^{\alpha}_{a+,x} + q(x)\big] u(t,x) \quad \text{for all } (t,x)\in(0,\infty)\times[a,b], \tag{4-1} $$

where $0 < \beta < 1$, $\frac{1}{2} < \alpha < 1$ and ${}^{C}D^{\beta}_{0+,t}$ and ${}^{C}D^{\alpha}_{b-,x}$, ${}^{C}D^{\alpha}_{a+,x}$ are partial fractional derivatives, with the boundary and initial conditions

$$ u(t,a) = u(t,b) = 0, \qquad t \in (0,\infty), \tag{4-2} $$

$$ u(0,x) = f(x), \qquad x \in [a,b], \tag{4-3} $$

has a continuous solution $u : [0, \infty) \times [a, b] \to \mathbb{R}$ given by the series

$$u(t, x) = \sum_{k=1}^{\infty} \langle y_k, f \rangle E_\beta(-\lambda_k t^\beta) y_k(x). \tag{4-4}$$

In (4-4), $\langle f, g \rangle := \int_a^b r_\alpha(x) f(x) g(x)\, dx$, $E_\beta$ is the one-parameter Mittag–Leffler function and $y_k$ and $\lambda_k$ ($k = 1, 2, \dots$) are the eigenfunctions and the eigenvalues of the fractional Sturm–Liouville problem (3-13)–(3-14). Thus, numerical methods, presented in this paper, for finding the eigenvalues and eigenfunctions of fractional Sturm–Liouville problems can also be used to approximate solutions to fractional diffusion problems of the form (4-1)–(4-3). We have presented a link between fractional Sturm–Liouville and fractional isoperimetric variational problems that provides a possible method for solution of the former. Discrete problems with the Grünwald–Letnikov difference were analyzed: we proved the existence of orthogonal solutions to the discrete fractional Sturm–Liouville eigenvalue problem and showed that its eigenvalues can be characterized as values of certain functionals. For continuous problems with the Caputo fractional derivatives, in order to examine the performance of the proposed method, the approximation based on the shifted Grünwald–Letnikov definition was used. This type of discretization is most popular in practical applications, when numerically solving fractional diffusion equations, due to the fact that such methods are mass-preserving [Defterli et al. 2015].

## Acknowledgements

Research was supported by Portuguese funds through the Center for Research and Development in Mathematics and Applications (CIDMA) and the Portuguese Foundation for Science and Technology (FCT) within UID/MAT/04106/2013 (Almeida) and UID/MAT/00013/2013 (Morgado), by the Bialystok University of Technology grant S/WI/1/2016 (Malinowska) and by the Warsaw School of Economics grant KAE/BMN16/18/16 (Odzijewicz).

## References

[Abdeljawad 2011] T. Abdeljawad, "On Riemann and Caputo fractional differences", *Comput. Math. Appl.* **62**:3 (2011), 1602–1611.

[Agrawal 2006] O. P. Agrawal, "Fractional variational calculus and the transversality conditions", *J. Phys. A* **39**:33 (2006), 10375–10384.

[Al-Mdallal 2009] Q. M. Al-Mdallal, "An efficient method for solving fractional Sturm–Liouville problems", *Chaos Solitons Fractals* **40**:1 (2009), 183–189.

[Al-Mdallal 2010] Q. M. Al-Mdallal, "On the numerical solution of fractional Sturm–Liouville problems", *Int. J. Comput. Math.* **87**:12 (2010), 2837–2845.

[Almeida and Torres 2011] R. Almeida and D. F. M. Torres, "Necessary and sufficient conditions for the fractional calculus of variations with Caputo derivatives", *Commun. Nonlinear Sci. Numer. Simul.* **16**:3 (2011), 1490–1500.

[Almeida et al. 2015] R. Almeida, S. Pooseh, and D. F. M. Torres, *Computational methods in the fractional calculus of variations*, Imperial College, London, 2015.

[Atıcı and Eloe 2009a] F. M. Atıcı and P. W. Eloe, "Discrete fractional calculus with the nabla operator", *Electron. J. Qual. Theory Differ. Equ.* **Spec. Ed. I** (2009), 3.

[Atıcı and Eloe 2009b] F. M. Atıcı and P. W. Eloe, "Initial value problems in discrete fractional calculus", *Proc. Amer. Math. Soc.* **137**:3 (2009), 981–989.

[Benson et al. 2001] D. A. Benson, R. Schumer, M. M. Meerschaert, and S. W. Wheatcraft, “Fractional dispersion, Lévy motion, and the MADE tracer tests”, *Transp. Porous Media* **42**:1–2 (2001), 211–240.

[Blaszczyk and Ciesielski 2014] T. Blaszczyk and M. Ciesielski, “Numerical solution of fractional Sturm–Liouville equation in integral form”, *Fract. Calc. Appl. Anal.* **17**:2 (2014), 307–320.

[Bourdin et al. 2013] L. Bourdin, J. Cresson, I. Greff, and P. Inizan, “Variational integrator for fractional Euler–Lagrange equations”, *Appl. Numer. Math.* **71** (2013), 14–23.

[Carpinteri and Mainardi 1997] A. Carpinteri and F. Mainardi (editors), *Fractals and fractional calculus in continuum mechanics* (Udine, Italy, 1996), CISM Courses and Lectures **378**, Springer, Vienna, 1997.

[Chen et al. 2012] Z.-Q. Chen, M. M. Meerschaert, and E. Nane, “Space-time fractional diffusion on bounded domains”, *J. Math. Anal. Appl.* **393**:2 (2012), 479–488.

[Defterli et al. 2015] O. Defterli, M. D’Elia, Q. Du, M. Gunzburger, R. Lehoucq, and M. M. Meerschaert, “Fractional diffusion on bounded domains”, *Fract. Calc. Appl. Anal.* **18**:2 (2015), 342–360.

[Díaz and Osler 1974] J. B. Díaz and T. J. Osler, “Differences of fractional order”, *Math. Comp.* **28**:125 (1974), 185–202.

[Domek and Pworak 2016] S. Domek and P. Pworak (editors), *Theoretical developments and applications of non-integer order systems: 7th Conference on Non-Integer Order Calculus and Its Applications* (Szczecin, Poland, 2015), Lecture Notes in Electrical Engineering **357**, Springer, Cham, Switzerland, 2016.

[D’Ovidio 2012] M. D’Ovidio, “From Sturm–Liouville problems to fractional and anomalous diffusions”, *Stochastic Process. Appl.* **122**:10 (2012), 3513–3544.

[He 1998] J.-H. He, “Approximate analytical solution for seepage flow with fractional derivatives in porous media”, *Comput. Methods Appl. Mech. Engrg.* **167**:1–2 (1998), 57–68.

[Hilfer 2000] R. Hilfer (editor), *Applications of fractional calculus in physics*, World Scientific, River Edge, NJ, 2000.

[Kaczorek 2011] T. Kaczorek, *Selected problems of fractional systems theory*, Lecture Notes in Control and Information Sciences **411**, Springer, Berlin, 2011.

[Kilbas et al. 2006] A. A. Kilbas, H. M. Srivastava, and J. J. Trujillo, *Theory and applications of fractional differential equations*, North-Holland Mathematics Studies **204**, Elsevier, Amsterdam, 2006.

[Klimek 2009] M. Klimek, *On solutions of linear fractional differential equations of a variational type*, Monographs **172**, Czestochowa University of Technology, 2009.

[Klimek 2015] M. Klimek, “Fractional Sturm–Liouville problem and 1D space-time fractional diffusion with mixed boundary conditions”, pp. 46808 in *Proceedings of the ASME 2015 International Design Engineering Technical Conferences and Computers and Information in Engineering Conference* (Boston, 2015), vol. 9: 2015 ASME/IEEE International Conference on Mechatronic and Embedded Systems and Applications, American Society of Mechanical Engineers, New York, 2015.

[Klimek 2016] M. Klimek, “Fractional Sturm–Liouville problem in terms of Riesz derivatives”, pp. 3–16 in *Theoretical developments and applications of non-integer order systems: 7th Conference on Non-Integer Order Calculus and Its Applications* (Szczecin, Poland, 2015), edited by S. Domek and P. Dworak, Lecture Notes in Electrical Engineering **357**, Springer, Cham, Switzerland, 2016.

[Klimek et al. 2014] M. Klimek, T. Odzijewicz, and A. B. Malinowska, “Variational methods for the fractional Sturm–Liouville problem”, *J. Math. Anal. Appl.* **416**:1 (2014), 402–426.

[Klimek et al. 2016] M. Klimek, A. B. Malinowska, and T. Odzijewicz, “Applications of fractional Sturm–Liouville problem to the space-time fractional diffusion in a finite domain”, *Fract. Calc. Appl. Anal.* **19**:2 (2016), 516–550.

[Leonenko et al. 2013] N. N. Leonenko, M. M. Meerschaert, and A. Sikorskii, “Fractional Pearson diffusions”, *J. Math. Anal. Appl.* **403**:2 (2013), 532–546.

[Li and Ostoja-Starzewski 2011] J. Li and M. Ostoja-Starzewski, “Micropolar continuum mechanics of fractal media”, *Internat. J. Engrg. Sci.* **49**:12 (2011), 1302–1310.

[Mainardi 2010] F. Mainardi, *Fractional calculus and waves in linear viscoelasticity: an introduction to mathematical models*, Imperial College, London, 2010.

[Malinowska and Odzijewicz 2016] A. B. Malinowska and T. Odzijewicz, “Multidimensional discrete-time fractional calculus of variations”, pp. 17–28 in *Theoretical developments and applications of non-integer order systems: 7th Conference on Non-Integer Order Calculus and Its Applications* (Szczecin, Poland, 2015), edited by S. Domek and P. Dworak, Lecture Notes in Electrical Engineering **357**, Springer, Cham, Switzerland, 2016.

[Malinowska and Torres 2012] A. B. Malinowska and D. F. M. Torres, *Introduction to the fractional calculus of variations*, Imperial College, London, 2012.

[Malinowska et al. 2015] A. B. Malinowska, T. Odzijewicz, and D. F. M. Torres, *Advanced methods in the fractional calculus of variations*, Springer, Cham, Switzerland, 2015.

[Meerschaert 2012] M. M. Meerschaert, "Fractional calculus, anomalous diffusion, and probability", pp. 265–284 in *Fractional dynamics*, edited by J. Klafter et al., World Scientific, Hackensack, NJ, 2012.

[Meerschaert and Sikorskii 2012] M. M. Meerschaert and A. Sikorskii, *Stochastic models for fractional calculus*, Studies in Mathematics **43**, Walter de Gruyter, Berlin, 2012.

[Metzler and Klafter 2000] R. Metzler and J. Klafter, "The random walk's guide to anomalous diffusion: a fractional dynamics approach", *Phys. Rep.* **339**:1 (2000), 1–77.

[Miller and Ross 1989] K. S. Miller and B. Ross, "Fractional difference calculus", pp. 139–152 in *Univalent functions, fractional calculus, and their applications* (Kōriyama, Japan, 1988), edited by H. M. Srivastava and S. Owa, Horwood, Chichester, United Kingdom, 1989.

[Nigmatullin 1986] R. R. Nigmatullin, "The realization of the generalized transfer equation in a medium with fractal geometry", *Phys. Stat. Sol. B* **133**:1 (1986), 425–430.

[Ostalczyk 2008] P. Ostalczyk, *Zarys rachunku różniczkowo-całkowego ułamkowych rzedów: teoria i zastosowanie w automatyce*, Politechniki Łódzkiej, Łódż, Poland, 2008.

[Podlubny 1999] I. Podlubny, *Fractional differential equations: an introduction to fractional derivatives, fractional differential equations, to methods of their solution and some of their applications*, Mathematics in Science and Engineering **198**, Academic, San Diego, 1999.

[Pooseh et al. 2013] S. Pooseh, R. Almeida, and D. F. M. Torres, "Discrete direct methods in the fractional calculus of variations", *Comput. Math. Appl.* **66**:5 (2013), 668–676.

[Riewe 1996] F. Riewe, "Nonconservative Lagrangian and Hamiltonian mechanics", *Phys. Rev. E* (3) **53**:2 (1996), 1890–1899.

[Riewe 1997] F. Riewe, "Mechanics with fractional derivatives", *Phys. Rev. E* (3) **55**:3 (1997), 3581–3592.

[Tarasov 2010] V. E. Tarasov, *Fractional dynamics: applications of fractional calculus to dynamics of particles, fields and media*, Springer, Heidelberg, Germany, 2010.

[Wyss 1986] W. Wyss, "The fractional diffusion equation", *J. Math. Phys.* **27**:11 (1986), 2782–2785.

[Zaslavsky 2005] G. M. Zaslavsky, *Hamiltonian chaos and fractional dynamics*, Oxford University, 2005.

[Zaslavsky and Edelman 2004] G. M. Zaslavsky and M. A. Edelman, "Fractional kinetics: from pseudochaotic dynamics to Maxwell's Demon", *Phys. D* **193**:1–4 (2004), 128–147.

[Zayernouri and Karniadakis 2013] M. Zayernouri and G. E. Karniadakis, "Fractional Sturm–Liouville eigen-problems: theory and numerical approximation", *J. Comput. Phys.* **252** (2013), 495–517.

Received 30 Jan 2016. Revised 29 Mar 2016. Accepted 4 Apr 2016.

RICARDO ALMEIDA: `ricardo.almeida@ua.pt`
*Center for Research and Development in Mathematics and Applications, Department of Mathematics, University of Aveiro, 3810-193 Aveiro, Portugal*

AGNIESZKA B. MALINOWSKA: `a.malinowska@pb.edu.pl`
*Faculty of Computer Science, Bialystok University of Technology, 15-351 Bialystok, Poland*

M. LUÍSA MORGADO: `luisam@utad.pt`
*Centro de Matemática, pólo CMAT-UTAD, Department of Mathematics, University of Trás-os-Montes e Alto Douro, 5000-801 Vila Real, Portugal*

TATIANA ODZIJEWICZ: `tatiana.odzijewicz@sgh.waw.pl`
*Department of Mathematics and Mathematical Economics, Warsaw School of Economics, 02-554 Warsaw, Poland*

JOURNAL OF MECHANICS OF MATERIALS AND STRUCTURES
Vol. 12, No. 1, 2017
dx.doi.org/10.2140/jomms.2017.12.23

# ANALYTICAL AND NUMERICAL SOLUTION OF THE FRACTIONAL EULER–BERNOULLI BEAM EQUATION

TOMASZ BLASZCZYK

In this paper a new formulation of the Euler–Bernoulli beam equation is proposed, which is based on fractional calculus. The fractional Euler–Bernoulli beam equation is derived by using a variational approach. Such formulation leads to an equation containing left and right fractional Caputo derivatives simultaneously. The obtained equation is transformed into an integral equation and then is solved analytically and numerically. Finally, examples of computations and error analysis are shown.

## 1. Introduction

Fractional calculus has recently played a very important role in various fields of science [Baleanu et al. 2015; Błasik and Klimek 2015; Leszczyński 2011; Klimek 2001; Kukla and Siedlecka 2015; Podlubny 1999; Torres 2015; Zhang et al. 2016; Zingales and Failla 2015]. It is caused largely by the fact that the fractional derivatives are nonlocal operators and depend on the past values of a function (the left derivative) or the future values of a function (the right derivative) [Baleanu et al. 2012; Kilbas et al. 2006; Podlubny 1999; Samko et al. 1993]. On the other hand, nonlocal formulations play an essential role in the description of the material deformation. Considering the description of the material deformation, including length scale, we are able to describe the phenomena (such as scale effects) where the classical approach is no longer valid [Sumelka 2014a; 2014b; Sumelka and Błaszczyk 2014].

The first approaches to link fractional calculus and nonlocal continuum mechanics come from Vazquez [2004], Lazopoulos [2006] and Di Paola and Zingales [2008]. Later, Carpinteri et al. [2011; 2014] used spatial fractional calculus to examine a material whose nonlocal stress is defined as the fractional integral of the strain field, highlighting its connection with Eringen [2002] nonlocal elasticity. A similar problem was analyzed by Atanackovic & Stankovic [2009] starting from fractional nonlocal strain measure. Other fractional approaches to nonlocal elasticity are presented in [Alotta et al. 2015; Drapaca and Sivaloganathan 2012; Tarasov 2006; Zingales and Failla 2015].

In addition, recent research [Paola et al. 2013; Pirrotta et al. 2015] on the response evaluation of a viscoelastic Euler–Bernoulli or Timoshenko beam under quasistatic and dynamic loads have shown that for a better understanding of the viscoelastic behavior, a fractional constitutive law should be considered. Vibration analysis of a simply supported beam with a fractional order viscoelastic material model is presented in [Freundlich 2013].

As a concluding remark, it should be emphasised that there are many concepts dealing with nonlocal formulations. However, they often require a large number of parameters. As a response to this inconvenience, new models based on fractional calculus have been proposed.

---

*Keywords:* Euler–Bernoulli beam equation, Fractional Euler–Lagrange equation, Analytical and numerical solution, Caputo derivatives.

A good example where fractional models have been successfully used is in a group of viscoelastic materials. These materials are very attractive for civil engineering applications, and thus need special attention. One of the advantages of the fractional models is that they require a smaller number of parameters than classical models containing operators of integer order.

Due to these facts, fractional calculus looks to be a promising tool for modeling scale-dependent material behavior.

In our previous work [Sumelka et al. 2015] the classical Euler–Bernoulli beam theory was reformulated utilizing fractional Riesz–Caputo derivatives. In this paper we propose a new formulation of the Euler–Bernoulli beam equation based on fractional variational calculus. The fractional derivative included in a functional (see Section 3) is applied as a spatial derivative, and influences spatial response. In this sense the fractional order of derivation $\alpha$ is a new material parameter. The parameter $\alpha$ controls the way in which information is governed from the region of influence [Sumelka 2014a; Sumelka and Błaszczyk 2014].

We transform the obtained Euler–Lagrange equation into an integral equation and we find the exact solution. Next, we present two numerical schemes. The first one is based on the discretization of the analytical solution and the second one is based on the discretization of the Euler–Lagrange equation. We calculated the errors generated by both schemes and estimated the rates of convergences of the presented methods for a particular case. Finally, we show a few computational examples of static deflections for various types of loads.

## 2. Fractional operators

In this section, we only recall necessary definitions of fractional operators and their properties [Kilbas et al. 2006; Podlubny 1999; Samko et al. 1993]. The left and right Caputo derivatives of order $1 < \alpha \le 2$ are respectively defined as

$$
{}^{C}D^{\alpha}_{0+}u(x) := \begin{cases} I^{2-\alpha}_{0+}D^2u(x) & \text{for } 1 < \alpha < 2, \\ D^2u(x) & \text{for } \alpha = 2, \end{cases} \tag{2-1}
$$

$$
{}^{C}D^{\alpha}_{L-}u(x) := \begin{cases} I^{2-\alpha}_{L-}D^2u(x) & \text{for } 1 < \alpha < 2, \\ D^2u(x) & \text{for } \alpha = 2, \end{cases} \tag{2-2}
$$

where $D^2$ is the operator of the second order derivative and operators $I^{\alpha}_{0+}$ and $I^{\alpha}_{b-}$ are respectively the left and right fractional Riemann–Liouville integrals of order $\alpha > 0$ defined by

$$
I^{\alpha}_{0+}u(x) := \frac{1}{\Gamma(\alpha)} \int_0^x \frac{u(\tau)}{(x-\tau)^{1-\alpha}}\, \mathrm{d}\tau \quad (x > 0), \tag{2-3}
$$

$$
I^{\alpha}_{L-}u(x) := \frac{1}{\Gamma(\alpha)} \int_x^L \frac{u(\tau)}{(\tau-x)^{1-\alpha}}\, \mathrm{d}\tau \quad (x < L), \tag{2-4}
$$

where $\Gamma$ is the Euler Gamma function.

The composition rules of the fractional operators (for $\alpha \in (1, 2]$) are

$$I^{\alpha}_{0+}\,{}^{C}D^{\alpha}_{0+}\,u(x) = u(x) - xu'(a) - u(a), \tag{2-5}$$

$$I^{\alpha}_{L^-}\,{}^{C}D^{\alpha}_{L^-}\,u(x) = u(x) + (L-x)u'(L) - u(L). \tag{2-6}$$

The left and right fractional integrals of the constant $C$ have the following form:

$$I^{\alpha}_{0+}\,C = Cx^{\alpha}/(\Gamma(1+\alpha)), \tag{2-7}$$

$$I^{\alpha}_{L^-}C = C(L-x)^{\alpha}/(\Gamma(\alpha+1)), \tag{2-8}$$

and for power functions we have

$$I^{\alpha}_{0+}x^{\beta} = \frac{\Gamma(\beta+1)}{\Gamma(\alpha+\beta+1)}x^{\alpha+\beta}, \tag{2-9}$$

$$I^{\alpha}_{0+}x^{\beta}(L-x)^{\gamma} = \frac{\Gamma(\beta+1)x^{\alpha+\beta}L^{\gamma}}{\Gamma(\alpha+\beta+1)}\,{}_2F_1(\beta+1, -\gamma; \alpha+\beta+1; x/L), \tag{2-10}$$

where ${}_2F_1$ is a hypergeometric function [Kilbas et al. 2006; Samko et al. 1993].

## 3. Mathematical model

In classical mechanics, the minimization of the potential energy for a fixed supported beam of length $L$ with a downward transverse load per unit length $f(x)$ requires that the functional

$$V = \int_0^L F(x, u, u'')\,\mathrm{d}x = \int_0^L \left[\frac{1}{2}EI(u''(x))^2 - f(x)\,u(x)\right]\mathrm{d}x \tag{3-1}$$

be minimized, where $EI$ is the bending stiffness (which is constant) and $u(x)$ is the static deflection of the beam.

The individual terms $1/2EI(u''(x))^2$ and $f(x)$, $u(x)$ represent potential (strain) energy due to bending and potential energy due to the lateral deflection, respectively. The boundary conditions for the fixed supported beam are

$$u(0) = u'(0) = u(L) = u'(L) = 0. \tag{3-2}$$

The corresponding Euler–Lagrange equation for the considered problem has the form

$$\frac{\partial F(x, u, u'')}{\partial u} + \frac{\mathrm{d}^2}{\mathrm{d}x^2}\left(\frac{\partial F(x, u, u'')}{\partial u''}\right) = 0, \tag{3-3}$$

which leads to the Euler–Bernoulli beam equation

$$EI\frac{\mathrm{d}^4}{\mathrm{d}x^4}u(x) - f(x) = 0. \tag{3-4}$$

In this manuscript, we propose a procedure for constructing the fractional Euler–Bernoulli beam equation utilizing the fractional variational calculus. We start from the functional (3-1) and we replace the second order derivative by the left Caputo derivative (2-1) in the following way:

$$\begin{aligned} V_{\text{frac}} &= \int_0^L F_{\text{frac}}(x, u, {}^C D^{\alpha}_{0^+} u)\, \mathrm{d}x \\ &= \int_0^L \left[ \frac{1}{2} EI (\ell^{\alpha-2} {}^C D^{\alpha}_{0^+} u(x))^2 - f(x)\, u(x) \right] \mathrm{d}x. \end{aligned} \tag{3-5}$$

Next, by using results presented in [Lazo and Torres 2013; Malinowska et al. 2015] we get the corresponding form of the fractional Euler–Lagrange equation for the problem (3-5), which has the form

$$\frac{\partial F_{\text{frac}}(x, u, {}^C D^{\alpha}_{0^+} u)}{\partial u} + {}^C D^{\alpha}_{L^-} \left( \frac{\partial F_{\text{frac}}(x, u, {}^C D^{\alpha}_{0^+} u)}{\partial {}^C D^{\alpha}_{0^+} u} \right) = 0, \tag{3-6}$$

and leads to the fractional Euler–Bernoulli beam equation

$$\ell^{2(\alpha-2)} EI {}^C D^{\alpha}_{L^-} {}^C D^{\alpha}_{0^+} u(x) - f(x) = 0, \tag{3-7}$$

where $\ell$ is a length scale [Sumelka 2014a; Sumelka and Błaszczyk 2014; Sumelka et al. 2015].

It should be highlighted that if we put $\alpha = 2$ into (3-6) and (3-7) we obtain equations (3-3) and (3-4), respectively. Therefore, we do not need to look at the proposed model as competitive to the classical Euler–Bernoulli model. Because of this, we should treat this model as a complement to the classical theory.

## 4. Analytical solution

Let us start with the denotation

$$f^*(x) = \frac{f(x)}{\ell^{2(\alpha-2)} EI}. \tag{4-1}$$

By using the denotation above we can rewrite the fractional Euler–Bernoulli Equation (3-7) as

$${}^C D^{\alpha}_{L^-} {}^C D^{\alpha}_{0^+} u(x) = f^*(x). \tag{4-2}$$

We start with the transformation of (4-2) into an integral equation [Błaszczyk and Ciesielski 2015; 2016; Ciesielski and Błaszczyk 2015]. We integrate (4-2) two times by using fractional integral operators (2-3) and (2-4), giving

$$I^{\alpha}_{0^+} I^{\alpha}_{L^-} {}^C D^{\alpha}_{L^-} {}^C D^{\alpha}_{0^+} u(x) = I^{\alpha}_{0^+} I^{\alpha}_{L^-} f^*(x). \tag{4-3}$$

After the first integration in regards to the property (2-6), we get

$$I^{\alpha}_{0^+} \left( {}^C D^{\alpha}_{0^+} u(x) - {}^C D^{\alpha}_{0^+} u(x)\big|_{x=L} + (L-x) {}^C D^{\alpha+1}_{0^+} u(x)\big|_{x=L} \right) = I^{\alpha}_{0^+} I^{\alpha}_{L^-} f^*(x). \tag{4-4}$$

Next, we have to calculate the expression

$$I^{\alpha}_{0^+}(L-x) = I^{\alpha}_{0^+} L - I^{\alpha}_{0^+} x = L \frac{x^{\alpha}}{\Gamma(1+\alpha)} - \frac{x^{\alpha+1}}{\Gamma(2+\alpha)}, \tag{4-5}$$

and on the basis of the composition rule (2-5) we obtain the following form of the considered equation:

$$\begin{aligned} &u(x) - u(0) - u'(0)x \\ &\qquad - \left( \frac{{}^{C}D_{0+}^{\alpha} u(x)|_{x=L}}{\Gamma(\alpha+1)} x^{\alpha} + \frac{{}^{C}D_{0+}^{\alpha+1} u(x)|_{x=L}}{\Gamma(\alpha+2)} \big( -(\alpha+1)Lx^{\alpha} + x^{\alpha+1} \big) \right) = I_{0+}^{\alpha} I_{L-}^{\alpha} f^{*}(x). \end{aligned} \tag{4-6}$$

Now, we can write the above equation in a simpler form:

$$u(x) - (C_1 x^{\alpha} + C_2 x^{\alpha}(x - (\alpha+1)L)) = I_{0+}^{\alpha} I_{L-}^{\alpha} f^{*}(x), \tag{4-7}$$

where $C_1 = {}^{C}D_{0+}^{\alpha} u(x)|_{x=L} / \Gamma(\alpha+1)$ and $C_2 = {}^{C}D_{0+}^{\alpha+1} u(x)|_{x=L} / \Gamma(\alpha+2)$.

In order to determine the constants $C_1$ and $C_2$, we differentiate (4-7):

$$u'(x) - C_1 \alpha x^{\alpha-1} + C_2(\alpha+1)x^{\alpha-1}(\alpha L - x) = I_{0+}^{\alpha-1} I_{L-}^{\alpha} f^{*}(x), \tag{4-8}$$

and in accordance with the boundary conditions (3-2) we write the adequate system of equations as

$$\begin{cases} C_1 L^{\alpha} - C_2 \alpha L^{\alpha+1} = u(L) - I_{0+}^{\alpha} I_{L-}^{\alpha} f^{*}(x)\big|_{x=L}, \\ C_1 \alpha L^{\alpha-1} - C_2(1-\alpha^2)L^{\alpha} = u'(L) - I_{0+}^{\alpha-1} I_{L-}^{\alpha} f^{*}(x)\big|_{x=L}. \end{cases} \tag{4-9}$$

We determine the values $C_1$ and $C_2$ as

$$C_1 = \frac{\big(u(L) - I_{0+}^{\alpha} I_{L-}^{\alpha} f^{*}(x)\big|_{x=L}\big)(1-\alpha^2) - \big(u'(L) - I_{0+}^{\alpha-1} I_{L-}^{\alpha} f^{*}(x)\big|_{x=L}\big)\alpha L}{L^{\alpha}}, \tag{4-10}$$

$$C_2 = \frac{\big(u'(L) - I_{0+}^{\alpha-1} I_{L-}^{\alpha} f^{*}(x)\big|_{x=L}\big) - \big(u(L) - I_{0+}^{\alpha} I_{L-}^{\alpha} f^{*}(x)\big|_{x=L}\big)\alpha / L}{L^{\alpha}}. \tag{4-11}$$

Substituting the right-hand side of the expressions (4-10) and (4-11) into (4-7) we get

$$\begin{aligned} &u(x) - (x/L)^{\alpha}\big[(\alpha x/L - \alpha - 1) I_{0+}^{\alpha} I_{L-}^{\alpha} f^{*}(x)\big|_{x=L} + (L-x) I_{0+}^{\alpha-1} I_{L-}^{\alpha} f^{*}(x)\big|_{x=L}\big] \\ &\qquad\qquad = -(x/L)^{\alpha}\big[(\alpha + 1 - \alpha x/L)u(L) + (L-x)u'(L)\big] + I_{0+}^{\alpha} I_{L-}^{\alpha} f^{*}(x). \end{aligned} \tag{4-12}$$

Taking into account the boundary conditions (3-2) we obtain the final integral form of the fractional Euler–Bernoulli Equation (4-2):

$$u(x) - (x/L)^{\alpha}\big[(\alpha x/L - \alpha - 1) I_{0+}^{\alpha} I_{L-}^{\alpha} f^{*}(x)\big|_{x=L} + (L-x) I_{0+}^{\alpha-1} I_{L-}^{\alpha} f^{*}(x)\big|_{x=L}\big] = I_{0+}^{\alpha} I_{L-}^{\alpha} f^{*}(x). \tag{4-13}$$

The analytical solution of the considered problem (4-2) and (3-2) is

$$u(x) = (x/L)^{\alpha}\big[(\alpha x/L - \alpha - 1) g_{\alpha,\alpha}(L) + (L-x) g_{\alpha-1,\alpha}(L)\big] + g_{\alpha,\alpha}(x), \tag{4-14}$$

where $g_{\eta,\mu}(x) = I_{0+}^{\eta} I_{L-}^{\mu} f^{*}(x)$.

## 5. Numerical solution

In this section we present two numerical schemes for (3-7). The first one is dedicated to the integral equation and the second one is dedicated to the differential equation (the Euler–Lagrange equation).

The ability to calculate the composition of left and right fractional operators is very important to get a graphical interpretation of solutions for fractional variational differential equations. The analytical evaluations for any function of such composition are difficult to achieve. In some cases, we can express them through special functions, but this causes difficulties in calculating the function values. Therefore, numerical methods are a useful tool to obtain an approximation of integral or differential operators with different types of kernel [Błaszczyk and Ciesielski 2015; 2016; Ciesielski and Błaszczyk 2015; Durajski 2014; 2015; Siedlecki et al. 2015].

Let us start with introducing the following grid of $n+1$ nodes with the constant step $\Delta x = L/n$: $0 = x_0 < x_1 < \ldots < x_i < x_{i+1} < \ldots < x_n = L$, and $x_i = i\,\Delta x$, $i = 0, 1, \ldots, n$.

**5.1.** ***Method I: discretization of the integral equation.*** In our previous works [Błaszczyk and Ciesielski 2015; 2016] we determined the discrete form of the composition of the left and right fractional integrals of order $\alpha$. On the basis of these results, we present the composition of fractional integrals (2-3) and (2-4):

$$I^{\alpha}_{0^+}\, I^{\alpha}_{L^-} f^*(x)\big|_{x=x_i} \approx \sum_{j=0}^{i} w^{(\alpha)}_{i,j} \sum_{k=j}^{n} v^{(\alpha)}_{j,k}\, f^*_k, \tag{5-1}$$

where coefficients $w^{(\alpha)}_{i,j}$ and $v^{(\alpha)}_{i,j}$ have the form

$$w^{(\alpha)}_{i,j} = \frac{(\Delta x)^{\alpha}}{\Gamma(\alpha+2)} \begin{cases} 0 & \text{for } i=0 \text{ and } j=0 \\ (i-1)^{\alpha+1} - i^{\alpha+1} + i^{\alpha}(\alpha+1) & \text{for } i>0 \text{ and } j=0, \\ (i-j+1)^{\alpha+1} - 2(i-j)^{\alpha+1} + (i-j-1)^{\alpha+1} & \text{for } i>0 \text{ and } 0<j<i, \\ 1 & \text{for } i>0 \text{ and } j=i \end{cases} \tag{5-2}$$

$$v^{(\alpha)}_{i,j} = \frac{(\Delta x)^{\alpha}}{\Gamma(\alpha+2)} \begin{cases} 0 & \text{for } i=n \text{ and } j=n, \\ (n-i-1)^{\alpha+1} - (n-i)^{\alpha+1} + (n-i)^{\alpha}(\alpha+1) & \text{for } i<n \text{ and } j=n, \\ (j-i+1)^{\alpha+1} - 2(j-i)^{\alpha+1} + (j-i-1)^{\alpha+1} & \text{for } i<n \text{ and } i<j<n, \\ 1 & \text{for } i<n \text{ and } j=i. \end{cases} \tag{5-3}$$

Now we present the discrete form of the integral equation (4-13). For every grid node $x_i$, $i = 0, 1, \ldots, n$, we write the equation as

$$u_i = -\left(\frac{i}{n}\right)^{\alpha} \left[ \left(\frac{\alpha i}{n} - \alpha - 1\right) \sum_{j=0}^{n} w^{(\alpha)}_{n,j} \sum_{k=j}^{n} v^{(\alpha)}_{j,k} f^*_k + (n-i)\Delta x \sum_{j=0}^{n} w^{(\alpha-1)}_{n,j} \sum_{k=j}^{n} v^{(\alpha)}_{j,k} f^*_k \right] + \sum_{j=0}^{i} w^{(\alpha)}_{i,j} \sum_{k=j}^{n} v^{(\alpha)}_{j,k} f^*_k. \tag{5-4}$$

One can observe that for the calculation of values $u_0, u_1, \ldots, u_n$ we do not need to solve the system of $n+1$ linear equations. This is an important advantage of the proposed scheme from the computational point of view.

**5.2.** ***Method II: discretization of the differential equation.*** In this case, we based the calculation on the discrete form of the composition of the left and right fractional derivatives of order $\alpha$ presented in

[Błaszczyk et al. 2011]:

$$ {}^{C}D^{\alpha}_{L^-}\, {}^{C}D^{\alpha}_{0^+} u(x)\big|_{x=x_i} \approx \sum_{j=i-1}^{n} q^{(\alpha)}_{i,j} \sum_{k=0}^{j+1} r^{(\alpha)}_{j,k}\, u_k, \tag{5-5}$$

where coefficients $r^{(\alpha)}_{i,j}$ and $q^{(\alpha)}_{i,j}$ have the following forms:

$$ r^{(\alpha)}_{i,j} = \frac{(\Delta x)^{-\alpha}}{\Gamma(3-\alpha)} \begin{cases} 0 & \text{for } i=0 \text{ and } j=0,\\ i^{2-\alpha}-(i-1)^{2-\alpha} & \text{for } i>0 \text{ and } j=0,\\ 3(i-1)^{2-\alpha}-2i^{2-\alpha}-(i-2)^{2-\alpha} & \text{for } i>0 \text{ and } j=1,\\ (i-j+2)^{2-\alpha}-3(i-j+1)^{2-\alpha} & \\ \qquad +3(i-j)^{2-\alpha}-(i-j)^{2-\alpha} & \text{for } i>0 \text{ and } 1<j<i,\\ 2^{2-\alpha}-3 & \text{for } i<n \text{ and } j=i,\\ 1 & \text{for } i<n \text{ and } j=i+1 \end{cases} \tag{5-6}$$

$$ q^{(\alpha)}_{i,j} = \frac{(\Delta x)^{-\alpha}}{\Gamma(3-\alpha)} \begin{cases} 0 & \text{for } i=n \text{ and } j=n,\\ (n-i)^{2-\alpha}-(n-i-1)^{2-\alpha} & \text{for } i<n \text{ and } j=n,\\ 3(n-i-1)^{2-\alpha}-2(n-i)^{2-\alpha}-(n-i-2)^{2-\alpha} & \text{for } i<n \text{ and } j=n-1,\\ (j-i+2)^{2-\alpha}-3(j-i+1)^{2-\alpha} & \\ \qquad +3(j-i)^{2-\alpha}-(j-i-1)^{2-\alpha} & \text{for } i<n \text{ and } i<j<n-1,\\ 2^{2-\alpha}-3 & \text{for } i<n \text{ and } j=i,\\ 1 & \text{for } i<n \text{ and } j=i-1 \end{cases} \tag{5-7}$$

Now we present the discrete form of the Euler–Lagrange Equation (3-7). For every grid node $x_i$, $i = 2, 3, \dots, n-2$, we write the equation as

$$ \sum_{j=i-1}^{n} q^{(\alpha)}_{i,j} \sum_{k=0}^{j+1} r^{(\alpha)}_{j,k}\, u_k = f^{*}_{i}. \tag{5-8}$$

In this approach we have to solve the system of $n+1$ algebraic equations.

## 6. Error analysis

Let us consider the following example:

$$ {}^{C}D^{\alpha}_{1^-}\, {}^{C}D^{\alpha}_{0^+} u(x) = C, \tag{6-1}$$

where $C$ is a constant. The exact solution of the considered problem (6-1) with boundary conditions (3-2) is

$$ u(x) = x^{\alpha}\big[(\alpha x-\alpha-1)g_{\alpha,\alpha}(1) + (1-x)g_{\alpha-1,\alpha}(1)\big] + g_{\alpha,\alpha}(x), \tag{6-2}$$

where

$$ g_{\eta,\mu}(x) = I^{\eta}_{0^+} I^{\mu}_{1^-} C = I^{\eta}_{0^+} \frac{C(1-x)^{\mu}}{\Gamma(\mu+1)} = \frac{Cx^{\eta}}{\Gamma(\eta+1)\Gamma(\mu+1)}\, {}_2F_1(1,-\mu;\eta+1;x). \tag{6-3}$$

| Method I | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| $n$ | $\alpha = 1.5$ | $p$ | $\alpha = 1.6$ | $p$ | $\alpha = 1.7$ | $p$ | $\alpha = 1.8$ | $p$ | $\alpha = 1.9$ | $p$ |
| 80 | $2.27 \cdot 10^{-6}$ | – | $1.08 \cdot 10^{-6}$ | – | $4.72 \cdot 10^{-7}$ | - | $1.84 \cdot 10^{-7}$ | – | $5.42 \cdot 10^{-8}$ | – |
| 160 | $7.01 \cdot 10^{-7}$ | 1.70 | $3.10 \cdot 10^{-7}$ | 1.80 | $1.29 \cdot 10^{-7}$ | 1.87 | $4.84 \cdot 10^{-8}$ | 1.93 | $1.39 \cdot 10^{-8}$ | 1.96 |
| 320 | $2.08 \cdot 10^{-7}$ | 1.75 | $8.61 \cdot 10^{-8}$ | 1.84 | $3.42 \cdot 10^{-8}$ | 1.91 | $1.25 \cdot 10^{-8}$ | 1.95 | $3.53 \cdot 10^{-9}$ | 1.98 |
| 640 | $6.04 \cdot 10^{-8}$ | 1.79 | $2.34 \cdot 10^{-8}$ | 1.88 | $8.93 \cdot 10^{-9}$ | 1.94 | $3.18 \cdot 10^{-9}$ | 1.97 | $8.91 \cdot 10^{-10}$ | 1.99 |
| 1280 | $1.72 \cdot 10^{-8}$ | 1.81 | $6.27 \cdot 10^{-9}$ | 1.90 | $2.30 \cdot 10^{-9}$ | 1.95 | $8.07 \cdot 10^{-10}$ | 1.98 | $2.24 \cdot 10^{-10}$ | 1.99 |

| Method II | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| $n$ | $\alpha = 1.5$ | $p$ | $\alpha = 1.6$ | $p$ | $\alpha = 1.7$ | $p$ | $\alpha = 1.8$ | $p$ | $\alpha = 1.9$ | $p$ |
| 80 | $8.74 \cdot 10^{-3}$ | – | $7.97 \cdot 10^{-3}$ | – | $7.13 \cdot 10^{-4}$ | - | $5.61 \cdot 10^{-4}$ | – | $3.05 \cdot 10^{-4}$ | – |
| 160 | $5.15 \cdot 10^{-4}$ | 0.75 | $5.07 \cdot 10^{-4}$ | 0.65 | $5.11 \cdot 10^{-4}$ | 0.48 | $4.65 \cdot 10^{-4}$ | 0.27 | $2.91 \cdot 10^{-4}$ | 0.07 |
| 320 | $3.09 \cdot 10^{-4}$ | 0.74 | $3.20 \cdot 10^{-4}$ | 0.66 | $3.59 \cdot 10^{-4}$ | 0.51 | $3.75 \cdot 10^{-4}$ | 0.31 | $2.72 \cdot 10^{-4}$ | 0.09 |
| 640 | $1.89 \cdot 10^{-4}$ | 0.70 | $2.01 \cdot 10^{-4}$ | 0.67 | $2.47 \cdot 10^{-4}$ | 0.54 | $2.96 \cdot 10^{-4}$ | 0.34 | $2.49 \cdot 10^{-4}$ | 0.13 |
| 1280 | $1.20 \cdot 10^{-4}$ | 0.65 | $1.27 \cdot 10^{-4}$ | 0.67 | $1.68 \cdot 10^{-4}$ | 0.55 | $2.31 \cdot 10^{-4}$ | 0.35 | $2.25 \cdot 10^{-4}$ | 0.15 |

**Table 1.** Maximum absolute errors and experimental estimation of rates of convergence $p$ for $\alpha \in \{1.5;\ 1.6;\ 1.7;\ 1.8;\ 1.9\}$ and $C = -1/\ell^{2(\alpha-2)}$.

In this case (when the exact solution is available) we can compute errors generated by the presented methods (5-4) and (5-8). Maximum errors were calculated based on the standard expression

$$\operatorname{err}(\Delta x) = \max_{i=0,\dots,n} (|u(x_i) - u_i|), \tag{6-4}$$

and the experimental rate of convergence was estimated by using the formula

$$p = \log_2 \left( \frac{\operatorname{err}(\Delta x)}{\operatorname{err}(\Delta x/2)} \right). \tag{6-5}$$

The maximum absolute errors and the experimental rate of convergence $p$ for various values of $\alpha$ and $n$ are included in Table 1. One can note that errors decrease by increasing the number $n$. We also conclude that the results obtained by using Method I are much better than those from Method II. From these numerical tests, one may see that the rate of convergence $p$ for scheme (5-4) is significantly higher than for scheme (5-8).

## 7. Example of computations

On the basis of the numerical scheme (5-4) presented in the paper, we implemented an algorithm in Maple and carried out computational simulations for various values of parameters $\alpha$, $\ell$, and a different type of function $f$. In all presented examples we consider a nondimensional case and assumed $L = 1$, $E = 1$, $I = 1$, $\varepsilon = 0.01$ and $\Delta x = 0.001$. Two cases are considered:

(i) The beam with central load

$$f(x)=\begin{cases}-1/\varepsilon, & \text{for } L/2-\varepsilon/2\le x\le L/2+\varepsilon/2\\ 0, & \text{otherwise}\end{cases}. \tag{7-1}$$

(ii) The beam with asymmetric load

$$f(x)=\begin{cases}-1/\varepsilon, & \text{for } 3L/4-\varepsilon/2\le x\le 3L/4+\varepsilon/2\\ 0, & \text{otherwise}\end{cases}. \tag{7-2}$$

The numerical results are presented in Figures 1 and 2.

Analyzing the results presented in Figures 1 and 2, we observe that when length scale $\ell$ increases in comparison to beam length $L$, the nonlocal effects are more noticeable. If $\ell$ is close to 0 or if the order of Caputo derivative $\alpha$ tends to 2, the fractional Euler–Bernoulli beam equation reduces to the classical model. It is also clearly seen that when values of $\alpha$ decrease the difference between the classical and fractional result increases.

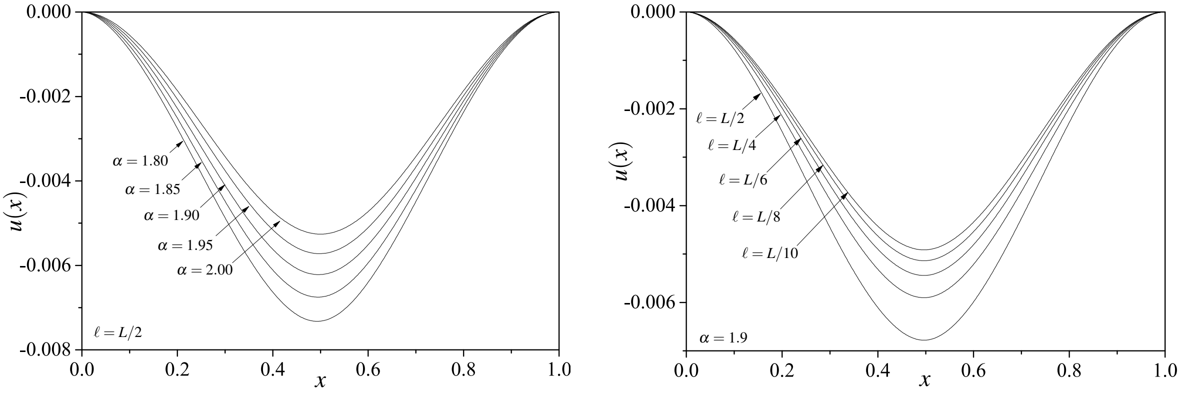

**Figure 1.** Comparison of the static deflections for a beam with central load (case (i)) for a various order of fractional derivatives $\alpha$ and the length scale $\ell$.

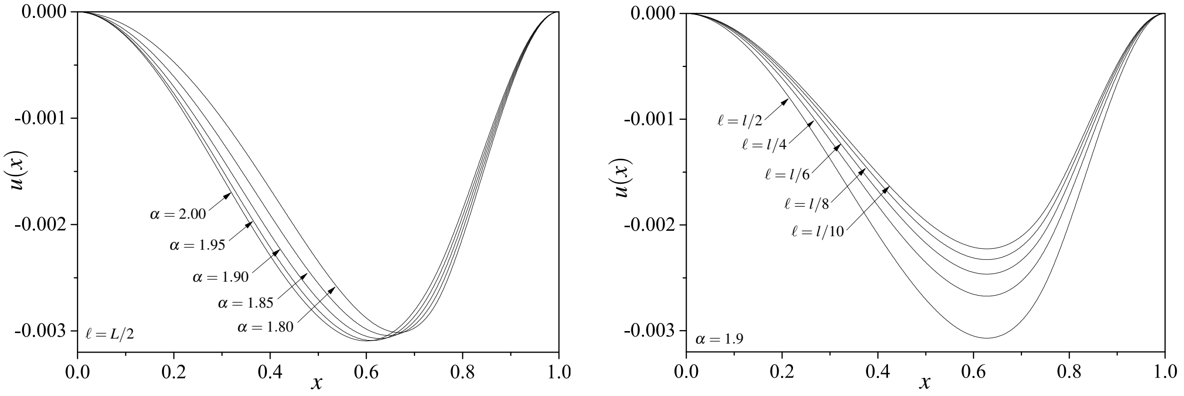

**Figure 2.** Comparison of the static deflections for a beam with asymmetric load (case (ii)) for a various order of fractional derivatives $\alpha$ and the length scale $\ell$.

## 8. Conclusions

In this paper a new formulation of the Euler–Bernoulli beam equation utilizing fractional calculus was presented. It should be stressed that this is the first time the fractional variational approach was used to obtain the Euler–Bernoulli beam equation. Therefore, the spatial fractional operator appearing in this differential equation is the composition of left and right Caputo derivatives and this operator is defined on the finite interval. The obtained equation was transformed into its equivalent integral form and then was solved. We received both the exact and numerical solution of the integral equation. The errors generated by the presented numerical scheme (Method I) and the experimental rate of convergence were calculated and compared with numerical results obtained by discretizing the Euler–Lagrange equation (Method II) for a particular case. We also carried out simulations to demonstrate how parameters $\ell$ and $\alpha$ affect the static deflections of the beam.

## References

[Alotta et al. 2015] G. Alotta, G. Failla, and M. Zingales, “Finite-element formulation of a nonlocal hereditary fractional-order timoshenko beam”, *J. Eng. Mech.-ASCE* (2015), D4015001.

[Atanackovic and Stankovic 2009] T. M. Atanackovic and B. Stankovic, “Generalized wave equation in nonlocal elasticity”, *Acta Mech.* **208**:1-2 (2009), 1–10.

[Baleanu et al. 2012] D. Baleanu, K. Diethelm, E. Scalas, and J. J. Trujillo, *Fractional calculus and numerical methods*, Series on Complexity, Nonlinearity and Chaos **3**, World Scientific, Hackensack, New Jersey, 2012.

[Baleanu et al. 2015] D. Baleanu, J. H. Asad, and I. Petras, “Numerical solution of the fractional Euler–Lagrange’s equations of a thin elastica model”, *Nonlinear Dynam.* **81**:1-2 (2015), 97–102.

[Błasik and Klimek 2015] M. Błasik and M. Klimek, “Numerical solution of the one phase 1D fractional Stefan problem using the front fixing method”, *Math. Methods Appl. Sci.* **38**:15 (2015), 3214–3228.

[Błaszczyk and Ciesielski 2015] T. Błaszczyk and M. Ciesielski, “Fractional oscillator equation—transformation into integral equation and numerical solution”, *Appl. Math. Comput.* **257** (2015), 428–435.

[Błaszczyk and Ciesielski 2016] T. Błaszczyk and M. Ciesielski, “Fractional oscillator equation—analytical solution and algorithm for its approximate computation”, *J. Vib. Control* **22**:8 (2016), 2045–2052.

[Błaszczyk et al. 2011] T. Błaszczyk, M. Ciesielski, M. Klimek, and J. Leszczynski, “Numerical solution of fractional oscillator equation”, *Appl. Math. Comput.* **218**:6 (2011), 2480–2488.

[Carpinteri et al. 2011] A. Carpinteri, P. Cornetti, and A. Sapora, “A fractional calculus approach to nonlocal elasticity”, *Eur. Phys. J. Spec. Top.* **193**:1 (2011), 193–204.

[Carpinteri et al. 2014] A. Carpinteri, P. Cornetti, and A. Sapora, “Nonlocal elasticity: an approach based on fractional calculus”, *Meccanica* **49**:11 (2014), 2551–2569.

[Ciesielski and Błaszczyk 2015] M. Ciesielski and T. Błaszczyk, “Numerical solution of non-homogenous fractional oscillator equation in integral form”, *J. Theor. Appl. Mech.* **53**:4 (2015), 959–968.

[Drapaca and Sivaloganathan 2012] C. S. Drapaca and S. Sivaloganathan, “A fractional model of continuum mechanics”, *J. Elasticity* **107**:2 (2012), 105–123.

[Durajski 2014] A. P. Durajski, “Phonon-mediated superconductivity in compressed NbH4 compound”, *Eur. Phys. J. B* **87**:9 (2014), 1–6.

[Durajski 2015] A. P. Durajski, “Influence of hole doping on the superconducting state in graphane”, *Supercond. Sci. Tech.* **28**:3 (2015), 35002–35009.

[Eringen 2002] A. C. Eringen, *Nonlocal continuum field theories*, Springer, New York, 2002.

[Freundlich 2013] J. Freundlich, “Vibrations of a simply supported beam with a fractional viscoelastic material model — supports movement excitation”, *Shock and Vibration* **20**:6 (2013), 1103–1112.

[Kilbas et al. 2006] A. A. Kilbas, H. M. Srivastava, and J. J. Trujillo, *Theory and applications of fractional differential equations*, North-Holland Mathematics Studies **204**, Elsevier Science B.V., Amsterdam, 2006.

[Klimek 2001] M. Klimek, "Fractional sequential mechanics—models with symmetric fractional derivative", *Czechoslovak J. Phys.* **51**:12 (2001), 1348–1354.

[Kukla and Siedlecka 2015] S. Kukla and U. Siedlecka, "Laplace transform solution of the problem of time-fractional heat conduction in a two-layered slab", *J. Appl. Math. Comput. Mech.* **14**:4 (2015), 105–113.

[Lazo and Torres 2013] M. J. Lazo and D. F. M. Torres, "The DuBois–Reymond fundamental lemma of the fractional calculus of variations and an Euler–Lagrange equation involving only derivatives of Caputo", *J. Optim. Theory Appl.* **156**:1 (2013), 56–67.

[Lazopoulos 2006] K. A. Lazopoulos, "Non-local continuum mechanics and fractional calculus", *Mech. Res. Comm.* **33**:6 (2006), 753–757.

[Leszczyński 2011] J. Leszczyński, *An introduction to fractional mechanics*, Publishing Office of Czestochowa University of Technology, 2011.

[Malinowska et al. 2015] A. B. Malinowska, T. Odzijewicz, and D. F. M. Torres, *Advanced methods in the fractional calculus of variations*, Springer Briefs in Applied Sciences and Technology, Springer, Cham, Switzerland, 2015.

[Paola and Zingales 2008] M. D. Paola and M. Zingales, "Long-range cohesive interactions of non-local continuum faced by fractional calculus", *Int. J. Solids Struct.* **45**:21 (2008), 5642–5659.

[Paola et al. 2013] M. D. Paola, R. Heuer, and A. Pirrotta, "Fractional visco-elastic Euler–Bernoulli beam", *Int. J. Solids Struct.* **50**:22–23 (2013), 3505–3510.

[Pirrotta et al. 2015] A. Pirrotta, S. Cutrona, and S. Di Lorenzo, "Fractional visco-elastic Timoshenko beam from elastic Euler–Bernoulli beam", *Acta Mech.* **226**:1 (2015), 179–189.

[Podlubny 1999] I. Podlubny, *Fractional differential equations*, Mathematics in Science and Engineering **198**, Academic Press, San Diego, California, 1999.

[Samko et al. 1993] S. G. Samko, A. A. Kilbas, and O. I. Marichev, *Fractional integrals and derivatives: theory and applications*, Gordon and Breach Science Publishers, Yverdon, Switzerland, 1993.

[Siedlecki et al. 2015] J. Siedlecki, M. Ciesielski, and T. Błaszczyk, "Transformation of the second order boundary value problem into integral form — different approaches and a numerical solution", *J. Appl. Math. Comput. Mech.* **14**:3 (2015), 103–108.

[Sumelka 2014a] W. Sumelka, "Application of fractional continuum mechanics to rate independent plasticity", *Acta Mech.* **225**:11 (2014), 3247–3264.

[Sumelka 2014b] W. Sumelka, "Thermoelasticity in the framework of the fractional continuum mechanics", *J. Therm. Stresses* **37**:6 (2014), 678–706.

[Sumelka and Błaszczyk 2014] W. Sumelka and T. Błaszczyk, "Fractional continua for linear elasticity", *Arch. Mech.* (*Arch. Mech. Stos.*) **66**:3 (2014), 147–172.

[Sumelka et al. 2015] W. Sumelka, T. Błaszczyk, and C. Liebold, "Fractional Euler–Bernoulli beams: theory, numerical study and experimental validation", *Eur. J. Mech. A Solids* **54** (2015), 243–251.

[Tarasov 2006] V. E. Tarasov, "Continuous limit of discrete systems with long-range interaction", *J. Phys. A* **39**:48 (2006), 14895–14910.

[Torres 2015] C. Torres, "Ground state solution for differential equations with left and right fractional derivatives", *Math. Methods Appl. Sci.* **38**:18 (2015), 5063–5073.

[Vázquez 2004] L. Vázquez, "A fruitful interplay: from nonlocality to fractional calculus", pp. 129–133 in *Nonlinear waves: Classical and quantum aspects. Proceedings of the NATO advanced research workshop* (Estoril, Portugal), edited by F. K. Abdullaev and V. V. Konotop, Dordrecht: Kluwer Academic Publishers, 2004.

[Zhang et al. 2016] Y. Zhang, L. Chen, D. M. Reeves, and H. Sun, "A fractional-order tempered-stable continuity model to capture surface water runoff", *J. Vib. Control* **22**:8 (2016), 1993–2003.

[Zingales and Failla 2015] M. Zingales and G. Failla, "The finite element method for fractional non-local thermal energy transfer in non-homogeneous rigid conductors", *Commun. Nonlinear Sci. Numer. Simul.* **29**:1-3 (2015), 116–127.

Received 5 Jan 2016. Revised 13 May 2016. Accepted 27 May 2016.

TOMASZ BLASZCZYK: `tomasz.blaszczyk@im.pcz.pl`
*Institute of Mathematics, Czestochowa University of Technology, Al. Armii Krajowej 21, 42-200 Czestochowa, Poland*

JOURNAL OF MECHANICS OF MATERIALS AND STRUCTURES
Vol. 12, No. 1, 2017
dx.doi.org/10.2140/jomms.2017.12.35

# FRACTIONAL CALCULUS IN NEURONAL ELECTROMECHANICS

CORINA S. DRAPACA

Traumatic brain injuries (TBI) are among the leading causes of death and permanent disability worldwide. Recent experimental observations suggest that damage in brain tissue involves complex local as well as nonlocal chemomechanical interactions that happen on multiple spatiotemporal scales. Biomechanical models of TBI existing in the literature do not incorporate either electrochemical or multiscaling features. Given that neurons are the brain cells responsible for electrochemical signaling on multiplexed temporal scales we propose a novel mathematical model of neuronal electromechanics that uses a constrained Lagrangian formulation and Hamilton's principle to couple Newton's law of motion for a linear viscoelastic Kelvin–Voigt solid-state neuron and the classic Hodgkin–Huxley equations of the electronic neuron. We will use fractional order derivatives of variable order to model multiple temporal scales. Numerical simulations of possible damage dynamics in neurons due to mechanical trauma will be presented and discussed.

*A list of symbols can be found on page 53.*

## 1. Introduction

Traumatic brain injuries (TBI) are among the main causes of death and disability worldwide, contributing to approximately 30% of all injury deaths in the United States in 2010 [CDC 2016]. The data collected in the United States during the period 2001-2010 [CDC 2016] show a dramatic increase of 70% in TBI-related visits to emergency rooms, while the death rates decreased only by 7%. This prompted an unprecedented unified effort from various US organizations (government, health and social services) to come up with a public health approach for TBI in [Bell et al. 2015]. The authors of that article concluded it by emphasizing the need for continued progress in brain science that can "inform and suggest solutions for a problem that is of significant concern to the public". In particular, mathematical models and corresponding computer simulations can increase our comprehension of brain responses to TBI and help us design better experiments for measurements and hypothesis testing that ultimately will lead to improved medical diagnostic and therapeutic protocols. In the last few decades a multitude of mathematical models have been proposed to study brain biomechanics and, independently, brain biochemistry at cell and tissue levels (see [Goldsmith 2001; Goriely et al. 2015] for comprehensive reviews of these models). However, most of these models have many physical parameters which are hard, if not impossible, to find experimentally and the high complexity of the corresponding computations makes these models hard to use in today's clinical applications. In addition, these biomechanical models were built at the tissue level and thus they cannot predict the mechanochemical responses of brain cells to mechanical and/or electrochemical events that happen at the tissue and organ scales.

---

*Keywords:* fractional derivatives of variable order, entangled scales, Hamilton's principle, fractional calculus, Hodgkin–Huxley model, neuronal electromechanics, TBI.

Following the recommendation of Goriely et al. [2015] for the development of *bottom-up* mathematical models that link brain mechanics and electrochemistry at each relevant length scale as well as across scales, we recently proposed a lower-dimensional electromechanical model of a neuron which is simple enough so that its predictions may be experimentally verified, and could be used as a foundation model for more advanced multiscaling mathematical models [Drapaca 2015]. By assuming that the electrochemical activity of a neuron is described by the classic Hodgkin–Huxley [1952] equations and that the neuron behaves mechanically like a linear viscoelastic Kelvin–Voigt solid, we showed through numerical simulations that very fast initially applied speeds (jabbing) inhibit the action potentials and thus might cause neuronal damage.

In this paper we generalize the model from [Drapaca 2015] by incorporating multiple time scales using fractional order temporal derivatives of variable orders. In the last few decades fractional calculus has been successfully used in a wide range of applications to model stochastic, multiscaling and nonlocal phenomena in various physical systems (some relevant books on fractional calculus and its applications are [Podlubny 1999; Samko et al. 1993; Oldham and Spanier 2006; Hilfer 2000; Baleanu et al. 2012; Milici and Draganescu 2015; West 2015; Tarasov 2010]). Given that neuronal electrochemical dynamics are stochastic [Schiff 2012] and fractional calculus is a natural mathematical representation of stochasticity [West 2015], Sherief et al [2012] generalized the classic Hodgkin–Huxley model by replacing the first order temporal derivatives with fractional order ones. The use of fractional order derivatives in mathematical models of neuronal dynamics is supported by the experimental observations made in [Lundstrom et al. 2008], and only last year Grevesse et al. [2015] showed empirically that mechanical creep of neurons follow a power law of fractional order. Thus we propose to replace the first order time derivatives in the model from [Drapaca 2015] with fractional temporal derivatives with variable orders. The time-dependency of the fractional orders represents the biological variability of neurons as well as the intrinsic entanglement of states existing in the complex mixture of physical components that makes up a neuron. We call this inseparability of time scales *entangled scales*.

In this paper we use the respective left and right Riemann–Liouville fractional derivatives of variable order which were introduced in [Atanackovic and Pilipovic 2011] as

$$ {}_0D_t^{\alpha(t)} f(t) = \frac{f(0)}{\Gamma(1-\alpha(t))t^{\alpha(t)}} + \int_0^t \frac{\mathrm{d}f(\tau)/\mathrm{d}\tau}{\Gamma(1-\alpha(t-\tau))(t-\tau)^{\alpha(t-\tau)}}\,\mathrm{d}\tau, \tag{1-1}$$

$$ {}_tD_T^{\alpha(t)} f(t) = \frac{f(T)}{\Gamma(1-\alpha(T-t))(T-t)^{\alpha(T-t)}} - \int_t^T \frac{\mathrm{d}f(\tau)/\mathrm{d}\tau}{\Gamma(1-\alpha(\tau-t))(\tau-t)^{\alpha(\tau-t)}}\,\mathrm{d}\tau, \tag{1-2}$$

where $f$ is an absolutely continuous[1] function on $[0, T]$ with $f(t) = 0$, $\forall t \in \mathbb{R} - [0, T]$, and the variable order $\alpha(t)$ is a continuous function on $[0, T)$ and $0 \le \alpha(t) < 1$. Definitions (1-1) and (1-2) allow us to use a nonconservative form of Hamilton's principle proposed in [Atanackovic and Pilipovic 2011] and obtain the generalized integro-differential Euler–Lagrange equations corresponding to our electromechanical model.

Although Definitions (1-1) and (1-2) can model the fading memory of materials with variable viscoelasticity [Lorenzo and Hartley 2002], it was shown in [Chicone and Mashhoon 2002] that in the case of

---

[1] The absolute continuity of a real-valued function $f$ on an interval $[0, T]$ is equivalent to the existence almost everywhere of the derivative $\mathrm{d}f/\mathrm{d}t$ which is Lebesgue integrable and $f(t) = f(0) + \int_0^t (\mathrm{d}f/\mathrm{d}s)\mathrm{d}s$, $\forall t \in [0, T]$.

piecewise uniform accelerated (linear and circular) motions these derivatives introduce additional memory effects that infringe causality. On the other hand, the respective left and right Marchaud fractional derivatives of variable order

$$ {}_0\tilde{D}_t^{\alpha(t)} f(t) = \frac{f(0)}{\Gamma(1-\alpha(t))t^{\alpha(t)}} + \frac{1}{\Gamma(1-\alpha(t))} \int_0^t \frac{\mathrm{d}f(\tau)/\mathrm{d}\tau}{(t-\tau)^{\alpha(t)}}\,\mathrm{d}\tau, \tag{1-3} $$

$$ {}_t\tilde{D}_T^{\alpha(t)} f(t) = \frac{f(T)}{\Gamma(1-\alpha(t))(T-t)^{\alpha(t)}} - \frac{1}{\Gamma(1-\alpha(t))} \int_t^T \frac{\mathrm{d}f(\tau)/\mathrm{d}\tau}{(\tau-t)^{\alpha(t)}}\,\mathrm{d}\tau, \tag{1-4} $$

satisfy the causality law [Chicone and Mashhoon 2002] and therefore their mathematical properties and applications have been increasingly studied [Coimbra and Kobayashi 2002; Coimbra 2003; Soon et al. 2005; Sun et al. 2012; Almeida and Torres 2013]. However, it is easy to check (using power series expansions) that for variable order functions $\alpha : [0, T] \to [0, 1)$ which are differentiable with continuous derivatives and $|\mathrm{d}\alpha/\mathrm{d}t(t)| \ll 1$, $t \in (0, T)$, the approximations

$$ {}_0D_t^{\alpha(t)} f(t) \approx {}_0\tilde{D}_t^{\alpha(t)} f(t) \quad \text{and} \quad {}_tD_T^{\alpha(t)} f(t) \approx {}_t\tilde{D}_T^{\alpha(t)} f(t) \tag{1-5} $$

hold, which are identically satisfied for $\alpha(t) = \text{constant} \in [0, 1)$. Thus, in this paper we will use the left and right Riemann–Liouville fractional derivatives (1-1) and (1-2) with variable orders belonging to the following class of functions: $\mathscr{C} = \{\alpha : [0, T] \to [0, 1)/(\mathrm{d}\alpha/\mathrm{d}t)$ exists and is continuous $|(\mathrm{d}\alpha/\mathrm{d}t)| \ll 1\}$ such that, thanks to the approximations (1-5), the causality and nonlocality criteria introduced in [Chicone and Mashhoon 2002] are satisfied and the expansion formulas with higher-order derivatives proposed in [Almeida and Torres 2013] can be applied.

We model a neuron as a linear viscoelastic Kelvin–Voigt solid with variable viscoelasticity whose electrochemical activity is described by fractional order Hodgkin–Huxley equations with variable order. In addition, we introduce three linear viscoelastic Maxwell fluid elements with variable viscoelasticity that provide a physical representation for the three ionic gates with gating variables $m$,$n$, and $h$ introduced by the Hodgkin–Huxley model. The physical analogy of the ionic gates is that of door closers. We use a Lagrangian formulation and Hamilton's principle to obtain the equations of motion that couple macroscopic (cell level) and microscopic (ionic level) mechanical and electrical information and therefore they can describe neuronal mechanotransduction. As in [Drapaca 2015], we assume that at the macroscopic level the membrane's capacitance depends on the mechanical displacement of the neuron and that the Young's modulus of the neuron depends on the gated variables $m$,$n$, and $h$. Our numerical simulations solve a simplified version of the proposed equations using Matlab. Our results are comparable to those in [Drapaca 2015]: when a constant external electric current is applied and the initial displacement and speed are of orders of magnitude comparable to the size of the membrane, the action potentials look similar to the ones seen in healthy neurons, while at very fast initial speeds (which could model a serious traumatic event) and in the presence of a constant applied external current, high persisting oscillations in the volume of the neuron are observed and the action potentials do not happen. Some points of note: these results were obtained for a variable fractional order of the macroscopic Kelvin–Voigt element, which was chosen based on mathematical simplicity rather than physical inspiration; and the classic Hodgkin–Huxley equations were used instead of those of the microscopic Maxwell elements. The classic Hodgkin–Huxley model has very finely adjusted equations and parameters which might have hidden the possible effects of the variable viscoelasticity modeled with fractional order time derivatives of variable order. It is also possible that there exist other variable fractional orders that enclose physical information that complements the Hodgkin–Huxley model. The lack of experimentally supported information on neuronal mechanics and mechanotransduction limits our ability to explore the full capabilities of the proposed model. However, the model is general enough and can be adapted to practical applications. For instance, the model is independent of the Hodgkin–Huxley equations and thus simpler equations could be used together with fractional order time derivatives of variable orders (and possibly the corresponding evolution equations of the variable order functions) to observe action potentials and neuronal mechanics. In this case the model might have fewer parameters with better prospects of finding experimental proof. Therefore the work presented here can be seen as a first step towards a simpler chemomechanical model of a neuron and its membrane.

In short, the main contributions of this paper are:

(1) Providing the physical structure of door closers to the ionic gates $m$,$n$, and $h$.

(2) Introducing the concept of entangled temporal scales for the stochastic nature of the action potential and for the inseparability of the multiple time scales involved in the neuronal mechanochemical processes.

(3) Use of fractional temporal derivatives of variable orders to model the entangled temporal scales.

(4) Showing through numerical simulations that after a serious traumatic event the elastic behavior of a neuron dominates over its viscoelastic response, which appears to be in agreement with the experimental observations reported in [Grevesse et al. 2015].

In the next section we present our mathematical model, followed by our results. The paper ends with a section of conclusions and future directions which contains a first attempt at modeling entanglement using an area law and level sets.

## 2. Mathematical model

We model the neuron as an axisymmetric circular cylindrical annulus whose inner core is filled with the intracellular space and the outer core is the cell's membrane (Figure 1). We assume that the intracellular space and the membrane are homogeneous and thus reduce the study of neuronal electromechanics to the study of a simple electromechanical element that we introduce here. Our low-dimensional electromechanical model couples spring-dashpot-mass mechanical elements for the intracellular space and for the ionic gates located in the cell's membrane to an electric circuit model of the cell's membrane (Figure 1). Motivated by the experimental findings in [Lu et al. 2006; Grevesse et al. 2015], we model the intracellular space as a linear viscoelastic Kelvin–Voigt solid. We use the Hodgkin–Huxley [1952] electric circuit to model the macroscopic electric dynamics of neuron's membrane. Besides providing a mathematical representation for neuronal electric dynamics, the Hodgkin–Huxley model introduces three ion gates, $m$,$n$, and $h$, that produce action potentials by controlling the ionic flow into and out of the neuron. Although $m$,$n$, and $h$ are seen as nondimensional gate positions whose open or close state is determined by phenomenologically established first order ordinary differential equations, no physical structure has been given to them so far. In this paper we propose to model the $m$,$n$, and $h$ gates as linear viscoelastic Maxwell fluid elements located in the cell's membrane (Figure 1). The physical analogy for these ionic gates is a door closer. It is important to notice here that the electric circuit and Maxwell

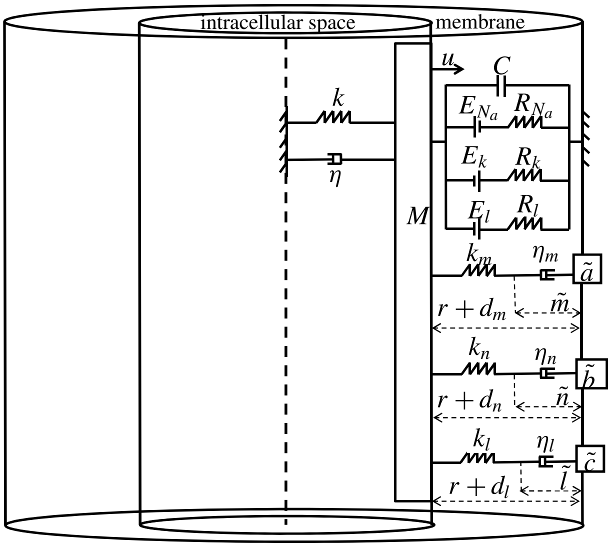

**Figure 1.** Schematic of the proposed model: the neuron is an axisymmetric homogeneous circular cylinder whose inner core is the intracellular space, and the outer layer is the membrane. Due to the symmetry (dashpot line) and material homogeneity, it is enough to study half of the neuron whose properties are encapsulated into a spring-dashpot-mass mechanical system with the spring and dashpot connected in parallel (Kelvin–Voigt model), and the membrane is represented as an electric circuit governed by the classic Hodgkin–Huxley equations.

elements shown in Figure 1 might not actually be independent and this could be further investigated using the method of mechanoelectric analogy, which is commonly employed in systems engineering [Koenig and Blackwell 1961] (a few insightful comments on the mechanoelectric analogy are stated after formula (2-32)). However, given the empirical nature of the Hodgkin–Huxley model and the current lack of knowledge of neuronal mechanotransduction, in this paper we will treat the electric circuit and the Maxwell elements as independent. The coupling of the Kelvin–Voigt and Maxwell mechanical elements to the Hodgkin–Huxley electric circuit is achieved by using a Lagrangian formulation and Hamilton's principle. We introduce a Lagrangian of the form [Drapaca 2015; Galley et al. 2014]

$$\begin{aligned}\mathcal{L}(q_{\mathrm{Na}}, q_{\mathrm{K}}, q_l, u, d_m, d_n, d_h, \tilde{m}, \tilde{n}, \tilde{h}) &= \tfrac{1}{2}M({}_0D_t^{\alpha(t)}u)^2 + \tfrac{1}{2}\tilde{a}({}_0D_t^{\beta(t)}(r+d_m))^2 \\ &\quad + \tfrac{1}{2}\tilde{b}({}_0D_t^{\beta(t)}(r+d_n))^2 + \tfrac{1}{2}\tilde{c}({}_0D_t^{\beta(t)}(r+d_h))^2 \\ &\quad - \frac{1}{2C(u)}q_C^2 - \tfrac{1}{2}k(\tilde{m},\tilde{n},\tilde{h})u^2 - \tfrac{1}{2}k_m(q_C,u)(r+d_m-\tilde{m})^2 \\ &\quad - \tfrac{1}{2}k_n(q_C,u)(r+d_n-\tilde{n})^2 - \tfrac{1}{2}k_h(q_C,u)(r+d_h-\tilde{h})^2, \qquad (2\text{-}1)\end{aligned}$$

where $M$ is half of the constant mass of the neuron with constant cross-sectional area $A$; $u(t)$ is the macroscopic (cell scale) displacement of mass $M$ that depends on time $t$; $\tilde{a}$, $\tilde{b}$, and $\tilde{c}$ are the masses of the gates $m$, $n$, and $h$, respectively; $r + d_m$, $r + d_n$, and $r + d_h$ are the relative displacements between mass $M$ and each of the masses $\tilde{a}$, $\tilde{b}$, and $\tilde{c}$, respectively; $r$ is the constant thickness of the membrane; and $\tilde{m}$, $\tilde{n}$, and $\tilde{h}$ are the microscopic (ionic scale) displacements of the dashpots in the Maxwell elements corresponding to the $m$, $n$, $h$ gates, respectively. In addition, $C(u)$ is the macroscopic capacitance of membrane's lipid bilayer modeled as a capacitor of electric charge $q_C$, $k(\tilde{m}, \tilde{n}, \tilde{h})$ is the spring constant of the Kelvin–Voigt element, while $k_m$, $k_n$, and $k_h$ are the spring constants of the corresponding Maxwell elements. Lastly, $q_{\mathrm{Na}}$, $q_{\mathrm{K}}$, and $q_l$ are the electric charges of $\mathrm{Na}^+$, $\mathrm{K}^+$, and leakage channels, respectively. For simplicity, we assume that the relative displacements $d_m$, $d_n$, and $d_h$ are independent of displacement $u$. The fractional derivative in formula (2-1) is given by (1-1) for a variable order $\alpha \in \mathscr{C}$. In formula (2-1) we identify the first term of the Lagrangian $\mathscr{L}$ with a macrokinetic mechanical energy, and the second, third and fourth terms with microkinetic mechanical energies. The fifth term in formula (2-1) represents a macropotential electric energy, the sixth term is a macropotential mechanical energy, and the last three terms are micropotential mechanical energies.

The conservation law of electric charges provides the constraint

$$q_C + q_{\mathrm{Na}} + q_{\mathrm{K}} + q_l = 0. \tag{2-2}$$

We take $q_{\mathrm{Na}}$, $q_{\mathrm{K}}$, $q_l$, $u$, $d_m$, $d_n$, $d_h$, $\tilde{m}$, $\tilde{n}$, and $\tilde{h}$ to be generalized coordinates and introduce corresponding independent variations $\delta q_{\mathrm{Na}}$, $\delta q_{\mathrm{K}}$, $\delta q_l$, $\delta u$, $\delta d_m$, $\delta d_n$, $\delta d_h$, $\delta\tilde{m}$, $\delta\tilde{n}$, and $\delta\tilde{h}$ that vanish at times $t = 0$ and $t = T$. Using formula (2-2) we calculate the first variation of the Lagrangian $\mathscr{L}$ as

$$\begin{aligned}
\delta\mathscr{L} &= \lim_{\epsilon\to 0}\frac{\mathrm{d}\mathscr{L}}{\mathrm{d}\epsilon}(q_{\mathrm{Na}}+\epsilon\delta q_{\mathrm{Na}}, q_{\mathrm{K}}+\epsilon\delta q_{\mathrm{K}}, q_l+\epsilon\delta q_l, u+\epsilon\delta u, d_m+\epsilon\delta d_m, d_n+\epsilon\delta d_n,\\
&\qquad\qquad\qquad\qquad d_h+\epsilon\delta d_h, \tilde{m}+\epsilon\delta\tilde{m}, \tilde{n}+\epsilon\tilde{n}, \tilde{h}+\epsilon\tilde{h})\\
&= \biggl[M\,{}_0D_t^{\alpha(t)}u\,{}_0D_t^{\alpha(t)}(\delta u) - ku - \frac{1}{2}\frac{\partial k_m}{\partial u}(r+d_m-\tilde{m})^2\\
&\qquad\qquad -\frac{1}{2}\frac{\partial k_n}{\partial u}(r+d_n-\tilde{n})^2 - \frac{1}{2}\frac{\partial k_h}{\partial u}(r+d_h-\tilde{h})^2 + \frac{1}{2C^2}q_C^2\biggr]\delta u\\
&\quad + \bigl[\tilde{a}\,{}_0D_t^{\beta(t)}(r+d_m)\,{}_0D_t^{\beta(t)}(\delta d_m) - k_m(r+d_m-\tilde{m})\bigr]\delta d_m\\
&\quad + \bigl[\tilde{b}\,{}_0D_t^{\beta(t)}(r+d_n)\,{}_0D_t^{\beta(t)}(\delta d_n) - k_n(r+d_n-\tilde{n})\bigr]\delta d_n\\
&\quad + \bigl[\tilde{c}\,{}_0D_t^{\beta(t)}(r+d_h)\,{}_0D_t^{\beta(t)}(\delta d_h) - k_h(r+d_h-\tilde{h})\bigr]\delta d_h + \biggl[-\frac{1}{2}\frac{\partial k}{\partial\tilde{m}}u^2 + k_m(r+d_m-\tilde{m})i\biggr]\delta\tilde{m}\\
&\quad + \biggl[-\frac{1}{2}\frac{\partial k}{\partial\tilde{n}}u^2 + k_n(r+d_n-\tilde{n})\biggr]\delta\tilde{n} + \biggl[-\frac{1}{2}\frac{\partial k}{\partial\tilde{h}}u^2 + k_h(r+d_h-\tilde{h})\biggr]\delta\tilde{h}\\
&\quad + \biggl[-\frac{1}{C}q_C - \frac{1}{2}\frac{\partial k_m}{\partial q_C}(r+d_m-\tilde{m})^2 - \frac{1}{2}\frac{\partial k_n}{\partial q_C}(r+d_n-\tilde{n})^2\\
&\qquad\qquad -\frac{1}{2}\frac{\partial k_h}{\partial q_C}(r+d_h-\tilde{h})^2\biggr](-\delta q_{\mathrm{Na}} - \delta q_{\mathrm{K}} - \delta q_l).
\end{aligned} \tag{2-3}$$

We define the virtual work done by nonconservative forces by [Drapaca 2015]

$$\begin{aligned}
\delta\mathscr{W} = &-\big[R_{\mathrm{Na}}({}_0D_t^{\gamma(t)}q_{\mathrm{Na}})\delta q_{\mathrm{Na}} + R_{\mathrm{K}}({}_0D_t^{\gamma(t)}q_{\mathrm{K}})\delta q_{\mathrm{K}} + R_l({}_0D_t^{\gamma(t)}q_l)\delta q_l\big] \\
&\quad - \big[\eta({}_0D_t^{\alpha(t)}u)\delta u + \eta_m({}_0D_t^{\beta(t)}\tilde{m})\delta\tilde{m} + \eta_n({}_0D_t^{\beta(t)}\tilde{n})\delta\tilde{n} + \eta_h({}_0D_t^{\beta(t)}\tilde{h})\delta\tilde{h}\big] \\
&\qquad + \big[-E_{\mathrm{Na}}\delta q_{\mathrm{Na}} - E_{\mathrm{K}}\delta q_{\mathrm{K}} - E_l\delta q_l + F_m\delta d_m + F_n\delta d_n + F_h\delta d_h + F\delta u\big].
\end{aligned} \tag{2-4}$$

In formula (2-4), the terms inside the first two sets of parentheses represent dissipative forces: the first is due to the resistors with resistances $R_{\mathrm{Na}}$,$R_{\mathrm{K}}$, and $R_l$ in the Hodgkin–Huxley electric circuit, and the second is due to the linear dashpots in the Kelvin–Voigt and Maxwell elements whose damping coefficients are $\eta$,$\eta_m$,$\eta_n$, and $\eta_h$. The last set of parentheses in (2-4) contain the reverse potentials $E_{\mathrm{Na}}$,$E_{\mathrm{K}}$, and $E_l$ of the Hodgkin–Huxley model and the forces $F_m$, $F_n$, $F_h$, and $F$ which are work conjugates for the Maxwell elements and the Kelvin–Voigt element, respectively. Again, the fractional derivatives in formula (2-4) are given by (1-1) with variable orders $\beta, \gamma \in \mathscr{C}$. The choice of signs in formula (2-4) guarantees that $\delta\mathscr{W}$ is thermodynamically consistent. The nomenclature of the physical and structural quantities used throughout the paper is given on page 53.

We use now the nonconservative form of Hamilton's principle

$$\int_0^T \delta\mathscr{L} + \delta\mathscr{W}\,\mathrm{d}t = 0. \tag{2-5}$$

We assume that the generalized coordinates and their variations are absolutely continuous functions on $[0, T]$ so that the following integration by parts formula can be applied [Atanackovic and Pilipovic 2011]:

$$\int_0^T \big({}_0D_t^{\alpha(t)}f(t)\big)\big({}_0D_t^{\alpha(t)}\delta f(t)\big)\,\mathrm{d}t = \int_0^T {}_tD_T^{\alpha(t)}({}_0D_t^{\alpha(t)}f(t))\delta f(t)\,\mathrm{d}t. \tag{2-6}$$

Thus, by replacing formulas (2-3) and (2-4) into (2-5), using the integration by parts formula (2-6), the independence of the variations of generalized coordinates and the fact that these are zero at 0 and $T$, we obtain the generalized Euler–Lagrange equations[2]

$$\begin{aligned}
M\,{}_tD_T^{\alpha(t)}({}_0D_t^{\alpha(t)}u) - \eta\,{}_0D_t^{\alpha(t)}u - ku + \frac{1}{2}\frac{\mathrm{d}C}{\mathrm{d}u}V^2 \\
-\frac{1}{2}\bigg[\frac{\partial k_m}{\partial u}(r + d_m - \tilde{m})^2 + \frac{\partial k_n}{\partial u}(r + d_n - \tilde{n})^2 + \frac{\partial k_h}{\partial u}(r + d_h - \tilde{h})^2\bigg] + F = 0,
\end{aligned} \tag{2-7}$$

$$\tilde{a}\,{}_tD_T^{\beta(t)}({}_0D_t^{\beta(t)}(r + d_m)) - k_m(r + d_m - \tilde{m}) + F_m = 0, \tag{2-8}$$

$$\tilde{b}\,{}_tD_T^{\beta(t)}({}_0D_t^{\beta(t)}(r + d_n)) - k_n(r + d_n - \tilde{n}) + F_n = 0, \tag{2-9}$$

$$\tilde{c}\,{}_tD_T^{\beta(t)}({}_0D_t^{\beta(t)}(r + d_h)) - k_h(r + d_h - \tilde{h}) + F_h = 0, \tag{2-10}$$

$$-\eta_m\,{}_0D_t^{\beta(t)}\tilde{m} + k_m(r + d_m - \tilde{m}) - \frac{1}{2}\frac{\partial k}{\partial\tilde{m}}u^2 = 0, \tag{2-11}$$

$$-\eta_n\,{}_0D_t^{\beta(t)}\tilde{n} + k_n(r + d_n - \tilde{n}) - \frac{1}{2}\frac{\partial k}{\partial\tilde{n}}u^2 = 0, \tag{2-12}$$

$$-\eta_h\,{}_0D_t^{\beta(t)}\tilde{h} + k_h(r + d_h - \tilde{h}) - \frac{1}{2}\frac{\partial k}{\partial\tilde{h}}u^2 = 0, \tag{2-13}$$

---

[2] These equations can also be obtained directly from [Atanackovic and Pilipovic 2011, Equation (35)].

$$-R_{\mathrm{Na}}({}_0D_t^{\gamma(t)}q_{\mathrm{Na}}) + V - E_{\mathrm{Na}} + \frac{1}{2C}\left[\frac{\partial k_m}{\partial V}(r + d_m - \tilde{m})^2 + \frac{\partial k_n}{\partial V}(r + d_n - \tilde{n})^2 + \frac{\partial k_h}{\partial V}(r + d_h - \tilde{h})^2\right] = 0, \quad (2\text{-}14)$$

$$-R_{\mathrm{K}}({}_0D_t^{\gamma(t)}q_{\mathrm{K}}) + V - E_{\mathrm{K}} + \frac{1}{2C}\left[\frac{\partial k_m}{\partial V}(r + d_m - \tilde{m})^2 + \frac{\partial k_n}{\partial V}(r + d_n - \tilde{n})^2 + \frac{\partial k_h}{\partial V}(r + d_h - \tilde{h})^2\right] = 0, \quad (2\text{-}15)$$

$$-R_l({}_0D_t^{\gamma(t)}q_l) + V - E_l + \frac{1}{2C}\left[\frac{\partial k_m}{\partial V}(r + d_m - \tilde{m})^2 + \frac{\partial k_n}{\partial V}(r + d_n - \tilde{n})^2 + \frac{\partial k_h}{\partial V}(r + d_h - \tilde{h})^2\right] = 0, \quad (2\text{-}16)$$

where $V = q_C/C$ is the electric potential of the capacitor.

As in [Sherief et al. 2012], we introduce a generalized Kirchhoff's current law of the form

$${}_0D_t^{\gamma(t)}\,(CV + q_{\mathrm{Na}} + q_{\mathrm{K}} + q_l) = I, \quad (2\text{-}17)$$

where $I$ is a known external current applied on the membrane. By replacing Equations (2-14), (2-15), and (2-16) into (2-17) we obtain an equation for the membrane potential given as

$${}_0D_t^{\gamma(t)}(CV) = I - \frac{1}{R_{\mathrm{Na}}}(V - E_{\mathrm{Na}}) - \frac{1}{R_{\mathrm{K}}}(V - E_{\mathrm{K}}) - \frac{1}{R_l}(V - E_l) - \frac{1}{2C}\left(\frac{1}{R_{\mathrm{Na}}} + \frac{1}{R_{\mathrm{K}}} + \frac{1}{R_l}\right)\left[\frac{\partial k_m}{\partial V}(r + d_m - \tilde{m})^2 + \frac{\partial k_n}{\partial V}(r + d_n - \tilde{n})^2 + \frac{\partial k_h}{\partial V}(r + d_h - \tilde{h})^2\right]. \quad (2\text{-}18)$$

The unknown functions $u$, $d_m$, $d_n$, $d_h$, $\tilde{m}$, $\tilde{n}$, $\tilde{h}$, and $V$ can be found by solving the coupled Equations (2-7), (2-8)–(2-13) and (2-18) with appropriate initial conditions. Given the insufficient knowledge of neuronal mechanotransduction processes, it is very difficult to provide expressions for $\tilde{a}$, $\tilde{b}$, $\tilde{c}$, $k$, $k_m$, $k_n$, $k_h$, $\eta_m$, $\eta_n$, $\eta_h$, $C$, $F_m$, $F_n$, and $F_h$ which are needed in order to solve these equations. Therefore, we now make some simplifying assumptions. We start by observing that the equations are not only coupling neuronal mechanical and electrical behaviors but also microscopic (ionic level) and macroscopic (cell level) length scales. This mixture of length scales will cause the system of differential equations to be stiff numerically. Considering how little is known about most of the parameters in our model, a proper analysis of the system and separation of the length scales using perturbation theory is not feasible at this time. Thus we will give a qualitative rather than quantitative analysis of the terms of the equations. We assume that the microscopic masses $\tilde{a}$, $\tilde{b}$, and $\tilde{c}$ are negligible with respect to the cell's mass $M$ and thus we remove Equations (2-8)–(2-10) from the system of equations and also remove $d_m$, $d_n$, and $d_h$ from the set of unknowns. We further assume that the first variations of the spring constants $k_m$, $k_n$, and $k_h$ of the microscopic Maxwell elements with the macroscopic functions $V$ and $u$ are very small and thus we neglect these terms from Equations (2-7) and (2-18). For simplicity, we take $F = 0$. Lastly, we introduce the nondimensional displacements

$$m = \tilde{m}/r, \quad n = \tilde{n}/r, \quad h = \tilde{h}/r. \quad (2\text{-}19)$$

With these assumptions, the system of equations reduces to

$$M\,{}_tD_T^{\alpha(t)}({}_0D_t^{\alpha(t)}u) - \eta\,{}_0D_t^{\alpha(t)}u - ku + \frac{1}{2}\frac{\mathrm{d}C}{\mathrm{d}u}V^2 = 0, \tag{2-20}$$

$${}_0D_t^{\beta(t)}m = -\frac{k_m}{\eta_m}m + \left[\frac{k_m}{\eta_m}(1 + d_m/r) - \frac{1}{2r^2\eta_m}\frac{\partial k}{\partial m}u^2\right], \tag{2-21}$$

$${}_0D_t^{\beta(t)}n = -\frac{k_n}{\eta_n}n + \left[\frac{k_n}{\eta_n}(1 + d_n/r) - \frac{1}{2r^2\eta_n}\frac{\partial k}{\partial n}u^2\right], \tag{2-22}$$

$${}_0D_t^{\beta(t)}h = -\frac{k_h}{\eta_h}h + \left[\frac{k_h}{\eta_h}(1 + d_h/r) - \frac{1}{2r^2\eta_h}\frac{\partial k}{\partial h}u^2\right], \tag{2-23}$$

$${}_0D_t^{\gamma(t)}(CV) = I - \frac{1}{R_{\mathrm{Na}}}(V - E_{\mathrm{Na}}) - \frac{1}{R_{\mathrm{K}}}(V - E_{\mathrm{K}}) - \frac{1}{R_l}(V - E_l). \tag{2-24}$$

We notice now that when $\beta(t) = 1$, Equations (2-21)–(2-23) have the same forms as the classic Hodgkin–Huxley equations for the gating variables $m$, $n$, and $h$ and thus we identify our nondimensional displacements $m$, $n$, and $h$ with the variables representing the activations of the $\mathrm{Na}^+$ and $\mathrm{K}^+$ channels and the inactivation of $\mathrm{Na}^+$ channel, respectively. Consequently, we have provided a physical meaning for $m$,$n$, and $h$ and their evolution equations which was, missing from the original ad hoc derivation of the classic Hodgkin–Huxley [1952] model. Since the physical parameters required by Equations (2-21)–(2-23) are not known we will replace the right-hand sides of these equations and Equation (2-24) by the expressions from the Hodgkin–Huxley model [Dayan and Abbott 2001] and thus obtain the system of equations

$$M\,{}_tD_T^{\alpha(t)}({}_0D_t^{\alpha(t)}u) - \eta\,{}_0D_t^{\alpha(t)}u - ku + \frac{1}{2}\frac{\mathrm{d}C}{\mathrm{d}u}V^2 = 0, \tag{2-25}$$

$${}_0D_t^{\beta(t)}m = \alpha_m(1 - m) - \beta_m m, \tag{2-26}$$

$${}_0D_t^{\beta(t)}n = \alpha_n(1 - n) - \beta_n n, \tag{2-27}$$

$${}_0D_t^{\beta(t)}h = \alpha_h(1 - h) - \beta_h h, \tag{2-28}$$

$${}_0D_t^{\gamma(t)}(CV) = I - g_{\mathrm{Na}}m^3h\tilde{A}(V - E_{\mathrm{Na}}) - g_{\mathrm{K}}n^4\tilde{A}(V - E_{\mathrm{K}}) - g_l\tilde{A}(V - E_l), \tag{2-29}$$

where $\tilde{A}$ is the surface area of the neuron, $g_{\mathrm{Na}}$, $g_{\mathrm{K}}$, and $g_l$ are respectively the maximal conductances of the $\mathrm{Na}^+$, $\mathrm{K}^+$, and leakage currents, and

$$\begin{aligned}
\alpha_m &= \frac{0.1(V+40)}{1-\exp(-0.1(V+40))}, & \beta_m &= 4\exp(-0.0556(V + 65)),\\
\alpha_n &= \frac{0.01(V+55)}{1-\exp(-0.1(V+55))}, & \beta_n &= 0.125\exp(-0.0125(V + 65)),\\
\alpha_h &= 0.07\exp(-0.05(V + 65)), & \beta_h &= \frac{1}{1+\exp(-0.1(V+35))}.
\end{aligned} \tag{2-30}$$

The physical units of the parameters in (2-30) are the same as those of the left-hand sides of the Equations (2-26)–(2-28), since all the constants multiplying potential $V$ are meant to remove its physical units (mV). Lastly, we notice that for a constant $C$, Equations (2-26)–(2-29) reduce to either the equations proposed in [Sherief et al. 2012] when $\beta(t)$ and $\gamma(t)$ are constants between 0 and 1, or the equations of the classic Hodgkin–Huxley model [Dayan and Abbott 2001] when $\beta(t) = \gamma(t) = 1$.

In addition, we need to provide expressions for $C(u)$, $k(m, n, h)$, $\alpha(t)$, $\beta(t)$, and $\gamma(t)$. As in [Drapaca 2015], we assume that the membrane acts like a parallel plate capacitor and introduce

$$C = c_m \tilde{A} = \frac{\epsilon \tilde{A}}{r(1+u/r)} \approx \frac{\epsilon \tilde{A}}{r}(1 - u/r), \tag{2-31}$$

where $c_m$ is the specific membrane capacitance and $\epsilon$ is membrane's permittivity. For the dynamic spring constant $k(m, n, h)$ we propose the expression [Drapaca 2015]

$$k(m, n, h) = k_0(1 + m^3(1 - h)n^4), \tag{2-32}$$

where $k_0$ is the spring constant of the neuron in the inactive state. Equation (2-32) suggests that the cell stiffens during an action potential which appears to be in agreement with the observations made in [Hille 2001; Zou et al. 2013]. Lastly, due to a lack of knowledge of relationships among the multiple time scales on which physical processes in a neuron take place, mathematical simplicity rather than physical inspiration was used to select expressions for $\alpha(t)$, $\beta(t)$, and $\gamma(t)$ and these will be provided in the results section.

The system of Equations (2-25)–(2-29) with parameters given by (2-30), (2-31), and (2-32) can be solved either by direct numerical discretization or, as we will see below, by using higher order expansion formulas that transforms the system into a system of first and second order differential equations which can be solved using existing software such as Matlab. Before proceeding further it is worthwhile to comment on the features of some of these equations. For the sake of argument we take $\beta(t) = \gamma(t) = 1$. Then, Equations (2-26)–(2-29) and (2-30) become the classic Hodgkin–Huxley equations. The amount of combined work, intuition and inspiration needed to obtain these very well tuned equations is obvious from their expressions and this is why we decided to keep these formulas in our model. However, given our interpretation of the gating variables $m$, $n$, and $h$ as nondimensional displacements of the dashpots in the microscopic Maxwell elements that model the ionic gates present in the membrane, we could use the mechanoelectric analogy [Koenig and Blackwell 1961] to provide a macroscopic mechanical description of the membrane corresponding to the Hodgkin–Huxley electric circuit. According to the mechanoelectric analogy, the capacitance $C$ is the average mass of the membrane seen as a multicomponent and multiphasic porous medium, the voltage $V$ is the average velocity of the membrane, the external current $I$ is an external force, the reverse potentials $E_{\mathrm{Na}}$, $E_{\mathrm{K}}$, and $E_l$ are flux velocities of $\mathrm{Na}^+$, $\mathrm{K}^+$, and $\mathrm{Cl}^-$ (leakage), and the inverses of the resistances $1/R_{\mathrm{Na}}$,$1/R_{\mathrm{K}}$, and $1/R_l$ are viscous damping coefficients describing the friction caused by the transport of ions through the membrane. In this analogy, Equation (2-24) (or (2-29)) is Newton's second law of motion where the only internal forces are the damping forces $(1/R_{\mathrm{Na}})(V - E_{\mathrm{Na}})$, $(1/R_{\mathrm{K}})(V - E_{\mathrm{K}})$, and $(1/R_l)(V - E_l)$ corresponding to the ion fluxes expressed relative to the average velocity of the membrane. One advantage of this mechanical analog is that thermodynamics theory can be used to relate the flux velocities of the considered ionic species to their concentrations via their chemical potentials [Doi and Edwards 1986] and thus recover the Nernst equations, which are commonly used to express the dependency of the reverse potentials on ion concentrations (see for instance [Wei et al. 2014]). Another possible advantage is that as experimentally supported information on the chemomechanical properties of the membrane's components becomes available, the validation of a simpler mechanical model with fewer parameters may become possible. In this context we notice that our model is not only independent of the Hodgkin–Huxley equations but also could be made to have fewer parameters thanks to the use of fractional order derivatives of variable order. For instance, according to the Hodgkin–Huxley model, the damping coefficients $1/R_{\mathrm{Na}}$, $1/R_{\mathrm{K}}$, and $1/R_l$ in the mechanical analog model depend on $m$, $n$, and $h$ and thus vary in time according to (2-26)–(2-29). This means that the mechanical analog model implements already variable viscoelasticity. Alternatively, using a fractional order derivative of variable order $\gamma(V(t)) \in [0, 1)$ in Equation (2-24) to model variable viscoelasticity could drastically simplify Equations (2-26)–(2-28) and (2-30) and reduce the number of parameters needed to be found experimentally. Thus the work presented in this paper can be seen as a first step towards a simpler chemomechanical model of a neuron and its membrane.

We propose to simplify the system of Equations (2-25)–(2-29) even more by using (1-5) and the expansion formulas (written here for a generic function $f$) [Almeida and Torres 2013]

$$\begin{aligned}{}_0D_t^{\alpha(t)} f(t) \approx{}& \frac{1}{\Gamma(1-\alpha(t))}\left[1+\sum_{p=2}^{N}\frac{\Gamma(p-1+\alpha(t))}{\Gamma(\alpha(t))(p-1)!}\right]t^{-\alpha(t)} f(t)\\ &+\frac{1}{\Gamma(2-\alpha(t))}\left[1+\sum_{p=1}^{N}\frac{\Gamma(p-1+\alpha(t))}{\Gamma(\alpha(t)-1)p!}\right]t^{1-\alpha(t)}\frac{\mathrm{d}f}{\mathrm{d}t}(t)\\ &+\sum_{k=2}^{N}\frac{\Gamma(k-1+\alpha(t))}{\Gamma(-\alpha(t))\Gamma(1+\alpha(t))(k-1)!}t^{1-k-\alpha(t)}F_k(t),\end{aligned}\tag{2-33}$$

$$\frac{\mathrm{d}F_k}{\mathrm{d}t}(t) = (k-1)t^{k-2} f(t), \quad F_k(0) = 0, \quad k = 2, 3, \dots, N,\tag{2-34}$$

$$\begin{aligned}{}_tD_T^{\alpha(t)} f(t) \approx{}& \frac{1}{\Gamma(1-\alpha(t))}\left[1+\sum_{p=2}^{N}\frac{\Gamma(p-1+\alpha(t))}{\Gamma(\alpha(t))(p-1)!}\right](T-t)^{-\alpha(t)} f(t)\\ &-\frac{1}{\Gamma(2-\alpha(t))}\left[1+\sum_{p=1}^{N}\frac{\Gamma(p-1+\alpha(t))}{\Gamma(\alpha(t)-1)p!}\right](T-t)^{1-\alpha(t)}\frac{\mathrm{d}f}{\mathrm{d}t}(t)\\ &+\sum_{k=2}^{N}\frac{\Gamma(k-1+\alpha(t))}{\Gamma(-\alpha(t))\Gamma(1+\alpha(t))(k-1)!}(T-t)^{1-k-\alpha(t)}G_k(t),\end{aligned}\tag{2-35}$$

$$\frac{\mathrm{d}G_k}{\mathrm{d}t}(t) = (1-k)(T-t)^{k-2} f(t), \quad G_k(T) = 0, \quad k = 2, 3, \dots, N.\tag{2-36}$$

By replacing formula (2-33) into formula (2-35) the expression

$$\begin{aligned}{}_tD_T^{\alpha(t)}({}_0D_t^{\alpha(t)} f(t)) \approx{}& -\frac{1}{\Gamma(2-\alpha(t))^2}\left[1+\sum_{p=1}^{N}\frac{\Gamma(p-1+\alpha(t))}{\Gamma(\alpha(t)-1)p!}\right]^2 (t(T-t))^{1-\alpha(t)}\frac{\mathrm{d}^2 f}{\mathrm{d}t^2}\\ &+\text{lower order terms}\end{aligned}\tag{2-37}$$

is obtained, where some of the lower order terms are linear in $f$ and $\mathrm{d}f/\mathrm{d}t$ and the rest of the terms are combinations of the extra functions $F_k$, $G_k$, $k = 2, 3, \dots N$. Formulas (2-33)–(2-36) suggest that the extra terms in (2-37) introduce additional memory effects which could invalidate the causality law.

Thus in Equation (2-20) we take ${}_tD_T^{\alpha(t)}({}_0D_t^{\alpha(t)}u) = -(\mathrm{d}^2u)/(\mathrm{d}t^2)$ and obtain a much simplified form of the equation of motion for displacement $u$. If we replace formula (2-33) into Equations (2-20)–(2-24) and add the extra $N-1$ equations (2-34) corresponding to each unknown function $u$, $m$, $n$, $h$, $V$ we obtain a system of integer order differential equations with $5N$ equations. It was shown in [Almeida and Torres 2013] that a value of $N = 3$ gives very accurate results, so we take $N = 3$ and obtain a total of 15 equations: 14 of them are first order and one is a second order differential equation. We notice that these mathematical approximations replace the variable fractional order time derivatives by integer (first or second) order derivatives and the effect of the variable orders is contained in extra memory terms (represented as power functions of time with variable fractional order) which are added to the proposed equations and their corresponding evolution equations.

## 3. Results

In our numerical simulations we used the following parameters [Dayan and Abbott 2001]:

$$\begin{aligned} E_{\mathrm{Na}} &= 50\,\mathrm{mV}, \\ E_{\mathrm{K}} &= -77\,\mathrm{mV}, \\ E_l &= -54.387\,\mathrm{mV}, \\ g_{\mathrm{Na}} &= 1.2\,\mathrm{mS/mm^2}, \\ g_{\mathrm{K}} &= 0.36\,\mathrm{mS/mm^2}, \\ g_l &= 0.003\,\mathrm{mS/mm^2}, \\ r &= 4\,\mathrm{nm}, \\ r_0 &= 2\,\mu\mathrm{m}, \\ \tilde{A} &= 0.01\,\mathrm{mm^2}, \end{aligned} \tag{3-1}$$

where $r_0$ is the radius of the neuron. At mechanical equilibrium ($u = 0$), the specific membrane capacitance is $0.01\,\mu\mathrm{F/mm^2}$, which combined with formula (2-31) gives

$$c_m = 0.01(1 - u/r)\,\mu\mathrm{F/mm^2}.$$

We also used $E_0 = 200\,\mathrm{Pa}$ as an average Young's modulus of a neuron [Lu et al. 2006; Zou et al. 2013], $M = 0.1\,\mathrm{ng}$ as half of the neuronal mass [Corbin et al. 2014], and $\mu = 4\,\mathrm{mPa}\cdot\mathrm{s}$ as neuronal dynamic viscosity [Park et al. 2010]. From these parameters and the assumption of circular cylindrical shape, the following parameters can be calculated [Drapaca 2015]:

$$k_0 = 0.0013\,\mathrm{mg/ms^2}, \quad \eta = 2.5\cdot 10^{-11}\,\mathrm{mg/ms}, \quad \mathrm{vol}_0 = 9.95\cdot 10^{-6}\,\mathrm{mm^3},$$

where $\mathrm{vol}_0$ is the volume of the neuron at mechanical equilibrium. In all numerical simulations we applied a constant external current per unit surface area $I = 0.1\,\mu\mathrm{A/mm^2}$. Lastly, we used a characteristic time of 25 ms, and $\alpha(t) = 0.001\exp(1 - t/25) \in \mathscr{C}$, $\beta(t) = \gamma(t) = 1$. We chose a simple function $\alpha(t)$ that belongs to the class $\mathscr{C}$, looks similar to the parameters from formulas (2-30), and gives apparently reasonable results when the neuronal electrochemistry is described by the classic Hodgkin–Huxley model ($\beta(t) = \gamma(t) = 1$).

In this case the system of Equations (2-20)–(2-24) with approximations (2-33)–(2-34) become

$$\frac{\mathrm{d}}{\mathrm{d}t}u = v, \tag{3-2}$$

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}t}v = &-\left[\frac{k}{M} + \frac{\eta}{M}\frac{1}{\Gamma(1-\alpha(t))}\left(1+\sum_{p=2}^{3}\frac{\Gamma(p-1+\alpha(t))}{\Gamma(\alpha(t))(p-1)!}\right)t^{-\alpha(t)}\right]u \\ &- \frac{\eta}{M}\frac{1}{\Gamma(2-\alpha(t))}\left(1+\sum_{p=1}^{3}\frac{\Gamma(p-1+\alpha(t))}{\Gamma(\alpha(t)-1)p!}\right)t^{1-\alpha(t)}v \\ &- \frac{\eta}{M}\sum_{k=2}^{3}\frac{\Gamma(k-1+\alpha(t))}{\Gamma(-\alpha(t))\Gamma(1+\alpha(t))(k-1)!}t^{1-k-\alpha(t)}F_k + \frac{1}{2}\frac{\mathrm{d}C}{\mathrm{d}u}V^2,\end{aligned} \tag{3-3}$$

$$\frac{\mathrm{d}}{\mathrm{d}t}F_k = (k-1)t^{k-2}u, \quad k = 2, 3, \tag{3-4}$$

$$\frac{\mathrm{d}}{\mathrm{d}t}m = \alpha_m(1-m) - \beta_m m, \tag{3-5}$$

$$\frac{\mathrm{d}}{\mathrm{d}t}n = \alpha_n(1-n) - \beta_n n, \tag{3-6}$$

$$\frac{\mathrm{d}}{\mathrm{d}t}h = \alpha_h(1-h) - \beta_h h, \tag{3-7}$$

$$\frac{\mathrm{d}}{\mathrm{d}t}(V) = \frac{1}{C}\big[I - g_{\mathrm{Na}}m^3h\tilde{A}(V-E_{\mathrm{Na}}) - g_{\mathrm{K}}n^4\tilde{A}(V-E_{\mathrm{K}}) - g_l\tilde{A}(V-E_l)\big] - \frac{1}{C}v\frac{\mathrm{d}C}{\mathrm{d}u}V. \tag{3-8}$$

We solved the system (3-2)–(3-8) with the initial conditions

$$V(0) = -65\,\mathrm{mV}, \qquad m(0) = \frac{\alpha_m(V(0))}{\alpha_m(V(0))+\beta_m(V(0))}, \tag{3-9}$$

$$n(0) = \frac{\alpha_n(V(0))}{\alpha_n(V(0))+\beta_n(V(0))}, \qquad h(0) = \frac{\alpha_h(V(0))}{\alpha_h(V(0))+\beta_h(V(0))}, \tag{3-10}$$

$$F_k(0) = 0, \quad k = 2, 3 \tag{3-11}$$

$$\text{Set 1: } u(0) = 1\,\mathrm{nm}, \quad v(0) = 10\,\mathrm{nm/ms},$$

$$\text{Set 2: } u(0) = 0, \quad v(0) = 1\,\mathrm{nm}/\mu\mathrm{s},$$

Some numerical experimentation with physically plausible values for $u(0)$ and $v(0)$ showed the existence of two trends in the behavior of stable solutions and thus the initial conditions given by sets 1 and 2 were chosen such that both of these situations could be presented.

We used Matlab's built-in function ode15s that solves stiff ordinary differential equations using

(1) a modified linear multistep backward difference formula of order up to 5 known to have good stability, and

(2) an adaptive step size that changes according to a numerical scheme that calculates relative and absolute error tolerances [Shampine and Reichelt 1997].

In our simulations we kept the default values of ode15s for the relative error tolerance ($10^{-3}$) and for the absolute error tolerance ($10^{-6}$). For the chosen parameters and initial conditions we noticed that the order of magnitude of $\mathrm{d}C/\mathrm{d}uV$ in Equations (3-3) and (3-8) is much bigger than the rest of the terms in

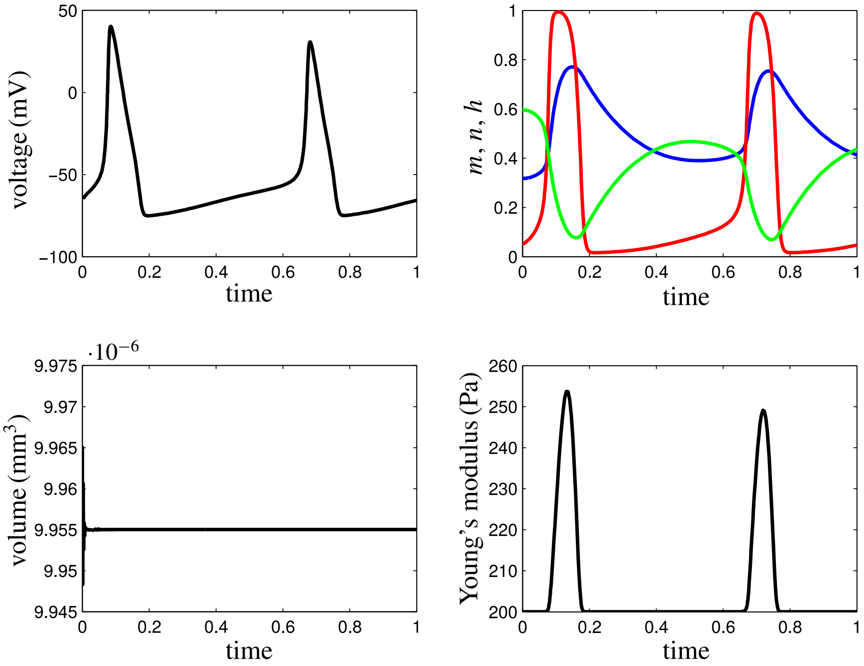

**Figure 2.** Results for initial displacement of 1 nm and initial speed of 10 nm/ms: voltage; functions $n$ (blue curve), $m$ (red), and $h$ (green); volume; and Young's modulus.

these equations and no action potentials are observed. Thus, as in [Drapaca 2015], we remove this term from Equations (3-3) and (3-8). This simplification preserves a weaker coupling between the mechanical and electrical behaviors of the neuron expressed by (2-31) and (2-32).

In Figures 2 and 3 we show the evolutions of the voltage, gating variables, volume and Young's modulus for the mechanical initial conditions in set 1 and set 2, respectively. For set 1, the action potentials occur and the Young's modulus and the volume variations appear to be physically admissible and possibly within a healthy range. The dynamics of the cell's stiffness (Figure 2, bottom right) seems in agreement with the experimental observations in the normally functioning regime reported in [Zou et al. 2013]. The initial conditions in set 2 mimic a more serious traumatic event and our simulations show that there are no action potentials (Figure 3, top left), and the displacements of the gates $m$, $n$, and $h$ (top right) as well as the Young's modulus (bottom right) remain at their corresponding initial values. The solutions obtained in this case might show damaging effects of a very fast initial speed (jabbing) on the material structure and electrochemical activity of a neuron. To better understand the simulated neuronal mechanotransduction, in Figures 4 and 5 we look closer at the voltages and corresponding volumes obtained using the mechanical initial conditions in set 1 and set 2. While oscillations in the cell's volume are quickly attenuated for set 1 of initial conditions (Figure 4, right) such that the action potential can develop soon afterwards (Figure 4, left); for the initial conditions of set 2 the amplitudes of the oscillations in volume are much higher than in the previous case and do not appear to diminish in time (Figure 5, right), and thus the membrane's depolarization does not happen. These results look similar to the ones we reported in [Drapaca 2015], even though here Equation (3-3) has extra, time-dependent terms

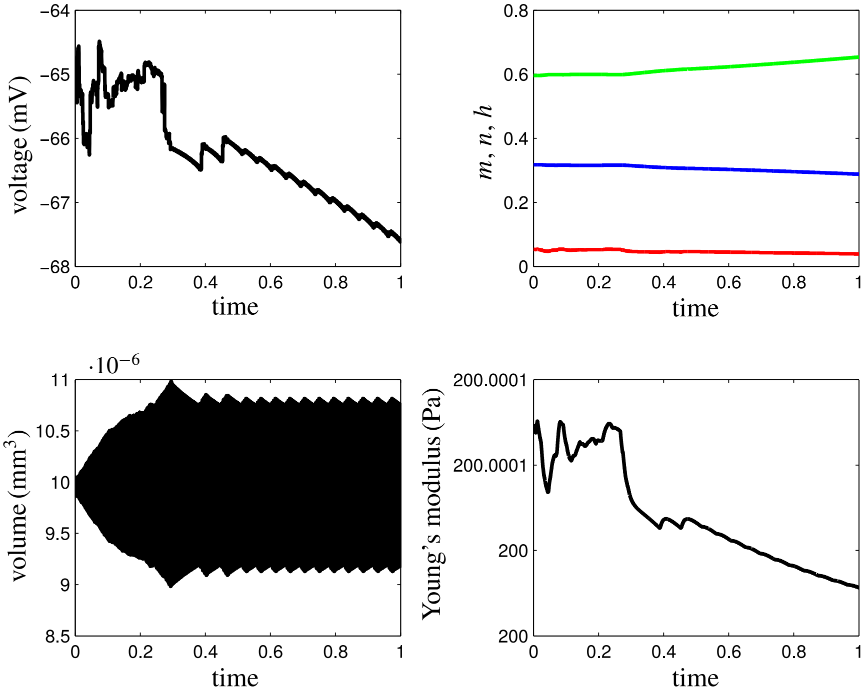

**Figure 3.** Results for zero initial displacement and initial speed of speed of 10 nm/ms: voltage; functions $n$ (blue curve), $m$ (red), and $h$ (green); volume; and Young's modulus.

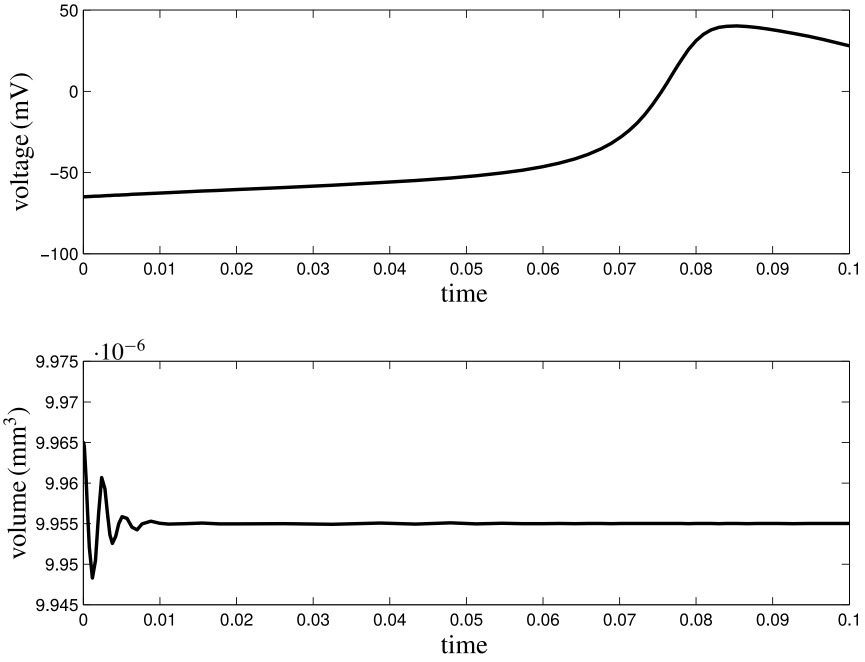

**Figure 4.** A zoom-in of Figure 2: voltage potential (top) and volume (bottom).

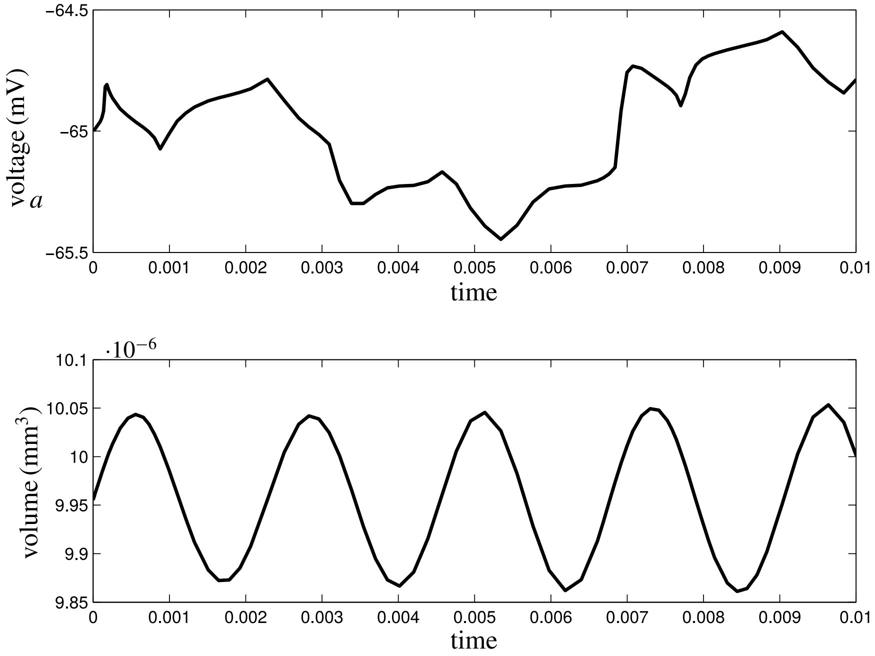

**Figure 5.** A zoom-in of Figure 3: voltage potential (top) and volume (bottom).

that account for the variable viscoelasticity. This might suggest that — for the chosen initial conditions, parameters and electromechanical couplings — after a serious traumatic event the elastic behavior of a neuron dominates over its viscoelastic response, which agrees with the experimentally supported claims made in [Grevesse et al. 2015].

## 4. Conclusions and future directions

In this paper we proposed a generalization of our electromechanical model from [Drapaca 2015] in which the temporal variations of the physical fields are represented using Riemann–Liouville fractional derivatives of variable orders. The neuron was modeled as a linear viscoelastic Kelvin–Voigt solid with variable viscoelasticity whose electrochemical activity was described by fractional order Hodgkin–Huxley equations. In addition, we introduced three linear viscoelastic Maxwell fluids with variable viscoelasticity to model the three ionic gates with gating variables $m$,$n$, and $h$. This provides a *physical structure for the ionic gates which can be interpreted as door closers*. The fractional orders of the derivatives model the multiple temporal scales used by a neuron for electrochemical signaling. The time dependency of the fractional orders models the biological variability of neurons as well as the intrinsic entanglement of states existing in the complex mixture that makes up a neuron. For this reason we call these time scales *entangled scales*. We used a Lagrangian formulation and Hamilton's principle to obtain the coupled equations of motion. This approach links macroscopic (cell level) and microscopic (ionic level) mechanical and electrical information and hence it can describe neuronal mechanotransduction [Drapaca 2015]. As in [Drapaca 2015], we assumed that at the macroscopic level the membrane's capacitance depends on the mechanical displacement of the neuron and the Young's modulus of the neuron depends on the nondimensional displacements of the dashpots in the Maxwell elements. In order to satisfy the physical criterion of causality, we restrict the class of allowable variable order functions to the class of differentiable functions with continuous first order derivatives and of slow variation. For such variable order functions, we can approximate the Riemann–Liouville derivatives to Marchaud fractional derivatives and further use the higher integer order expansions proposed in [Almeida and Torres 2013] to transform the integro-differential Euler–Lagrange equations into classic Euler–Lagrange equations with extra time-dependent terms, some with their own evolution equations. We further performed numerical simulations in Matlab using the built-in function ode15s to solve simplified versions of our differential equations. We used the same initial conditions and physical constants as in [Drapaca 2015] and we obtained results comparable with those in the same paper. Namely, when a constant external electric current was applied and the initial displacement and speed were of orders of magnitude comparable to the size of the membrane, the action potentials occurred and looked similar to the ones seen in healthy neurons. In this case the dynamics of the neuron's stiffness seemed to agree with experimental measurements done on healthy neurons [Zou et al. 2013]. At very fast initial speeds (which could model a serious traumatic event) and in the presence of a constant applied external current, our numerical simulations showed high persisting oscillations in the volume of the neuron and the action potentials did not happen. Also, the Young's modulus of the neuron and the displacements $m$, $n$, and $h$ of the dashpots in the Maxwell elements were almost constant, suggesting possible structural and functional damage of the neuron.

Some of the limitations of the proposed model and solution approach are as follows. One limitation is finding physically valid capacitance-displacement and, respectively, stiffness-gating variable relationships because there are no experimental observations that could guide us. Given the simplicity of the proposed model, we hope that our model will inspire future experimental work that will provide empirical relationships among the model's mechanical and electrical parameters. Another limitation of our approach is the use of the Matlab built-in function ode15s to solve the simplified system of stiff differential equations. Shampine and Bogacki [1989] advised caution in drastically reducing the step size in the discretization implemented in ode15s since this action could increase numerical error and cause instabilities in the solutions. In addition, the class of admissible variable order functions is too restrictive. In our future work we plan to develop our own numerical solver for the stiff system of differential equations that uses the Riemann–Liouville fractional derivatives of variable orders. We believe that for living neurons the causality law should not be imposed and therefore by removing this restriction on our model we could use the Riemann–Liouville derivatives with variable order functions belonging to a larger class of admissible (continuous) functions. A numerical solver will allow us to address another limitation of our model: providing a more physically meaningful expression for the variable fractional order $\alpha(t)$. One approach is to use the experimentally found dependence of a power law exponent on the stress reported in [Grevesse et al. 2015] to calculate a corresponding $\alpha(t)$. Another approach is to add the unknown function $\alpha(t)$ to the list of generalized coordinates and use Hamilton's principle to obtain one more Euler–Lagrange equation that presumably could be solved for $\alpha$ [Atanackovic and Pilipovic 2011]. Another possible limitation is the use in our numerical simulations of the classic Hodgkin–Huxley model, whose very well tuned equations and parameters might have hidden the possible effects of the variable viscoelasticity modeled with fractional order time derivatives of variable order. Since the proposed model is not dependent on the Hodgkin–Huxley equations, we could use the mechanoelectric analogy mentioned in Section 2 to obtain a macroscopic mechanical analog model of the neuronal membrane made of the membrane's mass and three dashpots connected in parallel and located between the neuronal mass $M$ and the membrane's mass. The corresponding equations of motion with classic time derivatives as well as with fractional order derivatives of variable orders will not only have much simpler forms but also may be easier to analyze and solve. In particular, a bifurcation analysis might be possible in this case and the role of variable fractional order derivatives in critical transitions could be investigated.

Lastly, we intend to generalize our model by including spatial variations. Let $\Omega \subset \mathbb{R}^3$ be the domain occupied by the intracellular space, the cell's membrane, and a very small $\epsilon$-neighborhood of the membrane located in the extracellular space that contains the ions exchanged with the neuron through its membrane. This domain is filled with a very complex mixture of solids, fluids, ionic components and other proteins and molecules that interact dynamically with each other on multiple time and space scales and thus there exists a subset $\omega \subset \Omega$ that contains the entangled state. This region $\omega$ could be seen as a sort of black hole in the sense that once components enter $\omega$, any information about them is lost (however, unlike black holes, components can "escape" $\omega$). Thus the energies contributing to the Lagrangian form are the kinetic and potential energies as well as an interfacial energy due to the presence of $\omega$. All these energies are in fact part of the total entropy of domain $\Omega$, and the entanglement entropy, which is the entropy of $\omega$, satisfies a so-called *area law* that says that the entanglement entropy is proportional to the surface area of $\omega$ [Eisert et al. 2010]. If we introduce the level set function

$$\phi(\mathbf{x}) = \begin{cases} d_x & \text{if } \mathbf{x} \in \Omega - \omega, \\ -d_x & \text{if } \mathbf{x} \in \omega, \\ 0 & \text{if } \mathbf{x} \in \partial\omega, \end{cases}$$

then we can introduce the Lagrangian

$$\begin{aligned} \mathscr{L} = \int_\Omega \tfrac{1}{2}\rho_M ({}_0D_t^{\alpha(t,\boldsymbol{x})} u)^2 H(\phi) &+ \tfrac{1}{2}\rho_{\tilde{a}} ({}_0D_t^{\beta(t,\boldsymbol{x})} (r + d_m))^2 H(\phi) \\ &+ \tfrac{1}{2}\rho_{\tilde{b}} ({}_0D_t^{\beta(t,\boldsymbol{x})} (r + d_n))^2 H(\phi) + \tfrac{1}{2}\rho_{\tilde{c}} ({}_0D_t^{\beta(t,\boldsymbol{x})} (r + d_h))^2 H(\phi) \\ &- \frac{1}{2C(u)} q_C^2 H(\phi) - \sigma_u \cdot \epsilon_u H(\phi) - \sigma_{dm} \cdot \epsilon_{dm} H(\phi) \\ &- \sigma_{dn} \cdot \epsilon_{dn} H(\phi) - \sigma_{dh} \cdot \epsilon_{dh} H(\phi) + \lambda \delta(\phi) ||\nabla\phi|| \, \mathrm{d}\boldsymbol{x}, \end{aligned} \tag{4-1}$$

where $H(\phi)$ is the Heaviside function; $\delta(\phi)$ is the Dirac distribution; $\rho_M$, $\rho_{\tilde{a}}$, $\rho_{\tilde{b}}$, and $\rho_{\tilde{c}}$ are the densities of the masses represented as subindices; $\sigma_u, \sigma_{dm}, \sigma_{dn}, \sigma_{dh}$ are the stress tensors for the displacements $u, r + d_m - \tilde{m}, r + d_n - \tilde{n}, r + d_h - \tilde{h}$ with $\epsilon_u, \epsilon_{dm}, \epsilon_{dn}$, and $\epsilon_{dh}$ the corresponding strain tensors. We denoted by $||\nabla\phi||$ the Euclidean norm of the Jacobian matrix of $\phi$. The last term in formula (4-1) is the surface area of $\omega$. An expression for the corresponding virtual work can be obtained from formula (2-4) in a similar manner. Hamilton's principle will provide again the system of nonlinear integro-differential Euler–Lagrange equations, which will have to be completed with the Kirchhoff current law and an evolution equation for the level set function $\phi$. Appropriate initial and boundary conditions will complete this model which we intend to fully describe and study in the near future.

## List of Symbols

| | |
|---|---|
| $A$ | neuronal cross-sectional area |
| $\tilde{A}$ | neuronal surface area |
| $r_0$ | neuronal radius |
| $r$ | thickness of neuronal membrane |
| $t$ | time (independent variable) |
| $\alpha, \beta, \gamma \in [0, 1)$ | time-dependent fractional orders (entangled time scales) |
| $M$ | half of the neuronal mass |
| $\tilde{a}, \tilde{b}, \tilde{c}$ | masses attached to the Maxwell elements |
| $u$ | neuronal displacement |
| $r+d_m, r+d_n, r+d_h$ | relative displacements between $M$ and each of $\tilde{a}, \tilde{b}, \tilde{c}$ |
| $\tilde{m}, \tilde{n}, \tilde{h}$ | displacements of the dashpots in the Maxwell elements |
| $k$ | spring constant of the Kelvin–Voigt element |
| $\eta$ | damping coefficient of the Kelvin–Voigt element |
| $F$ | force (work conjugate) for the Kelvin–Voigt element |
| $k_m, k_n, k_h$ | spring constants of the Maxwell elements |
| $\eta_m, \eta_n, \eta_h$ | damping coefficients of the Maxwell elements |
| $F_m, F_n, F_h$ | forces (work conjugates) for the Maxwell elements |
| $C$ | capacitance of neuronal membrane |
| $c_m$ | specific membrane capacitance |
| $q_C$ | electric charge of the capacitor |
| $V$ | electric potential (voltage) of the capacitor |
| $\epsilon$ | membrane's permittivity |
| $I$ | external current |
| $q_{\mathrm{Na}}, q_{\mathrm{K}}, q_l$ | electric charges of $\mathrm{Na}^+$,$\mathrm{K}^+$, and leakage channels |
| $R_{\mathrm{Na}}, R_{\mathrm{K}}, R_l$ | resistances of the resistors |
| $E_{\mathrm{Na}}, E_{\mathrm{K}}, E_l$ | reverse potentials |
| $g_{\mathrm{Na}}, g_{\mathrm{K}}, g_l$ | maximal conductances of the $\mathrm{Na}^+$, $\mathrm{K}^+$, and leakage currents |
| $\alpha_m, \alpha_n, \alpha_h, \beta_m, \beta_n, \beta_h$ | voltage-dependent parameters of the Hodgkin–Huxley model |
| $F_k, G_k, k = 2, 3$ | extra memory terms due to mathematical approximations |

## References

[Almeida and Torres 2013] R. Almeida and D. F. M. Torres, "An expansion formula with higher-order derivatives for fractional operators of variable order", *Scientific World J.* (2013).

[Atanackovic and Pilipovic 2011] T. M. Atanackovic and S. Pilipovic, "Hamilton's principle with variable order fractional derivatives", *Fract. Calc. Appl. Anal.* **14**:1 (2011), 94–109.

[Baleanu et al. 2012] D. Baleanu, K. Diethelm, E. Scalas, and J. J. Trujillo, *Fractional calculus: models and numerical methods*, Series on Complexity, Nonlinearity and Chaos **3**, World Scientific, Hackensack, NJ, 2012.

[Bell et al. 2015] J. M. Bell, C. A. Taylor, and M. J. Breiding, "The public health approach to TBI", *J. Head Trauma Rehabil.* **30**:3 (2015).

[CDC 2016] "TBI: get the facts", Centers for Disease Control and Prevention, 2016, http://www.cdc.gov/traumaticbraininjury/get_the_facts.html.

[Chicone and Mashhoon 2002] C. Chicone and B. Mashhoon, "Acceleration-induced nonlocality: kinetic memory versus dynamic memory", *Ann. Phys.* **11**:4 (2002), 309–332.

[Coimbra 2003] C. F. M. Coimbra, "Mechanics with variable-order differential operators", *Ann. Phys.* (8) **12**:11-12 (2003), 692–703.

[Coimbra and Kobayashi 2002] C. F. M. Coimbra and M. H. Kobayashi, "On the viscous motion of a small particle in a rotating cylinder", *J. Fluid Mech.* **469** (2002), 257–286.

[Corbin et al. 2014] E. A. Corbin, L. J. Millet, K. R. Keller, W. P. King, and R. Bashir, "Measuring physical properties of neuronal and glial cells with resonant microsensors", *Anal. Chem.* **86**:10 (2014), 4864–4872.

[Dayan and Abbott 2001] P. Dayan and L. F. Abbott, *Theoretical neuroscience: computational and mathematical modeling of neural systems*, MIT Press, Cambridge, MA, 2001.

[Doi and Edwards 1986] M. Doi and S. Edwards, *The theory of polymer dynamics*, Oxford University Press, Oxford, 1986.

[Drapaca 2015] C. S. Drapaca, "An electromechanical model of neuronal dynamics using Hamilton's principle", *Front. Cell. Neurosci.* **9**:271 (2015).

[Eisert et al. 2010] J. Eisert, M. Cramer, and M. B. Plenio, "Colloquium: area laws for the entanglement entropy", *Rev. Modern Phys.* **82**:1 (2010), 277–306.

[Galley et al. 2014] C. R. Galley, D. Tsang, and L. C. Stein, "The principle of stationary nonconservative action for classical mechanics and field theories", 2014. arXiv

[Goldsmith 2001] W. Goldsmith, "The state of head injury biomechanics: past, present, and future: part 1", *Crit Rev Biomed Eng* **29**:5-6 (2001), 441–600.

[Goriely et al. 2015] A. Goriely, M. G. Geers, G. A. Holzapfel, J. Jayamohan, A. Jerusalem, S. Sivaloganathan, W. Squier, J. A. van Dommelen, S. Waters, and E. Kuhl, "Mechanics of the brain: perspectives, challenges, and opportunities", *Biomech. Model. Mechanobiol* **14**:5 (2015), 931–965.

[Grevesse et al. 2015] T. Grevesse, B. E. Dabiri, K. K. Parker, and S. Gabriele, "Opposite rheological properties of neuronal microcompartments predict axonal vulnerability in brain injury", *Scientific Reports* **5** (2015), 9475.

[Hilfer 2000] R. Hilfer, *Applications of fractional calculus in physics*, World Scientific, River Edge, NJ, 2000.

[Hille 2001] B. Hille, *Ion channels of excitable membranes*, Sinauer Associates, Sunderland, MA, 2001.

[Hodgkin and Huxley 1952] A. L. Hodgkin and A. F. Huxley, "A quantitative description of membrane current and its application to conduction and excitation in nerve", *J. Physiol.* (*Lond.*) **117**:4 (1952), 500–544.

[Koenig and Blackwell 1961] H. E. Koenig and W. A. Blackwell, *Electromechanical system theory*, McGraw-Hill, New York, 1961.

[Lorenzo and Hartley 2002] C. F. Lorenzo and T. T. Hartley, "Variable order and distributed order fractional operators", *Nonlinear Dynam.* **29**:1-4 (2002), 57–98.

[Lu et al. 2006] Y. B. Lu, K. Franze, G. Seifert, C. Steinhauser, F. Kirchhoff, H. Wolburg, J. Guck, P. Janmey, E. Q. Wei, J. Kas, and A. Reichenbach, "Viscoelastic properties of individual glial cells and neurons in the CNS", *Proc. Natl. Acad. Sci. U.S.A.* **103**:47 (2006), 17759–17764.

[Lundstrom et al. 2008] B. N. Lundstrom, M. H. Higgs, W. J. Spain, and A. L. Fairhall, "Fractional differentiation by neocortical pyramidal neurons", *Nat. Neurosci.* **11**:11 (2008), 1335–1342.

[Milici and Draganescu 2015] C. Milici and G. Draganescu, *New methods and problems in fractional calculus*, LAP LAMBERT Academic Publishing, 2015.

[Oldham and Spanier 2006] K. B. Oldham and J. Spanier, *The fractional calculus: theory and applications of differentiation and integration to arbitrary order*, Dover, Mineola, NY, 2006.

[Park et al. 2010] K. Park, L. J. Millet, N. Kim, H. Li, X. Jin, G. Popescu, N. R. Aluru, K. J. Hsia, and R. Bashir, "Measurement of adherent cell mass and growth", *Proc. Natl. Acad. Sci. U.S.A.* **107**:48 (2010), 20691–20696.

[Podlubny 1999] I. Podlubny, *Fractional differential equations*, Mathematics in Science and Engineering **198**, Academic Press, San Diego, 1999.

[Samko et al. 1993] S. G. Samko, A. A. Kilbas, and O. I. Marichev, *Fractional integrals and derivatives*, Gordon and Breach, Yverdon, Switzerland, 1993.

[Schiff 2012] S. J. Schiff, *Neural control engineering: the emerging intersection between control theory and neuroscience*, MIT Press, Cambridge, MA, 2012.

[Shampine and Bogacki 1989] L. F. Shampine and P. Bogacki, "The effect of changing the stepsize in linear multistep codes", *SIAM J. Sci. Statist. Comput.* **10**:5 (1989), 1010–1023.

[Shampine and Reichelt 1997] L. F. Shampine and M. W. Reichelt, "The MATLAB ODE suite", *SIAM J. Sci. Comput.* **18**:1 (1997), 1–22.

[Sherief et al. 2012] H. H. Sherief, A. M. A. El-Sayed, S. H. Behiry, and W. E. Raslan, "Using fractional derivatives to generalize the Hodgkin–Huxley model", pp. 275–282 in *Fractional dynamics and control*, edited by D. Baleanu et al., Springer, New York, 2012.

[Soon et al. 2005] C. Soon, C. Coimbra, and M. Kobayashi, "The variable viscoelasticity oscillator", *Ann. Phys.* (8) **14**:6 (2005), 378–389.

[Sun et al. 2012] H. Sun, W. Chen, C. Li, and Y. Chen, "Finite difference schemes for variable-order time fractional diffusion equation", *Internat. J. Bifur. Chaos Appl. Sci. Engrg.* **22**:4 (2012), 1250085, 16.

[Tarasov 2010] V. E. Tarasov, *Fractional dynamics: applications of fractional calculus to dynamics of particles, fields and media*, Springer, Heidelberg, 2010.

[Wei et al. 2014] Y. Wei, G. Ullah, and S. J. Schiff, "Unification of neuronal spikes, seizures, and spreading depression", *J. Neurosci.* **34**:35 (2014), 11733–11743.

[West 2015] B. J. West, *Fractional calculus view of complexity: tomorrow's science*, CRC Press, Boca Raton, FL, 2015.

[Zou et al. 2013] S. Zou, R. Chisholm, J. Tauskela, G. Mealing, L. Johnston, and C. Morris, "Force spectroscopy measurements show that cortical neurons exposed to excitotoxic agonists stiffen before showing evidence of bleb damage", *PLoS ONE* **8**:8 (2013), e73499.

Received 28 Feb 2016. Revised 26 May 2016. Accepted 30 May 2016.

CORINA S. DRAPACA: `csd12@psu.edu`
*Department of Engineering Science and Mechanics, Pennsylvania State University, University Park, PA 16802, United States*

JOURNAL OF MECHANICS OF MATERIALS AND STRUCTURES
Vol. 12, No. 1, 2017
dx.doi.org/10.2140/jomms.2017.12.57

# TIME-ADAPTIVE FINITE ELEMENT SIMULATIONS OF DYNAMICAL PROBLEMS FOR TEMPERATURE-DEPENDENT MATERIALS

MATTHIAS GRAFENHORST, JOACHIM RANG AND STEFAN HARTMANN

Dynamical systems in finite elements yield systems of second-order differential equations. Incorporating inelastic material properties, thermomechanical coupling and particular Dirichlet boundary conditions essentially changes the underlying mathematical problem. In this respect, we investigate the behavior of a number of subproblems such as reaction force computation, high-order time-integration, time-adaptivity, etc., which yield (depending on the underlying problem) systems of differential-algebraic equations or a mixture of systems of second-order and first-order ordinary differential equations (especially if the constitutive equations are of evolutionary-type, as in the case of viscoelasticity and viscoplasticity). The main goals are to provide higher-order time integration schemes using diagonally implicit Runge–Kutta methods and the generalized-$\alpha$ method so that they may be applied to the constitutive equations, and to apply time-adaptivity via embedded schemes so that step-sizes are chosen automatically. The constitutive equations are given by a thermoviscoplasticity model of Perzyna/Chaboche-type with nonlinear kinematic hardening.

## 1. Introduction

The modeling of thermomechanically coupled material properties is an important and challenging task for both constitutive modeling and for numerical treatment. In the case of polymers, small temperature variations essentially change the mechanical response. Cyclic processes induce heat and must be considered under such circumstances. Other applications are metal forming processes either via heat treatment or by recurrent processes.

The numerical treatment of thermomechanically coupled problems using finite elements is connected to the historical evolution of considering the constitutive equations of rate-type for both quasistatic computations and in the case of dynamical situations. To begin with, let us discuss quasistatic mechanical but transient thermal problems where the constitutive equations are defined by evolution equations of rate-type. The unsteady heat equation yields after its spatial discretization using finite elements to a system of ordinary differential equations (ODEs), which is coupled with the algebraic equations defining the mechanical equilibrium [Lewis et al. 1996; Reddy and Gartling 2000; Quint et al. 2011]. The latter is coupled with the constitutive model of rate-type. In our case, it forms a first-order ODE system as well (with case distinction in the case of viscoplasticity with yield function). This system can be solved using partitioned or monolithic approaches [Schrefler 2004]. In [Felippa and Park 1980; Felippa et al. 2001], the authors proposed a partitioned approach where the mechanical and the thermal parts are solved separately (commonly two different programs and/or meshes are chosen [Erbts and Düster 2012; Erbts et al. 2015; Wendt et al. 2015]). Particular investigations on partitioned systems and their stability

---

*Keywords:* dynamics, DIRK methods, finite elements, generalized-alpha method, time-adaptivity, thermoviscoplasticity.

are provided in [Simo and Miehe 1992; Armero and Simo 1992; 1993]. The monolithic approach was proposed in [Glaser 1991; Miehe 1988] and extended to high-order time-integration schemes in [Quint 2012; Rothe 2015; Rothe et al. 2015a]. In the latter case, the resulting equations define differential-algebraic equations (DAE) after spatial discretization, using finite elements that lead to changes in the numerical and theoretical properties of the mathematical model. In view of thermoviscoelasticity and the combination of high-order methods in space and time, we refer to [Netz and Hartmann 2015; Hamkar 2013].

Concerning inertia effects, the case of linear elastodynamics for small strains represents a huge scientific area which is not discussed here, as we focus on proposals treating inelastic material properties. The case of finite strain viscoelasticity is treated in [Conde Martín et al. 2014], where the constitutive model must have a particular structure (a so-called GENERIC scheme) to fit within the numerical scheme. This is extended in [Krüger et al. 2016] to the case of dynamical systems for finite strain thermoviscoelasticity. The application of dynamical systems in the context of viscoplasticity is proposed in [Noels et al. 2006; 2008], where variational update formulations are considered.

Our goal is the application of higher-order time-integration schemes (at least second-order accurate) so that all quantities — displacements, velocities, temperatures, and internal variables — reach second order. Commonly, first-order accurate methods are used. Even the Newmark-method has only a maximum order of two in the context of dynamical systems [Newmark 1959; Hughes 1987; Zienkiewicz and Taylor 2000]. However, it is common to treat the evolution equations on Gauss-point level with a first-order method so that the entire system cannot reach the expected order (regarding the relation of the Newmark-method in the context of Nyström-methods, see [Fritzen 1997]). Since we have a coupled system of second-order and first-order ODEs, the question arises as to what the appropriate method might be, since there are methods to treat second-order ODEs directly or the second-order ODEs are transferred into a system of first-order ODEs. Thus, two methods are considered here: first, we convert the second-order ODEs to a first-order ODE system and apply diagonally implicit Runge–Kutta (DIRK) methods to the entire system (for details concerning DIRK methods, see [Hairer et al. 1989; 1993; Hairer and Wanner 1996]). These methods have been successfully applied in the context of finite elements for isothermal and quasistatic problems in [Ellsiepen and Hartmann 2001; Hartmann 2002; Hartmann et al. 2008a] and quasistatic coupled problems in [Birken et al. 2010; Hartmann and Rothe 2013; Rothe et al. 2015b]. In the present article, the quasistatic problems are recapped as well, where a DAE system has to be solved. One intention is to show the underlying difference in the equations.

For problems in the field of linear structural dynamics, the generalized-$\alpha$ method formulated as a one-step method was originally introduced in [Chung and Hulbert 1993]. For first-order ODEs the method was introduced in [Jansen et al. 2000]. These methods show second-order accuracy in most numerical experiments, minimal numerical dissipation of lower modes, and maximal numerical dissipation of higher modes in the linear structural dynamics regime, where this numerical dissipation mechanism can be controlled by certain parameters. For first-order ODEs, an analysis of the generalized-$\alpha$ scheme was performed in [Dettmer and Perić 2003]. For the corresponding consistency and stability analysis of the generalized-$\alpha$ method for the second-order ODE system in structural dynamics, the reader is referred to [Chung and Hulbert 1993; Erlicher et al. 2002]. Furthermore, the generalized-$\alpha$ method includes the most popular classical numerical dissipative and nondissipative time integration schemes (the Newmark-family [Newmark 1959], the HHT-$\alpha$ method [Hilber et al. 1977] and the WBZ-$\alpha$ method [Wood et al.

1980]) in the field of structural dynamics. For this reason, it is a widely applied time integration scheme which has been successfully applied, for example, in [Hartmann 2007; Popp 2012; Kuhl 1996; Kuhl and Crisfield 1999; Kuhl and Ramm 1999]. Rang [2013a] proposed a simple adaptive time-step control for the one-step versions of the generalized-$\alpha$ method. A coupling between the generalized-$\alpha$ method for first-order ODEs and for second-order ODEs is presented in [Rang 2013b].

This paper is structured as follows: first, the basic equations are summarized; starting with a model problem of thermoviscoplasticity, the weak forms and their spatial discretization. Second, the time discretization is explained using DIRK and a modified generalized-$\alpha$ method. Finally, we investigate the order and efficiency of these methods by drawing on some examples.

The notation in use is defined in the following manner: geometrical vectors are symbolized by $\vec{a}$ and second-order tensors $\mathbf{A}$ by boldface Roman letters. We introduce matrices at the global level symbolized by boldface italic sans-serif letters $\boldsymbol{\mathsf{A}}$ and matrices on local level using the upright counterpart, $\mathsf{A}$.

## 2. Basic equations

In the following, the basic equations are summarized, which are: the constitutive model under consideration; the weak forms of the balance of linear momentum; the unsteady, nonlinear heat equation; and the result of the spatial discretization using finite elements.

**2.1.** ***Constitutive model.*** In the scope of our application in thermoviscoplasticity, we restrict ourselves to the case of small strains (large strains are straight forward and are embedded in the numerical schemes as well). Let $\mathbf{E}(\vec{x}, t) = \frac{1}{2}(\operatorname{grad}\vec{u}(\vec{x}, t) + \operatorname{grad}^T \vec{u}(\vec{x}, t))$ be the linearized strain tensor, $\vec{u}(\vec{x}, t)$ the displacement vector, $\vec{x}$ the material point and $t$ the time. Most phenomenological constitutive models are expressed by

$$\mathbf{T} = \mathbf{h}(\mathbf{E}, \Theta, \mathsf{q}), \tag{1}$$

$$\dot{\mathsf{q}}(t) = \mathsf{r}_\mathrm{q}(\mathbf{E}, \Theta, \mathsf{q}), \tag{2}$$

where $\mathsf{q} \in \mathbb{R}^{n_\mathrm{q}}$ is the vector of internal variables, $\mathbf{T}(\vec{x}, t)$ is the stress tensor and $\Theta(\vec{x}, t)$ is the absolute temperature. In Table 1, $\mathsf{q}^T = \{\mathsf{E}_\mathrm{v}^T \, \mathsf{E}_\mathrm{r}^T\}$ is given by two different strain-like internal variables in Voigt-notation, i.e., $n_\mathrm{q} = 12$, $\mathsf{E}_\mathrm{r} \in \mathbb{R}^6$, $\mathsf{E}_\mathrm{v} \in \mathbb{R}^6$. For details of the model, see [Tsakmakis and Willuweit 2004; Rothe et al. 2015a]. Pure (thermo)elasticity is embedded as well ($\mathsf{q} = \dot{\mathsf{q}} = \mathsf{0}$).

**2.2.** ***Spatial discretization.*** We investigate the partial differential equations

$$\operatorname{div}\mathbf{T}(\vec{x}, t) + \rho(\vec{x})\vec{k} = \rho(\vec{x})\ddot{\vec{u}}(\vec{x}, t), \tag{3}$$

$$\hat{c}_\mathrm{p}(\mathbf{E}, \Theta)\dot{\Theta} = \frac{1}{\rho}\operatorname{div}(\kappa(\Theta)\operatorname{grad}\Theta) + \hat{p}(\mathbf{E}, \dot{\mathbf{E}}, \Theta, \mathbf{E}_\mathrm{v}, \mathbf{E}_\mathrm{r}), \tag{4}$$

where $\rho$ is the density, $\vec{k}$ is the acceleration of gravity,

$$c_\mathrm{p} = \hat{c}_\mathrm{p}(\mathbf{E}, \Theta) = \left(c_\Theta - \frac{9K\alpha_\Theta^2}{\rho} - \frac{3c_K\alpha_\Theta}{\rho}\gamma(\mathbf{E}, \Theta)\right)\Theta \tag{5}$$

is the temperature- and strain-dependent heat capacity (with the abbreviation $\gamma(\mathbf{E}, \Theta) = \mathbf{I}\cdot\mathbf{E} - 3\alpha_\Theta\Theta$),

$$
\begin{array}{rl}
\underline{\text{yield function:}} & f(\mathbf{T},\mathbf{X},\Theta)=\sqrt{\tfrac{3}{2}(\mathbf{T}-\mathbf{X})^D\cdot(\mathbf{T}-\mathbf{X})^D}-\hat{k}(\Theta)\\
\underline{\text{elasticity relation:}} & \mathbf{T}=K(\Theta)(\operatorname{tr}\mathbf{E})\mathbf{I}+2G(\Theta)(\mathbf{E}-\mathbf{E}_{\mathrm{v}})^D-3K(\Theta)\alpha_\Theta\vartheta\mathbf{I}\\
\underline{\text{loading condition:}} & \begin{cases}\text{elasticity} & f<0\\ \text{viscoplasticity} & f\geq 0\end{cases}\\
\underline{\text{flow rules:}} & \begin{cases}\text{elasticity} & \dot{\mathbf{E}}_{\mathrm{v}}=\mathbf{0},\ \dot{\mathbf{E}}_{\mathrm{r}}=\mathbf{0}\\ \text{viscoplasticity} & \dot{\mathbf{E}}_{\mathrm{v}}=\lambda\dfrac{(\mathbf{T}-\mathbf{X})^{\mathrm{D}}}{\|(\mathbf{T}-\mathbf{X})^{\mathrm{D}}\|},\ \dot{\mathbf{E}}_{\mathrm{r}}=\lambda\beta\mathbf{X}^D\end{cases}\\
\underline{\text{abbreviations:}} & \mathbf{X}=c_X\mathbf{E}_{\mathrm{k}}^D=c_X(\mathbf{E}_{\mathrm{v}}-\mathbf{E}_{\mathrm{r}})^D,\\
 & \lambda=1/\eta\left\langle f(\mathbf{T},\mathbf{X},\Theta)/\sigma_0\right\rangle^m,\\
 & \hat{k}(\Theta)=(k_0-k_H)e^{-b(\Theta-\Theta_0)}+k_H,\\
 & K(\Theta)=K_0-c_K(\Theta-\Theta_0),\\
 & G(\Theta)=G_0-c_G(\Theta-\Theta_0)
\end{array}
$$

**Table 1.** Summary of constitutive equations (thermoviscoplasticity small strains).

and the heat production term due to dissipation and thermoelastic coupling is

$$
p=\hat{p}(\mathbf{E},\dot{\mathbf{E}},\Theta,\mathbf{E}_{\mathrm{v}},\mathbf{E}_{\mathrm{r}})=\delta-\frac{\Theta}{\rho}\big((3K\alpha_\Theta+c_K\gamma)\mathbf{I}+2c_G(\mathbf{E}-\mathbf{E}_{\mathrm{v}})^D\big)\cdot\dot{\mathbf{E}}+\frac{\Theta}{\rho}2c_G(\mathbf{E}-\mathbf{E}_{\mathrm{v}})^D\cdot\dot{\mathbf{E}}_{\mathrm{v}} \tag{6}
$$

with

$$
\delta=\frac{1}{\rho\eta}\left\langle\frac{\hat{f}(\mathbf{E},\Theta,\mathbf{E}_{\mathrm{v}},\mathbf{E}_{\mathrm{r}})}{\sigma_0}\right\rangle^m\big(\|2G(\Theta)(\mathbf{E}-\mathbf{E}_{\mathrm{v}})^D-c_X(\mathbf{E}_{\mathrm{v}}-\mathbf{E}_{\mathrm{r}})^D\|-\beta c_X^2\|(\mathbf{E}_{\mathrm{v}}-\mathbf{E}_{\mathrm{r}})^D\|^2\big). \tag{7}
$$

The angle brackets define Macauley brackets symbolizing a case distinction: $\langle x\rangle=0$ if $x<0$ and $\langle x\rangle=x$ if $x\geq 0$. Equation (3) represents the balance of linear momentum. In view of the balance of energy (5), Fourier's model ($\vec{q}=-\kappa(\Theta)\operatorname{grad}\Theta$) is assumed, where $\vec{q}$ defines the heat flux vector.

The partial differential equation (3) is multiplied with virtual displacements $\delta\vec{u}(\vec{x})$ and is integrated over the volume $V$. The divergence theorem is applied, which leads to d'Alembert's principle

$$
\int_V\underbrace{\mathbf{h}(\mathbf{E},\Theta,\mathbf{q})}_{\mathbf{T}}\cdot\delta\mathbf{E}\,\mathrm{d}V+\int_V\rho\ddot{\vec{u}}\cdot\delta\vec{u}\,\mathrm{d}V=\int_A\vec{t}\cdot\delta\vec{u}\,\mathrm{d}A+\int_V\rho\vec{k}\cdot\delta\vec{u}\,\mathrm{d}V, \tag{8}
$$

where $\delta\mathbf{E}(\vec{x})=(\operatorname{grad}\delta\vec{u}(\vec{x})+\operatorname{grad}^T\delta\vec{u}(\vec{x}))/2$ defines the virtual strain tensor. In this context $\delta\vec{u}(\vec{x})=\vec{0}$ has to hold on the boundary $A^u$ of the material body, where the displacements are prescribed; $A=A^\sigma\cup A^u$; $\vec{u}(\vec{x},t)=\overline{\vec{u}}(t)$ on $A^u$ (Dirichlet boundary conditions). Here, $\vec{t}(\vec{x},t)=\mathbf{T}(\vec{x},t)\vec{n}(\vec{x},t)$ defines the stress vector on the surface $A^\sigma$, where $\vec{n}$ represents the surface normal and $\vec{t}(\vec{x},t)=\overline{\vec{t}}(\vec{x},t)$ on $A^\sigma$ defines the Neumann boundary conditions.

We multiply the unsteady, nonlinear heat equation (4) with the virtual temperature $\delta\Theta$ in a similar manner, and perform both the integration over the volume as well as the application of the divergence theorem, yielding the principle of virtual temperatures

$$\int_V \rho\hat{c}_{\mathrm{p}}(\mathbf{E},\Theta)\dot{\Theta}\delta\Theta\,\mathrm{d}V = \int_A q\,\delta\Theta\,\mathrm{d}A - \int_V \kappa(\Theta)\,\mathrm{grad}\,\Theta\cdot\mathrm{grad}\,\delta\Theta\,\mathrm{d}V + \int_V \rho\hat{p}(\mathbf{E},\dot{\mathbf{E}},\Theta,\mathbf{q})\delta\Theta\,\mathrm{d}V. \quad (9)$$

The boundary conditions are: $\Theta(\vec{x},t)=\overline{\Theta}(\vec{x},t)$ on the surface, $A^{\Theta}$ and $q(\vec{x},t)=\overline{q}(\vec{x},t)$ on $A^{q}$. Here, we assume $q=-\vec{q}\cdot\vec{n}$ with the normal vector $\vec{n}$ on the outer surface. Of course, mixed boundary conditions for temperature-dependent heat fluxes such as convection or radiation must be considered as well, $q=\hat{q}(\Theta)$. Here, we refer to [Quint et al. 2011] and the literature cited therein. The entire mathematical model considers the heat production due to mechanical dissipation in an exact manner.

In order to solve the initial boundary-value problem (3) and (4) together with constitutive models of type (1) and (2), we draw on the method of vertical lines. First, the spatial discretization is performed, followed by the time-discretization. The spatial discretization is carried out using the finite element method, i.e., an ansatz for the displacements, virtual displacements, absolute temperatures and virtual temperatures are inserted into the weak forms (8) and (9). As summarized in more detail in Appendix A, this leads to the mixed first and second-order ODE system

$$\begin{aligned} \boldsymbol{M}\ddot{\boldsymbol{u}}(t) &= -\boldsymbol{g}_{\mathrm{u}}(t,\boldsymbol{u},\boldsymbol{\Theta},\boldsymbol{q}) - \boldsymbol{M}_{\mathrm{up}}\ddot{\bar{\boldsymbol{u}}}(t),\\ \boldsymbol{C}_{\Theta}(t,\boldsymbol{u},\boldsymbol{\Theta})\dot{\boldsymbol{\Theta}}(t) &= \boldsymbol{r}_{\Theta}(t,\boldsymbol{u},\dot{\boldsymbol{u}},\boldsymbol{\Theta},\boldsymbol{q}),\\ \dot{\boldsymbol{q}}(t) &= \boldsymbol{r}_{\mathrm{Q}}(t,\boldsymbol{u},\boldsymbol{\Theta},\boldsymbol{q}), \end{aligned} \quad (10)$$

with the initial conditions

$$\boldsymbol{u}(0)=\boldsymbol{u}_0,\quad \dot{\boldsymbol{u}}(0)=\boldsymbol{v}_0,\quad \boldsymbol{\Theta}(0)=\boldsymbol{\Theta}_0,\quad \boldsymbol{q}(0)=\boldsymbol{q}_0. \quad (11)$$

Equation $(10)_1$ represents the (semidiscrete) equation of motion with a constant and consistent mass matrix $\boldsymbol{M}\in\mathbb{R}^{n_{\mathrm{uu}}\times n_{\mathrm{uu}}}$ (see (76)) which is assigned to the unknown displacement degrees of freedom (DOF) $\boldsymbol{u}\in\mathbb{R}^{n_{\mathrm{uu}}}$. Here, $\boldsymbol{M}_{\mathrm{up}}\in\mathbb{R}^{n_{\mathrm{uu}}\times n_{\mathrm{up}}}$ symbolizes the mass matrix assigned to the known (prescribed) displacement DOF $\bar{\boldsymbol{u}}\in\mathbb{R}^{n_{\mathrm{up}}}$. Thus, $\ddot{\bar{\boldsymbol{u}}}(t)\in\mathbb{R}^{n_{\mathrm{up}}}$ are the prescribed nodal accelerations, which are explicit functions of time $t$. The matrix $\boldsymbol{C}_{\Theta}(t,\boldsymbol{u},\boldsymbol{\Theta})\in\mathbb{R}^{n_{\Theta\mathrm{u}}\times n_{\Theta\mathrm{u}}}$ defines the deformation and temperature-dependent heat-capacity matrix (see (82)) where $\boldsymbol{\Theta}\in\mathbb{R}^{n_{\Theta\mathrm{u}}}$ are the unknown nodal temperatures. The ODE system $(10)_3$ is the result of assembling all evolution equations of the internal variables into a large vector $\boldsymbol{q}\in\mathbb{R}^{n_{\mathrm{Q}}}$, see the discussion in [Ellsiepen and Hartmann 2001; Hartmann 2005].

**2.2.1.** *Subproblems.* There are a number of subproblems which are connected to the dynamical system (10) if particular quantities do not appear, i.e., if they are neglected. In the case of $\dot{\boldsymbol{q}}(t)=\boldsymbol{q}(t)=\boldsymbol{0}$, the case of thermoelasticity under dynamical agencies is given, see (12) of Table 2. Additionally, if $\boldsymbol{\Theta}_{\mathrm{a}}$ is constant spatially and temporally (where $\boldsymbol{\Theta}_{\mathrm{a}}\in\mathbb{R}^{n_{\Theta\mathrm{a}}}$), and $n_{\Theta\mathrm{a}}=n_{\Theta\mathrm{u}}+n_{\Theta\mathrm{p}}$ defines all nodal temperatures, pure elasto-dynamic has to be solved (see (13)). Of course, further subproblems can appear, such as: $\boldsymbol{C}_{\Theta}=$ constant or $\boldsymbol{g}_{\mathrm{u}}(t,\boldsymbol{u},\boldsymbol{\Theta})$ represents a linear function in $\boldsymbol{u}$ and $\boldsymbol{\Theta}$, or one-sided coupling is assumed, etc. This can be solved with specific numerical methods.

The quasistatic case holds for $\ddot{\boldsymbol{u}}(t)\approx\boldsymbol{0}$ leading to DAE systems, see the case of thermoinelasticity (14), isothermal inelasticity (viscoelasticity, viscoplasticity) (15), and pure elasticity (16). The case of rate-independent elasto-plasticity with a yield function leads to the evolution equations $\boldsymbol{A}\dot{\boldsymbol{q}}(t)=\boldsymbol{r}_{\mathrm{Q}}(t,\boldsymbol{u},\boldsymbol{\Theta},\boldsymbol{q})$,

$$\text{thermoelasticity:}\quad \begin{aligned} \boldsymbol{M}\ddot{\boldsymbol{u}}(t) &= -\boldsymbol{g}_{\mathrm{u}}(t,\boldsymbol{u},\boldsymbol{\Theta}) - \boldsymbol{M}_{\mathrm{up}}\ddot{\bar{\boldsymbol{u}}}(t) \\ \boldsymbol{C}_{\Theta}(t,\boldsymbol{u},\boldsymbol{\Theta})\dot{\boldsymbol{\Theta}}(t) &= \boldsymbol{r}_{\Theta}(t,\boldsymbol{u},\dot{\boldsymbol{u}},\boldsymbol{\Theta}) \end{aligned} \tag{12}$$

$$\text{elasto-dynamics:}\quad \boldsymbol{M}\ddot{\boldsymbol{u}}(t) = -\boldsymbol{g}_{\mathrm{u}}(t,\boldsymbol{u}) - \boldsymbol{M}_{\mathrm{up}}\ddot{\bar{\boldsymbol{u}}}(t) \tag{13}$$

$$\text{quasistatic thermoinelasticity:}\quad \begin{aligned} \boldsymbol{0} &= \boldsymbol{g}_{\mathrm{u}}(t,\boldsymbol{u},\boldsymbol{\Theta},\boldsymbol{q}) \\ \boldsymbol{C}_{\Theta}(t,\boldsymbol{u},\boldsymbol{\Theta})\dot{\boldsymbol{\Theta}}(t) &= \boldsymbol{r}_{\Theta}(t,\boldsymbol{u},\dot{\boldsymbol{u}},\boldsymbol{\Theta},\boldsymbol{q}) \\ \dot{\boldsymbol{q}}(t) &= \boldsymbol{r}_{\mathrm{Q}}(t,\boldsymbol{u},\boldsymbol{\Theta},\boldsymbol{q}) \end{aligned} \tag{14}$$

$$\text{quasistatic inelasticity:}\quad \begin{aligned} \boldsymbol{0} &= \boldsymbol{g}_{\mathrm{u}}(t,\boldsymbol{u},\boldsymbol{q}) \\ \dot{\boldsymbol{q}}(t) &= \boldsymbol{r}_{\mathrm{Q}}(t,\boldsymbol{u},\boldsymbol{q}) \end{aligned} \tag{15}$$

$$\text{elasticity:}\quad \boldsymbol{0} = \boldsymbol{g}_{\mathrm{u}}(t,\boldsymbol{u}) \tag{16}$$

**Table 2.** Dynamical and quasistatic subproblems.

where $\boldsymbol{A}$ is a singular matrix and $\boldsymbol{r}$ contains case distinctions (see the discussion in [Ellsiepen and Hartmann 2001]).

**2.2.2.** *Reaction force computation.* Another question connected to the dynamical systems under consideration is related to the computation of the reaction forces at those degrees of freedom, where the displacements are prescribed. This is not provided by d'Alembert's principle (8) since the virtual displacements vanish at those DOF for which the displacements are prescribed (and no virtual work is produced). The same situation holds for the heat flux computation for those nodes with prescribed temperatures.

This is consistently discussed for isothermal quasistatic problems in [Hartmann 2005] and the literature cited therein. Hamkar [2013] treats this issue for quasistatic thermomechanical problems since other time-integration methods require these investigations. He applies a concept developed by Gear [1986]. The basic idea is as follows: we assume that all displacement DOF $\boldsymbol{u}_{\mathrm{a}} \in \mathbb{R}^{n_{\mathrm{uu}}+n_{\mathrm{up}}}$ are unknown. The vector is decomposed into DOF that are connected to Dirichlet boundary conditions and the remaining quantities,

$$\boldsymbol{u}_{\mathrm{a}}(t) = \begin{cases} \boldsymbol{u}(t) \\ \hat{\boldsymbol{u}}(t) \end{cases} \quad \text{with } \boldsymbol{u} \in \mathbb{R}^{n_{\mathrm{uu}}} \text{ and } \hat{\boldsymbol{u}} \in \mathbb{R}^{n_{\mathrm{up}}}.$$

Consistently, all virtual displacements are arbitrary, $\delta\boldsymbol{u}_{\mathrm{a}} \in \mathbb{R}^{n_{\mathrm{uu}}+n_{\mathrm{up}}}$. However, there is the constraint equation

$$\boldsymbol{c}_{\mathrm{u}}(t,\boldsymbol{a}_{\mathrm{a}}) = \boldsymbol{Z}_{\mathrm{u}}^{T}\boldsymbol{u}_{\mathrm{a}}(t) - \bar{\boldsymbol{u}}(t) = \hat{\boldsymbol{u}}(t) - \bar{\boldsymbol{u}}(t) = \boldsymbol{0} \tag{17}$$

at the DOF for which displacements are prescribed, with

$$\boldsymbol{Z}_{\mathrm{u}} = \begin{bmatrix} \mathbf{0}_{n_{\mathrm{uu}}\times n_{\mathrm{up}}} \\ \boldsymbol{I}_{n_{\mathrm{up}}\times n_{\mathrm{up}}} \end{bmatrix}, \boldsymbol{Z}_{\mathrm{u}} \in \mathbb{R}^{(n_{\mathrm{uu}}+n_{\mathrm{up}})\times n_{\mathrm{up}}}$$

and the prescribed functions $\overline{\boldsymbol{u}}(t) \in \mathbb{R}^{n_{\mathrm{up}}}$. Additionally, the weak form changes to

$$\delta \boldsymbol{u}_{\mathrm{a}}^{T} \left\{ \boldsymbol{M}_{\mathrm{a}} \ddot{\boldsymbol{u}}_{\mathrm{a}}(t) + \boldsymbol{g}_{\mathrm{ua}}(t, \boldsymbol{u}_{\mathrm{a}}, \boldsymbol{\Theta}_{\mathrm{a}}, \boldsymbol{q}) - \boldsymbol{Z}_{\mathrm{u}} \boldsymbol{\lambda}_{\mathrm{u}}(t) \right\} = 0,$$

i.e., there is an additional term defined by the Lagrange-multiplier $\boldsymbol{\lambda}_{\mathrm{u}} \in \mathbb{R}^{n_{\mathrm{up}}}$. In other words,

$$\begin{aligned} &\boldsymbol{M}_{\mathrm{a}} \ddot{\boldsymbol{u}}_{\mathrm{a}}(t) + \boldsymbol{g}_{\mathrm{ua}}(t, \boldsymbol{u}_{\mathrm{a}}, \boldsymbol{\Theta}_{\mathrm{a}}, \boldsymbol{q}) - \boldsymbol{Z}_{\mathrm{u}} \boldsymbol{\lambda}_{\mathrm{u}}(t) \\ &\qquad = \begin{bmatrix} \boldsymbol{M} & \boldsymbol{M}_{\mathrm{up}} \\ \boldsymbol{M}_{\mathrm{pu}} & \boldsymbol{M}_{\mathrm{pu}} \end{bmatrix} \begin{Bmatrix} \ddot{\boldsymbol{u}}(t) \\ \ddot{\hat{\boldsymbol{u}}}(t) \end{Bmatrix} + \begin{Bmatrix} \boldsymbol{g}_{\mathrm{u}}(t, \boldsymbol{u}, \hat{\boldsymbol{u}}, \boldsymbol{\Theta}, \hat{\boldsymbol{\Theta}}, \boldsymbol{q}) \\ \hat{\boldsymbol{g}}_{\mathrm{u}}(t, \boldsymbol{u}, \hat{\boldsymbol{u}}, \boldsymbol{\Theta}, \hat{\boldsymbol{\Theta}}, \boldsymbol{q}) - \boldsymbol{\lambda}_{\mathrm{u}}(t) \end{Bmatrix} = \boldsymbol{0} \end{aligned} \tag{18}$$

has to be solved in combination with the constraint (17). The same idea holds for the heat equation having the constraint

$$\boldsymbol{c}_{\Theta}(t, \boldsymbol{\Theta}_{\mathrm{a}}) = \boldsymbol{Z}_{\Theta}^{T} \boldsymbol{\Theta}_{\mathrm{a}}(t) - \overline{\boldsymbol{\Theta}}(t) = \hat{\boldsymbol{\Theta}}(t) - \overline{\boldsymbol{\Theta}}(t) = \boldsymbol{0} \tag{19}$$

for the given nodal temperature functions $\overline{\boldsymbol{\Theta}}(t) \in \mathbb{R}^{n_{\Theta\mathrm{p}}}$. The modified weak form reads

$$\delta \boldsymbol{\Theta}_{\mathrm{a}}^{T} \left\{ \boldsymbol{C}_{\Theta\mathrm{a}}(t, \boldsymbol{u}_{\mathrm{a}}, \boldsymbol{\Theta}_{\mathrm{a}}) \dot{\boldsymbol{\Theta}}_{\mathrm{a}}(t) - \boldsymbol{r}_{\Theta\mathrm{a}}(t, \boldsymbol{u}_{\mathrm{a}}, \dot{\boldsymbol{u}}_{\mathrm{a}}, \boldsymbol{\Theta}_{\mathrm{a}}, \boldsymbol{q}) - \boldsymbol{Z}_{\Theta} \boldsymbol{\lambda}_{\Theta}(t) \right\} = 0,$$

with the Lagrange multiplier $\boldsymbol{\lambda}_{\Theta}(t) \in \mathbb{R}^{n_{\Theta\mathrm{p}}}$ and the matrix

$$\boldsymbol{Z}_{\Theta} = \begin{bmatrix} \boldsymbol{0}_{n_{\Theta\mathrm{u}} \times n_{\Theta\mathrm{p}}} \\ \boldsymbol{I}_{\, n_{\Theta\mathrm{p}} \times n_{\Theta\mathrm{p}}} \end{bmatrix}, \boldsymbol{Z}_{\Theta} \in \mathbb{R}^{(n_{\Theta\mathrm{u}} + n_{\Theta\mathrm{p}}) \times n_{\Theta\mathrm{p}}}.$$

The decomposition yields the ODE system

$$\begin{bmatrix} \boldsymbol{C}_{\Theta} & \boldsymbol{C}_{\Theta\mathrm{up}} \\ \boldsymbol{C}_{\Theta\mathrm{pu}} & \boldsymbol{C}_{\Theta\mathrm{pp}} \end{bmatrix} \begin{Bmatrix} \dot{\boldsymbol{\Theta}}(t) \\ \dot{\hat{\boldsymbol{\Theta}}}(t) \end{Bmatrix} = \begin{Bmatrix} \boldsymbol{r}_{\Theta}(t, \boldsymbol{u}, \hat{\boldsymbol{u}}, \dot{\boldsymbol{u}}, \dot{\hat{\boldsymbol{u}}}, \boldsymbol{\Theta}, \hat{\boldsymbol{\Theta}}, \boldsymbol{q}) \\ \hat{\boldsymbol{r}}_{\Theta}(t, \boldsymbol{u}, \hat{\boldsymbol{u}}, \dot{\boldsymbol{u}}, \dot{\hat{\boldsymbol{u}}}, \boldsymbol{\Theta}, \hat{\boldsymbol{\Theta}}, \boldsymbol{q}) + \boldsymbol{\lambda}_{\Theta}(t) \end{Bmatrix}. \tag{20}$$

Table 3 summarizes the entire DAE system. For the case of quasistatic and isothermal problems and implicit finite elements, see [Hartmann et al. 2008b].

## 3. Time discretization schemes

In the following we propose two implicit time integration methods to treat the ODE system (10) since we are interested in a low frequency response. Explicit methods are widely used in the context of high frequency responses and wave-like phenomena — or in high velocity impact situations, where contact

$$\begin{aligned} \boldsymbol{M}_{\mathrm{a}} \ddot{\boldsymbol{u}}_{\mathrm{a}}(t) &= -\boldsymbol{g}_{\mathrm{ua}}(t, \boldsymbol{u}_{\mathrm{a}}, \boldsymbol{\Theta}_{\mathrm{a}}, \boldsymbol{q}) + \boldsymbol{Z}_{\mathrm{u}} \boldsymbol{\lambda}_{\mathrm{u}}(t) \\ \boldsymbol{Z}_{\mathrm{u}}^{T} \boldsymbol{u}_{\mathrm{a}}(t) - \overline{\boldsymbol{u}}(t) &= \boldsymbol{0} \\ \boldsymbol{C}_{\Theta\mathrm{a}}(t, \boldsymbol{cr}_{\mathrm{a}}, \boldsymbol{\Theta}_{\mathrm{a}}) \dot{\boldsymbol{\Theta}}_{\mathrm{a}}(t) &= \boldsymbol{r}_{\Theta\mathrm{a}}(t, \boldsymbol{u}_{\mathrm{a}}, \dot{\boldsymbol{u}}_{\mathrm{a}}, \boldsymbol{\Theta}_{\mathrm{a}}, \boldsymbol{q}) + \boldsymbol{Z}_{\Theta} \boldsymbol{\lambda}_{\Theta}(t) \\ \boldsymbol{Z}_{\Theta}^{T} \boldsymbol{\Theta}_{\mathrm{a}}(t) - \overline{\boldsymbol{\Theta}}(t) &= \boldsymbol{0} \\ \dot{\boldsymbol{q}}(t) &= \boldsymbol{r}_{\mathrm{Q}}(\boldsymbol{u}_{\mathrm{a}}, \boldsymbol{\Theta}_{\mathrm{a}}, \boldsymbol{q}) \end{aligned} \tag{21}$$

**Table 3.** Entire DAE system if reaction force and reaction heat flux have to be determined

conditions dominate the problem. Implicit methods have the advantage in that it is possible to apply relatively large time-step sizes in comparison to explicit schemes. However, the implementation of implicit methods is more challenging due to the fact that nonlinear solution methods have to be considered, see in this context [Belytschko et al. 2000].

Since we are interested in one-step methods, we follow the numerical schemes in [Ellsiepen and Hartmann 2001] (see [Fritzen 1997] for problems concerning the aspect of dynamics). Accordingly, the ODE part of second order is transferred into a system of first-order ODEs and stiffly accurate diagonally implicit Runge–Kutta methods are applied. A second method is an extension of the generalized-$\alpha$ method seen in [Chung and Hulbert 1993]. Here, we draw on a one-step formulation.

In contrast to these methods, there are multistep methods that require a special starting procedure as well as access to time-step information reaching back even further than $t_n$.

**3.1.** ***DIRK methods.*** In order to obtain a basic background of the proposed time integration scheme, we recall the general class of implicit Runge–Kutta methods (IRK) applied to (10) [Strehmel et al. 2012; Hairer et al. 1993]. In the next step, we introduce the class of diagonally implicit Runge–Kutta (DIRK) methods using a special choice of algorithmic parameters in order to obtain the important goal of computational efficiency. To this end, we transform the coupled system of second-order ODEs in (10) to a first-order ODE system by exploiting the trivial equation $\dot{\boldsymbol{u}} = \boldsymbol{v}(t)$. Accordingly, the system

$$\begin{aligned} \boldsymbol{M}\dot{\boldsymbol{v}}(t) &= -\boldsymbol{g}_{\mathrm{u}}(t, \boldsymbol{u}, \boldsymbol{\Theta}, \boldsymbol{q}) - \boldsymbol{M}_{\mathrm{up}}\dot{\bar{\boldsymbol{v}}}(t), \\ \dot{\boldsymbol{u}}(t) &= \boldsymbol{v}(t), \\ \boldsymbol{C}_{\Theta}(t, \boldsymbol{u}, \boldsymbol{\Theta})\dot{\boldsymbol{\Theta}}(t) &= \boldsymbol{r}_{\Theta}(t, \boldsymbol{u}, \boldsymbol{v}, \boldsymbol{\Theta}, \boldsymbol{q}), \\ \dot{\boldsymbol{q}}(t) &= \boldsymbol{r}_{\mathrm{Q}}(t, \boldsymbol{u}, \boldsymbol{\Theta}, \boldsymbol{q}), \end{aligned} \tag{22}$$

must be solved. To achieve a high-order method, we divide the integration step from time $t_n$ to $t_{n+1}$ in further points in time

$$T_{ni} = t_n + c_i \Delta t_n, \quad i = 1, \dots, s, \tag{23}$$

with coefficients $c_i$, $i = 1, \dots, s$. These points are called *stages*, where $s$ defines the number of stages. At every stage, the stage variables in the corresponding field are defined according to

$$\begin{aligned} \boldsymbol{U}_{ni} &= \boldsymbol{u}_n + \Delta t_n \sum_{j=1}^{s} a_{ij} \boldsymbol{V}_{nj}, & \boldsymbol{V}_{ni} &= \boldsymbol{v}_n + \Delta t_n \sum_{j=1}^{s} a_{ij} \boldsymbol{A}_{nj}, \\ \boldsymbol{\Theta}_{ni} &= \boldsymbol{\Theta}_n + \Delta t_n \sum_{j=1}^{s} a_{ij} \dot{\boldsymbol{\Theta}}_{nj}, & \boldsymbol{Q}_{ni} &= \boldsymbol{q}_n + \Delta t_n \sum_{j=1}^{s} a_{ij} \dot{\boldsymbol{Q}}_{nj}, \end{aligned} \tag{24}$$

with weighting factors $a_{ij}$. The index $ni$ denotes the $i$-th stage at time $T_{ni}$. Formally, we now have two sets of unknowns, namely the so-called *stage values* $\boldsymbol{U}_{ni}, \boldsymbol{V}_{ni}, \boldsymbol{\Theta}_{ni}, \boldsymbol{Q}_{ni}$ (displacements, velocities, temperatures and internal variables) and the *stage derivatives* $\boldsymbol{V}_{ni}, \boldsymbol{A}_{ni}, \dot{\boldsymbol{\Theta}}_{ni}, \dot{\boldsymbol{Q}}_{ni}$. However, since they are not independent, due to (24), we choose the stage derivatives $\boldsymbol{V}_{ni}, \boldsymbol{A}_{ni}, \dot{\boldsymbol{\Theta}}_{ni}$ for the following explanations, and the stage value $\boldsymbol{Q}_{ni}$ as the primary unknown for the internal variables. Since we are looking for the approximate solution at $t_{n+1}$, we need a relation which maps the stage derivatives to the unknown

$$\begin{array}{c|cccc} c_1 & a_{11} & a_{12} & \cdots & a_{1s} \\ c_2 & a_{21} & a_{22} & \cdots & a_{2s} \\ \vdots & \vdots & & \ddots & \vdots \\ c_s & a_{1s} & a_{2s} & \cdots & a_{ss} \\ \hline & b_1 & b_2 & \cdots & b_s \end{array}$$

(a) IRK method

$$\begin{array}{c|cccc} c_1 & a_{11} & 0 & \cdots & 0 \\ c_2 & a_{21} & a_{22} & \ddots & \vdots \\ \vdots & \vdots & & \ddots & 0 \\ c_s & b_1 & b_2 & \cdots & b_s \\ \hline & b_1 & b_2 & \cdots & b_s \end{array}$$

(b) SDIRK method

$$\begin{array}{c|c} \boldsymbol{c} & \boldsymbol{A} \\ \hline & \boldsymbol{b}^T \end{array}$$

(c) RK scheme

$$\begin{array}{c|c|c} \boldsymbol{c} & \bar{\boldsymbol{A}} & \boldsymbol{A} \\ \hline & \bar{\boldsymbol{b}}^T & \boldsymbol{b}^T \end{array}$$

(d) Indirect RKN scheme

**Figure 1.** Butcher-tableaus.

quantities at $t_{n+1}$. These integration formulas are based on quadrature rules in [Hairer et al. 1993; Hairer and Wanner 1996; Strehmel et al. 2012] and are defined by

$$\begin{aligned} \boldsymbol{u}_{n+1} &= \boldsymbol{u}_n + \Delta t_n \sum_{i=1}^{s} b_i \boldsymbol{V}_{ni}, & \boldsymbol{v}_{n+1} &= \boldsymbol{v}_n + \Delta t_n \sum_{i=1}^{s} b_i \boldsymbol{A}_{ni}, \\ \boldsymbol{\Theta}_{n+1} &= \boldsymbol{\Theta}_n + \Delta t_n \sum_{i=1}^{s} b_i \dot{\boldsymbol{\Theta}}_{ni}, & \boldsymbol{q}_{n+1} &= \boldsymbol{q}_n + \Delta t_n \sum_{i=1}^{s} b_i \dot{\boldsymbol{Q}}_{ni}, \end{aligned} \tag{25}$$

with the weighting factors $b_i$, $i = 1, \dots, s$. The weighting factors $a_{ij}$, $b_i$, and $c_i$ are usually summarized in a Butcher array to distinguish different classes of Runge–Kutta methods, see Figure 1(a). These factors are defined and calculated in order to obtain efficient, stable and accurate methods. By applying the general implicit Runge–Kutta scheme to (22) (i.e., inserting the primary variables and the relations of (24)), we obtain (for $i = 1, \dots, s$)

$$\begin{aligned} \boldsymbol{M}\boldsymbol{A}_{ni} &= -\boldsymbol{g}_{\mathrm{u}}(T_{ni}, \boldsymbol{U}_{ni}(\boldsymbol{A}_{nj}), \boldsymbol{\Theta}_{ni}(\dot{\boldsymbol{\Theta}}_{nj}), \boldsymbol{Q}_{ni}) - \boldsymbol{M}_{\mathrm{up}}\ddot{\bar{\boldsymbol{u}}}(T_{ni}), \\ \boldsymbol{V}_{ni}(\boldsymbol{A}_{nj}) &= \boldsymbol{v}_n + \Delta t_n \sum_{j=1}^{s} a_{ij} \boldsymbol{A}_{nj}, \quad j = 1, \dots, s, \\ \boldsymbol{C}_{\Theta}(\boldsymbol{U}_{ni}, \boldsymbol{\Theta}_{ni})\dot{\boldsymbol{\Theta}}_{ni} &= \boldsymbol{r}_{\Theta}(T_{ni}, \boldsymbol{U}_{ni}(\boldsymbol{A}_{nj}), \boldsymbol{V}_{ni}(\boldsymbol{A}_{nj}), \boldsymbol{\Theta}_{ni}(\dot{\boldsymbol{\Theta}}_{nj}), \boldsymbol{Q}_{ni}), \\ \dot{\boldsymbol{Q}}_{ni}(\boldsymbol{Q}_{nj}) &= \boldsymbol{r}_{\mathrm{Q}}(T_{ni}, \boldsymbol{U}_{ni}(\boldsymbol{A}_{nj}), \boldsymbol{\Theta}_{ni}(\dot{\boldsymbol{\Theta}}_{nj}), \boldsymbol{Q}_{ni}), \end{aligned} \tag{26}$$

a coupled nonlinear system of equations with the dimension $(2n_{\mathrm{uu}} + n_{\Theta\mathrm{u}} + n_{\mathrm{Q}}) \times s$ in each time-step $\Delta t_n$, where all stage unknowns are coupled. Next, we can eliminate $\boldsymbol{V}_{ni}$ in $(25)_1$ and $(24)_1$ by exploiting $(26)_2$. Furthermore, we assume existing stability requirements in the form of

$$c_i = \sum_{j=1}^{s} a_{ij}, \quad \sum_{i=1}^{s} b_i = 1, \tag{27}$$

$$\begin{array}{c|cc} 0 & 0 & 0 \\ 1 & \frac{1}{2}-\beta & \beta \\ \hline & \frac{1}{2}-\beta & \beta \\ \hline & 1-\gamma & \gamma \end{array} \qquad\qquad \begin{array}{c|c} \boldsymbol{c} & \bar{\boldsymbol{A}} \\ \hline & \bar{\boldsymbol{b}}^T \\ \hline & \boldsymbol{b}^T \end{array}$$

**Figure 2.** Butcher tableau for Newmark-method and presentation of schematic direct RKN scheme: Newmark-methods embedded as Nyström-method (left) and Nyström-scheme (right).

which ensures a consistent method and applicability for nonautonomous ODEs. This leads to

$$\boldsymbol{U}_{ni} = \boldsymbol{u}_{n+1} + c_i \Delta t_n \boldsymbol{v}_n + \Delta t_n^2 \sum_{j=1}^{s} \bar{a}_{ij} \boldsymbol{A}_{nj}, \tag{28}$$

$$\boldsymbol{u}_{n+1} = \boldsymbol{u}_n + \Delta t_n \boldsymbol{v}_n + \Delta t_n^2 \sum_{i=1}^{s} \bar{b}_i \boldsymbol{A}_{ni}, \tag{29}$$

with the new coefficients

$$\bar{a}_{ij} = \sum_{k=1}^{s} a_{ik} a_{kj}, \quad \text{and} \quad \bar{b}_i = \sum_{j=1}^{s} b_j a_{ji}. \tag{30}$$

This approach is identical to the construction of an indirect Runge–Kutta–Nyström method (RKN) by consistent application of an underlying Runge–Kutta scheme [Hairer et al. 1993; Strehmel et al. 2012; Fritzen 1997]. A schematic Butcher-tableau for an indirect RKN method is depicted in Figure 1(d). The stage quantities $\boldsymbol{U}_{ni}$ and $\boldsymbol{V}_{ni}$ can now be expressed as functions of the stage acceleration $\boldsymbol{A}_{nj}$, so that it is sufficient to store only the stage derivatives. For a pure structural dynamical problem, it is possible to save half the storage.

There are further schemes apart from these indirectly constructed Runge–Kutta–Nyström methods. We call them direct Runge–Kutta–Nyström methods if the conditions in (30) do not exist and if the weighting factors $\bar{a}_{ij}$ and $\bar{b}_i$ are independent from $a_{ij}$ and $b_i$. The direct RKN methods are principally applicable to first-order ODEs and to coupled systems, such as the one presented in (10). These methods were originally developed for general second-order ODEs. Furthermore, in the case of second-order ODEs which are not coupled with ODEs of first order and do not depend on $\dot{\boldsymbol{u}}$ (i.e., $\boldsymbol{g}_{\mathrm{u}}$ does not depend on $\dot{\boldsymbol{u}}$, see the case of structural dynamics in (13)), the coefficients $a_{ij}$ are no longer needed. Methods for which no coefficient matrix $\boldsymbol{A}$ is given are called *Nyström methods* [Hairer et al. 1993], see Figure 2 (right). However, taking $(24)_2$–$(24)_4$ in to account, these coefficients are indispensable. Fritzen [1997] showed that the popular Newmark-family is included in the class of diagonally implicit, direct Runge–Kutta–Nyström (DIRKN) methods in form of the Butcher-tableau in Figure 2 (left), with the well-known algorithmic parameters $\gamma \in [0, 1]$ and $\beta \in [0, 1/2]$ determining the stability and accuracy behavior (see Section 3.2). Because of the restrictions regarding $\dot{\boldsymbol{u}}$ mentioned above and the treatment of ODEs, portions of first-order general Newmark methods with arbitrary algorithmic parameter combinations are not applicable to (10). There exists one exception to this statement. The trapezoidal rule, which is known as the constant average acceleration method (CAA) in structural dynamics literature, is considered the most effective if the parameters $\{\gamma = 1/2, \beta = 1/4\}$ are chosen [Hughes 1987]. This second-order method fits into the class of DIRK methods [Ellsiepen and Hartmann 2001]. For this reason, the trapezoidal rule is the method of choice if we solve (10) with a Newmark-scheme (see [Ortiz and Popov 1985] for additional information regarding the trapezoidal rule).

The computational expense of a solution using IRK methods is high and thus is a drawback compared to other time integration methods. To decouple the stage quantities and to preserve the sparse structure of the linearized finite element system, we use DIRK methods where $a_{ij} = 0$ and — with regard to $(30)_1$ — $\bar{a}_{ij} = 0$ for $j > i$ holds (compare Figures 1(a) and 1(b)). In this case, the integration step for each field variable reduces to

$$\begin{aligned} \boldsymbol{U}_{ni} &= \boldsymbol{U}^{\mathrm{S}}_{ni} + \Delta t_n^2 \bar{a}_{ii} \boldsymbol{A}_{ni}, & \boldsymbol{V}_{ni} &= \boldsymbol{V}^{\mathrm{S}}_{ni} + \Delta t_n a_{ii} \boldsymbol{A}_{ni}, \\ \boldsymbol{\Theta}_{ni} &= \boldsymbol{\Theta}^{\mathrm{S}}_{ni} + \Delta t_n a_{ii} \dot{\boldsymbol{\Theta}}_{ni}, & \boldsymbol{Q}_{ni} &= \boldsymbol{Q}^{\mathrm{S}}_{ni} + \Delta t_n a_{ii} \dot{\boldsymbol{Q}}_{ni}, \end{aligned} \tag{31}$$

with starting values

$$\begin{aligned} \boldsymbol{U}^{\mathrm{S}}_{ni} &= \boldsymbol{u}_n + c_i \Delta t_n \boldsymbol{v}_n + \Delta t_n^2 \sum_{j=1}^{i-1} \bar{a}_{ij} \boldsymbol{A}_{nj}, & \boldsymbol{V}^{\mathrm{S}}_{ni} &= \boldsymbol{v}_n + \Delta t_n \sum_{j=1}^{i-1} a_{ij} \boldsymbol{A}_{nj}, \\ \boldsymbol{\Theta}^{\mathrm{S}}_{ni} &= \boldsymbol{\Theta}_n + \Delta t_n \sum_{j=1}^{i-1} a_{ij} \dot{\boldsymbol{\Theta}}_{nj}, & \boldsymbol{Q}^{\mathrm{S}}_{ni} &= \boldsymbol{q}_n + \Delta t_n \sum_{j=1}^{i-1} a_{ij} \dot{\boldsymbol{Q}}_{nj}, \end{aligned} \tag{32}$$

depending only on stage derivatives already calculated in previous stages. In each time-step — due to the decoupling of the stage quantities — we have to consecutively solve a sequence of $s$ nonlinear systems of equations with $n_{\mathrm{uu}} + n_{\Theta\mathrm{u}} + n_{\mathrm{Q}}$ unknowns $(\boldsymbol{A}_{ni}, \dot{\boldsymbol{\Theta}}_{ni}, \boldsymbol{Q}_{ni})$ of the form

$$\begin{aligned} \boldsymbol{G}_{\mathrm{u}}(T_{ni}, \boldsymbol{A}_{ni}, \dot{\boldsymbol{\Theta}}_{ni}, \boldsymbol{Q}_{ni}) &= \boldsymbol{M}\boldsymbol{A}_{ni} + \boldsymbol{g}_{\mathrm{u}}(T_{ni}, \boldsymbol{U}_{ni}(\boldsymbol{A}_{ni}), \boldsymbol{\Theta}_{ni}(\dot{\boldsymbol{\Theta}}_{ni}), \boldsymbol{Q}_{ni}) + \boldsymbol{M}_{\mathrm{up}} \ddot{\bar{\boldsymbol{u}}}(T_{ni}) = \boldsymbol{0}, \\ \boldsymbol{G}_{\Theta}(T_{ni}, \boldsymbol{A}_{ni}, \dot{\boldsymbol{\Theta}}_{ni}, \boldsymbol{Q}_{ni}) &= \boldsymbol{C}_{\Theta}(\boldsymbol{U}_{ni}, \boldsymbol{\Theta}_{ni})\, \dot{\boldsymbol{\Theta}}_{ni} \\ &\qquad - \boldsymbol{r}_{\Theta}(T_{ni}, \boldsymbol{U}_{ni}(\boldsymbol{A}_{ni}), \boldsymbol{V}_{ni}(\boldsymbol{A}_{ni}), \boldsymbol{\Theta}_{ni}(\dot{\boldsymbol{\Theta}}_{ni}), \boldsymbol{Q}_{ni}) = \boldsymbol{0}, \\ \boldsymbol{L}(T_{ni}, \boldsymbol{A}_{ni}, \dot{\boldsymbol{\Theta}}_{ni}, \boldsymbol{Q}_{ni}) &= \boldsymbol{Q}_{ni} - \boldsymbol{Q}^{\mathrm{S}}_{ni} - \Delta t_n a_{ii}\, \boldsymbol{r}_{\mathrm{Q}}(T_{ni}, \boldsymbol{U}_{ni}(\boldsymbol{A}_{ni}), \boldsymbol{\Theta}_{ni}(\dot{\boldsymbol{\Theta}}_{ni}), \boldsymbol{Q}_{ni}) = \boldsymbol{0}. \end{aligned} \tag{33}$$

Equation $(31)_4$ can be rearranged to

$$\dot{\boldsymbol{Q}}_{ni} = \frac{\boldsymbol{Q}_{ni} - \boldsymbol{Q}^{\mathrm{S}}_{ni}}{\Delta t_n a_{ii}}. \tag{34}$$

In other words, we have to solve the nonlinear system of equations

$$\begin{aligned} \boldsymbol{G}_{\mathrm{u}}(\boldsymbol{A}_{ni}, \dot{\boldsymbol{\Theta}}_{ni}, \boldsymbol{Q}_{ni}) &= \boldsymbol{0}, \\ \boldsymbol{G}_{\Theta}(\boldsymbol{A}_{ni}, \dot{\boldsymbol{\Theta}}_{ni}, \boldsymbol{Q}_{ni}) &= \boldsymbol{0}, \\ \boldsymbol{L}(\boldsymbol{A}_{ni}, \dot{\boldsymbol{\Theta}}_{ni}, \boldsymbol{Q}_{ni}) &= \boldsymbol{0}, \end{aligned} \tag{35}$$

at each stage $T_{ni}$ to obtain the stage quantities (i.e., the acceleration $\boldsymbol{A}_{ni}$, the temperature-rate $\dot{\boldsymbol{\Theta}}_{ni}$ and the internal variables $\boldsymbol{Q}_{ni}$). Here, we have omitted the stage time $T_{ni}$ for brevity. This is done using the multilevel Newton algorithm (MLNA) in [Rabbat et al. 1979]; regarding finite elements, see [Ellsiepen and Hartmann 2001; Hartmann 2005].

After solving these systems, the stage derivatives are known and the final solution at time $t_{n+1}$ is computable by the sum (29) and $(25)_2$–$(25)_4$.

At this point we would like to remark that for DAE systems resulting from the quasistatic case, stiffly accurate methods are preferable [Ellsiepen and Hartmann 2001; Hartmann 2002; Hartmann et al. 2008a; Birken et al. 2010; Hartmann and Rothe 2013; Rothe et al. 2015b]. Here, the solution variables $\boldsymbol{A}_{ns}$, $\dot{\boldsymbol{\Theta}}_{ns}$, $\boldsymbol{Q}_{ns}$ in the last stage coincide with the new solution at time $t_{n+1}$, $\boldsymbol{a}_{n+1} = \boldsymbol{A}_{ns}$, $\dot{\boldsymbol{\Theta}}_{n+1} = \dot{\boldsymbol{\Theta}}_{ns}$, $\boldsymbol{q}_{n+1} = \boldsymbol{Q}_{ns}$. This property guarantees that the algebraic constraints are fulfilled at the new time-step [Prothero and Robinson 1974]. With regard to the Butcher-tableaus in Figure 1(b), stiff accuracy implies a regular coefficient matrix $\mathsf{A}$ satisfying $a_{si} = b_i$ and $\bar{a}_{si} = \bar{b}_i$. From (24) and (25) it is obvious that this results in the mentioned equivalence. Thus, the latter equation can be omitted, saving additional computational time. In Section 4, we draw on stiffly accurate, diagonally implicit Runge–Kutta (SDIRK) methods for our numerical examples in order to treat the quasistatic as well as the dynamic case.

### 3.2. *Generalized-α method.*

In this section we apply the generalized-$\alpha$ method, formulated as a one-step method, which was originally introduced in [Chung and Hulbert 1993] for second-order ODEs in the field of computational solid dynamics. This unconditionally stable method allows us to introduce controllable numerical dissipation into the considered system. For a special choice of the integration parameter set consisting of the algorithmic parameters $\alpha_f$, $\alpha_m$, $\beta$ and $\gamma$, the generalized-$\alpha$ method includes the most popular classical numerical dissipative and nondissipative time integration schemes (for more details regarding the Newmark-family [Newmark 1959], the HHT-$\alpha$ method [Hilber et al. 1977] and the WBZ-$\alpha$ method [Wood et al. 1980], see [Kuhl and Crisfield 1999]). In many publications [Chung and Hulbert 1993; Jansen et al. 2000; Erlicher et al. 2002; Popp 2012], the proposed single-step version of the integrator, which comprises a special parameter set, is assumed to be second-order accurate at all times. In [Rang 2013a; 2013b], however, it was shown that this statement holds only for a particular parameter set, which is not identical to the commonly used one, since there is a distinction between the one-step and the multistep version [Erlicher et al. 2002]. These parameter sets $\{\alpha_f, \alpha_m, \beta, \gamma\}$ are often expressed in terms of a spectral radius $\rho_\infty \in [0, 1]$ as the sole free parameter

$$\alpha_f = \frac{1}{1+\rho_\infty}, \quad \alpha_m = \frac{2-\rho_\infty}{1+\rho_\infty}, \quad \gamma = \frac{1}{2} + \alpha_m - \alpha_f, \quad \beta = \frac{1}{4}(1+\alpha_m-\alpha_f)^2. \tag{36}$$

The given setting follows directly from requirements of unconditional stability, optimized numerical dissipation, and second-order accuracy for the multistep method. The spectral radius $\rho_\infty$ controls the high-frequency dissipation, whereas $\rho_\infty = 1$ designates the no dissipation case, and $\rho_\infty = 0$ means full annihilation [Chung and Hulbert 1993; Kuhl and Crisfield 1999; Jansen et al. 2000; Popp 2012]. Furthermore, $\rho_\infty$ ensures algorithmic parameters leading to optimal time integration schemes. In the linear structural dynamic regime, this leads to maximized high-frequency dissipation, while damping for the important lower modes is kept at a minimum. However, if these parameter sets are used for the one-step method, it is theoretically only possible to reach first order. The error constant in this case is very small, so the observed numerical order of convergence is two. Moreover, in [Rang 2013b], the one-step version of the generalized-$\alpha$ method applied to second and first-order ODEs obtains better results than the corresponding multistep version.

The chosen notation in this paper for the generalized-$\alpha$ method is based on the works of Rang [2013a; 2013b] and Jansen et al. [2000]. In the following, we apply the one-step version of the generalized-$\alpha$

method for second-order ODEs to the semidiscrete equations of motion $(10)_1$ and use the presented algorithmic parameter set $\{\alpha_f, \alpha_m, \beta, \gamma\}$ as a function of $\rho_\infty$. We start with linear interpolation rules for the generalized mid-point quantities which are commonly established for the generalized-$\alpha$ method,

$$\boldsymbol{U}_{n+\alpha_f} = \alpha_f \boldsymbol{u}_{n+1} + (1-\alpha_f)\boldsymbol{u}_n, \quad \dot{\boldsymbol{U}}_{n+\alpha_f} = \alpha_f \boldsymbol{v}_{n+1} + (1-\alpha_f)\boldsymbol{v}_n, \quad \ddot{\boldsymbol{U}}_{n+\alpha_m} = \alpha_m \boldsymbol{a}_{n+1} + (1-\alpha_m)\boldsymbol{a}_n. \tag{37}$$

Now, we introduce the well-known Newmark [1959] approximations to describe the discrete velocities $\boldsymbol{v}_{n+1}$ and accelerations $\boldsymbol{a}_{n+1}$ at $t_{n+1}$ as functions of the unknown displacements $\boldsymbol{u}_{n+1}$ and in terms of already known quantities at time $t_n$,

$$\boldsymbol{v}_{n+1} = \frac{\gamma}{\beta \Delta t_n}(\boldsymbol{u}_{n+1} - \boldsymbol{u}_n) - \frac{\gamma - \beta}{\beta}\boldsymbol{v}_n - \frac{\gamma - 2\beta}{2\beta}\Delta t_n \boldsymbol{a}_n, \tag{38}$$

$$\boldsymbol{a}_{n+1} = \frac{1}{\beta \Delta t_n^2}(\boldsymbol{u}_{n+1} - \boldsymbol{u}_n) - \frac{1}{\beta \Delta t_n}\boldsymbol{v}_n - \frac{1-2\beta}{2\beta}\boldsymbol{a}_n. \tag{39}$$

The generalized midpoint velocities and accelerations in (37) can be expressed as function of the displacements $\boldsymbol{u}_{n+1}$,

$$\dot{\boldsymbol{U}}_{n+\alpha_f} = \frac{\alpha_f \gamma}{\beta \Delta t_n}(\boldsymbol{u}_{n+1} - \boldsymbol{u}_n) - \frac{\alpha_f \gamma - \beta}{\beta}\boldsymbol{v}_n - \frac{(\gamma - 2\beta)\alpha_f}{2\beta}\Delta t_n \boldsymbol{a}_n, \tag{40}$$

$$\ddot{\boldsymbol{U}}_{n+\alpha_m} = \frac{\alpha_m}{\beta \Delta t_n^2}(\boldsymbol{u}_{n+1} - \boldsymbol{u}_n) - \frac{\alpha_m}{\beta \Delta t_n}\boldsymbol{v}_n - \frac{\alpha_m - 2\beta}{2\beta}\boldsymbol{a}_n. \tag{41}$$

By applying the generalized-$\alpha$ method to the semidiscrete equation of motion $(10)_1$ with

$$T_{n+\alpha_f} = t_n + \alpha_f \Delta t_n = \alpha_f t_{n+1} + (1-\alpha_f)t_n \tag{42}$$

and

$$T_{n+\alpha_m} = t_n + \alpha_m \Delta t_n = \alpha_m t_{n+1} + (1-\alpha_m)t_n, \tag{43}$$

we obtain the discrete linear momentum balance, i.e., the fully (meaning in space and time) discretized finite element formulation of nonlinear structural dynamics. This modified structural equation of motion, evaluated at some instant inside $[t_n, t_{n+1}]$, reads

$$\boldsymbol{G}_\mathrm{u}(t_{n+1}, \boldsymbol{u}_{n+1}, \boldsymbol{\Theta}_{n+1}, \boldsymbol{q}_{n+1}) = \boldsymbol{M}\ddot{\boldsymbol{U}}_{n+\alpha_m} + \boldsymbol{g}_\mathrm{u}(T_{n+\alpha_f}, \boldsymbol{U}_{n+\alpha_f}, \boldsymbol{\Theta}_{n+\alpha_f}, \boldsymbol{Q}_{n+\alpha_f}) + \boldsymbol{M}_\mathrm{up}\ddot{\overline{\boldsymbol{U}}}(T_{n+\alpha_m}) = \boldsymbol{0}, \tag{44}$$

where we have to evaluate the internal forces and external forces at the midpoint $n + \alpha_f$, which occur in $\boldsymbol{g}_\mathrm{u}$ defined in (75). This is a midpoint-type approach. In this way we obtain an equivalent equation structure and can use the same material and element routines as for the DIRK integrators. Alternatively, a linear interpolation for $\boldsymbol{g}_\mathrm{u}$ based on a trapezoidal rule can be applied [Kuhl and Crisfield 1999; Erlicher et al. 2002; Popp 2012]. For the linear case (small deformation and linear elastic material behavior) both approaches are equivalent.

At this point, we would like to briefly mention the so-called energy-momentum conserving time integration schemes [Simo and Tarnow 1992; Simo et al. 1992; Gonzalez 2000; Kuhl and Crisfield 1999; Kuhl and Ramm 1999], which are all included in an enhanced version of the generalized energy momentum method (the so-called GEMM+$\xi$ method) developed by Kuhl and Crisfield [1999]. Basically, the only difference between the GEMM+$\xi$ method and the generalized-$\alpha$ method is the way the internal forces are evaluated — or more specifically, the way for computing a generalized midpoint strain quantity.

In the following, we apply the generalized-$\alpha$ scheme for first-order ODEs formulated as a one-step method to the semidiscrete heat equation $(10)_2$. Based on the algorithmic parameter set introduced in (36), whereas $\beta$ is not used anymore, we violate the criterion of an optimized numerical dissipation in case of the generalized-$\alpha$ for first-order ODEs [Jansen et al. 2000; Rang 2013a; 2013b]. Due to this coupling, the optimal condition for $\alpha_m$ cannot be satisfied anymore. However, unconditional stability and second-order accuracy for the multistep version are still maintained. We start again with identical linear interpolation rules for the generalized mid-point temperature quantities

$$\boldsymbol{\Theta}_{n+\alpha_f} = \alpha_f \boldsymbol{\Theta}_{n+1} + (1-\alpha_f)\boldsymbol{\Theta}_n, \qquad \dot{\boldsymbol{\Theta}}_{n+\alpha_m} = \alpha_m \dot{\boldsymbol{\Theta}}_{n+1} + (1-\alpha_m)\dot{\boldsymbol{\Theta}}_n. \tag{45}$$

The discrete temperature velocities $\dot{\boldsymbol{\Theta}}_{n+1}$ at $t = t_{n+1}$ can be expressed as

$$\dot{\boldsymbol{\Theta}}_{n+1} = \frac{1}{\gamma \Delta t_n}(\boldsymbol{\Theta}_{n+1} - \boldsymbol{\Theta}_n) - \frac{1-\gamma}{\gamma}\dot{\boldsymbol{\Theta}}_n. \tag{46}$$

Due to the properties of the heat capacity matrix $\boldsymbol{C}_\Theta$, the semidiscrete heat equation $(10)_2$ can be transformed into explicit form. Applying the generalized-$\alpha$ method yields

$$\dot{\boldsymbol{\Theta}}_{n+\alpha_m} = \boldsymbol{C}_\Theta^{-1}(\boldsymbol{U}_{n+\alpha_f}, \boldsymbol{\Theta}_{n+\alpha_f})\, \boldsymbol{r}_\Theta(T_{n+\alpha_f}, \boldsymbol{U}_{n+\alpha_f}, \dot{\boldsymbol{U}}_{n+\alpha_f}, \boldsymbol{\Theta}_{n+\alpha_f}, \boldsymbol{Q}_{n+\alpha_f}), \tag{47}$$

with a capacity matrix evaluated at $T_{n+\alpha_f}$ (see [Jansen et al. 2000] for further explanation).

In order to apply the same material and element subroutines, we transfer (47) to a DIRK similar structure. By transforming $(45)_1$ to

$$\boldsymbol{\Theta}_{n+1} = \frac{1}{\alpha_f}\boldsymbol{\Theta}_{n+\alpha_f} - \frac{1-\alpha_f}{\alpha_f}\boldsymbol{\Theta}_n, \tag{48}$$

and substituting into $(45)_2$, we arrive at

$$\dot{\boldsymbol{\Theta}}_{n+\alpha_m} = \frac{1}{\Delta t_\alpha}(\boldsymbol{\Theta}_{n+\alpha_f} - \boldsymbol{\Theta}^{\mathrm{S}}_{n+\alpha_f}), \tag{49}$$

with the abbreviation $\Delta t_\alpha = \alpha_f \gamma \Delta t_n / \alpha_m$ for a time-step quantity and the starting value

$$\boldsymbol{\Theta}^{\mathrm{S}}_{n+\alpha_f} = \boldsymbol{\Theta}_n + \left(\alpha_f - \gamma\frac{\alpha_f}{\alpha_m}\right)\Delta t_n \dot{\boldsymbol{\Theta}}_n = \boldsymbol{\Theta}_n + \frac{\alpha_m - \gamma}{\gamma}\Delta t_\alpha\, \dot{\boldsymbol{\Theta}}_n, \tag{50}$$

which depends only on already calculated quantities at time $t_n$. Finally, we obtain the discrete formulation

$$\begin{aligned}\boldsymbol{G}_\Theta(t_{n+1}, \boldsymbol{u}_{n+1}, \boldsymbol{\Theta}_{n+1}, \boldsymbol{q}_{n+1}) &= \boldsymbol{C}_\Theta(\boldsymbol{U}_{n+\alpha_f}, \boldsymbol{\Theta}_{n+\alpha_f})(\boldsymbol{\Theta}_{n+\alpha_f} - \boldsymbol{\Theta}^{\mathrm{S}}_{n+\alpha_f}) \\ &\quad - \Delta t_\alpha\, \boldsymbol{r}_\Theta(T_{n+\alpha_f}, \boldsymbol{U}_{n+\alpha_f}, \dot{\boldsymbol{U}}_{n+\alpha_f}, \boldsymbol{\Theta}_{n+\alpha_f}, \boldsymbol{Q}_{n+\alpha_f}) = \boldsymbol{0}\end{aligned} \tag{51}$$

of the weak formulation in (9). Analogously to the just considered first-order ODE, the generalized-$\alpha$ method applied to the evolution equations for the internal variables $(10)_3$ can be written as

$$\boldsymbol{L}(t_{n+1}, \boldsymbol{u}_{n+1}, \boldsymbol{\Theta}_{n+1}, \boldsymbol{q}_{n+1}) = \boldsymbol{Q}_{n+\alpha_f} - \boldsymbol{Q}^{\mathrm{S}}_{n+\alpha_f} - \Delta t_\alpha\, \boldsymbol{r}_{\mathrm{Q}}(T_{n+\alpha_f}, \boldsymbol{U}_{n+\alpha_f}, \boldsymbol{\Theta}_{n+\alpha_f}, \boldsymbol{Q}_{n+\alpha_f}) = \boldsymbol{0}, \tag{52}$$

with the starting value

$$\boldsymbol{Q}^{\mathrm{S}}_{n+\alpha_f} = \boldsymbol{q}_n + \left(\alpha_f - \gamma\frac{\alpha_f}{\alpha_m}\right)\Delta t_n \dot{\boldsymbol{q}}_n = \boldsymbol{q}_n + \frac{\alpha_m - \gamma}{\gamma}\Delta t_\alpha\, \dot{\boldsymbol{q}}_n. \tag{53}$$

The derivative $\dot{\boldsymbol{q}}_{n+1}$ at $t = t_{n+1}$ is defined, analogously to (46), by

$$\dot{\boldsymbol{q}}_{n+1} = \frac{1}{\gamma\,\Delta t_n}(\boldsymbol{q}_{n+1} - \boldsymbol{q}_n) - \frac{1-\gamma}{\gamma}\dot{\boldsymbol{q}}_n. \tag{54}$$

In other words, at each point in time, we have to solve the nonlinear system of equations

$$\begin{aligned} \boldsymbol{G}_{\mathrm{u}}(\boldsymbol{u}, \boldsymbol{\Theta}, \boldsymbol{q}) &= \boldsymbol{0}, \\ \boldsymbol{G}_{\Theta}(\boldsymbol{u}, \boldsymbol{\Theta}, \boldsymbol{q}) &= \boldsymbol{0}, \\ \boldsymbol{L}_{\mathrm{Q}}(\boldsymbol{u}, \boldsymbol{\Theta}, \boldsymbol{q}) &= \boldsymbol{0}, \end{aligned} \tag{55}$$

where we have omitted the time $t_{n+1}$ and the index $n+1$ for brevity. One drawback of the procedure can be seen in (54), because we need the derivatives of the quantities at time $t = 0$. These can be obtained by evaluating (10) at the beginning of the entire computation.

**3.3.** ***Time adaptivity.*** Step-size control is an essential issue to obtain both accurate results and efficient computations. We draw on the approach in [Gustafsson 1994; Hairer et al. 1993; Hairer and Wanner 1996], by using a PI-controller approach. Furthermore, the physical meaning of the different quantities is taken into account. For details, especially for the case of the generalized-$\alpha$ method, see Appendix B.

## 4. Examples

In this section, the performance of the presented algorithms is studied by means of several numerical examples. To demonstrate the advantages and capabilities of the approach in a thermoviscoplastic finite element analysis, two different three-dimensional cases are investigated in detail. The finite element analyses are carried out using the in-house code Tasafem for monolithic, coupled and time-adaptive simulations.

**4.1.** ***Order analysis.*** In the first example, we perform an order analysis drawing on a uniaxial tensile specimen, where only one-eighth of the thermoviscoplastic steel specimen is discretized (symmetry conditions are assumed for the displacement as well as temperature distribution, see Figure 3 (left)). For the spatial discretization, we employ a mesh with $n_{\mathrm{el}} = 360$ twenty-noded hexahedral finite elements ($p = 2$) consisting of a total number $n_p = 2089$ nodes ($n_{\mathrm{uu}} = 4803$, $n_{\Theta\mathrm{u}} = 2089$, $n_{\mathrm{Q}} = n_{\mathrm{el}} \times 27 \times 12 = 116640$), see Figure 3 (left) for the mesh.

The specimen is loaded by a sinusoidal displacement-control at the middle cross-section ($y = 0$) with $\bar{u}_y = \hat{u}\sin(2\pi f t)$, $\hat{u} = 0.2\,\mathrm{mm}$, $f = 10\,\mathrm{Hz}$ within the time interval $t \in [0\,\mathrm{s}, 0.1\mathrm{s}]$. Starting from the rest position, we choose initial conditions $\mathbf{u}^h(\mathbf{x}, 0) = \mathbf{0}$, and $\Theta^h(\mathbf{x}, 0) = \Theta_0$, with an initial temperature of $\Theta_0 = 25\,°\mathrm{C}$. To ensure consistent initial conditions in the coupled field equations for the velocity and acceleration fields, we apply a smoothing step function (polynomial of fifth order) for $\bar{u}_y(t)$ within $t \in [0\,\mathrm{s}, 0.0125\,\mathrm{s}]$ for which $\mathbf{v}^h(\mathbf{x}, 0) = \mathbf{0}$ and $\mathbf{a}^h(\mathbf{x}, 0) = \mathbf{0}$ holds, see Figure 3 (right).

The chosen unit system consists of the basic units mm, t, s, and K. The material parameters used in this example can be found in Table 4, where the heat capacity $c_{\mathrm{p}}$ is given according to the unit

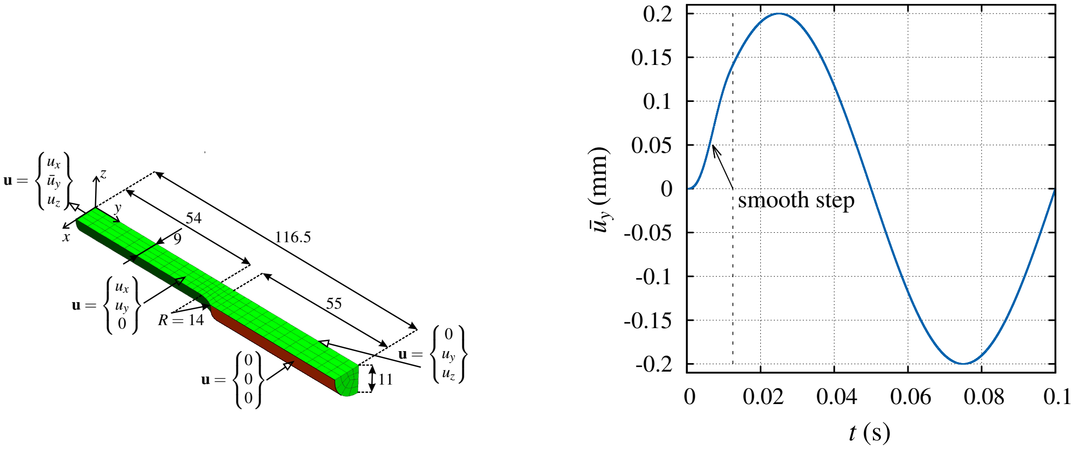

**Figure 3.** Model setup: (left) geometry, mesh and boundary conditions (units in mm); (right) loading path.

| material parameters | symbol | value | unit |
|---|---|---|---|
| bulk modulus | $K$ | $1.6666 \cdot 10^5$ | N/mm$^2$ |
| change of $K$ with temperature | $c_K$ | 91 | N/(mm$^2$K) |
| shear modulus | $G$ | $0.769 \cdot 10^5$ | N/mm$^2$ |
| change of $G$ with temperature | $c_G$ | 42 | N/(mm$^2$K) |
| thermal exp. coef. | $\alpha_\Theta$ | $1.2 \cdot 10^{-5}$ | K$^{-1}$ |
| hardening parameter | $c_X$ | 4230 | N/mm$^2$ |
| hardening parameter | $\beta$ | $3 \cdot 10^{-3}$ | N$^{-1}$mm$^{-2}$ |
| initial yield stress | $k_0$ | 450 | N/mm$^2$ |
| yield stress at high temperature | $k_H$ | 100 | N/mm$^2$ |
| yield stress slope | $b$ | $4.2 \cdot 10^{-3}$ | C$^{-1}$ |
| viscosity | $\eta$ | $6 \cdot 10^4$ | s |
| exponent | $m$ | 1 | - |
| normalization stress | $\sigma_0$ | 1 | N/mm$^2$ |
| density | $\rho$ | $7.836 \cdot 10^{-9}$ | Ns$^2$/mm$^4$ |

**Table 4.** Thermoviscoplastic material parameters.

system in mm$^2$/(s$^2$K). Thus, the conversion factor between SI units J/(kgK) and the chosen system is $1\,\text{mm}^2/(\text{s}^2\text{K}) = 10^{-6}\,\text{J}/(\text{kgK})$. Instead of choosing the heat capacity given by (5), we choose a temperature-dependent approach that originates from differential scanning calorimetry measurements. In order to take into consideration the phase transformation at higher temperatures, a log-interpolation of the two functions

$$c_{p1}(\Theta) = a_1 e^{a_2 \Theta} + a_3, \qquad c_{p2}(\Theta) = a_4 e^{-a_5(\Theta - \tilde{\Theta}_0)} + a_6 \Theta, \tag{56}$$

| DIRK | | | | |
|---|---|---|---|---|
| **abbreviation** | **method** | **order $p$** | **stages $s$** | **reference** |
| BE | Backward Euler | 1 | 1 | [Butcher 2008] |
| Alex2/Ell | Alexander/Ellsiepen | 2 | 2 | [Alexander 1977; Ellsiepen 1999] |
| CAA | trapezoidal rule or CAA | 2 | 2 | [Hughes 1987; Ellsiepen and Hartmann 2001] |
| Alex3/Cash | Alexander/Cash | 3 | 3 | [Alexander 1977; Cash 1979] |
| CAA+BE | CAA + Backward Euler | 2/1 | 2/1 | - |

| generalized-$\alpha$ | | | | |
|---|---|---|---|---|
| **abbreviation** | **method** | **order $p$** | **stages $s$** | **reference** |
| Gen-$\alpha$ | generalized-$\alpha$, $\rho_\infty = 0.8$ | 2 | - | [Chung and Hulbert 1993] |

**Table 5.** Applied one-step methods with references in the context of ODE solution and applications within finite elements.

is chosen according to Kreisselmeier and Steinhauser [1979]

$$c_{\mathrm{p}}(\Theta) = -c_W \ln\left(\frac{e^{-c_{p1}(\Theta)/c_W} + e^{-c_{p2}(\Theta)/c_W}}{2}\right). \tag{57}$$

The calibrated parameters are as follows: $a_1 = 34.2\,\mathrm{J/(kgK)}$, $a_2 = 0.0026\,\mathrm{K}^{-1}$, $a_3 = 421.15\,\mathrm{J/(kgK)}$, $a_4 = 956.5\,\mathrm{J/(kgK)}$, $a_5 = 0.012\,\mathrm{K}^{-1}$, $a_6 = 0.45\,\mathrm{K/(kgK^2)}$, and $\tilde{\Theta}_0 = 900\,\mathrm{K}$. The weighting factor of the interpolation function is chosen to be $c_W = 30\,\mathrm{J/(kgK)}$.

The thermal conductivity $\kappa$ is given in tmm/(s$^3$K), which is equivalent to W/(mK). According to Quint et al. [2011] we assume

$$\kappa(\Theta) = b_0 + b_1\Theta + b_2\Theta^2 + b_3\Theta^3, \tag{58}$$

with $b_0 = 40.1\,\mathrm{W/(mK)}$, $b_1 = 0.05\,\mathrm{W/(mK^2)}$, $b_2 = -10^{-4}\,\mathrm{W/(mK^3)}$, and $b_3 = 4.9 \cdot 10^{-8}\,\mathrm{W/(mK^4)}$, which are taken from the steel 51CrV4.

For the order considerations, the relative error measures of the displacement, velocity, temperature field and internal variables are computed. The relative error quantities are defined by the maximum relative error over all points in time $t_n$, given as

$$\mathbf{err}_{\boldsymbol{u}} = \max_n \left(\frac{\|\boldsymbol{u}_n^{\mathrm{ref}} - \boldsymbol{u}_n\|}{\max_n\left(\|\boldsymbol{u}_n^{\mathrm{ref}}\|\right)}\right), \qquad \mathbf{err}_{\boldsymbol{v}} = \max_n \left(\frac{\|\boldsymbol{v}_n^{\mathrm{ref}} - \boldsymbol{v}_n\|}{\max_n\left(\|\boldsymbol{v}_n^{\mathrm{ref}}\|\right)}\right), \tag{59}$$

$$\mathbf{err}_{\boldsymbol{\Theta}} = \max_n \left(\frac{\|\boldsymbol{\Theta}_n^{\mathrm{ref}} - \boldsymbol{\Theta}_n\|}{\max_n\left(\|\boldsymbol{\Theta}_n^{\mathrm{ref}}\|\right)}\right), \qquad \mathbf{err}_{\boldsymbol{q}} = \max_n \left(\frac{\|\boldsymbol{q}_n^{\mathrm{ref}} - \boldsymbol{q}_n\|}{\max_n\left(\|\boldsymbol{q}_n^{\mathrm{ref}}\|\right)}\right), \tag{60}$$

For the illustrated problem the integration methods compiled in Table 5 are analyzed in view of their expected and achieved temporal convergence order for different fields. For further information regarding the applied DIRK methods, we refer to [Ellsiepen and Hartmann 2001] and the literature cited in Table 5. Cash's [1979] method is the time-adaptive extension using an embedded scheme of Alexander's [1977]

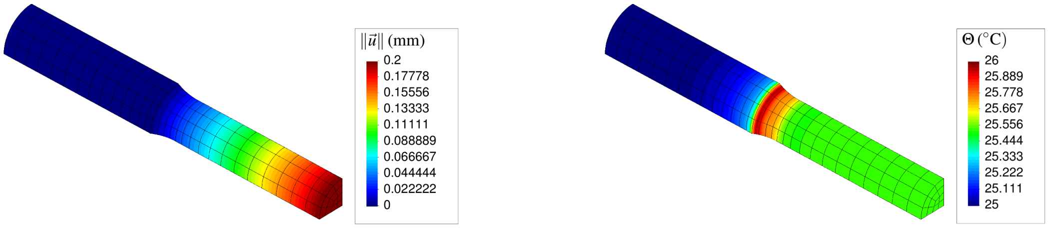

**Figure 4.** Displacement norm (left) and temperature distribution (right) for reference solution (Alex3/Cash, $\Delta t_n = 10^{-5}$ s) at $t = 0.075$ s.

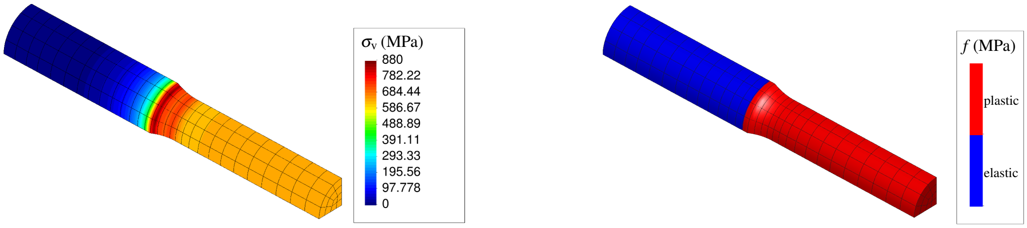

**Figure 5.** von Mises-stress distribution (left) and yield function state (right) for reference solution (Alex3/Cash, $\Delta t_n = 10^{-5}$ s) at $t = 0.075$ s.

method of third order. Similarly, Ellsiepen's [1999] method is the time-adaptive extension of Alexander's [1977] second-order method. In the following we call them Alex3/Cash and Alex2/Ell.

The algorithmic parameters of the generalized-$\alpha$ method — $\alpha_f$, $\alpha_m$, $\beta$ and $\gamma$ — are computed using (36) as function of the spectral radius $\rho_\infty$. The order of convergence is given by the slope of the curves in a double logarithmic plot. The reference solution is given by Alex3/Cash (highest-order method of our investigations), with a constant step-size of $\Delta t_n = 10^{-5}$ s. The resulting deformation and temperature distribution for the reference solution at $t = 0.075$ s are shown in Figure 4. Mechanical agencies of the structure change the temperature distribution due to plastic dissipation, and cause plastic deformations, see Figure 5.

Regarding the order considerations for the displacement field, only the Backward Euler (which is the simplest DIRK method), the applied Newmark-scheme in the parametrization of the trapezoidal rule (CAA), and the generalized-$\alpha$ method reach their theoretical orders in Figure 6 (left). The convergence order for Alex2/Ell and Alex3/Cash degenerates after passing an optimal step-size. The reasons for this are not known.

Due to the strong coupling of the displacement field to the internal variables, the combination of the CAA and Backward Euler methods yields only order one, i.e., the order of the global trapezoidal rule (Newmark-scheme) will be reduced, and the artificially created order reduction in the internal variables affects the order in the displacement field. Thus, commercial programs using the Newmark-method combined with a Backward Euler-like scheme on Gauss-point level for integrating the internal variables cannot reach second order. This is different in the pure CAA approach.

For the convergence behavior in the velocity field, the higher-order DIRK methods — excluding the Newmark-type schemes (CAA, generalized-$\alpha$ method) — reach only first order (see Figure 6 (right)).

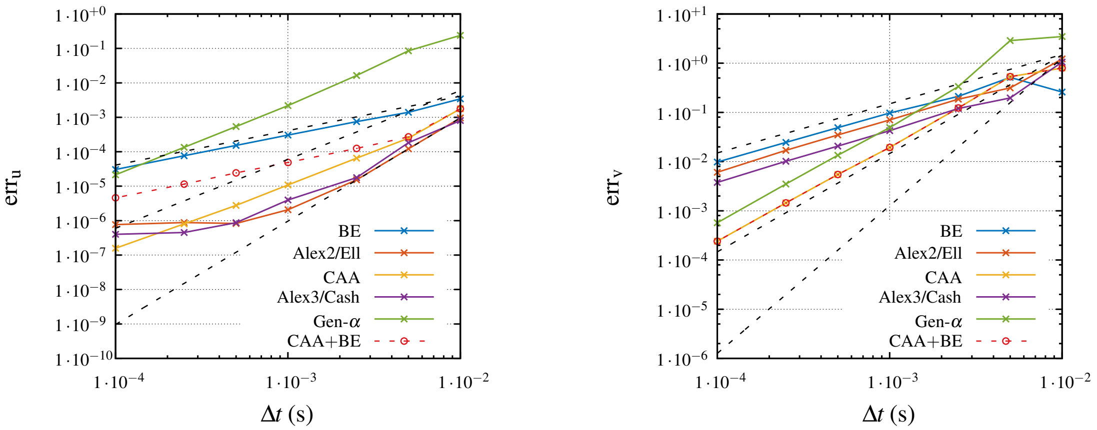

**Figure 6.** Convergence behavior for order analysis: (left) global error in displacement field and (right) global error in velocity field.

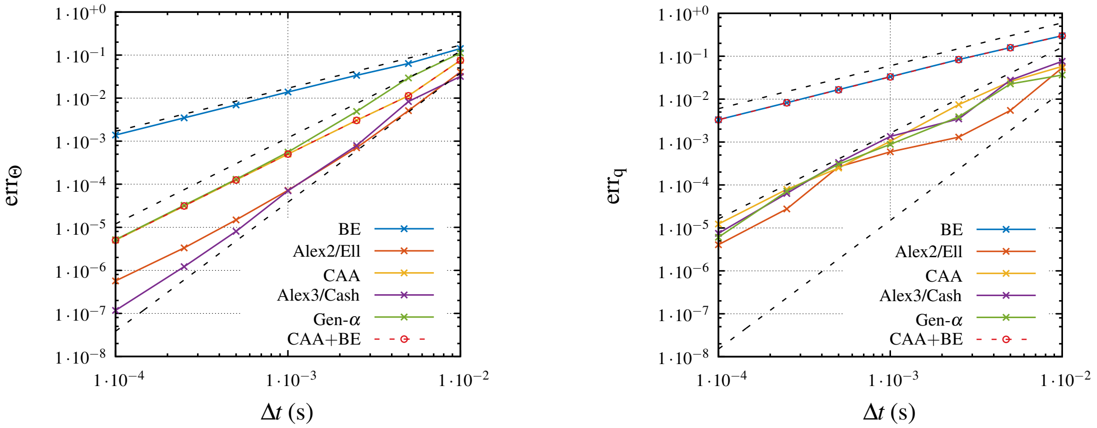

**Figure 7.** Convergence behavior for order analysis: (left) global error in temperature field and (right) global error for internal variables.

However, the field variable velocity is only of minor importance compared to the others. On the other hand, the convergence behavior for the temperature field in Figure 7 (left) for the different time integration methods matches the theory. As reported in [Ellsiepen 1999; Ellsiepen and Hartmann 2001], we observe in the convergence graph for the internal variables (Figure 7 (right)) — due to the lack of smoothness in the time domain — an order reduction where the third order of the Alex3/Cash method is not attained. This is known from quasistatic and isothermal computations [Ellsiepen and Hartmann 2001]. As already mentioned, the combination of Backward Euler for the internal variables (stress-algorithm) and the Newmark-method yields only order one.

It is not only the rate of convergence that is of practical importance, but the computational costs of the methods as well. Figures 8 and 9 show the relative error in each field variable versus the computational time.

In these plots, the second-order DIRK methods (Alex2/Ell, CAA), with the exception of Figure 8 (right), show an identical slope, which is steeper than the more accurate Alex3/Cash-method. They are more

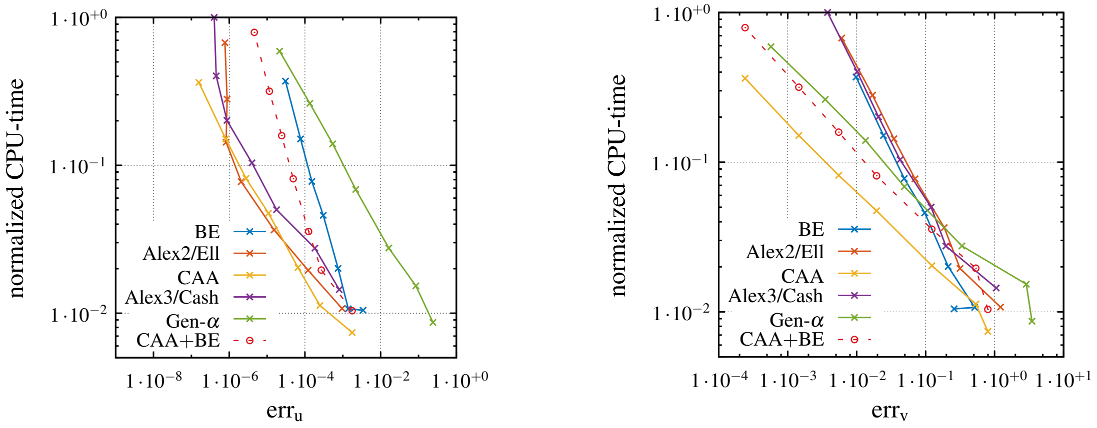

**Figure 8.** Efficiency analyses for displacement (left) and velocity (right) fields.

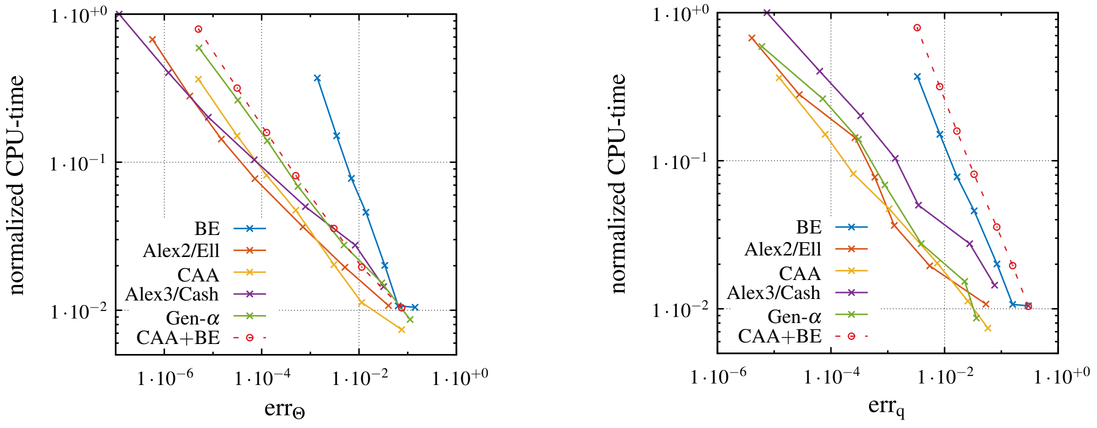

**Figure 9.** Efficiency analyses for temperature field (left) and internal variables (right).

efficient for high accuracy requirements, i.e., the fastest methods for a given error tolerance. The Backward Euler method yields only comparatively rough results — even for high computational times. The considered generalized-$\alpha$ method behaves only moderately: acceptable results are only obtained for the smallest time-step, see Figure 8 (left). All in all, the combination of Newmark-method and Backward Euler seems to be superior to the pure Backward Euler method, but is not as attractive as the generalized-$\alpha$ method due to the small slope in the displacements.

We would like to remark that the increase of elements by a factor of two in all directions, which yields more accurate results in the spatial domain, does not influence the order diagrams, especially the order reduction phenomenon of CAA and Alex3/Cash in Figure 6 (left). This can be explained by the accuracy of the error tolerance of the global MLNA step, where the tolerances (norm of $\Delta\boldsymbol{u}$ and $\Delta\boldsymbol{\Theta}$) could not be smaller than $10^{-7}$ in the computations. Otherwise, the MLNA will not converge.

**4.2.** ***H-Beam.*** In the second example, we study the dynamical behavior of a thermoviscoplastic H-Beam profile by performing time-adaptive computations. The error of the current time-step is estimated by using embedding methods as described in Section 3.3. The dimensions of the profile are depicted in

| $\varepsilon_{\mathrm{a}}^{u}$ | $\varepsilon_{\mathrm{r}}^{u}$ | $\varepsilon_{\mathrm{a}}^{\Theta}$ | $\varepsilon_{\mathrm{r}}^{\Theta}$ | $\varepsilon_{\mathrm{a}}^{q}$ | $\varepsilon_{\mathrm{r}}^{q}$ |
|---|---|---|---|---|---|
| $1.0 \cdot 10^{-2}$ mm | $1.0 \cdot 10^{-3}$ | $1.0 \cdot 10^{-2}$ °C | $1.0 \cdot 10^{-3}$ | $1.0 \cdot 10^{-5}$ | $1.0 \cdot 10^{-5}$ |

**Table 6.** Absolute and relative error tolerances for step-size estimation.

Figure 10 (left) and the material parameters are given in Table 4 with the nonlinear material functions (57) and (58). The left end of the beam is clamped, i.e., all displacements are fixed, and the right side is loaded by a time-dependent cyclic pressure load in $z$-direction,

$$\bar{p}(t) = \begin{cases} K_p t & \text{for } 0\,\mathrm{s} \le t \le 1\,\mathrm{s} \text{ with } K_p = 130\,\mathrm{MPa/s} \\ p_0 \sin(2\pi f t) & \text{for } 1\,\mathrm{s} < t \le 2\,\mathrm{s} \text{ with } p_0 = 20\,\mathrm{MPa},\ f = 233\,\mathrm{Hz} \end{cases}, \tag{61}$$

see Figure 10 (right). The frequency is chosen to be close to an eigenfrequency so that the influence of the inertia terms becomes larger. Similar computations — which are not shown here — at a loading frequency of $f = 60$ Hz show very similar amplitudes and phases for all time integration schemes. This holds for the step-size behavior as well. The whole profile is discretized with $n_{\mathrm{el}} = 11200$ twenty-noded hexahedral finite elements. Consequently, this implies a total number of 68591 nodes ($n_{\mathrm{uu}} = 204300$, $n_{\Theta\mathrm{u}} = 68591$) with $n_{\mathrm{Q}} = n_{\mathrm{el}} \times 27 \times 12 = 3628800$ internal variables. The problem setup is completed by the initial conditions $\mathbf{u}^h(\mathbf{x}, 0) = \mathbf{0}$, $\mathbf{v}^h(\mathbf{x}, 0) = \mathbf{0}$, $\mathbf{q}(0) = \mathbf{0}$, and initial temperature distribution $\Theta^h(\mathbf{x}, 0) = \Theta_0$ with reference temperature $\Theta_0 = 25\,°\mathrm{C}$.

For all investigated time-integration methods, namely the generalized-$\alpha$ (Gen-$\alpha$), Newmark (CAA), and Ellsiepen (Alex2/Ell) time integrators, the same absolute and relative error tolerances for the field variable's displacement, temperature and internal variables are used (see Table 6) in order to estimate the new step-sizes. The proposed adaptive time-step control uses the error measure $\mathrm{e}_m = \max(\mathrm{e}_u, \mathrm{e}_v, \mathrm{e}_\theta, \mathrm{e}_q)$ for the step-size selection rule (95). As a result of the observed poor convergence behavior in the velocity field for the DIRK methods Alex2/Ell and Alex3/Cash (see Figure 6 (right)), we do not take $\mathrm{e}_v$ into account to compute $\mathrm{e}_m$. This — in combination with the chosen factors $f_{\min} = 0.3$, $f_{\max} = 2.5$, and

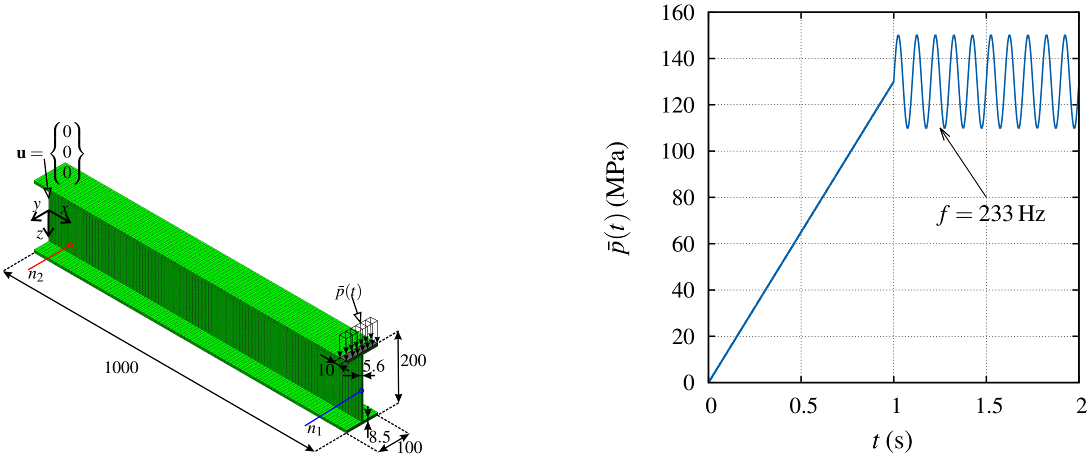

**Figure 10.** Model setup and loading process: dimensions of the H-Beam, mesh and boundary conditions (units in mm) (left); schematic loading path (right).

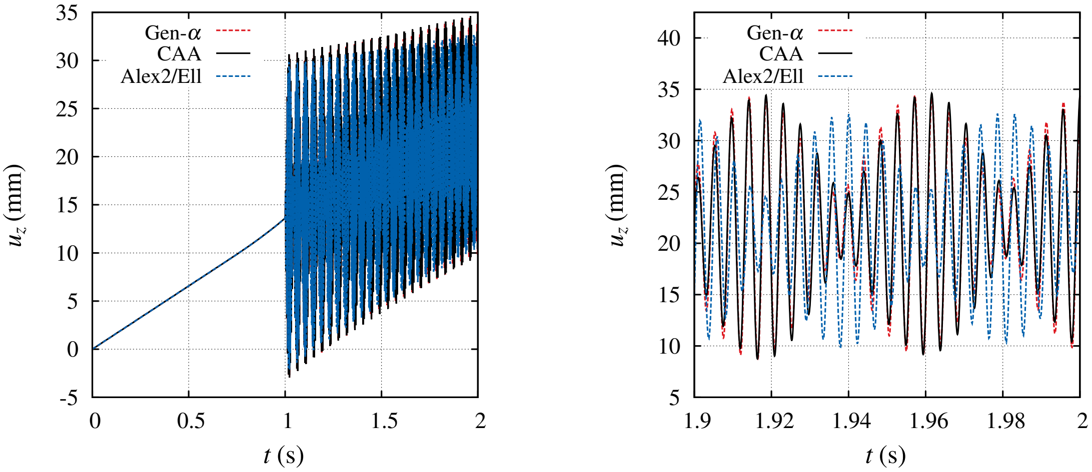

**Figure 11.** Displacement result for evaluation point $n_1$ of Figure 10: deflection in $z$-direction ($0 \leq t \leq 2\,\mathrm{s}$) (left); deflection in $z$-direction (subinterval $1.9\,\mathrm{s} \leq t \leq 2\,\mathrm{s}$) (right).

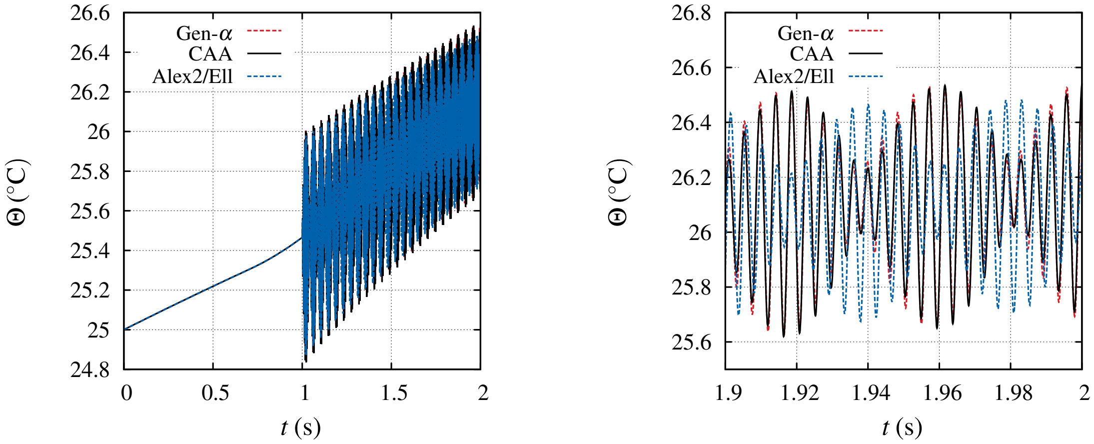

**Figure 12.** Temperature results for evaluation point $n_2$ shown in Figure 10: temperature evolution ($0 \leq t \leq 2\,\mathrm{s}$) (left); temperature evolution (subinterval $1.9\,\mathrm{s} \leq t \leq 2\,\mathrm{s}$) (right).

$f_{\text{safety}} = 0.85$ of the step-size controller — leads to the step-size behavior shown in Figure 15 (left). An initial step-size of $\Delta t_0 = 10^{-2}\,\mathrm{s}$ is chosen and applied at each external load-change, i.e., at $t_n = 0\,\mathrm{s}$ and $t_n = 1\,\mathrm{s}$. In the quasistatic loading range the step-size increases until it is limited by the stability of the employed Multilevel Newton algorithm. If there are time-steps that are too large, a failure of the applied Newton algorithm on the local level is obtained. Thus, step-size rejections are observable. As the dynamics of higher modes of the structure are stimulated by the cyclic loading, the step-size in this time interval is reduced significantly by the time-adaptive scheme.

For a comparison of the integration schemes, we consider the structural response at two nodes $n_1$ and $n_2$. Figure 11 shows the deflection/time behavior at node $n_1$. At the second evaluation point $n_2$, however, we observe the temperature evolution depicted in Figure 12. The generalized-$\alpha$ method and the applied Newmark-scheme (CAA) yield approximately the same structural response, see Figure 11 (right) and

| Method | No. of time-steps | CPU time (s) | Efficiency factor |
|---|---|---|---|
| Gen-$\alpha$ | 12592 | 253136 | 0.54 |
| CAA | 12601 | 467034 | 1.00 |
| Alex2/Ell | 4107 | 106077 | 0.23 |

**Table 7.** Computational costs of applied time integration schemes.

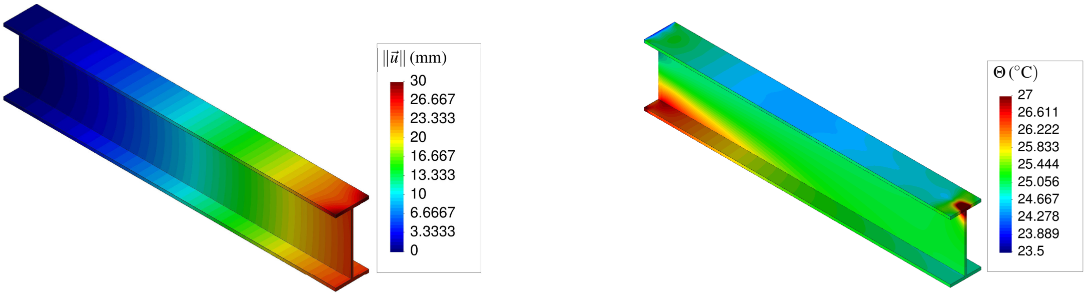

**Figure 13.** Displacement (left) and temperature (right) distribution at the end of the computation ($t = 2$ s), computed with Ellsiepen's method.

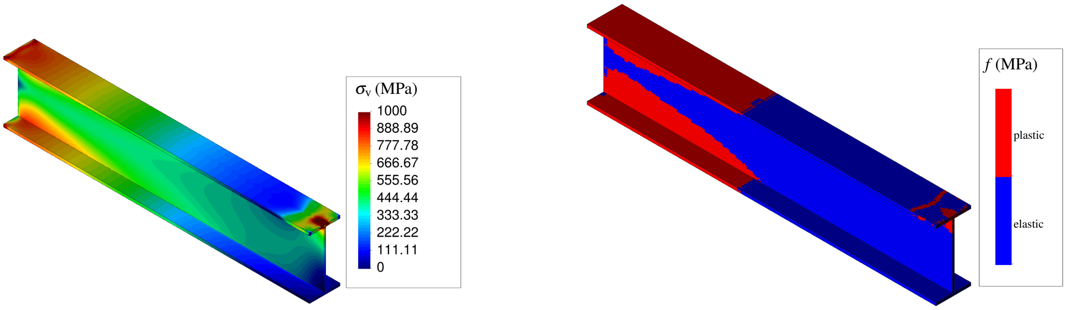

**Figure 14.** von Mises stress (left) and yield function (right) state at the end of the computation ($t = 2$ s), computed with Ellsiepen's method.

Figure 12 (right). Moreover, both show a very similar step-size behavior which can be observed in Figure 15. However, as shown by the CPU times in Table 7, the generalized-$\alpha$ requires only 54% of the computational time of the trapezoidal rule (CAA). This fact is due to the use of a further stage in the embedded scheme for the trapezoidal rule (a Backward Euler stage) for applying the proposed time-adaptivity procedure in Section 3. The generalized-$\alpha$ method uses an estimation introduced by Rang [2013a; 2013b]. Ellsiepen's method yields larger step-sizes in the cyclic loading range and, accordingly, leads to the fastest computation.

Figures 13 and 14 (left) show the displacement, temperature and von Mises stress distribution at the final stage of a time-adaptive computation with the Ellsiepen integrator. In Figure 14 (right), we consider the regions where yielding appears. When the cyclic load is applied, a plastic zone starts to evolve at the upper and the lower surface of the clamping. One can clearly see that in the area of clamping — above and

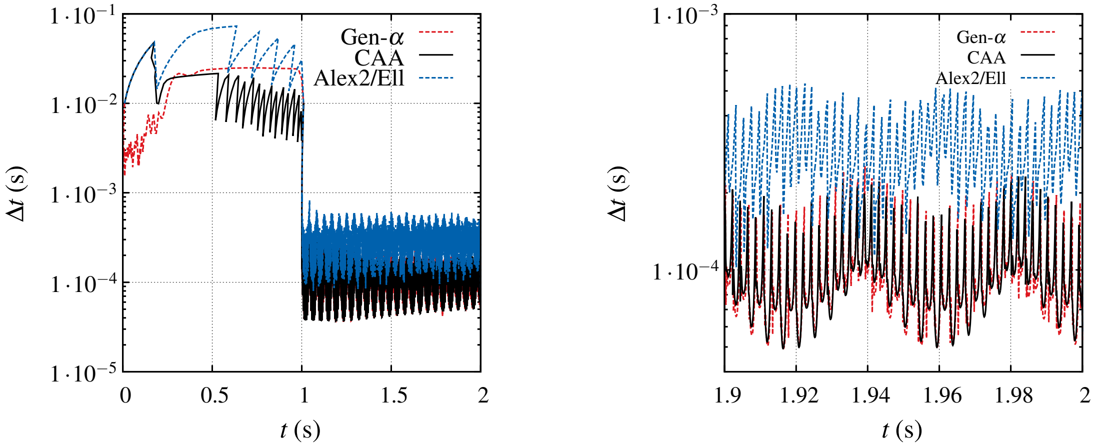

**Figure 15.** Step-size behavior: (left) $0\,\mathrm{s} \le t \le 2\,\mathrm{s}$; (right) subinterval $1.9\,\mathrm{s} \le t \le 2\,\mathrm{s}$.

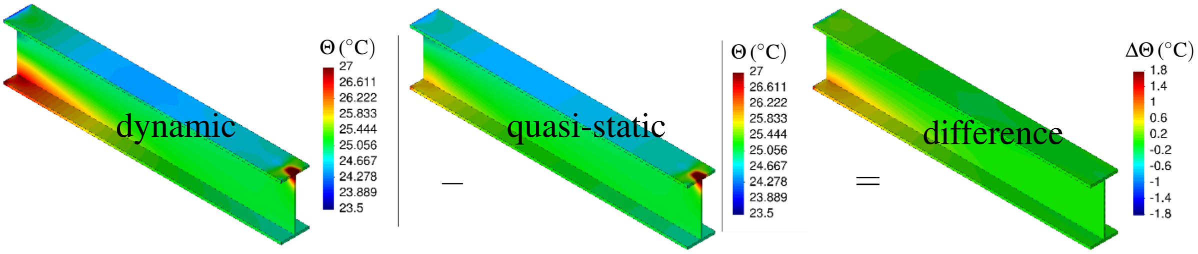

**Figure 16.** Temperature difference between dynamic and quasistatic computation at the end of the computation ($t = 2\,\mathrm{s}$), computed with Ellsiepen's method.

below, where a singularity exists — the highest stresses occur. Furthermore, due to the elastic coupling effect in the constitutive model, the maximum temperature occurs in the lower part of the clamping.

Figure 16 shows the influence of the inertia terms on the temperature distribution by comparing a dynamical simulation with a quasistatic one. For this purpose, we perform an identical quasistatic simulation by ignoring the inertia term and computing the difference between both simulations. The resulting DAE system of the quasistatic simulation is listed in (14). It can clearly be seen that in the lower part of the clamping, the temperature differs and reveals a temperature increase. In the remaining part of the structure, there is no drastic change in temperature.

## 5. Conclusions

We studied the resulting ODE system of dynamical problems which were combined with heat conduction problems and inelastic material behavior. All equations are coupled. The problem under consideration is related to the computation of a system of second-order ODEs with first-order ODEs. The classical Newmark-method, which is a method of second-order combined with stress-algorithms of first order on Gauss-point level, does not lead to a second-order scheme. Thus, other methods must be treated. Another question is connected to an automatic step-size control technique based on local error estimations, due to which the physical quantities and their temporal evolution determine the step-size.

Apart from the compilation of different semidiscrete forms, we investigated stiffly accurate, diagonally implicit Runge–Kutta methods and the generalized-$\alpha$ method (single-step version). In the case of DIRK methods, the second-order ODE part is transformed into a system of first-order ODEs. A particular discussion of the relation to Runge–Kutta–Nyström methods was offered as well. The generalized-$\alpha$ method, originally developed for second-order ODEs, was combined with a generalized-$\alpha$ method for first-order ODEs. The different schemes were studied in view of their order of accuracy revealing similar behavior. Since an order-reduction is known for the case of yield function-based models, only second-order accuracy was expected. This was obtained for all methods in the displacements, temperatures and internal variables. Only the fully consistently applied trapezoidal scheme (CAA) — consistent with the integration scheme on Gauss-point level to determine the internal variables — the expected order is achieved in the velocities as well.

Additional investigations of the embedded time-step control of the DIRK methods and the generalized-$\alpha$ methods proved both the applicability of the time-adaptivity of the two methods as well as the greater efficiency of DIRK methods. Unfortunately, the fully consistent CAA method is computationally four times slower than the DIRK approach in our examples. Thus, either Ellsiepen's method, which has turned out to be very efficient for quasistatic and coupled problems, or the generalized-$\alpha$ should be applied. The latter method is also applicable for problems with internal variables if it is applied in a consistent manner. However, there is the drawback that the function evaluations are at stage-times within a time-interval so that constraint problems cannot be satisfied.

## Appendix A. Derivation of matrices

In this section, we draw on the notation used in [Rothe et al. 2015a]. In finite elements, the displacements $\vec{u}(\vec{x}, t)$, the virtual displacements $\delta\vec{u}(\vec{x})$, the temperatures $\Theta(\vec{x}, t)$, and the virtual temperatures $\delta\Theta(\vec{x})$ are approximated by shape functions within an element $e$,

$$\mathsf{u}^h(\mathsf{x}, t) = \mathsf{N}^e_{\mathrm{u}}(\boldsymbol{\varphi}^e_{\mathrm{u}}(\mathsf{x}))\mathsf{u}^e(t) = \mathsf{N}^e_{\mathrm{u}}(\boldsymbol{\varphi}^e_{\mathrm{u}}(\mathsf{x}))\boldsymbol{Z}^{\,e}_{\mathrm{ua}}\boldsymbol{U}_{\mathrm{a}} = \mathsf{N}^e_{\mathrm{u}}(\boldsymbol{\varphi}^e_{\mathrm{u}}(\mathsf{x}))\big\{\boldsymbol{Z}^{\,e}_{\mathrm{u}}\boldsymbol{u} + \overline{\boldsymbol{Z}}{}^{\,e}_{\mathrm{u}}\bar{\boldsymbol{u}}(t)\big\}, \tag{62}$$

$$\dot{\mathsf{u}}^h(\mathsf{x}, t) = \mathsf{N}^e_{\mathrm{u}}(\boldsymbol{\varphi}^e_{\mathrm{u}}(\mathsf{x}))\dot{\mathsf{u}}^e(t) = \mathsf{N}^e_{\mathrm{u}}(\boldsymbol{\varphi}^e_{\mathrm{u}}(\mathsf{x}))\boldsymbol{Z}^{\,e}_{\mathrm{ua}}\dot{\boldsymbol{U}}_{\mathrm{a}} = \mathsf{N}^e_{\mathrm{u}}(\boldsymbol{\varphi}^e_{\mathrm{u}}(\mathsf{x}))\big\{\boldsymbol{Z}^{\,e}_{\mathrm{u}}\dot{\boldsymbol{u}} + \overline{\boldsymbol{Z}}{}^{\,e}_{\mathrm{u}}\dot{\bar{\boldsymbol{u}}}(t)\big\}, \tag{63}$$

$$\ddot{\mathsf{u}}^h(\mathsf{x}, t) = \mathsf{N}^e_{\mathrm{u}}(\boldsymbol{\varphi}^e_{\mathrm{u}}(\mathsf{x}))\ddot{\mathsf{u}}^e(t) = \mathsf{N}^e_{\mathrm{u}}(\boldsymbol{\varphi}^e_{\mathrm{u}}(\mathsf{x}))\boldsymbol{Z}^{\,e}_{\mathrm{ua}}\ddot{\boldsymbol{U}}_{\mathrm{a}} = \mathsf{N}^e_{\mathrm{u}}(\boldsymbol{\varphi}^e_{\mathrm{u}}(\mathsf{x}))\big\{\boldsymbol{Z}^{\,e}_{\mathrm{u}}\ddot{\boldsymbol{u}} + \overline{\boldsymbol{Z}}{}^{\,e}_{\mathrm{u}}\ddot{\bar{\boldsymbol{u}}}(t)\big\}, \tag{64}$$

$$\delta\mathsf{u}^h(\mathsf{x}) = \mathsf{N}^e_{\mathrm{u}}(\boldsymbol{\varphi}^e_{\mathrm{u}}(\mathsf{x}))\delta\mathsf{u}^e = \mathsf{N}^e_{\mathrm{u}}(\boldsymbol{\varphi}^e_{\mathrm{u}}(\mathsf{x}))\boldsymbol{Z}^{\,e}_{\mathrm{ua}}\delta\boldsymbol{U}_{\mathrm{a}} = \mathsf{N}^e_{\mathrm{u}}(\boldsymbol{\varphi}^e_{\mathrm{u}}(\mathsf{x}))\boldsymbol{Z}^{\,e}_{\mathrm{u}}\delta\boldsymbol{u}, \tag{65}$$

$$\Theta^h(\mathsf{x}, t) = \mathsf{N}^{e\,T}_{\Theta}(\boldsymbol{\varphi}^e_{\Theta}(\mathsf{x}))\boldsymbol{\Theta}^e(t) = \mathsf{N}^{e\,T}_{\Theta}(\boldsymbol{\varphi}^e_{\Theta}(\mathsf{x}))\boldsymbol{Z}^{\,e}_{\Theta\mathrm{a}}\boldsymbol{\Theta}_{\mathrm{a}} = \mathsf{N}^{e\,T}_{\Theta}(\boldsymbol{\varphi}^e_{\Theta}(\mathsf{x}))\big\{\boldsymbol{Z}^{\,e}_{\Theta}\boldsymbol{\Theta} + \overline{\boldsymbol{Z}}{}^{\,e}_{\Theta}\overline{\boldsymbol{\Theta}}(t)\big\}, \tag{66}$$

$$\dot{\Theta}^h(\mathsf{x}, t) = \mathsf{N}^{e\,T}_{\Theta}(\boldsymbol{\varphi}^e_{\Theta}(\mathsf{x}))\dot{\boldsymbol{\Theta}}^e(t) = \mathsf{N}^{e\,T}_{\Theta}(\boldsymbol{\varphi}^e_{\Theta}(\mathsf{x}))\boldsymbol{Z}^{\,e}_{\Theta\mathrm{a}}\dot{\boldsymbol{\Theta}}_{\mathrm{a}} = \mathsf{N}^{e\,T}_{\Theta}(\boldsymbol{\varphi}^e_{\Theta}(\mathsf{x}))\big\{\boldsymbol{Z}^{\,e}_{\Theta}\dot{\boldsymbol{\Theta}} + \overline{\boldsymbol{Z}}{}^{\,e}_{\Theta}\dot{\overline{\boldsymbol{\Theta}}}(t)\big\}, \tag{67}$$

$$\delta\Theta^h(\mathsf{x}) = \mathsf{N}^{e\,T}_{\Theta}(\boldsymbol{\varphi}^e_{\Theta}(\mathsf{x}))\delta\boldsymbol{\Theta}^e = \mathsf{N}^{e\,T}_{\Theta}(\boldsymbol{\varphi}^e_{\Theta}(\mathsf{x}))\boldsymbol{Z}^{\,e}_{\Theta\mathrm{a}}\delta\boldsymbol{\Theta}_{\mathrm{a}} = \mathsf{N}^{e\,T}_{\Theta}(\boldsymbol{\varphi}^e_{\Theta}(\mathsf{x}))\boldsymbol{Z}^{\,e}_{\Theta}\delta\boldsymbol{\Theta}, \tag{68}$$

where $\mathsf{u}^h \in \mathbb{R}^3$, $\delta\mathsf{u}^h \in \mathbb{R}^3$. For the three-dimensional case we have the matrix and vector shape functions $\mathsf{N}^e_{\mathrm{u}} \in \mathbb{R}^{3\times n^e_{\mathrm{u}}}$, and $\mathsf{N}^e_{\Theta} \in \mathbb{R}^{n^e_{\Theta}}$, where $n^e_{\mathrm{u}}$ and $n^e_{\Theta}$ are the displacement and temperature element degrees of freedom (DOF). Here, $\mathsf{u}^e \in \mathbb{R}^{n^e_{\mathrm{u}}}$ and $\boldsymbol{\Theta}^e \in \mathbb{R}^{n^e_{\Theta}}$ describe the element DOF. In the case of $h$-elements, these are displacements and nodal temperatures. The expression $\boldsymbol{\xi} = \boldsymbol{\varphi}^e_{\mathrm{u}}(\mathsf{x})$ defines the inverse coordinate transformation of the global coordinates $\mathsf{x}$ to the local coordinates $\boldsymbol{\xi}$ in the reference element. Here, we have to distinguish between the coordinate transformation $\boldsymbol{\varphi}_{\mathrm{u}}^e$ in the mechanical problem and the coordinate transformation in the thermal problem $\boldsymbol{\xi} = \boldsymbol{\varphi}_{\Theta}^e(\mathsf{x})$. In a monolithic approach, both transformations are assumed to be identical, $\boldsymbol{\varphi}_{\mathrm{u}}^e(\mathsf{x}) = \boldsymbol{\varphi}_{\Theta}^e(\mathsf{x})$ (of course, a generalization might be possible). In a partitioned approach both meshes can be different. Thus, the given coordinate transformations $\mathsf{x} = \boldsymbol{\chi}_{\mathrm{u}}^e(\boldsymbol{\xi})$ and $\mathsf{x} = \boldsymbol{\chi}_{\Theta}^e(\boldsymbol{\xi})$ with $\boldsymbol{\chi}_{\mathrm{u}}^e = \boldsymbol{\varphi}_{\mathrm{u}}^{e\,-1}$ and $\boldsymbol{\chi}_{\Theta}^e = \boldsymbol{\varphi}_{\Theta}^{e\,-1}$ are distinct in the general case. In this context, $\boldsymbol{u}_{\mathrm{a}} \in \mathbb{R}^{n_{\mathrm{ua}}}$ symbolizes all displacement DOF for the entire structure containing the unknown quantities $\boldsymbol{u} \in \mathbb{R}^{n_{\mathrm{uu}}}$ and the known (prescribed) displacement DOF $\overline{\boldsymbol{u}} \in \mathbb{R}^{n_{\mathrm{up}}}$. Obviously, $n_{\mathrm{ua}} = n_{\mathrm{uu}} + n_{\mathrm{up}}$ for $\boldsymbol{u}_{\mathrm{a}}^T = \{\boldsymbol{u}^T \overline{\boldsymbol{u}}^T\}$ holds. The same decomposition is carried out for the temperature DOF, $\boldsymbol{\Theta}_{\mathrm{a}} \in \mathbb{R}^{n_{\Theta\mathrm{a}}}$, $\boldsymbol{\Theta} \in \mathbb{R}^{n_{\Theta\mathrm{u}}}$, $\overline{\boldsymbol{\Theta}} \in \mathbb{R}^{n_{\Theta\mathrm{p}}}$ (i.e., $\boldsymbol{\Theta}_{\mathrm{a}}^T = \{\boldsymbol{\Theta}^T \overline{\boldsymbol{\Theta}}^T\}$ and $n_{\Theta\mathrm{a}} = n_{\Theta\mathrm{u}} + n_{\Theta\mathrm{p}}$). For $n_{ij}$ the first index denotes the mechanical and the thermal quantities, $i = \mathrm{u}, \Theta$. The second subscript, $j = \mathrm{a}, \mathrm{u}, \mathrm{p}$, stands for all, unknown and prescribed. The terms $\overline{\boldsymbol{u}}(t)$ and $\overline{\boldsymbol{\Theta}}(t)$ are the given functions at time $t$ representing displacement-control and prescribed temperature boundary conditions. The incidence matrices $\boldsymbol{Z}_{\mathrm{u}}^e \in \mathbb{R}^{n_{\mathrm{u}}^e \times n_{\mathrm{uu}}}$, $\overline{\boldsymbol{Z}}_{\mathrm{u}}^e \in \mathbb{R}^{n_{\mathrm{u}}^e \times n_{\mathrm{up}}}$, $\boldsymbol{Z}_{\Theta}^e \in \mathbb{R}^{n_{\Theta}^e \times n_{\Theta\mathrm{u}}}$, and $\overline{\boldsymbol{Z}}_{\Theta}^e \in \mathbb{R}^{n_{\Theta}^e \times n_{\Theta\mathrm{p}}}$ are related to the assemblage procedures, which are frequently symbolized by either $\bigcup_{e=1}^{n_{\mathrm{el}}^{\Theta}}$ or $\mathsf{A}_{e=1}^{n_{\mathrm{el}}^{\Theta}}$ in [Wriggers 2001; Hughes 1987]. Accordingly, the incidence matrices $\boldsymbol{Z}$ are not programmed explicitly, but symbolize the assemblage procedure of "local" quantities into "global" vectors or matrices. They are also helpful in developing new algorithms [Hartmann 2005; Hartmann and Hamkar 2010] or for describing the transition from local to global quantities.

In view of d'Alembert's principle (8) and temperatures (9), respectively, both the strain and virtual strain vectors as well as the temperature gradient and virtual temperature gradient are required:

$$\mathsf{E}^e(t, \boldsymbol{\xi}, \boldsymbol{u}) = \mathsf{B}_{\mathrm{u}}^e(\boldsymbol{\xi})\mathsf{u}^e(t) = \mathsf{B}_{\mathrm{u}}^e(\boldsymbol{\xi})\big\{\boldsymbol{Z}_{\mathrm{u}}^e \boldsymbol{u} + \overline{\boldsymbol{Z}}_{\mathrm{u}}^e \overline{\boldsymbol{u}}(t)\big\}, \tag{69}$$

$$\dot{\mathsf{E}}^e(t, \boldsymbol{\xi}, \dot{\boldsymbol{u}}) = \mathsf{B}_{\mathrm{u}}^e(\boldsymbol{\xi})\dot{\mathsf{u}}^e(t) = \mathsf{B}_{\mathrm{u}}^e(\boldsymbol{\xi})\big\{\boldsymbol{Z}_{\mathrm{u}}^e \dot{\boldsymbol{u}} + \overline{\boldsymbol{Z}}_{\mathrm{u}}^e \dot{\overline{\boldsymbol{u}}}(t)\big\}, \tag{70}$$

$$\delta\mathsf{E}^e(\boldsymbol{\xi}, \delta\boldsymbol{u}) = \mathsf{B}_{\mathrm{u}}^e(\boldsymbol{\xi})\delta\mathsf{u}^e = \mathsf{B}_{\mathrm{u}}^e(\boldsymbol{\xi})\boldsymbol{Z}_{\mathrm{u}}^e \delta\boldsymbol{u}, \tag{71}$$

$$\boldsymbol{\gamma}^e(t, \boldsymbol{\xi}, \boldsymbol{\Theta}) = \operatorname{grad}\Theta^h = \mathsf{B}_{\Theta}^e(\boldsymbol{\xi})\mathbf{\Theta}^e(t) = \mathsf{B}_{\Theta}^e(\boldsymbol{\xi})\big\{\boldsymbol{Z}_{\Theta}^e \boldsymbol{\Theta} + \overline{\boldsymbol{Z}}_{\Theta}^e \overline{\boldsymbol{\Theta}}(t)\big\}, \tag{72}$$

$$\delta\boldsymbol{\gamma}^e(\boldsymbol{\xi}, \delta\boldsymbol{\Theta}) = \operatorname{grad}\delta\Theta^h = \mathsf{B}_{\Theta}^e(\boldsymbol{\xi})\delta\mathbf{\Theta}^e = \mathsf{B}_{\Theta}^e(\boldsymbol{\xi})\boldsymbol{Z}_{\Theta}^e \delta\boldsymbol{\Theta}, \tag{73}$$

where $\mathsf{B}_{\mathrm{u}}^e(\boldsymbol{\xi}) \in \mathbb{R}^{6 \times n_{\mathrm{u}}^e}$ symbolizes the strain-displacement matrix, and $\mathsf{B}_{\Theta}^e(\boldsymbol{\xi}) \in \mathbb{R}^{3 \times n_{\Theta}^e}$ represents the temperature gradient-temperature matrix. In Voigt-notation, the stresses (1) read $\mathsf{T} = \mathsf{h}(\mathsf{E}, \Theta, \mathsf{q})$, $\mathsf{T} \in \mathbb{R}^6$, $\mathsf{E} \in \mathbb{R}^6$, which are inserted into the weak form (8) leading to the second-order ODE

$$\boldsymbol{M}\ddot{\boldsymbol{u}}(t) = -\boldsymbol{g}_{\mathrm{u}}(t, \boldsymbol{u}, \boldsymbol{\Theta}, \boldsymbol{q}) - \boldsymbol{M}_{\mathrm{up}}\ddot{\overline{\boldsymbol{u}}}(t) \tag{74}$$

with

$$\boldsymbol{g}_{\mathrm{u}}(t, \boldsymbol{u}, \boldsymbol{\Theta}, \boldsymbol{q}) = \sum_{e=1}^{n_{\mathrm{el}}^{\mathrm{u}}} \boldsymbol{Z}_{\mathrm{u}}^{eT} \left\{ \sum_{j=1}^{n_{\mathrm{GPu}}^e} w_{\mathrm{u}}^{e(j)} \mathsf{B}_{\mathrm{u}}^{e(j)T} \mathsf{h}(\mathsf{E}^{e(j)}, \Theta^{e(j)}, \mathsf{q}^{e(j)}) \det \mathsf{J}_{\mathrm{u}}^{e(j)} \right\} - \overline{\boldsymbol{p}}(t), \tag{75}$$

$$\boldsymbol{M} = \sum_{e=1}^{n_{\mathrm{el}}^{\mathrm{u}}} \boldsymbol{Z}_{\mathrm{u}}^{eT} \left[ \sum_{j=1}^{n_{\mathrm{GPu}}^e} w_{\mathrm{u}}^{e(j)} \mathsf{N}_{\mathrm{u}}^{e(j)T} \mathsf{N}_{\mathrm{u}}^{e(j)} \rho^{e(j)} \det \mathsf{J}_{\mathrm{u}}^{e(j)} \right] \boldsymbol{Z}_{\mathrm{u}}^e, \tag{76}$$

$$\boldsymbol{M}_{\mathrm{up}} = \sum_{e=1}^{n_{\mathrm{el}}^{\mathrm{u}}} \boldsymbol{Z}_{\mathrm{u}}^{eT} \left[ \sum_{j=1}^{n_{\mathrm{GPu}}^e} w_{\mathrm{u}}^{e(j)} \mathsf{N}_{\mathrm{u}}^{e(j)T} \mathsf{N}_{\mathrm{u}}^{e(j)} \rho^{e(j)} \det \mathsf{J}_{\mathrm{u}}^{e(j)} \right] \overline{\boldsymbol{Z}}_{\mathrm{u}}^e. \tag{77}$$

Here $n^e_{\mathrm{GPu}}$ are the number of Gauss-points in the mechanical element $e$; $w_{\mathrm{u}}^{e(j)}$ symbolizes the weighting factor of the spatial integration scheme, where we assume a Gauss-integration formula; $\mathsf{B}_{\mathrm{u}}^{e(j)} = \mathsf{B}_{\mathrm{u}}(\boldsymbol{\xi}^{(j)})$ is the strain-displacement matrix used in (69) and evaluated at Gauss-point $\boldsymbol{\xi}^{(j)}$; $\mathsf{E}^{e(j)} = \mathsf{E}^{e}(t, \boldsymbol{\xi}^{(j)}, \boldsymbol{u})$ are the strains (69); $\Theta^{e(j)} = \Theta^{e}(t, \boldsymbol{\xi}^{(j)}, \boldsymbol{\Theta})$ defines the temperature (66); and $\mathsf{q}^{e(j)} = \mathsf{q}^{e}(t, \boldsymbol{\xi}^{(j)})$ the internal variables at Gauss-point $\boldsymbol{\xi}^{(j)}$. The Jacobian of the coordinate transformation is defined as $\mathsf{J}_{\mathrm{u}}^{e(j)} = \mathrm{d}\boldsymbol{\chi}^{e}_{\mathrm{u}}(\boldsymbol{\xi})/\mathrm{d}\boldsymbol{\xi}|_{\boldsymbol{\xi}=\boldsymbol{\xi}^{(j)}}$ and $\overline{\boldsymbol{p}}(t)$ symbolizes the equivalent nodal force vector.

Following [Ellsiepen and Hartmann 2001], we formally assemble all internal variables of all Gauss-points into a large vector

$$\boldsymbol{q}(t) = \sum_{e=1}^{n^{\mathrm{u}}_{\mathrm{el}}} \sum_{j=1}^{n^{e}_{\mathrm{GPu}}} \boldsymbol{Z}_{\mathrm{Q}}^{e(j)^T} \mathsf{q}^{e(j)}(t) \quad (\text{or } \mathsf{q}^{e(j)}(t) = \boldsymbol{Z}_{\mathrm{Q}}^{e(j)} \boldsymbol{q}(t)), \tag{78}$$

and the ordinary differential equations (evolution equations of the internal variables) are treated in the same manner ($\boldsymbol{r}_{\mathrm{Q}} = \sum_{e=1}^{n^{\mathrm{u}}_{\mathrm{el}}} \sum_{j=1}^{n^{e}_{\mathrm{GPu}}} \boldsymbol{Z}_{\mathrm{Q}}^{e(j)^T} \mathsf{r}_{\mathrm{Q}}(\mathsf{E}^{e(j)}, \Theta^{e(j)}, \mathsf{q}^{e(j)})$), which leads to

$$\dot{\boldsymbol{q}}(t) = \boldsymbol{r}_{\mathrm{Q}}(t, \boldsymbol{u}, \boldsymbol{\Theta}, \boldsymbol{q}). \tag{79}$$

In our case $\mathsf{q}^{e(j)} \in \mathbb{R}^{n_{\mathrm{q}}}$ with $n_{\mathrm{q}} = 12$ (symmetry of the tensors $\mathbf{E}_{\mathrm{v}}$ and $\mathbf{E}_{\mathrm{r}}$ is assumed), $\boldsymbol{Z}_{\mathrm{Q}}^{e(j)} \in \mathbb{R}^{n_{\mathrm{q}} \times n_{\mathrm{Q}}}$, and $\boldsymbol{q} \in \mathbb{R}^{n_{\mathrm{Q}}}$ hold.

In its spatially discretized representation, the weak form of the heat equation (9) reads

$$\begin{aligned} \boldsymbol{g}_{\Theta}(t, \boldsymbol{u}, \dot{\boldsymbol{u}}, \boldsymbol{\Theta}, \dot{\boldsymbol{\Theta}}, \boldsymbol{q}) = \sum_{e=1}^{n^{\Theta}_{\mathrm{el}}} \boldsymbol{Z}_{\Theta}^{e\,T} \left\{ \sum_{j=1}^{n^{e}_{\mathrm{GP}\Theta}} w_{\Theta}^{(j)} \hat{c}_{\mathrm{p}}(\mathsf{E}^{e(j)}, \Theta^{e(j)})\, \mathsf{N}_{\Theta}^{e(j)} \dot{\Theta}^{e(j)}(t) \det \mathsf{J}_{\Theta}^{e(j)} \right\} \\ + \boldsymbol{p}_{\kappa}(t, \boldsymbol{\Theta}) - \overline{\boldsymbol{p}}_{\Theta}(t, \boldsymbol{\Theta}) - \hat{\boldsymbol{p}}_{\Theta}(t, \boldsymbol{u}, \dot{\boldsymbol{u}}, \boldsymbol{\Theta}, \boldsymbol{q}), \end{aligned} \tag{80}$$

or by inserting (66)–(68) and (72), (73) one obtains the more brief form

$$\boldsymbol{C}_{\Theta}(t, \boldsymbol{u}, \boldsymbol{\Theta}) \dot{\boldsymbol{\Theta}}(t) = \boldsymbol{r}_{\Theta}(t, \boldsymbol{u}, \dot{\boldsymbol{u}}, \boldsymbol{\Theta}, \boldsymbol{q}). \tag{81}$$

In this equation, the temperature-dependent heat capacity matrix is obtained:

$$\boldsymbol{C}_{\Theta}(t, \boldsymbol{u}, \boldsymbol{\Theta}) = \sum_{e=1}^{n^{\Theta}_{\mathrm{el}}} \boldsymbol{Z}_{\Theta}^{e\,T} \left[ \sum_{j=1}^{n^{e}_{\mathrm{GP}\Theta}} w_{\Theta}^{(j)} \hat{c}_{\mathrm{p}}(\mathsf{E}^{e(j)}, \Theta^{e(j)})\, \mathsf{N}_{\Theta}^{e(j)} \mathsf{N}_{\Theta}^{e(j)^T} \det \mathsf{J}_{\Theta}^{e(j)} \right] \boldsymbol{Z}_{\Theta}^{e} \tag{82}$$

($\boldsymbol{C}_{\Theta} \in \mathbb{R}^{n_{\Theta\mathrm{u}} \times n_{\Theta\mathrm{u}}}$) and

$$\boldsymbol{r}_{\Theta}(t, \boldsymbol{u}, \dot{\boldsymbol{u}}, \boldsymbol{\Theta}, \boldsymbol{q}) = -\boldsymbol{p}_{\kappa}(t, \boldsymbol{\Theta}) + \overline{\boldsymbol{p}}_{\Theta}(t, \boldsymbol{\Theta}) + \hat{\boldsymbol{p}}_{\Theta}(t, \boldsymbol{u}, \dot{\boldsymbol{u}}, \boldsymbol{\Theta}, \boldsymbol{q}), \tag{83}$$

with the conductivity term

$$\boldsymbol{p}_{\kappa}(t, \boldsymbol{\Theta}) = \sum_{e=1}^{n^{\Theta}_{\mathrm{el}}} \boldsymbol{Z}_{\Theta}^{e\,T} \left\{ \sum_{j=1}^{n^{e}_{\mathrm{GP}\Theta}} w_{\Theta}^{(j)} \kappa(\Theta^{e(j)})\, \mathsf{B}_{\Theta}^{e(j)^T} \boldsymbol{\gamma}^{e}(\Theta^{e(j)}(t)) \det \mathsf{J}_{\Theta}^{e(j)} \right\} \tag{84}$$

and the dissipation term

$$\hat{\boldsymbol{p}}_\Theta(t, \boldsymbol{u}, \dot{\boldsymbol{u}}, \boldsymbol{\Theta}, \boldsymbol{q}) = \sum_{e=1}^{n_{\mathrm{el}}^{\Theta}} \boldsymbol{Z}_\Theta^{e\,T} \left\{ \sum_{j=1}^{n^e_{\mathrm{GP}\Theta}} w_\Theta^{(j)}\, \hat{p}(\mathsf{E}^{e(j)}, \dot{\mathsf{E}}^{e(j)}, \Theta^{e(j)}, \mathsf{q}^{e(j)})\, \mathsf{N}_\Theta^{e(j)} \det \mathsf{J}_\Theta^{e(j)} \right\}. \tag{85}$$

For reasons of brevity, the heat flux and heat source on the surface are not explained in detail:

$$\overline{\mathsf{p}}_\Theta(t, \boldsymbol{\Theta}) = \int_A \mathsf{N}_\Theta q^h(\Theta^h)\, \mathrm{d}A + \int_V \mathsf{N}_\Theta r^h\, \mathrm{d}V. \tag{86}$$

## Appendix B. Time adaptivity

Regarding computational efficiency and accuracy, methods with a constant step-size perform poorly if the solution varies rapidly in some parts of the time interval and slowly in other parts. For this reason the step-size should be chosen in such a way that it is large in smooth parts and small in transient parts. This calls for a step-size control technique which relies on the approximation of the local integration error and adjusts the time-step so that the error measure remains within a prescribed tolerance. Since an adaptive step-size control takes the behavior of the underlying equations into account, it stabilizes the global procedure and keeps the global error within certain limits as well. The first widely used adaptive time-stepping strategies in the field of structural dynamics were proposed in [Zienkiewicz et al. 1984; Zienkiewicz and Xie 1991; Zeng et al. 1992; Li et al. 1993; Riccius and Schweizerhof 1996; Riccius 1997]. These approaches are based on an a posteriori estimation of the local integration error (see (89)) for the displacement field which results from a difference between numerical solution and an improved solution calculated in a Taylor series at $t_n$. Further details regarding this a posteriori error estimation technique and specific numerical examples in combination with generalized-$\alpha$ related methods can be found in [Kuhl 1996; Kuhl and Ramm 1999]. An a posteriori error estimation technique for displacements and velocities is proposed in [Hulbert and Jang 1995]. Step-size control algorithms based on the "apparent highest frequency" as well other classical approaches built on general frequency information lie out of our scope. In this section, we focus on the possibility of incorporating an efficient step-size control, which is based on the estimation of the local integration error (local truncation error) and can be achieved with a so-called embedding technique [Hairer et al. 1993; Hairer and Wanner 1996; Strehmel et al. 2012]. A further approach for the time-adaptivity of one-step methods — the Richardson extrapolation — is presented in [Hairer et al. 1993]. Due to the necessity of repeating the time-step with $\Delta t_n/2$ and comparing the results to the computations with the time-step $\Delta t_n$, a higher computational effort is required. For that reason, local error control using the Richardson extrapolation is not of particular interest to us.

We start with recapping the coupled ODE system in (10) in explicit form

$$\dot{\boldsymbol{y}}(t) = \boldsymbol{f}\,(t, \boldsymbol{y}(t)), \quad \boldsymbol{y}(t_0) = \boldsymbol{y}_0, \tag{87}$$

with $\boldsymbol{y} = \{\boldsymbol{u}^T, \boldsymbol{v}^T, \boldsymbol{\Theta}^T, \boldsymbol{q}^T\} \in \mathbb{R}^{(2n_{\mathrm{uu}}+n_{\Theta\mathrm{u}}+n_{\mathrm{Q}})}$.

Based on an exact value $\boldsymbol{y}_n = \boldsymbol{y}(t_n)$ the local integration error $\boldsymbol{\delta}$ in the time interval $\Delta t_n$ is defined by the difference between the exact solution $\boldsymbol{y}(t_{n+1})$ and the numerical solution

$$\boldsymbol{y}_{n+1} = \boldsymbol{y}(t_n) + \Delta t_n \boldsymbol{\Phi}(t_n, \boldsymbol{y}(t_n), \Delta t_n), \tag{88}$$

using an arbitrary integration method of order $p$ as

$$\begin{aligned}\boldsymbol{\delta} &= \boldsymbol{y}(t_{n+1}) - \boldsymbol{y}_{n+1} = \boldsymbol{y}(t_{n+1}) - \big(\boldsymbol{y}(t_n) + \Delta t_n \boldsymbol{\Phi}(t_n, \boldsymbol{y}(t_n), \Delta t_n)\big)\\ &= \Delta t_n^{p+1} \boldsymbol{\Psi}(t_n, \boldsymbol{y}) + \mathcal{O}(\Delta t_n^{p+2}).\end{aligned} \tag{89}$$

Here, $\boldsymbol{\Phi}$ marks the so-called increment function of the underlying integrator [Hairer et al. 1993; Strehmel et al. 2012]. Based on Taylor series expansions of the exact solution and the numerical solution, we can determine the order of consistency where the local truncation error can be split up into a main part $\Delta t_n^{p+1} \boldsymbol{\Psi}$ (principal error) and a remainder of order $p+2$. In the following, the main part of the local integration error $\boldsymbol{\delta}$ is estimated so as to be able to control the error quantity. For this purpose, two methods of different order $p$ and $\hat{p} = p - 1$ are assumed:

$$\begin{aligned}\boldsymbol{y}_{n+1} &= \boldsymbol{y}(t_n) + \Delta t_n \boldsymbol{\Phi}(t_n, \boldsymbol{y}(t_n), \Delta t_n),\\ \hat{\boldsymbol{y}}_{n+1} &= \boldsymbol{y}(t_n) + \Delta t_n \hat{\boldsymbol{\Phi}}(t_n, \boldsymbol{y}(t_n), \Delta t_n).\end{aligned} \tag{90}$$

Each of these methods yield an expression for the local integration error

$$\begin{aligned}\boldsymbol{\delta} &= \boldsymbol{y}(t_{n+1}) - \boldsymbol{y}_{n+1} = \Delta t_n^{p+1} \boldsymbol{\Psi}(t_n, \boldsymbol{y}) + \mathcal{O}(\Delta t_n^{p+2}),\\ \hat{\boldsymbol{\delta}} &= \boldsymbol{y}(t_{n+1}) - \hat{\boldsymbol{y}}_{n+1} = \Delta t_n^{\hat{p}+1} \hat{\boldsymbol{\Psi}}(t_n, \boldsymbol{y}) + \mathcal{O}(\Delta t_n^{\hat{p}+2}).\end{aligned} \tag{91}$$

Using the difference of both errors

$$\boldsymbol{\delta} - \hat{\boldsymbol{\delta}} = \boldsymbol{y}_{\text{err}} := \hat{\boldsymbol{y}}_{n+1} - \boldsymbol{y}_{n+1} = \Delta t_n^{\hat{p}+1} \hat{\boldsymbol{\Psi}}(t_n, \boldsymbol{y}) + \mathcal{O}(\Delta t_n^{\hat{p}+2}) \approx \Delta t_n^{\hat{p}+1} \hat{\boldsymbol{\Psi}}(t_n, \boldsymbol{y}), \tag{92}$$

we are able to estimate the main part $\Delta t_n^{\hat{p}+1} \hat{\boldsymbol{\Psi}}$ of the local integration error of the lower order method, i.e., we get the "second best" approximation of $\boldsymbol{y}(t_{n+1})$. Furthermore, we assume that the function $\hat{\boldsymbol{\Psi}}$ varies only slowly so that $\|\hat{\boldsymbol{\Psi}}(t_n, \boldsymbol{y})\| \approx C$ holds within $\Delta t_n$. With this error estimate in hand we are able to decide whether the error in each time step is lower than a user-specified tolerance

$$\|\boldsymbol{y}_{\text{err}}\| \approx C \Delta t_n^{\hat{p}+1} \leq \varepsilon_{\text{r}} \|\hat{\boldsymbol{y}}_n\| + \varepsilon_{\text{a}}, \tag{93}$$

where generally the tolerance value is given as combination of an absolute $\varepsilon_{\text{a}}$ and a relative $\varepsilon_{\text{r}}$ tolerance. For the computation of an optimal new step-size $\Delta t_{\text{new}}$, we demand that the error equals the prescribed mixed tolerance

$$C \Delta t_{\text{new}}^{\hat{p}+1} = \varepsilon_{\text{r}} \|\hat{\boldsymbol{y}}_n\| + \varepsilon_{\text{a}}. \tag{94}$$

Eliminating the constant $C$ by using the relationship in (93) and substituting it into (94), we arrive at the desired result

$$\Delta t_{\text{new}} = \Delta t_n \left( \frac{\varepsilon_{\text{r}} \|\hat{\boldsymbol{y}}_n\| + \varepsilon_{\text{a}}}{\|\boldsymbol{y}_{\text{err}}\|} \right)^{1/(\hat{p}+1)}. \tag{95}$$

In the numerical mathematics community it is common to advance the solution in time by using the higher-order approximation $\boldsymbol{y}_{n+1}$ instead of $\hat{\boldsymbol{y}}_{n+1}$. We do so as well. The case in which the concept of local error control is slightly abandoned is called *local extrapolation* (see [Hairer et al. 1993] for further explanations). A theoretical justification for this violation is given in [Deuflhard and Bornemann 1994].

For the approximation of the local error in the case of DIRK schemes, we use a highly efficient method which is based on an embedded Runge–Kutta scheme. For this purpose two $s$-stage methods of different

$$\begin{array}{c|c} \boldsymbol{c} & \boldsymbol{A} \\ \hline & \boldsymbol{b}^T \\ & \hat{\boldsymbol{b}}^T \end{array} \qquad\qquad \begin{array}{c|c|c} \boldsymbol{c} & \bar{\boldsymbol{A}} & \boldsymbol{A} \\ \hline & \bar{\boldsymbol{b}}^T & \boldsymbol{b}^T \\ & \hat{\bar{\boldsymbol{b}}}^T & \hat{\boldsymbol{b}}^T \end{array}$$

**Figure 17.** Butcher-tableaus with embedding: embedded RK scheme (left) and embedded RKN scheme (right).

order $p$ and $\hat{p}$ are constructed with identical stages $c_i$, the same coefficient matrix $\boldsymbol{A}$, but using different weighting factors $b_i$ and $\hat{b}_i$. The pair of embedded RK methods share the same stage computations. Thus, a further solution at $t_{n+1}$ can be computed by using (29), $(25)_2$–$(25)_4$

$$\begin{aligned} \hat{\boldsymbol{u}}_{n+1} &= \boldsymbol{u}_n + \Delta t_n \boldsymbol{v}_n + \Delta t_n^2 \sum_{i=1}^{s} \hat{\bar{b}}_i \boldsymbol{A}_{ni}, & \hat{\boldsymbol{v}}_{n+1} &= \boldsymbol{v}_n + \Delta t_n \sum_{i=1}^{s} \hat{b}_i \boldsymbol{A}_{ni}, \\ \hat{\boldsymbol{\Theta}}_{n+1} &= \boldsymbol{\Theta}_n + \Delta t_n \sum_{i=1}^{s} \hat{b}_i \dot{\boldsymbol{\Theta}}_{ni}, & \hat{\boldsymbol{q}}_{n+1} &= \boldsymbol{q}_n + \Delta t_n \sum_{i=1}^{s} \hat{b}_i \dot{\boldsymbol{Q}}_{ni}, \end{aligned} \tag{96}$$

with the newly introduced algorithmic parameters $\hat{b}_i$ and $\hat{\bar{b}}_i = \sum_{j=1}^{s} \hat{b}_j a_{ji}$. As in [Ehlers and Ellsiepen 1998; Diebels et al. 1999], we decompose $\boldsymbol{y}_{\text{err}}$ into the local integration error for each field variable in order to take the different physical properties and the order of magnitude into account. Commonly, the new weighting factors $\hat{b}_i$ and $\hat{\bar{b}}_i$ are appended to the Butcher array to form an embedded RK scheme (see Figure 17).

The consistency orders of the two methods differ by one (i.e., $\hat{p} = p - 1$). According to (92), a simple hand calculation results in the following estimation formula for RK schemes:

$$\begin{aligned} \boldsymbol{u}_{\text{err}} &= \hat{\boldsymbol{u}}_{n+1} - \boldsymbol{u}_{n+1} = \Delta t_n^2 \sum_{i=1}^{s} (\hat{\bar{b}}_i - \bar{b}_i) \boldsymbol{A}_{ni}, & \boldsymbol{v}_{\text{err}} &= \hat{\boldsymbol{v}}_{n+1} - \boldsymbol{V}_{n+1} = \Delta t_n \sum_{i=1}^{s} (\hat{b}_i - b_i) \boldsymbol{A}_{ni}, \\ \boldsymbol{\Theta}_{\text{err}} &= \hat{\boldsymbol{\Theta}}_{n+1} - \boldsymbol{\Theta}_{n+1} = \Delta t_n \sum_{i=1}^{s} (\hat{b}_i - b_i) \dot{\boldsymbol{\Theta}}_{ni}, & \boldsymbol{q}_{\text{err}} &= \hat{\boldsymbol{q}}_{n+1} - \boldsymbol{q}_{n+1} = \Delta t_n \sum_{i=1}^{s} (\hat{b}_i - b_i) \dot{\boldsymbol{Q}}_{ni}, \end{aligned} \tag{97}$$

where the stage derivatives of both integration schemes are equal on the basis of the same coefficients $a_{ij}$ and $\bar{a}_{ij}$.

In the case of the generalized-$\alpha$ method we follow a proposal in [Rang 2013a]. There, the approximation of the generalized-$\alpha$ scheme was used as a second-order approximation ($p = 2$) since the error constant is very small and since the methods in our numerical experiments (see Section 4) behave as a second-order method [Chung and Hulbert 1993; Erlicher et al. 2002; Rang 2013a]. The Backward Euler method can be used for a second solution of lower order $\hat{p} = 1$. To circumvent an entire calculation of (10) by using the Backward Euler method, the necessary derivatives will be approximated by using (38), (39), (46) and (54). The computation for the local truncation error quantities in each field variable yield

$$\begin{aligned} \boldsymbol{u}_{\text{err}} &= \hat{\boldsymbol{u}}_{n+1} - \boldsymbol{u}_{n+1} = \boldsymbol{u}_n + \Delta t_n \boldsymbol{v}_{n+1} - \boldsymbol{u}_{n+1}, & \boldsymbol{v}_{\text{err}} &= \hat{\boldsymbol{v}}_{n+1} - \boldsymbol{v}_{n+1} = \boldsymbol{v}_n + \Delta t_n \boldsymbol{a}_{n+1} - \boldsymbol{v}_{n+1}, \\ \boldsymbol{\Theta}_{\text{err}} &= \hat{\boldsymbol{\Theta}}_{n+1} - \boldsymbol{\Theta}_{n+1} = \boldsymbol{\Theta}_n + \Delta t_n \dot{\boldsymbol{\Theta}}_{n+1} - \boldsymbol{\Theta}_{n+1}, & \boldsymbol{q}_{\text{err}} &= \hat{\boldsymbol{q}}_{n+1} - \boldsymbol{q}_{n+1} = \boldsymbol{q}_n + \Delta t_n \dot{\boldsymbol{q}}_{n+1} - \boldsymbol{q}_{n+1}. \end{aligned} \tag{98}$$

Commonly, the estimation of $\Delta t_{\mathrm{new}}$ is carried out using a more efficient and computationally more stable procedure if it is described by (95). Following the proposal in [Hairer et al. 1993; Hairer and Wanner 1996; Gustafsson 1994], we employ the relative error measures

$$
\begin{gathered}
e_u := \sqrt{\frac{1}{n_{\mathrm{uu}}} \sum_{k=1}^{n_{\mathrm{uu}}} \left( \frac{u_{\mathrm{err}}^k}{\varepsilon_{\mathrm{r}}^u |u_n^k| + \varepsilon_{\mathrm{a}}^u} \right)^2}, \quad e_v := \sqrt{\frac{1}{n_{\mathrm{uu}}} \sum_{k=1}^{n_{\mathrm{uu}}} \left( \frac{v_{\mathrm{err}}^k}{\varepsilon_{\mathrm{r}}^v |v_n^k| + \varepsilon_{\mathrm{a}}^v} \right)^2}, \\
e_\Theta := \sqrt{\frac{1}{n_{\Theta\mathrm{u}}} \sum_{k=1}^{n_{\Theta\mathrm{u}}} \left( \frac{\Theta_{\mathrm{err}}^k}{\varepsilon_{\mathrm{r}}^\Theta |\Theta_n^k| + \varepsilon_{\mathrm{a}}^\Theta} \right)^2},
\end{gathered}
\tag{99}
$$

where $u_{\mathrm{err}}^k$, $v_{\mathrm{err}}^k$ and $\Theta_{\mathrm{err}}^k$ are the $k$-th components of the local integration error approximations $\boldsymbol{u}_{\mathrm{err}}$, $\boldsymbol{v}_{\mathrm{err}}$, $\boldsymbol{\Theta}_{\mathrm{err}}$, respectively. These error measures perform a componentwise weighting instead of the simple global weighting $\|\boldsymbol{y}_{\mathrm{err}}\| / (\varepsilon_{\mathrm{r}} \|\boldsymbol{y}_n\| + \varepsilon_{\mathrm{a}})$ introduced in (94). For the local truncation error $\boldsymbol{q}_{\mathrm{err}}$ in the interval variables, Diebels et al. [1999] proposed the maximum norm

$$
e_q := \max_{1 \le k \le n_{\mathrm{Q}}} \left| \frac{q_{\mathrm{err}}^k}{\varepsilon_{\mathrm{r}}^q |q_n^k| + \varepsilon_{\mathrm{a}}^q} \right|. \tag{100}
$$

Furthermore, the user-defined relative $\varepsilon_{\mathrm{r}}^u, \varepsilon_{\mathrm{r}}^v, \varepsilon_{\mathrm{r}}^\Theta, \varepsilon_{\mathrm{r}}^q$ and absolute $\varepsilon_{\mathrm{a}}^u, \varepsilon_{\mathrm{a}}^v, \varepsilon_{\mathrm{a}}^\Theta, \varepsilon_{\mathrm{a}}^q$ error tolerances may depend on the components of the field variables reflecting different magnitude and physical meaning of these quantities [Hairer et al. 1993]. For the sake of simplicity they are chosen to be constant. The maximum $\mathrm{e}_m = \max(\mathrm{e}_u, \mathrm{e}_v, \mathrm{e}_\theta, \mathrm{e}_q)$ of the weighted error measures is used to determine the new step-size,

$$
\Delta t_{\mathrm{new}} = \Delta t_n \cdot \begin{cases} \max(f_{\min}, f_{\mathrm{safety}} \mathrm{e}_m^{-1/(\hat{p}+1)}) & \text{if } \mathrm{e}_m > 1 \\ \min(f_{\max}, f_{\mathrm{safety}} \mathrm{e}_m^{-1/(\hat{p}+1)}) & \text{if } \mathrm{e}_m \le 1 \end{cases}, \tag{101}
$$

comparing $\mathrm{e}_m$ to one. In the case $\mathrm{e}_m \le 1$ the computed step is accepted and the integration can march forward with $\Delta t_{\mathrm{new}}$, otherwise the step has to be repeated with a smaller step-size $\Delta t_{\mathrm{new}}$. The safety factor $0 < f_{\mathrm{safety}} < 1$ prevents oscillations in the step-size controller while $f_{\min}$ and $f_{\max}$ keep the step-size from increasing and decreasing too fast. In practice, typical values for these factors are: $0.8 \le f_{\mathrm{safety}} \le 0.9$, $0.2 \le f_{\min} \le 0.5$, $2 \le f_{\max} \le 3$ [Ellsiepen and Hartmann 2001].

Although the proposed standard controller [Hairer et al. 1993] works quite well, the local error control algorithm yields strong oscillations in the step-size behavior in phases of quickly changing dynamics [Gustafsson et al. 1988; Gustafsson 1991; 1994; Strehmel et al. 2012]. With regard to a further stabilization and an increased time integration performance of the step-size controller, several approaches from a control theoretical point of view were studied in [Gustafsson et al. 1988; Gustafsson 1991; 1994; Söderlind 2002]. In Section 4 we employ a modified hybrid PI-controller based on [Lang 2001], which works according to the following step-size selection rule

$$
\Delta t_{\mathrm{new}} = \Delta t_n \cdot \begin{cases} \min\bigl(f_{\max}, \max(f_{\min}, f_{\mathrm{safety}}\ (1/\mathrm{e}_m^{n+1})^{K_I} (\mathrm{e}_m^n / \mathrm{e}_m^{n+1})^{K_P})\bigr) & \text{if } \mathrm{e}_m > 1 \\ \min\bigl(f_{\max}, \max(f_{\min}, f_{\mathrm{safety}}\ (1/\mathrm{e}_m^{n+1})^{K_I})\bigr) & \text{if } \mathrm{e}_m \le 1 \end{cases}, \tag{102}
$$

using time-step and error data $\mathrm{e}_m^n$ from the previous time-step. $K_P$ and $K_I$ define the proportional and the integral gain of the controller, which can also be dependent on previous time-steps sizes and on previously computed error data. For further details about the implementation of the proposed automatic step-length control we refer to [Gustafsson 1994; Hairer and Wanner 1996; Lang 2001].

## References

[Alexander 1977] R. Alexander, "Diagonally implicit Runge–Kutta methods for stiff ODE's", *SIAM J. Numer. Anal.* **14**:6 (1977), 1006–1021.

[Armero and Simo 1992] F. Armero and J. C. Simo, "A new unconditionally stable fractional step method for nonlinear coupled thermomechanical problems", *Internat. J. Numer. Methods Eng.* **35**:4 (1992), 737–766.

[Armero and Simo 1993] F. Armero and J. Simo, "A priori stability estimates and unconditionally stable product formula algorithms for nonlinear coupled thermoplasticity", *Int. J. Plasticity* **9**:6 (1993), 749–782.

[Belytschko et al. 2000] T. Belytschko, W. K. Liu, and B. Moran, *Nonlinear finite elements for continua and structures*, Wiley, Chichester, England, 2000.

[Birken et al. 2010] P. Birken, K. J. Quint, S. Hartmann, and A. Meister, "A time-adaptive fluid-structure interaction method for thermal coupling", *Comput. Vis. Sci.* **13**:7 (2010), 331–340.

[Butcher 2008] J. C. Butcher, *Numerical methods for ordinary differential equations*, 2nd ed., Wiley, Chichester, England, 2008.

[Cash 1979] J. Cash, "Diagonally implicit Runge–Kutta formulae with error estimates", *J. Inst. Math. Appl.* **24** (1979), 293–301.

[Chung and Hulbert 1993] J. Chung and G. M. Hulbert, "A time integration algorithm for structural dynamics with improved numerical dissipation: the generalized-$\alpha$ method", *Trans. ASME J. Appl. Mech.* **60**:2 (1993), 371–375.

[Conde Martín et al. 2014] S. Conde Martín, J. C. García Orden, and I. Romero, "Energy-consistent time integration for nonlinear viscoelasticity", *Comput. Mech.* **54**:2 (2014), 473–488.

[Dettmer and Perić 2003] W. Dettmer and D. Perić, "An analysis of the time integration algorithms for the finite element solutions of incompressible Navier–Stokes equations based on a stabilised formulation", *Comput. Methods Appl. Mech. Eng.* **192**:9-10 (2003), 1177–1226.

[Deuflhard and Bornemann 1994] P. Deuflhard and F. Bornemann, *Numerische Mathematik II: Integration gewöhnlicher Differentialgleichungen*, Walter de Gruyter & Co., Berlin, 1994.

[Diebels et al. 1999] S. Diebels, P. Ellsiepen, and W. Ehlers, "Error-controlled Runge–Kutta Time Integration of a Viscoplastic Hybrid Two-phases Model", *Technische Mechanik* **19** (1999), 19–27.

[Ehlers and Ellsiepen 1998] W. Ehlers and P. Ellsiepen, "Adaptive Zeitintegration–Verfahren für ein elastisch-viskoplastisches Zweiphasenmodell", *ZAMM Zeitschrift für Angewandte Mathematik und Mechanik* **78** (1998), S361–S362.

[Ellsiepen 1999] P. Ellsiepen, *Zeit- und ortsadaptive Verfahren angewandt auf Mehrphasenprobleme poröser Medien*, Doctoral Thesis, Institute of Mechanics II, University of Stuttgart, 1999.

[Ellsiepen and Hartmann 2001] P. Ellsiepen and S. Hartmann, "Remarks on the interpretation of current non-linear finite-element-analyses as differential-algebraic equations", *Int. J. Numer. Meth. Eng.* **51** (2001), 679–707.

[Erbts and Düster 2012] P. Erbts and A. Düster, "Accelerated staggered coupling schemes for problems of thermoelasticity at finite strains", *Comput. Math. Appl.* **64**:8 (2012), 2408–2430.

[Erbts et al. 2015] P. Erbts, S. Hartmann, and A. Düster, "A partitioned solution approach for electro-thermo-mechanical problems", *Archive of Applied Mechanics* **85**:8 (2015), 1075–1101.

[Erlicher et al. 2002] S. Erlicher, L. Bonaventura, and O. S. Bursi, "The analysis of the generalized-$\alpha$ method for non-linear dynamic problems", *Comput. Mech.* **28**:2 (2002), 83–104.

[Felippa and Park 1980] C. Felippa and K. Park, "Staggered transient analysis procedures for coupled mechanical systems: formulation", *Comput. Methods Appl. Mech. Eng.* **24** (1980), 61–111.

[Felippa et al. 2001] C. A. Felippa, K. Park, and C. Farhat, "Partitioned analysis of coupled mechanical systems", *Comput. Methods Appl. Mech. Eng.* **190**:24-25 (2001), 3247–3270.

[Fritzen 1997] P. Fritzen, *Numerische behandlung nichtlinearer probleme der elastizitäts- und plastizitätstheorie*, Ph.D. thesis, Universität Darmstadt, 1997.

[Gear 1986] C. W. Gear, "Maintaining solution invariants in the numerical solution of ODEs", *SIAM J. Sci. Statist. Comput.* **7**:3 (1986), 734–743.

[Glaser 1991] S. Glaser, "Berechnung gekoppelter thermomechanischer Prozesse", Institut für Statik und Dynamik der Luft- und Raumfahrtkonstruktionen, Universität Stuttgart, Stuttgart (Germany), 1991.

[Gonzalez 2000] O. Gonzalez, "Exact energy and momentum conserving algorithms for general models in nonlinear elasticity", *Comput. Methods Appl. Mech. Eng.* **190**:13-14 (2000), 1763–1783.

[Gustafsson 1991] K. Gustafsson, "Control-theoretic techniques for stepsize selection in explicit Runge–Kutta methods", *ACM Trans. Math. Software* **17**:4 (1991), 533–554.

[Gustafsson 1994] K. Gustafsson, "Control-theoretic techniques for stepsize selection in implicit Runge–Kutta methods", *ACM Trans. Math. Software* **20**:4 (1994), 496–517.

[Gustafsson et al. 1988] K. Gustafsson, M. Lundh, and G. Söderlind, "A PI stepsize control for the numerical solution of ordinary differential equations", *BIT* **28**:2 (1988), 270–287.

[Hairer and Wanner 1996] E. Hairer and G. Wanner, *Solving ordinary differential equations II*, 2nd ed., Springer Series in Computational Mathematics **14**, Springer, Berlin, 1996.

[Hairer et al. 1989] E. Hairer, C. Lubich, and M. Roche, *The numerical solution of differential-algebraic systems by Runge–Kutta methods*, Lecture Notes in Mathematics **1409**, Springer, Berlin, 1989.

[Hairer et al. 1993] E. Hairer, S. P. Nørsett, and G. Wanner, *Solving ordinary differential equations I*, 2nd ed., Springer Series in Computational Mathematics **8**, Springer, Berlin, 1993.

[Hamkar 2013] A.-W. Hamkar, *Eine iterationsfreie finite-elemente methode im Rahmen der finiten thermoviskoelastizität*, PhD-thesis, Report No. 1/2013, Institute of Applied Mechanics, Clausthal University of Technology, Clausthal-Zellerfeld, 2013.

[Hartmann 2002] S. Hartmann, "Computation in finite-strain viscoelasticity: finite elements based on the interpretation as differential-algebraic equations", *Comput. Methods Appl. Mech. Eng.* **191**:13-14 (2002), 1439–1470.

[Hartmann 2005] S. Hartmann, "A remark on the application of the Newton–Raphson method in non-linear finite element analysis", *Comput. Mech.* **36**:2 (2005), 100–116.

[Hartmann 2007] S. Hartmann, *Kontaktanalyse dünnwandiger Strukturen bei großen Deformationen*, Ph.D. thesis, Universität Stuttgart, 2007, Available at http://elib.uni-stuttgart.de/opus/volltexte/2007/3109/.

[Hartmann and Hamkar 2010] S. Hartmann and A.-W. Hamkar, "Rosenbrock-type methods applied to finite element computations within finite strain viscoelasticity", *Comput. Methods Appl. Mech. Eng.* **199**:23-24 (2010), 1455–1470.

[Hartmann and Rothe 2013] S. Hartmann and S. Rothe, "A rigorous application of the method of vertical lines to coupled systems in finite element analysis", pp. 161–175 in *Recent developments in the numerics of nonlinear hyperbolic conservation laws*, edited by R. Ansorge et al., Notes Numer. Fluid Mech. Multidiscip. Des. **120**, Springer, Heidelberg, Germany, 2013.

[Hartmann et al. 2008a] S. Hartmann, K. J. Quint, and M. Arnold, "On plastic incompressibility within time-adaptive finite elements combined with projection techniques", *Comput. Methods Appl. Mech. Eng.* **198**:2 (2008), 178–193.

[Hartmann et al. 2008b] S. Hartmann, K. J. Quint, and A.-W. Hamkar, "Displacement control in time-adaptive non-linear finite-element analysis", *ZAMM Z. Angew. Math. Mech.* **88**:5 (2008), 342–364.

[Hilber et al. 1977] H. M. Hilber, T. J. R. Hughes, and R. L. Taylor, "Improved numerical dissipation for time integration algorithms in structural dynamics", *Earthquake Eng. Struct.* **5**:3 (1977), 283–292.

[Hughes 1987] T. J. R. Hughes, *The finite element method*, Prentice Hall, Englewood Cliffs, New Jersey, 1987.

[Hulbert and Jang 1995] G. M. Hulbert and I. Jang, "Automatic time step control algorithms for structural dynamics", *Comput. Methods Appl. Mech. Eng.* **126**:1-2 (1995), 155–178.

[Jansen et al. 2000] K. E. Jansen, C. H. Whiting, and G. M. Hulbert, "A generalized-$\alpha$ method for integrating the filtered Navier–Stokes equations with a stabilized finite element method", *Comput. Methods Appl. Mech. Eng.* **190**:3-4 (2000), 305–319.

[Kreisselmeier and Steinhauser 1979] G. Kreisselmeier and R. Steinhauser, "Systematische Auslegung von Reglern durch Optimierung eines vektoriellen Gütekriteriums", *Regelungstechnik* **3** (1979), 76–79.

[Krüger et al. 2016] M. Krüger, M. Groß, and P. Betsch, "An energy-entropy-consistent time stepping scheme for nonlinear thermo-viscoelastic continua", *ZAMM Z. Angew. Math. Mech.* **96**:2 (2016), 141–178.

[Kuhl 1996] D. Kuhl, *Stabile Zeitintegrationsalgorithmen in der nichtlinearen Elastodynamik dünnwandiger Tragwerke*, PhD-thesis, Report No. 2 (1996), University of Stuttgart (Germany), Institute for Structural Mechanics, 1996.

[Kuhl and Crisfield 1999] D. Kuhl and M. A. Crisfield, "Energy-conserving and decaying algorithms in non-linear structural dynamics", *Int. J. Numer. Methods Eng.* **45**:5 (1999), 569–599.

[Kuhl and Ramm 1999] D. Kuhl and E. Ramm, "Generalized energy-momentum method for non-linear adaptive shell dynamics", *Comput. Methods Appl. Mech. Eng.* **178**:3-4 (1999), 343–366.

[Lang 2001] J. Lang, *Adaptive multilevel solution of nonlinear parabolic PDE systems*, Lecture Notes in Computational Science and Engineering **16**, Springer, Berlin, 2001.

[Lewis et al. 1996] R. W. Lewis, K. Morgan, H. R. Thomas, and K. N. Seetharamu, *The finite element method in heat transfer analysis*, Wiley, 1996.

[Li et al. 1993] X. Li, L. Zeng, and N.-E. Wiberg, "A simple local error estimator and an adaptive time-stepping procedure for direct integration method in dynamic analysis", *Commun. Numer. Methods Eng.* **9**:4 (1993), 273–292.

[Miehe 1988] C. Miehe, *Zur numerischen Behandlung thermomechanischer Prozesse*, Report No. F88/6, University of Hannover, Institut für Baumechanik und Numerische Mechanik, 1988.

[Netz and Hartmann 2015] T. Netz and S. Hartmann, "A monolithic finite element approach using high-order schemes in time and space applied to finite strain thermo-viscoelasticity", *Comput. Math. Appl.* **70**:7 (2015), 1457–1480.

[Newmark 1959] N. M. Newmark, "A method of computation for structural dynamics", *J. Eng. Mech. Div.* (*ASCE*) **85** (1959), 67–94.

[Noels et al. 2006] L. Noels, L. Stainier, and J.-P. Ponthot, "An energy momentum conserving algorithm using the variational formulation of visco-plastic updates", *Int. J. Numer. Methods Eng.* **65**:6 (2006), 904–942.

[Noels et al. 2008] L. Noels, L. Stainier, and J.-P. Ponthot, "A first-order energy-dissipative momentum-conserving scheme for elasto-plasticity using the variational updates formulation", *Comput. Methods Appl. Mech. Eng.* **197**:6-8 (2008), 706–726.

[Ortiz and Popov 1985] M. Ortiz and E. P. Popov, "Accuracy and stability of integration algorithms for elastoplastic constitutive relations", *Int. J. Numer. Methods Eng.* **21**:9 (1985), 1561–1576.

[Popp 2012] A. Popp, *Mortar methods for computational contact mechanics and general interface problems*, Ph.D. thesis, Technische Universität München, Institute for Computational Mechanics, 2012.

[Prothero and Robinson 1974] A. Prothero and A. Robinson, "On the stability and accuracy of one-step methods for solving stiff systems of ordinary differential equations", *Math. Comp.* **28** (1974), 145–162.

[Quint 2012] K. J. Quint, *Thermomechanically coupled processes for functionally graded materials: experiments, modelling, and finite element analysis using high-order DIRK-methods*, PhD-thesis, Report No. 2/2012, Institute of Applied Mechanics, Clausthal University of Technology, Clausthal-Zellerfeld, 2012.

[Quint et al. 2011] K. J. Quint, S. Hartmann, S. Rothe, N. Saba, and K. Steinhoff, "Experimental validation of high-order time integration for non-linear heat transfer problems", *Comput. Mech.* **48**:1 (2011), 81–96.

[Rabbat et al. 1979] N. Rabbat, A. L. Sangiovanni-Vincentelli, and H. Y. Hsieh, "A multilevel Newton algorithm with macromodeling and latency for the analysis of large-scale nonlinear circuits in the time domain", *IEEE Trans. Circuits Syst.* **26** (1979), 733–741.

[Rang 2013a] J. Rang, "Adaptive timestep control for the generalised-$\alpha$ method", pp. 559–570 in *Adaptive Modeling and Simulation*, edited by D. Aubry et al., 2013.

[Rang 2013b] J. Rang, "Coupling generalised-$\alpha$ methods: analysis, adaptivity, and numerics", pp. 399–410 in *Computational methods for coupled problems in science and engineering V: a conference celebrating the 60th birthday of Eugenio Oñate* (Santa Eulalia, Ibiza, Spain), edited by S. Idelsohn et al., Barcelona, 2013.

[Reddy and Gartling 2000] J. Reddy and D. K. Gartling, *The finite element method in heat transfer and fluid dynamics*, 2nd ed., CRC Press, Boca Raton, Florida, 2000.

[Riccius 1997] J. Riccius, *Adaptive Methoden zur statischen und dynamischen Analyse von Flächentragwerken mit linearen Finiten Elementen*, Ph.D. thesis, Institute of Engineering Mechanics, University of Karlsruhe, 1997.

[Riccius and Schweizerhof 1996] J. Riccius and K. Schweizerhof, "Aspects of hierarchical h-adaptive dynamic analysis", pp. 61–70 in *Advances in Finite Element Techniques*, edited by B. H. V. Topping, Civil-Comp Press, Edinburgh, Scotland, 1996.

[Rothe 2015] S. Rothe, *Electro-thermo-mechanical modeling of field assisted sintering technology: experiments, constitutive modeling and finite element analysis*, PhD-thesis, Report No. 1/2015, Institute of Applied Mechanics, Clausthal University of Technology, Clausthal-Zellerfeld, 2015.

[Rothe et al. 2015a] S. Rothe, P. Erbts, A. Düster, and S. Hartmann, "Monolithic and partitioned coupling schemes for thermo-viscoplasticity", *Comput. Methods Appl. Mech. Eng.* **293** (2015), 375–410.

[Rothe et al. 2015b] S. Rothe, J. H. Schmidt, and S. Hartmann, "Analytical and numerical treatment of electro-thermo-mechanical coupling", *Archive of Applied Mechanics* **85**:9 (2015), 1245–1264.

[Schrefler 2004] B. A. Schrefler, *Encyclopedia of computational mechanics: multifield problems*, vol. 2, Chapter 17, pp. 575–603, Wiley, Chichester, England, 2004.

[Simo and Miehe 1992] J. Simo and C. Miehe, "Associative coupled thermoplasticity at finite strains: formulation, numerical analysis and implementation", *Comput. Methods Appl. Mech. Eng.* **98**:1 (1992), 41–104.

[Simo and Tarnow 1992] J. C. Simo and N. Tarnow, "The discrete energy-momentum method. Conserving algorithms for nonlinear elastodynamics", *Z. Angew. Math. Phys.* **43**:5 (1992), 757–792.

[Simo et al. 1992] J. C. Simo, N. Tarnow, and K. K. Wong, "Exact energy-momentum conserving algorithms and symplectic schemes for nonlinear dynamics", *Comput. Methods Appl. Mech. Eng.* **100**:1 (1992), 63–116.

[Söderlind 2002] G. Söderlind, "Automatic control and adaptive time-stepping", *Numer. Algorithms* **31**:1-4 (2002), 281–310.

[Strehmel et al. 2012] K. Strehmel, R. Weiner, and H. Podhaisky, *Numerik gewöhnlicher differentialgleichungen*, 2nd ed., Vieweg + Teubner Verlag, Wiesbaden, Germany, 2012.

[Tsakmakis and Willuweit 2004] C. Tsakmakis and A. Willuweit, "A comparative study of kinematic hardening rules at finite deformations", *Int. J. Non-Linear Mech.* **39**:4 (2004), 539–554.

[Wendt et al. 2015] G. Wendt, P. Erbts, and A. Düster, "Partitioned coupling strategies for multi-physically coupled radiative heat transfer problems", *J. Comput. Phys.* **300** (2015), 327–351.

[Wood et al. 1980] W. L. Wood, M. Bossak, and O. C. Zienkiewicz, "An alpha modification of Newmark's method", *Internat. J. Numer. Methods Eng.* **15**:10 (1980), 1562–1566.

[Wriggers 2001] P. Wriggers, *Nichtlineare finite-elemente Methoden*, Springer, Berlin, 2001.

[Zeng et al. 1992] L. F. Zeng, N.-E. Wiberg, X. D. Li, and Y. M. Xie, "A posteriori local error estimation and adaptive time-stepping for newmark integration in dynamic analysis", *Earthquake Eng. Struct.* **21**:7 (1992), 555–571.

[Zienkiewicz and Taylor 2000] O. C. Zienkiewicz and R. L. Taylor, *The finite element method*, vol. 2, 5th ed., Butterworth-Heinemann, Oxford, England, 2000.

[Zienkiewicz and Xie 1991] O. C. Zienkiewicz and Y. M. Xie, "A simple error estimator and adaptive time stepping procedure for dynamic analysis", *Earthquake Eng. Struc.* **20**:9 (1991), 871–887.

[Zienkiewicz et al. 1984] O. C. Zienkiewicz, W. L. Wood, N. W. Hine, and R. L. Taylor, "A unified set of single step algorithms I: General formulation and applications", *Internat. J. Numer. Methods Eng.* **20**:8 (1984), 1529–1552.

Received 28 Feb 2016. Revised 21 May 2016. Accepted 13 Jun 2016.

MATTHIAS GRAFENHORST: `matthias.grafenhorst@tu-clausthal.de`
*Institute of Applied Mechanics, Clausthal University of Technology, Adolph-Roemer-straße 2A, D-38678 Clausthal–Zellerfeld, Germany*

JOACHIM RANG: `j.rang@tu-bs.de`
*Institute of Scientific Computing, Technical University Braunschweig, Hans–Sommer–Str. 65, D-38106 Braunschweig, Germany*

STEFAN HARTMANN: `stefan.hartmann@tu-clausthal.de`
*Institute of Applied Mechanics, Clausthal University of Technology, Adolph–Roemer–Straße 2A,*
*D-38678 Clausthal–Zellerfeld, Germany*

JOURNAL OF MECHANICS OF MATERIALS AND STRUCTURES
Vol. 12, No. 1, 2017
dx.doi.org/10.2140/jomms.2017.12.93

# COMPUTER SIMULATION OF THE EFFECTIVE VISCOSITY IN BRINKMAN FILTRATION EQUATION USING THE TREFFTZ METHOD

JAN ADAM KOŁODZIEJ, MAGDALENA MIERZWICZAK AND JAKUB KRZYSZTOF GRABSKI

This paper presents a determination of the effective viscosity in the Brinkman equation by a numerical simulation of an imaginary physical experiment with a viscometer. The model of a porous medium and the applied method of solution are very simple. In the idealized problem, we consider axial flow through an infinite array of cylindrical rods. Assuming that the flow in such a porous medium is described by the Brinkman filtration equation, the effective viscosity can be calculated as a function of the porosity. In this paper, a relation between the volume fraction and the effective viscosity is given for triangular and square arrays of rods.

## 1. Introduction

Usually for slow viscous fluid flow in porous media, the Darcy equation,

$$\boldsymbol{q} = -\frac{k}{\mu}\nabla P, \tag{1}$$

is used, where $\boldsymbol{q}$ is the macroscopic velocity [m/s], $P$ is the pressure [Pa], $\mu$ is the viscosity of fluid [Pa·s], and $k$ is the permeability of the porous medium [m$^2$].

For a porous medium with very high porosity in the presence of a free fluid region or wall-bounded porous medium (see Figure 1), sometimes the Brinkman equation is applied [Chen and Wang 2014; Cortez et al. 2010; Hill and Straughan 2009; Parvazinia et al. 2006; Tan and Pillai 2009]:

$$\nabla P = -\frac{\mu}{k}\boldsymbol{q} + \tilde{\mu}\nabla^2\boldsymbol{q}, \tag{2}$$

where $\tilde{\mu}$ is the effective viscosity of the porous medium [Pa·s].

Brinkman [1949] considered the viscous force exerted on a dense swarm of particles by fluid flowing through them. The validity and theoretical justification of the Brinkman equation were presented in the papers [Durlofsky and Brady 1987; Kim and Russel 1985; Lundgren 1972; Rubinstein 1986; Tam 1969; Vafai and Kim 1995]. However, there are also publications in which the authors questioned the application of the Brinkman equation as a proper mathematical description for filtration flow [James and Davis 2001; Nield 1983]. In their opinion, the main problem is related to the effective viscosity (which depends on the flow) and the unknown boundary condition between the porous medium and the pure fluid area.

Almost all authors using the Brinkman equation assumed that the effective viscosity $\tilde{\mu}$ was equal to the pure fluid viscosity $\mu$.

---

*Keywords:* porous media, Brinkman equation, effective viscosity, Trefftz method.

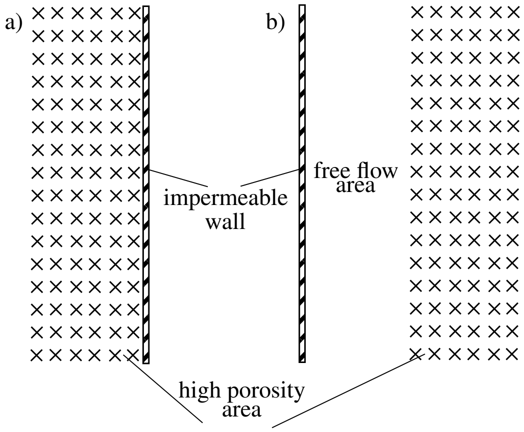

**Figure 1.** Porous medium with very high porosity in the presence of free fluid region (b) or only wall-bounded (a).

There are a limited number of papers in which the effective viscosity in the Brinkman filtration equation is determined. Brinkman [1949] suggested a possible use of Einstein's formula to describe the viscosity for a suspension, given as

$$\frac{\tilde{\mu}}{\mu} = 1 + 2.5\phi, \tag{3}$$

where $\phi$ is the volume fraction of the skeleton of the porous medium (the porosity of the medium equals $1 - \phi$). According to (3), the effective viscosity is greater than the fluid viscosity. However, it should be noticed that the porous medium differs significantly from the suspension. In the porous medium, the skeleton is fixed; while in the suspension, particles move during fluid motion. The fact that the effective viscosity is greater than the viscosity of the pure fluid is correct for the suspension. It does not have to be correct for the porous media. In the present paper, we show that the effective viscosity of the Brinkman equation is less than the fluid viscosity. However, it is a fact that in the literature some authors obtained an effective viscosity that is greater than the fluid viscosity.

Lundgren [1972] determined the ratio $\tilde{\mu}/\mu$ as a function of the volume fraction of the skeleton of the porous medium, which was modeled by the fixed particles. The term $\tilde{\mu}/\mu$ rises slightly above one at the beginning as the volume fraction increases. This ratio reaches a maximum at $\phi = 0.2$. Koplik et al. [1983] obtained the effective viscosity that is less than the pure fluid viscosity (calculating the dissipation energy around the fixed particle):

$$\frac{\tilde{\mu}}{\mu} = 1 - 0.5\phi. \tag{4}$$

Ochoa-Tapia and Whitaker [1995] used the volume averaging method in order to derive the differential equation for flow in the porous medium for the balance of momentum. They obtained the Brinkman equation with the effective viscosity in the form

$$\frac{\tilde{\mu}}{\mu} = \frac{1}{1 - \phi}. \tag{5}$$

That means that the effective viscosity is greater than the fluid viscosity. The effective medium theory was used by Freed and Muthukumar [1978]. They obtained the effective viscosity for a swarm of particles in the form

$$\frac{\tilde{\mu}}{\mu} = 1 + \tfrac{5}{2}\phi - \frac{9 \cdot \phi^{3/2}}{2\sqrt{2}}. \tag{6}$$

In this case the effective viscosity is greater than the viscosity of the pure fluid but is smaller than the effective viscosity resulting from (3).

In [Martys et al. 1994], the authors investigated flow of the fluid in the porous medium bounded by the free fluid region using a computer simulation. The porous medium was modeled as spherical particles, whose size and location was determined randomly. As a result, they obtained the ratio $\tilde{\mu}/\mu = 1.9 - 4.2$ at $\phi = 0.2 - 0.5$.

The effective viscosity in the Brinkman–Forchheimer equation was determined experimentally by Givler and Altobelli [1994]. In this work, the filtration equation had the form

$$\nabla P = -\frac{\mu}{k}\boldsymbol{q} - \frac{\rho \cdot c \cdot |\boldsymbol{q}|}{\sqrt{k}}\boldsymbol{q} + \tilde{\mu}\nabla^2\boldsymbol{q}, \tag{7}$$

where $\rho$ is the density of fluid and $c$ is an experimentally determined constant. The authors used the magnetic resonance in order to determine the velocity profile of flow in the porous medium bounded by the walls. They obtained $\tilde{\mu}/\mu = 7.5$ for a Reynolds number equal to 17 and the volume fraction $\phi = 0.028$.

The effective viscosity of the porous medium modeled by regular displaced particles of the same diameter was determined by Starov and Zhdanov [2001] using the finite difference method. The authors proposed the formula

$$\frac{\tilde{\mu}}{\mu} = \frac{1}{(1-\phi)^{5/2}}. \tag{8}$$

According to the above formula, the effective viscosity is greater than the fluid viscosity.

Beavers and Joseph [1967] published a paper about the boundary condition on the boundary between the porous medium and the fluid. In their work, an experiment was conducted in which the volume flow was measured in two connected channels — one with a pure fluid region and the other with fluid in the porous medium (Figure 1a). On the basis of this experimental data, the following boundary condition was proposed by the authors:

$$\left.\frac{\mathrm{d}u}{\mathrm{d}y}\right|_{y=0} = \frac{\alpha}{\sqrt{k}}(u_i - q), \tag{9}$$

where $u$ is the velocity in the fluid layer, $u_i$ is the slip velocity on the boundary between the porous medium and the free flow area (at $y = 0$), $k$ is the permeability of the porous medium, $q$ is the filtration velocity on the boundary between layers, and $\alpha$ is the dimensionless slip coefficient which depends on the porosity and the structure of the porous medium.

Taylor [1971] noticed that the Beavers–Joseph boundary condition can be related to the Brinkman filtration equation, but he did not develop this idea. Neale and Nader [1974] demonstrated that if the porous layer flow is described by the Brinkman equation and the pure fluid layer flow is described by the Poiseuille law, then the relationship $\alpha^2 = \tilde{\mu}/\mu$ is satisfied. On the basis of this relationship and data from experiments and calculations, the value of the constant $\alpha$ can be used to determine the effective viscosity of the porous medium. In the Beavers–Joseph experiment, the constant $\alpha$ was greater or smaller than one and was dependent on the porous material. That implies that the effective viscosity is greater or smaller than the fluid viscosity.

The purpose of the present paper is to determine the effective viscosity in the Brinkman filtration equation for a fibrous porous medium. This viscosity is determined by the numeric simulation of imaginary physical experiments. The porous medium is modeled by a parallel bundle of straight fibers arranged in regular square or triangular arrays. In this way, the relationship between the effective viscosity and the porosity and the relationship between the effective viscosity and the structure of the porous medium are investigated. In the Brinkman filtration equation, there are two parameters of the porous medium: the permeability and the effective viscosity. Because of this, the determination of these parameters is realized in two steps. In the first step, the permeability is determined by assuming an infinite bundle of fibers. In the second step, the effective viscosity of such a medium is determined by knowing the permeability of the porous medium. The numerical investigations are conducted for a Newtonian fluid in a direction parallel to the fibers.

## 2. Imagined physical experiment for the measurement of the effective viscosity

Measurement with a rotational viscometer is one of the methods used to determine the pure fluid viscosity. In this instrument, a fluid layer is located between two cylindrical surfaces. One of the surfaces is immovable and the second one rotates at a constant angular speed. The fluid viscosity is determined by measuring the torque of the rotating cylindrical surface. In this way, it is simple shear flow. It is supposed that the effective viscosity in the Brinkman equation can be determined in a similar way.

Let us consider a layer of porous medium of width $l$, which is filled by a fluid and located between two flat plates. One of the plates is immovable and the second one moves at a constant velocity $U$ (see Figure 2). The Brinkman equation for one-dimensional shear flow in the absence of a pressure gradient through the porous layer has the form

$$\tilde{\mu}\frac{\mathrm{d}^2 q_z}{\mathrm{d}x^2} - \frac{\mu}{k} q_z = 0, \tag{10}$$

where $q_z$ is the filtration velocity in the direction of the fiber axis [m/s] and $k$ is the permeability of the porous medium [m$^2$]. The following boundary conditions are formulated for the problem:

$$q_z = 0 \text{ for } x = 0, \quad \text{and} \quad q_z = U \text{ for } x = l. \tag{11}$$

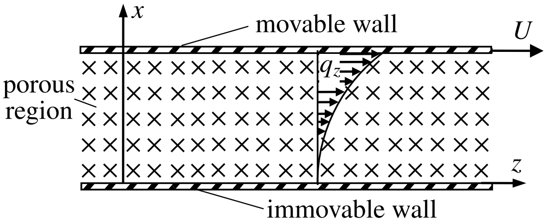

Figure 2. A porous region between two parallel plates.

Let us introduce the nondimensional variables

$$Q = \frac{q_z}{U}, \quad X = \frac{x}{b}, \quad \alpha = \frac{\tilde{\mu}}{\mu}, \quad F = \frac{k}{S}, \quad L = \frac{l}{b}, \tag{12}$$

where $S = \beta b^2$ [m$^2$] and $\beta$ is a dimensionless parameter of the porous medium. Then (10) can be written in the dimensionless form

$$\frac{\mathrm{d}^2 Q}{\mathrm{d}X^2} - \frac{1}{\alpha\beta F} Q = 0 \tag{13}$$

with the boundary conditions

$$Q = 0 \text{ for } X = 0, \quad \text{and} \quad Q = 1 \text{ for } X = L. \tag{14}$$

The exact solution of the problem takes the form

$$Q = \frac{\exp(X/\sqrt{\alpha\beta F}) - \exp(-X/\sqrt{\alpha\beta F})}{\exp(L/\sqrt{\alpha\beta F}) - \exp(-L/\sqrt{\alpha\beta F})}. \tag{15}$$

The tangential stress on the movable plate can be expressed as

$$\tau = \tilde{\mu} \frac{\mathrm{d}q_z}{\mathrm{d}x}\bigg|_{x=L} = \tilde{\mu} \frac{U}{b} \frac{1}{\sqrt{\alpha\beta F}} \coth\left(\frac{L}{\sqrt{\alpha\beta F}}\right). \tag{16}$$

If the dimensionless permeability $F$ and the tangential stress $\tau$ are known, then the constant $\alpha$ can be obtained from (16).

**2.1.** ***Determination of the permeability.*** In this section, a porous medium modeled by a regular array of parallel fibers is considered. Triangular ($\beta = 2\sqrt{3}$) and square ($\beta = 4$) arrays of fibers (see Figure 3) are analyzed.

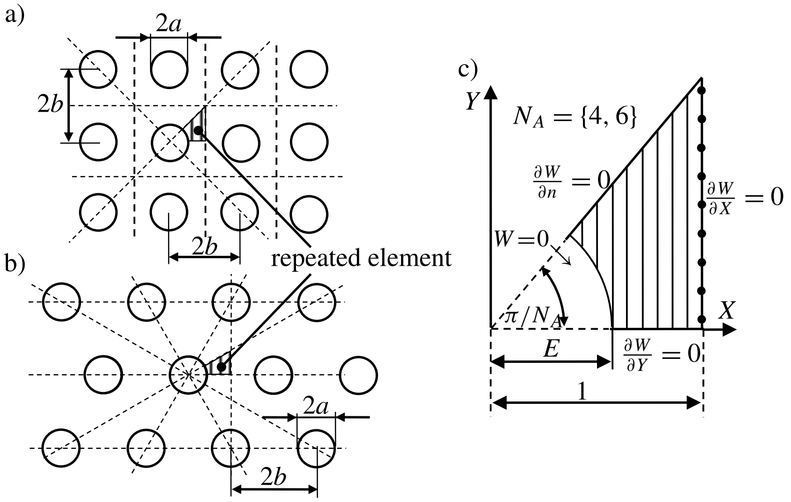

**Figure 3.** Unbounded porous medium: (a) square array, (b) triangular array, (c) repeated element of the array.

One of the basic methods for determining the permeability of the porous medium is an experiment in which the flow rate (the filtration velocity at the known pressure gradient) is measured. Based on this measurement, the permeability can be calculated.

Various authors considered longitudinal flow with respect to the fibers [Banerjee and Hadaller 1973; DeValve and Pitchumani 2012; Drummond and Tahir 1984; Gebart 1992; Happel 1959; Larson and Higdon 1986; Sparrow and Loeffler 1959; Wang 2002]. In the present paper, the same problem is considered but the solution is obtained using the special-purpose Trefftz functions [Mierzwiczak and Kołodziej 2012]. The method is semianalytical. That means that application of the method for the problem considered in the paper gives the analytical form of the dimensionless permeability of the porous medium, in which only the unknown coefficients of the solution are obtained numerically.

Let us consider steady, fully developed, laminar, isothermal flow of an incompressible viscous fluid driven by a constant pressure in a system of regular parallel fibers. The flow is longitudinal with respect to the fibers which are arranged in a regular square (Figure 3a) and triangular (Figure 3b) arrays. The radius of the fibers is equal to $a$, and the distance between the fibers is equal to $2b$. The fluid domain is the whole space $\mathbb{R}^3$. In this case, the equation of motion is reduced to a single partial differential equation in the form (in the polar coordinate system on the $xy$-plane)

$$\frac{\partial^2 w}{\partial r^2}+\frac{1}{r}\frac{\partial w}{\partial r}+\frac{1}{r^2}\frac{\partial^2 w}{\partial \theta^2}=\frac{1}{\mu}\frac{\mathrm{d}p}{\mathrm{d}z} \quad \text{in } \Omega_R, \tag{17}$$

where $w$ is the velocity component in the $z$-axis direction [m/s], $\mathrm{d}p/\mathrm{d}z$ is the constant pressure gradient [Pa/m], $\mu$ is the viscosity of the fluid [Pa·s], and $\Omega_R$ is the repeated element of array (Figure 3). It is convenient to introduce the dimensionless variables

$$W=-\frac{w}{(b^2/\mu)(\mathrm{d}p/\mathrm{d}z)}, \quad R=\frac{r}{b}, \quad E=\frac{a}{b}. \tag{18}$$

Now the governing equation (17) has the dimensionless form

$$\frac{\partial^2 W}{\partial R^2}+\frac{1}{R}\frac{\partial W}{\partial R}+\frac{1}{R^2}\frac{\partial^2 W}{\partial \theta^2}=-1, \tag{19}$$

with the boundary conditions

$$W=0 \quad \text{for } R=E, \tag{20}$$

$$\frac{\partial W}{\partial \theta}=0 \quad \text{for } \theta=\begin{cases}0,\\ \pi/N_A,\end{cases} \tag{21}$$

$$\frac{\partial W}{\partial X}=0 \quad \text{for } X=1, \tag{22}$$

where $W$ is the dimensionless axial velocity, $N_A=6$ for a triangular array and $N_A=4$ for a square array.

The fiber volume fraction is given by

$$\phi=\frac{A_P}{A_T}=\frac{\pi\cdot E^2}{N_A\cdot\tan(\pi/N_A)}, \tag{23}$$

where $A_T=0.5\tan(\pi/N_A)$ is the total area of the repeated element (areas of the fluid and the fiber together) and $A_P=0.5E^2\cdot\pi/N_A$ is the area of the fiber in the repeated element.

The exact solution of (19) can be expressed using the special-purpose Trefftz functions in the form

$$W(R,\theta) = -\tfrac{1}{4}(R^2 - E^2) + \frac{1}{2\cdot\phi_{\mathrm{MAX}}}\ln(R/E) + \sum_{k=1}^{N} B_k\left(R^{N_A k} - \frac{E^{2N_A k}}{R^{N_A k}}\right)\cos(N_A k\theta), \tag{24}$$

where $\phi_{\mathrm{MAX}}$ is the maximal volume fraction (obtained when neighboring fibers are in contact with each other, $E = 1:\ \phi_{\mathrm{MAX}} = \pi\sqrt{3}/6$ for a triangular array $N_A = 6$ and $\phi_{\mathrm{MAX}} = \pi/4$ for a square array $N_A = 4$).

The solution (24) satisfies exactly the boundary conditions (20)–(21) and the balance of the pressure and shear stress. The unknown coefficients $B_k (k = 1, \ldots, N)$ are determined by solving the system of linear equations that results from satisfying the boundary condition (22) using the boundary collocation technique. Imposing these boundary conditions at $N$ collocation points, the following set of equations is to be solved:

$$\sum_{j=1}^{N} A_{kj}\cdot B_k = C_k, \qquad k = 1, 2, \ldots, N,$$

$$A_{kj} = j\cdot N_A\cdot(\cos\theta_k)^{1-N_A j}\big\{\cos[(N_A j - 1)\theta_k] + (\phi/\phi_{\mathrm{MAX}})^{N_A j}\cdot(\cos\theta_k)^{2N_A j}\cos[(N_A j + 1)\theta_k]\big\},$$

$$C_k = \tfrac{1}{2} - \frac{1}{2\cdot\phi_{\mathrm{MAX}}}\cos^2\theta_k, \qquad \theta_k = \frac{\pi(k-1)}{N_A(N-1)}. \tag{25}$$

Using Darcy's law (1), the longitudinal permeability can be related to the average velocity through the repeated element of the fiber system. The longitudinal component of the filtration velocity has the form

$$q_z = \frac{1}{A_T}\iint_{A_F} w\cdot \mathrm{d}A_F, \tag{26}$$

where $A_F = A_T - A_P$ is the region occupied by the fluid in the repeated element.

Using the definition of the nondimensional velocity given by (18), and introducing the fiber volume fraction, the longitudinal component of the filtration velocity can be expressed as

$$q_z = -\frac{b^2}{\mu}F(\phi)\frac{\mathrm{d}p}{\mathrm{d}z}, \tag{27}$$

where

$$F(\phi) = \frac{\iint_{A_F} W(R,\theta)\,\mathrm{d}A_F}{\beta\cdot A_T} \tag{28}$$

is the nondimensional component of the permeability tensor in the direction parallel to the fibers.

After analytical integration in (28), the dimensionless permeability becomes a function of the number of collocation points $N$ and can be calculated from

$$F = \frac{1}{4\pi}\left\{\ln\frac{1}{\phi} + C + 2\phi - \frac{\phi^2}{2} + D\cdot\sum_{k=1}^{N} B_k\left[\frac{H_k}{N_A k + 2} + \Big(\frac{\phi}{\phi_{\mathrm{MAX}}}\Big)^{N_A k}\frac{G_k}{N_A k - 2}\right]\right\}, \tag{29}$$

where

$$H_k = \frac{\sin\big[(N_A k + 1)\pi/N_A\big]}{(N_A k + 1)\big[\cos(\pi/N_A)\big]^{N_A k + 1}}, \qquad G_k = \frac{\sin\big[(1 - N_A k)\pi/N_A\big]}{(1 - N_A k)\big[\cos(\pi/N_A)\big]^{1 - N_A k}}. \tag{30}$$

**2.2.** ***Determination of the effective viscosity.*** After determining the permeability of the porous medium, microstructural flow in the layer of porous medium modeled by a bundle of regular arranged (triangular or square) arrays of fibers can be considered.

Let us consider a layer of the porous medium located between two plates — the first one is fixed and the second one is movable (Figure 4). Because the flow is driven only by the movable wall (the pressure gradient equals zero), it is described by the Laplace equation,

$$\frac{\partial^2 W}{\partial R^2}+\frac{1}{R}\frac{\partial W}{\partial R}+\frac{1}{R^2}\frac{\partial^2 W}{\partial \theta^2}=0, \tag{31}$$

with no-slip boundary conditions: $W=0$ on the immovable wall and fibers and $W=1$ on the movable wall.

Since the array of fibers is streaked and periodic in one direction, it is sufficient to consider the problem only in one repeated strip; depicted in Figure 5 — for the square (Figure 5a) and triangular (Figure 5b) arrays. In both cases, the repeated strip is divided into smaller elements associated with each of the fibers, which are called large finite elements in rest of the paper.

The flow problem is solved by means of the Trefftz method using the special-purpose Trefftz functions. These functions are associated with the large finite elements (Figure 5). For each large finite element, the approximate solution is expressed as

$$W(R,\theta)=C_1\ln(R/E)+\sum_{k=2}^{N}C_k\left[R^{(k-1)}-\frac{E^{2(k-1)}}{R^{(k-1)}}\right]\cos[(k-1)\theta], \tag{32}$$

which satisfies exactly the governing equation (31) and some of the boundary conditions (see Figure 6). For each large finite element, the origin of the coordinates $(R,\theta)$ is placed at the center of fiber. The unknown coefficients $C_k (k=1,2,\dots,N)$ are determined using the boundary collocation technique by satisfying the remaining boundary conditions (in particular, the splitting boundary conditions between the large finite elements).

After determining the unknown coefficients $C_k (k=1,2,\dots,N)$ of the approximate solution, the tangential stress on the movable wall can be determined from

$$\tau=\mu\frac{U}{b}\int_0^{\tan(\pi/N_A)}\frac{\mathrm{d}W}{\mathrm{d}X}\,\mathrm{d}Y, \tag{33}$$

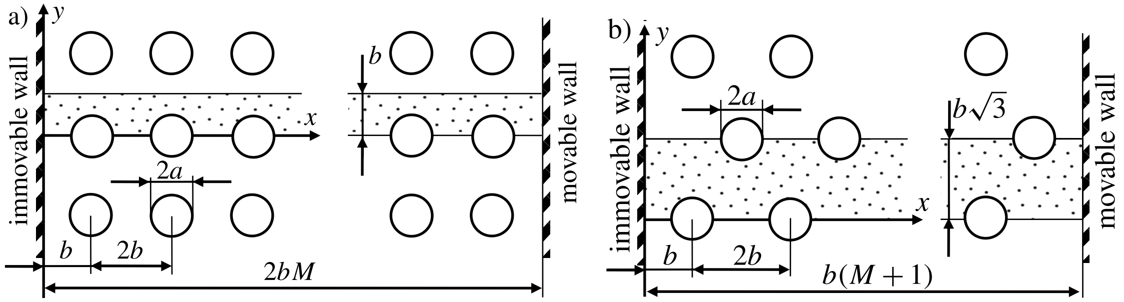

**Figure 4.** Wall-bounded porous medium: (a) square array (b) triangular array.

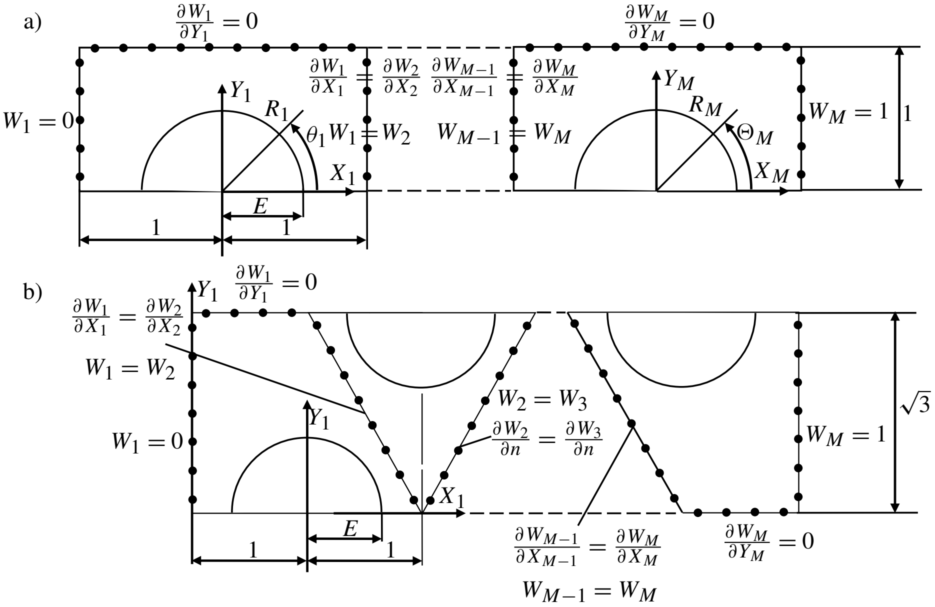

**Figure 5.** The symmetry lane of porous medium: square (a) and triangular array (b).

where the derivative $\mathrm{d}W/\mathrm{d}X$ can be easily obtained from the formula

$$\frac{\mathrm{d}W}{\mathrm{d}X}=\frac{\partial W}{\partial R}\cos(\theta)-\frac{1}{R}\frac{\partial W}{\partial \theta}\sin(\theta). \tag{34}$$

Comparing (16) and (33), the following relationship can be written:

$$\alpha=\frac{c(\phi,M,N)\cdot\sqrt{\alpha\beta F}}{\coth(L/\sqrt{\alpha\beta F})}, \tag{35}$$

where $c(\phi,M,N)=\int_0^{\tan(\pi/N_A)}(\mathrm{d}W/\mathrm{d}X)\,\mathrm{d}X$ and $M$ denotes the number of fiber rows.

## 3. Numerical results and discussion

The numerical calculations were performed for the fibrous porous medium with cylindrical fibers arranged according to the square and triangular arrays. The value of the viscosity and the dimensionless effective permeability as a function of the fiber volume fraction are presented in Figures 7a and 7b, respectively. The ratio of the effective viscosity to the viscosity of the pure fluid decreases as the volume fraction of fibers increases, and for the square array, the ratio is greater than for the triangular array. Permeability $F$ for both types of fiber arrays is similar and substantially different for $\phi>0.3$. The problem was solved for $N=10$ collocation points on the boundary. The dimensionless effective viscosity $\alpha$ was calculated for 5 rows of fibers for the square array and 8 rows for the triangular array.

The effect of the number of rows of the fibers $M=\{1,2,3,5\}$ on the value of the effective permeability for the square array was examined. In this experiment, 10 collocation points on the boundary were used. Table 1 shows the numerical results. For a small value of the volume fraction $\phi$, the influence of the number of fiber rows is significant. For $\phi=0.1$ with 2 or more rows of fibers, the same value of $\alpha$ was obtained. For $\phi=0.01$ the same value of $\alpha$ was obtained for 3 or more rows of fibers. The calculation

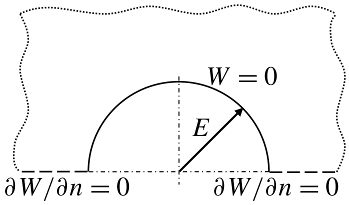

**Figure 6.** Boundary conditions satisfied by the special-purpose Trefftz functions applied in Equation (32).

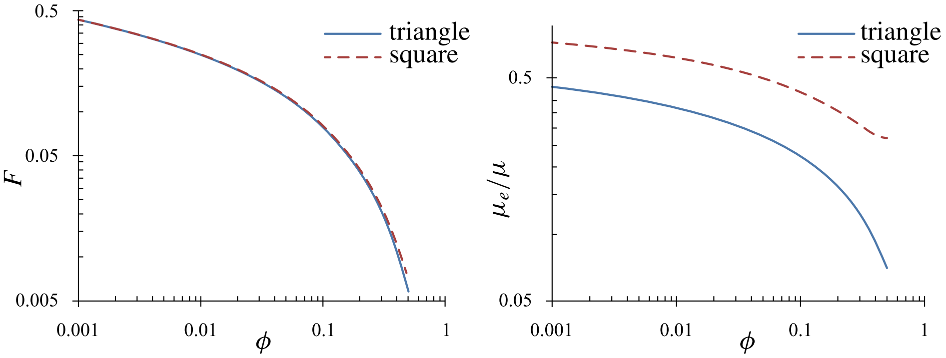

**Figure 7.** Values of the dimensionless effective viscosity $\alpha$ and the dimensionless permeability $F$, versus the values of the volume fraction $\phi$ for the square and triangular array.

| $\phi \downarrow M \rightarrow$ | 1 | 2 | 3 | 5 |
|---|---|---|---|---|
| $1.0 \cdot 10^{-8}$ | 1.8182 | 1.0103 | 0.9103 | 0.8906 |
| $1.0 \cdot 10^{-7}$ | 1.6394 | 0.9649 | 0.8886 | 0.8754 |
| $1.0 \cdot 10^{-6}$ | 1.4551 | 0.9173 | 0.8633 | 0.8554 |
| $1.0 \cdot 10^{-5}$ | 1.2664 | 0.8659 | 0.8319 | 0.8280 |
| $1.0 \cdot 10^{-4}$ | 1.0729 | 0.8068 | 0.7892 | 0.7878 |
| $1.0 \cdot 10^{-3}$ | 0.8722 | 0.7306 | 0.7246 | 0.7243 |
| $1.0 \cdot 10^{-2}$ | 0.6604 | 0.6170 | 0.6163 | 0.6163 |
| $1.0 \cdot 10^{-1}$ | 0.4329 | 0.4312 | 0.4312 | 0.4312 |

**Table 1.** Influence of the number of the rows of fibers {1, 2, 3, 5} for the square array on the values of the dimensionless effective viscosity $\alpha = \mu_e/\mu$ for different values of the volume fraction $\phi$.

of the effective permeability of the porous medium for $\phi \geq 0.01$ can be reduced to the numerical model with two rows of cylindrical fibers.

Since some of the boundary conditions are satisfied in an approximate sense for the collocation points located on the edge of the considered small repeated element, in the next step the effect of the number of these points on the value of the permeability $F$ was checked. Table 2 shows the results for various values of the fiber volume fraction $\phi = \{0.0001, 0.001, 0.01, 0.1\}$ for the square array of fibers. Whatever

| $N \downarrow \phi \rightarrow$ | 0.0001 | 0.001 | 0.01 | 0.1 |
|---|---|---|---|---|
| 2 | 0.61551076 | 0.43242005 | 0.25061461 | 0.08131019 |
| 3 | 0.61546913 | 0.43237843 | 0.25057299 | 0.08126870 |
| 4 | 0.61546816 | 0.43237745 | 0.25057202 | 0.08126775 |
| 5 | 0.61546826 | 0.43237756 | 0.25057212 | 0.08126785 |
| 6 | 0.61546825 | 0.43237755 | 0.25057211 | 0.08126784 |
| 7 | 0.61546825 | 0.43237755 | 0.25057211 | 0.08126784 |
| 8 | 0.61546825 | 0.43237755 | 0.25057211 | 0.08126784 |

**Table 2.** Influence the number $N$ of collocation points (square array) on the dimensionless permeability $F$, for different values of the volume fraction $\phi$.

| $N \downarrow \phi \rightarrow$ | $1.0 \cdot 10^{-8}$ | $1.0 \cdot 10^{-6}$ | 0.0001 | 0.001 | 0.01 | 0.1 |
|---|---|---|---|---|---|---|
| 2 | 0.410243 | 0.468108 | 0.565460 | 0.646570 | 0.783514 | 1.150752 |
| 3 | 0.411038 | 0.468957 | 0.566328 | 0.647407 | 0.784279 | 1.151888 |
| 4 | 0.410907 | 0.468816 | 0.566182 | 0.647263 | 0.784143 | 1.151706 |
| 5 | 0.410930 | 0.468841 | 0.566207 | 0.647288 | 0.784166 | 1.151726 |
| 6 | 0.410926 | 0.468836 | 0.566202 | 0.647283 | 0.784162 | 1.151724 |
| 7 | 0.410926 | 0.468837 | 0.566203 | 0.647284 | 0.784162 | 1.151724 |
| 8 | 0.410926 | 0.468837 | 0.566203 | 0.647284 | 0.784162 | 1.151724 |
| 9 | 0.410926 | 0.468837 | 0.566203 | 0.647284 | 0.784162 | 1.151724 |

**Table 3.** Values of $c(\phi, M, N)$ for the square array, versus the number $N$ of collocation points, for different values of the volume fraction $\phi$.

the value of $\phi$, the converged results were obtained for $N \geq 6$ collocation points. The present method requires a small number of collocation points; hence, it is very effective and is not time consuming.

Table 3 presents the effect of the number of collocation points $N$ on the constant $c(\phi, M, N)$ for different values of the volume fraction of fibers $\phi$ for the square array with five rows of fibers. As in the previous case, the converged solutions for the constant $c$ are obtained for $N \geq 6$.

Comparison of the results obtained in this study by means of the Trefftz method with other works is presented in Figure 9. The results clearly show that the ratio $\tilde{\mu}/\mu$ is less than one, which means that the effective viscosity $\tilde{\mu}$ is less than the fluid viscosity $\mu$.

## 4. Conclusions

In the Brinkman filtration equation there are two parameters of the porous medium: the permeability and the effective viscosity. These two parameters were determined by numerical simulation of an imaginary physical experiment. The porous medium was modeled by a parallel bundle of straight fibers arranged in a regular triangular or square array. In order to determine the permeability, the flow driven by the pressure gradient in an unbounded porous medium was considered. To determine the effective viscosity,

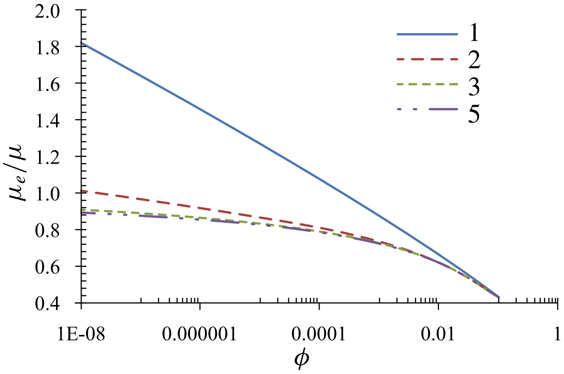

**Figure 8.** Influence of the number of the rows of fibers $\{1, 2, 3, 5\}$ for the square array on the values of the dimensionless effective viscosity $\alpha$ for different values of the volume fraction $\phi$.

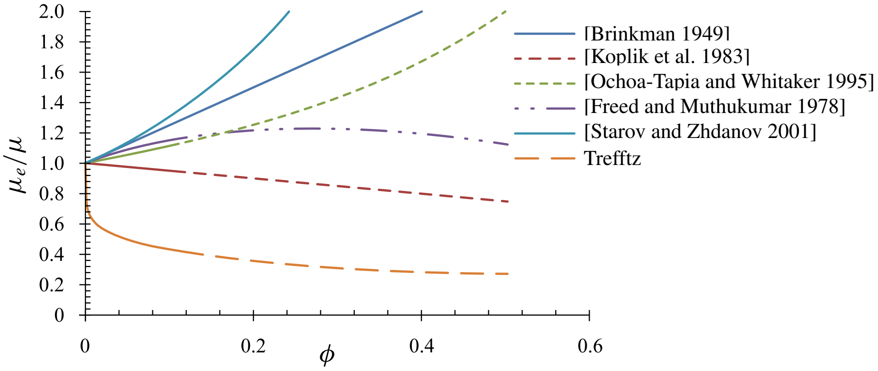

**Figure 9.** Comparison of the effective viscosity obtained in the present study with the others results.

the shear flow in a flat layer of porous medium was considered. In both numerical simulations, the Trefftz method with the special-purpose Trefftz functions was applied.

In all considered cases: the square and the triangular array for different porosity, in the Brinkman filtration equation the effective viscosity is lower than the viscosity of the pure fluid. The ratio of the effective viscosity to the viscosity of the pure fluid decreases as the volume fraction of fibers increases (decreasing porosity). For the same volume fraction of the fibers, both the permeability and the effective viscosity for the triangular array is smaller than for the square array. This difference is not significant in the permeability case, but is very important for the effective viscosity.

Parameters of the numerical method (for example the number of collocation points or the number of rows of fibers in the layer) have inconsiderable effect on the simulation results. To determine the permeability, a few collocation points are enough to get an accurate result ($N = 6$). For the effective viscosity, the result does not change with the number of rows greater than four.

The obtained results clearly show that the effective viscosity of the porous medium should be less than the fluid viscosity.

## Acknowledgments

This work was supported by the grant 2012/07/B/ST8/03449 from the National Science Center, Poland.

## References

[Banerjee and Hadaller 1973] S. Banerjee and G. I. Hadaller, "Longitudinal laminar flow between cylinders arranged in a triangular array by a variational technique", *J. Appl. Mech.* **40**:4 (1973), 1136–1138.

[Beavers and Joseph 1967] G. S. Beavers and D. D. Joseph, "Boundary conditions at a naturally permeable wall", *J. Fluid Mech.* **30**:1 (1967), 197–207.

[Brinkman 1949] H. C. Brinkman, "A calculation of the viscous force exerted by a flowing fluid on a dense swarm of particles", *Appl. Sci. Res.* **1**:1 (1949), 27–34.

[Chen and Wang 2014] H. Chen and X.-P. Wang, "A one-domain approach for modeling and simulation of free fluid over a porous medium", *J. Comput. Phys.* **259** (2014), 650–671.

[Cortez et al. 2010] R. Cortez, B. Cummins, K. Leiderman, and D. Varela, "Computation of three-dimensional Brinkman flows using regularized methods", *J. Comput. Phys.* **229**:20 (2010), 7609–7624.

[DeValve and Pitchumani 2012] C. DeValve and R. Pitchumani, "An analytical model for the longitudinal permeability of aligned fibrous media", *Compos. Sci. Techno.* **72**:13 (2012), 1500–1507.

[Drummond and Tahir 1984] J. E. Drummond and M. I. Tahir, "Laminar viscous flow through regular arrays of parallel solid cylinders", *Int. J. Multiph. Flow* **10**:5 (1984), 515–540.

[Durlofsky and Brady 1987] L. Durlofsky and J. F. Brady, "Analysis of the Brinkman equation as a model for flow in porous media", *Phys. Fluids* **30**:11 (1987), 3329–3341.

[Freed and Muthukumar 1978] K. F. Freed and M. Muthukumar, "On the Stokes problem for a suspension of spheres at finite concentrations", *J. Chem. Phys.* **68**:5 (1978), 2088–2096.

[Gebart 1992] B. R. Gebart, "Permeability of unidirectional reinforcements for RTM", *J. Compos. Mater.* **26**:8 (1992), 1100–1133.

[Givler and Altobelli 1994] R. C. Givler and S. A. Altobelli, "A determination of the effective viscosity for the Brinkman–Forchheimer flow model", *J. Fluid Mech.* **258** (1994), 355–370.

[Happel 1959] J. Happel, "Viscous flow relative to arrays of cylinders", *AIChE J.* **5**:2 (1959), 174–177.

[Hill and Straughan 2009] A. A. Hill and B. Straughan, "Poiseuille flow in a fluid overlying a highly porous material", *Adv. Water Resour.* **32**:11 (2009), 1609–1614.

[James and Davis 2001] D. F. James and A. M. J. Davis, "Flow at the interface of a model fibrous porous medium", *J. Fluid Mech.* **426** (2001), 47–72.

[Kim and Russel 1985] S. Kim and W. B. Russel, "Modelling of porous media by renormalization of the Stokes equations", *J. Fluid Mech.* **154** (1985), 269–286.

[Koplik et al. 1983] J. Koplik, H. Levine, and A. Zee, "Viscosity renormalization in the Brinkman equation", *Phys. Fluids* **26**:10 (1983), 2864–2870.

[Larson and Higdon 1986] R. E. Larson and J. J. L. Higdon, "Microscopic flow near the surface of two-dimensional porous media, part 1: Axial flow", *J. Fluid Mech.* **166** (1986), 449–472.

[Lundgren 1972] T. S. Lundgren, "Slow flow through stationary random beds and suspensions of spheres", *J. Fluid Mech.* **51**:2 (1972), 273–299.

[Martys et al. 1994] N. Martys, D. P. Bentz, and E. J. Garboczi, "Computer simulation study of the effective viscosity in Brinkman's equation", *Phys. Fluids* **6**:4 (1994), 1434–1439.

[Mierzwiczak and Kołodziej 2012] M. Mierzwiczak and J. A. Kołodziej, "Comparison of different methods for choosing the collocation points in the boundary collocation method for 2D-harmonic problems with special purpose Trefftz functions", *Eng. Anal. Bound. Elem.* **36**:12 (2012), 1883–1893.

[Neale and Nader 1974] G. Neale and W. Nader, "Practical significance of Brinkman's extension of Darcy's law: coupled parallel flows within a channel and a bounding porous medium", *Can. J. Chem. Eng.* **52**:4 (1974), 475–478.

[Nield 1983] D. A. Nield, "The boundary correction for the Rayleigh–Darcy problem: limitations of the Brinkman equation", *J. Fluid Mech.* **128** (1983), 37–46.

[Ochoa-Tapia and Whitaker 1995] J. A. Ochoa-Tapia and S. Whitaker, "Momentum transfer at the boundary between a porous medium and a homogeneous fluid, I: Theoretical development", *Int. J. Heat Mass Transfer* **38**:14 (1995), 2635–2646.

[Parvazinia et al. 2006] M. Parvazinia, V. Nassehi, and R. J. Wakeman, "Multi-scale finite element modelling of laminar steady flow through highly permeable porous media", *Chem. Eng. Sci.* **61**:2 (2006), 586–596.

[Rubinstein 1986] J. Rubinstein, "Effective equations for flow in random porous media with a large number of scales", *J. Fluid Mech.* **170** (1986), 379–383.

[Sparrow and Loeffler 1959] E. M. Sparrow and A. L. Loeffler, "Longitudinal laminar flow between cylinders arranged in regular array", *AIChE J.* **5**:3 (1959), 325–330.

[Starov and Zhdanov 2001] V. M. Starov and V. G. Zhdanov, "Effective viscosity and permeability of porous media", *Colloids Surf., A* **192**:1–3 (2001), 363–375.

[Tam 1969] C. K. W. Tam, "The drag on a cloud of spherical particles in low Reynolds number flow", *J. Fluid Mech.* **38**:3 (1969), 537–546.

[Tan and Pillai 2009] H. Tan and K. M. Pillai, "Finite element implementation of stress-jump and stress-continuity conditions at porous-medium, clear-fluid interface", *Comput. Fluids* **38**:6 (2009), 1118–1131.

[Taylor 1971] G. I. Taylor, "A model for the boundary condition of a porous material, part 1", *J. Fluid Mech.* **49**:2 (1971), 319–326.

[Vafai and Kim 1995] K. Vafai and S. J. Kim, "On the limitations of the Brinkman–Forchheimer-extended Darcy equation", *Int. J. Heat Fluid Flow* **16**:1 (1995), 11–15.

[Wang 2002] C. Y. Wang, "Slow viscous flow between hexagonal cylinders", *Transp. Porous Media* **47**:1 (2002), 67–80.

Received 29 Feb 2016. Revised 22 Aug 2016. Accepted 25 Sep 2016.

JAN ADAM KOŁODZIEJ: `jan.kolodziej@put.poznan.pl`
*Institute of Applied Mechanics, Poznan University of Technology, ul. Jana Pawła II 24, 60-965 Poznań, Poland*

MAGDALENA MIERZWICZAK: `magdalena.mierzwiczak@put.poznan.pl`
*Institute of Applied Mechanics, Poznan University of Technology, ul. Jana Pawła II 24, 60-965 Poznań, Poland*

JAKUB KRZYSZTOF GRABSKI: `jakub.grabski@put.poznan.pl`
*Institute of Applied Mechanics, Poznan University of Technology, ul. Jana Pawła II 24, 60-965 Poznań, Poland*

JOURNAL OF MECHANICS OF MATERIALS AND STRUCTURES
Vol. 12, No. 1, 2017
dx.doi.org/10.2140/jomms.2017.12.107

# NUMERICAL SIMULATIONS OF MECHANICAL PROPERTIES OF ALUMINA FOAMS BASED ON COMPUTED TOMOGRAPHY

ZDZISŁAW NOWAK, MARCIN NOWAK, RYSZARD PĘCHERSKI,
MAREK POTOCZEK AND ROMANA ŚLIWA

The aim of this paper is to apply the results of microtomography of alumina foam to create a numerical model and perform numerical simulations of compression tests. The geometric characteristics of real foam samples are estimated from tomographic and scanning electron microscopy images. The performance of the reconstructed models is compared to experimental values of elastic moduli. A preliminary analysis of failure strength simulations under compression of alumina foam is also provided.

## 1. Introduction

Computed tomography is one of several rapidly developing methods of noninvasive testing that plays an important role in medicine and in related diverse engineering applications (e.g., defect detection or the local characterization of a material's microstructure). Starting with an early paper by Bartholomew and Casagrande [1957], which reported the first images characterizing the density of particles in fluidized systems, industrial applications appeared in the 1980s. The development of high speed computers allowed large amounts of data to be processed, which enabled the creation of a new methodology called computed tomography (CT), leading to its widespread use in phase contrast tomography; the reconstruction of the microstructure of diverse cellular materials of polyurethane, metallic and ceramic skeletons, and metal/ceramic composites; and the characterization of void and reinforcement distributions in engineering materials (see [Baruchel et al. 2000] for one of the first comprehensive overviews of the subject). In [Alié et al. 2006], x-ray microtomography measurements coupled with image analysis were applied to study the quality of alumina foams after gel drying and calcinations.

Later, a new methodology was proposed to estimate 3D displacement fields from pairs of images obtained from x-ray computed microtomography. The method was illustrated with an analysis of a compression test on a polypropylene solid foam [Roux et al. 2008]. The presented methodology is also applicable to other kinds of foams (e.g., ceramic foams).

Complex ceramic shapes can be prepared through several methods, such as injection molding, slip casting, or gelcasting combined with a foaming method. If gelcasting techniques were combined with an emulsion template, the fabrication process of highly porous ceramics (60–95%) could be simplified and their strength could be improved. By the gelcasting technique, porous ceramics could satisfy the criteria

---

Financial support of the Structural Funds in the Operational Program Innovative Economy (IE OP) financed from the European Regional Development Fund Project "Modern material technologies in aerospace industry", Nr POIG.01.01.02-00-015/08-00 is gratefully acknowledged.

*Keywords:* Alumina open-cell foam, computed tomography microstructure, Young's modulus, compressive strength of alumina foams.

of homogeneity, reproducibility, reliability, and processability required for complex commercial ceramic foam shapes. The papers by Yang et al. [2011] and Tulliani et al. [2013] review the development and applications of the gelcasting technique.

In [Nowak et al. 2013; 2015] and [Fey et al. 2015], the application of CT is used for the analysis of microstructural morphology and the mechanical and thermal characterization of alumina gelcast foams manufactured with the use of environmentally friendly gelling agents. In the literature, there is little data of the elastic properties of such kinds of foams. Therefore, the aim of this paper is to present a numerical model of an open-cell foam with different porosities and discuss the estimation methodology of Young's moduli. The applied CT approach based on finite element analysis is hard and time consuming; therefore, the work on an optimized generation of the calculation model is essential.

The structure of cellular alumina can be characterized by its cell topology (open), morphology (i.e., cell size and cell spherical shape), and relative density (i.e., volume of the solid cell wall material divided by the volume of the cellular material). The open-cell foam is a structure where the pores form interconnected networks [Potoczek 2008].

It should be noticed that cellular ceramic materials are characterized by a low density combined with outstanding mechanical, thermal and acoustic properties [Gibson and Ashby 1999]. As a consequence of their interesting properties, open-cell foams can be used in lightweight constructions, combining energy absorption, structural damping, sound absorption and heat insulation. In the literature, the models of cellular materials are based on simplified skeleton geometries, notably the Gibson and Ashby cubic cell [Gibson and Ashby 1999]. The simplification of the real cellular material structure enables an analytical approach, reducing the requirements on computer hardware, (e.g., [Michalska and Pęcherski 2003], where further references can be found). However, the last few years have opened new research possibilities, such as numerical analysis based on microcomputed tomography images applied to the mechanical investigation of cellular materials.

The numerical parameters which are needed to build the unit cell model are based on the data obtained from microtomography images of real foam. Using the procedures described in [Nowak et al. 2013], the analysis of the microtomography images shows that the alumina foams are composed of approximately spherical cells interconnected by circular windows [Potoczek 2008].

In this paper, a novel methodology is proposed to establish the cell and windows distribution from 2D microcomputed tomography ($\mu$CT) scanning images of real alumina foam and how to reconstruct the virtual model of foam geometry. Then the two-step mesh method is employed to discretize the geometrical model by selecting an appropriate node, and finally the FEM of alumina foam is established directly. This approach is used to calculate the compression performance of alumina foam with porosity 74%, 86% and 90%. The calculation of alumina foam can reflect the mechanical behavior in the compression process of open-cell foams.

Elastic properties are predicted and compared with experimental data. The skeleton material of the alumina foam is assumed to be isotropic and linearly elastic. In numerical simulations of the compression test on alumina foam, the bottom surface of the sample is fully constrained and the top surface of the sample is moved parallel to the z-axis. The Young's modulus is estimated by numerically simulating uniaxial compression of the alumina foam for different values of porosity. A comparison of experimental data [Potoczek 2008] with numerical and analytical predictions [Gibson and Ashby 1999] of Young's modulus for $Al_2O_3$ ceramic foams of different porosity is presented.

## 2. Analysis of foam structure using microtomography

A picture of foam, which was produced by gelcasting combined with a foaming method [Potoczek 2008; Ortega et al. 2006] with 86% porosity is presented in Figure 1. The morphology of alumina foams manufactured by gelcasting is composed of approximately spherical cells interconnected with windows.

The microstructure studies of alumina foam with a porosity of 86% are carried out using the computed x-ray microtomograph SkyScan 1174. In order to obtain images with an optimal ratio of resolution to the dimensions of the foam sample, the pixel size is assumed to be equal to 8.08 $\mu$m. As a result, the series of cross-section images of the foam structure (approximately 1100) with dimensions $1536 \times 1164$ pixels were obtained (see Figure 2). The size of the pixel is also compared to the smallest thickness of cell wall. In the alumina foam with a porosity of 86%, this thickness is about 30 $\mu$m, which in terms of pixels gives the value of 4 px. This assumption gives the detailed structure of the foam without neglecting its relevant characteristics.

**2A.** ***Computation of porosity.*** In the first step, all of the cross-section images of the foam are converted to grayscale images where 0 represents the black color, while the value of 255 defines the white color. Next, the threshold value of phase separation (equal to 75) is determined and applied to the series of images. Each pixel with a value less than 75 was included in the pores of the foam (black), while pixels with a value greater than or equal to 75 was included in the skeleton of the foam (white), see Figure 3).

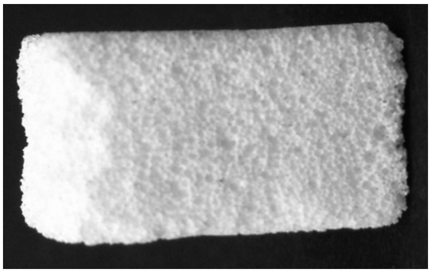

**Figure 1.** An example of real alumina 86% porosity foam produced by gelcasting.

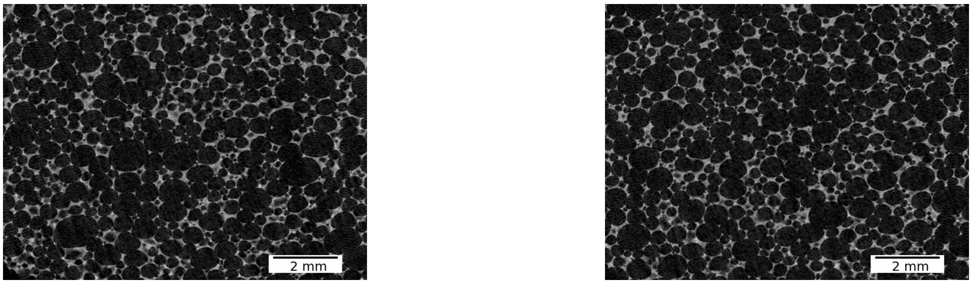

**Figure 2.** Images of cross-sections of the alumina foam with porosity of 86% obtained using computer microtomography: cross-section of the foam at the level $\frac{1}{4}$ of its height (left), cross-section of the foam at the level $\frac{1}{2}$ of its height (right) [Nowak 2014].

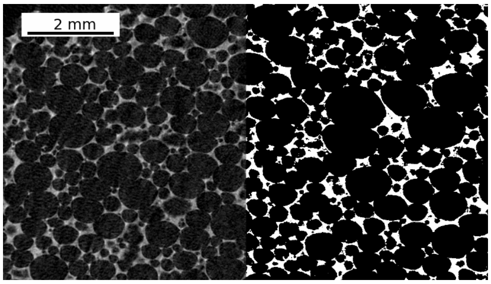

**Figure 3.** Image of cross-section of alumina foam with porosity of 86% after process of separation of two phases: white color — foam skeleton, black color — foam pores.

The threshold of phase separation is determined by comparing the calculated porosity of the foam sample according to (2-1) to the value measured experimentally. If those two porosities are equal, the value of threshold was chosen correctly.

Equation (2-1) gives the porosity of a cuboid shape sample constructed of cross-sections images and takes the form

$$\phi = \frac{\sum_{k=1}^{N} \sum_{i=a_1}^{a_2} \sum_{j=b_1}^{b_2} \operatorname{pix}(k,i,j)}{\sum_{k=1}^{N} \sum_{i=a_1}^{a_2} \sum_{j=b_1}^{b_2}}, \tag{2-1}$$

where $N$ is equal to the number of analyzed images, $i$ and $j$ are the coordinates of the center of pixel, and $(a_1, a_2)$ and $(b_1, b_2)$ are the minimum and maximum coordinates of pixel in horizontal and vertical direction, respectively. The function $\operatorname{pix}(k,i,j)$ is expressed by the formula

$$\operatorname{pix}(k,i,j) = \begin{cases} 1 & \text{if} \quad \operatorname{pix}(k,i,j) < 75, \\ 0 & \text{if} \quad \operatorname{pix}(k,i,j) \geq 75. \end{cases} \tag{2-2}$$

The presented formula of porosity (2-1) depends on the given values of the parameters $a_1$, $a_2$, $b_1$, and $b_2$, which determine the minimum and maximum coordinates of the pixels and the number of analyzed images $N$. In other words, the porosity depends on the calculated volume of the sample. When the volume is larger, the calculated porosity is closer to the value obtained experimentally (Figure 4). For values of $a > 2.0$ mm, where $a$ represents the dimension of the side of the cube, it can be concluded that the effect of the volume change is negligible. Thus, the minimum size of representative volume element (RVE) for foam with a porosity of 86% should be greater than or equal to 2.0 mm.

**2B.** ***Computation of cell and window radii.*** The radii of the cells and windows in the analyzed foam based on cross-section images were obtained with a detection algorithm [Atherton and Kerbyson 1999], available in the MATLAB image processing toolbox. The results of the performed calculations using the program MATLAB are depicted in Figure 5.

We also developed our own procedure for detecting the distribution of the radii of cells. Calculations were carried out in two steps. The first step involves the detection of circular areas with a maximum radius equal to 40 pixels. The second step involves the remaining size of cells. Detection of cells using our procedure is shown in Figure 6. Calculations were made using the open source library PIL (Python Imaging Library) for the Python programming language.

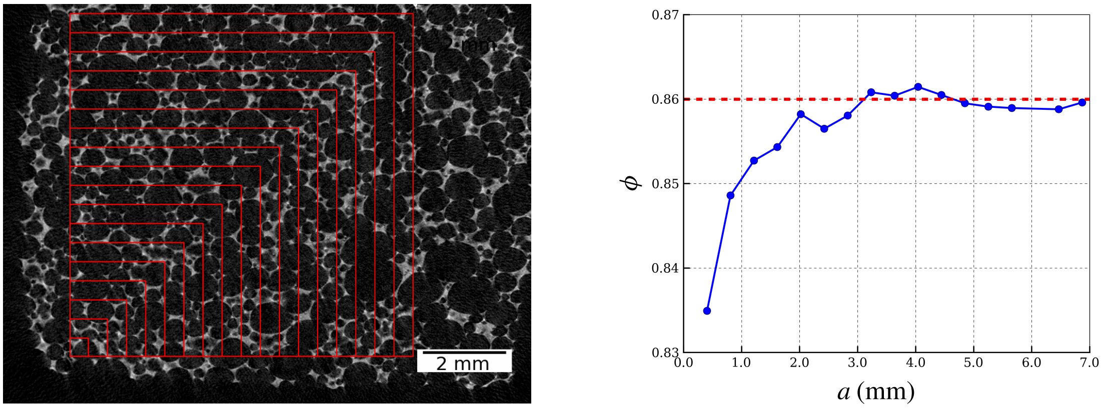

**Figure 4.** The edges of the square areas for computation of porosities (left). Effect of the size of a cubic volume with side $a$ on the calculated porosities in case of sample with average porosity 86% (right).

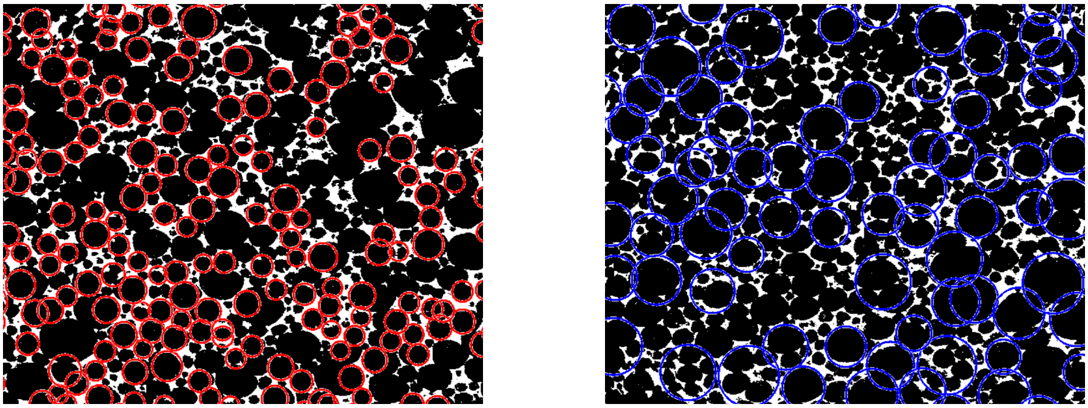

**Figure 5.** Detection of cells in a foam sample for a flat picture $\mu$CT with the use of MATLAB for radii smaller than 40 px (left) and greater than 40 px (right).

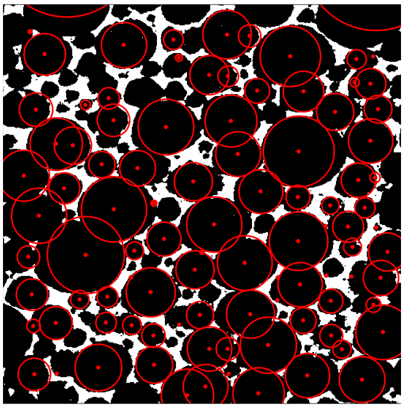

**Figure 6.** Detection of cells in a foam sample for a picture $\mu$CT using the developed algorithm.

**Figure 7.** Two situations of two circles: (left) disjoint circles and (right) circles intersecting at two points.

The computed edges of the cells give a set of overlapping circles, for which two situations can be distinguished (Figure 7). The circles may be separate or intersect at two points. The size of the overlap for the two circles describes the parameter $k$.

By knowing the coordinates of the center of each circle and its radius, the radius of the window $r_w$ can be described using the formula

$$r_w = \frac{\sqrt{2k\left(-\frac{1}{2}k + r_c^a\right)\left(-\frac{1}{2}k + r_c^b\right)(-k + r_c^a + r_c^b)}}{-k + r_c^a + r_c^b}, \tag{2-3}$$

where $r_c^a$ and $r_c^b$ are the radii of circles.

When the circles intersect, the parameter $k$ determines the size of the overlap. For neighboring circles that do not intersect, the parameter $k$ determines the thickness of the cell wall. The parameter $k$ is given by

$$k = r_c^a + r_c^b - \sqrt{(x_b - x_a)^2 + (y_b - y_a)^2}, \tag{2-4}$$

where $(x_a, y_a)$ and $(x_b, y_b)$ are the coordinates of the center of circles.

As a result, the distribution of the cell radius, window radius and the wall thickness were obtained (Figure 8).

A similar study was conducted for foam having a porosity of 74% and 90%. For the obtained distribution of the cell radius and window radius, the probability density functions were proposed. Due to the lack of symmetry in the presented results (Figure 8), the log normal distribution function is assumed. The log normal distribution is a continuous distribution in which the logarithm of a variable has a normal distribution. Thus, its probability density function, $f$, can be written as

$$f(x) = \frac{1}{x\sigma_{\log}\sqrt{2\pi}} \exp\left[-\frac{(\ln(x) - \mu_{\log})^2}{2\sigma_{\log}^2}\right]. \tag{2-5}$$

The parameters for this distribution ($\sigma_{\log}$ and $\mu_{\log}$) are determined by least squares method. Figure 9 shows the density of the probability distribution for the cell radius, $r_c$, and the window radius, $r_w$. When the porosity of the foam increases the mean value of the cell, the windows radius of the foam also increases.

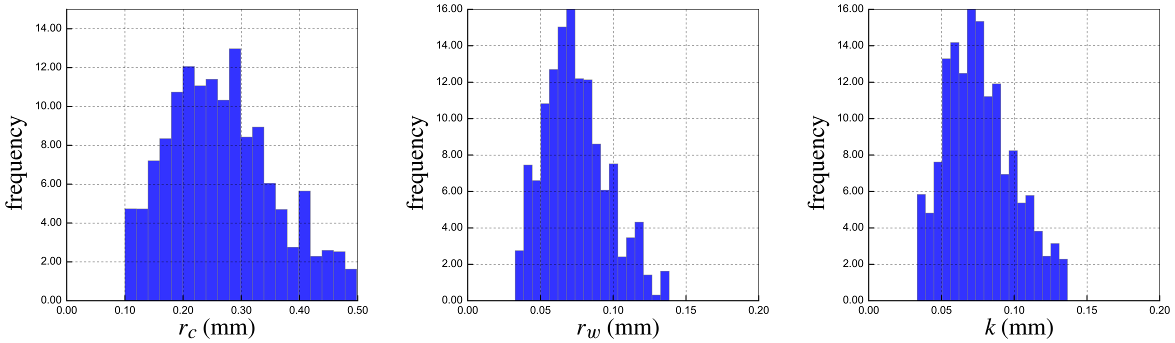

**Figure 8.** The frequency distribution for the foam sample with 86% porosity as a function of cells radius (left), window radius (center), and wall thickness (right).

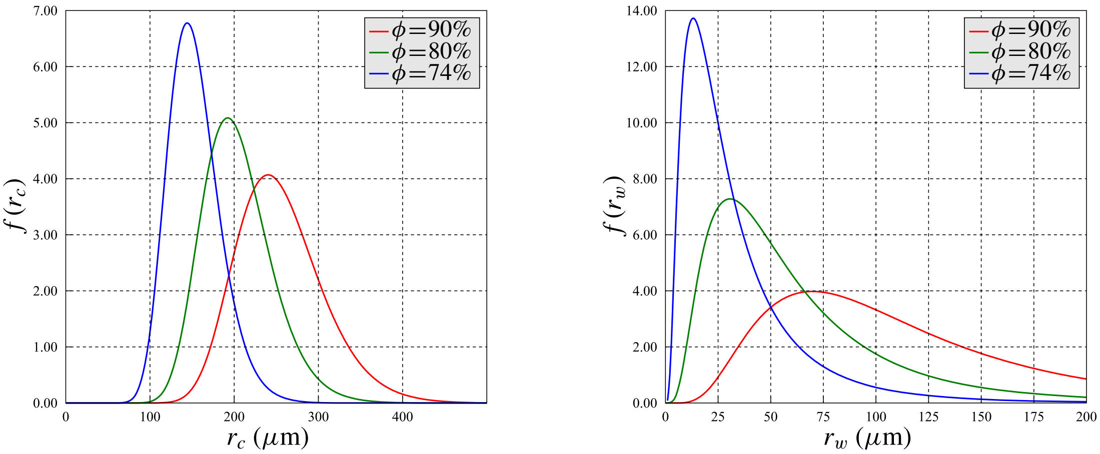

**Figure 9.** The density of the log-normal probability function for the radius of the cells of the foam (left) and the radius of the windows of the foam (right). Distribution parameters were determined using the least squares method.

| porosity (%) | $\mu_{\log}$ ($\mu$m) | $\sigma_{\log}$ ($\mu$m) |
|---|---|---|
| 90 | 256 | 80 |
| 80 | 190 | 33 |
| 74 | 78 | 32 |

**Table 1.** The summary of results for the mean ($\mu_{\log}$) and standard deviation ($\sigma_{\log}$) parameters of the log-normal density function as a function of foam porosity.

A detailed list of the computed parameters for the density of the log-normal probability function is shown in Table 1.

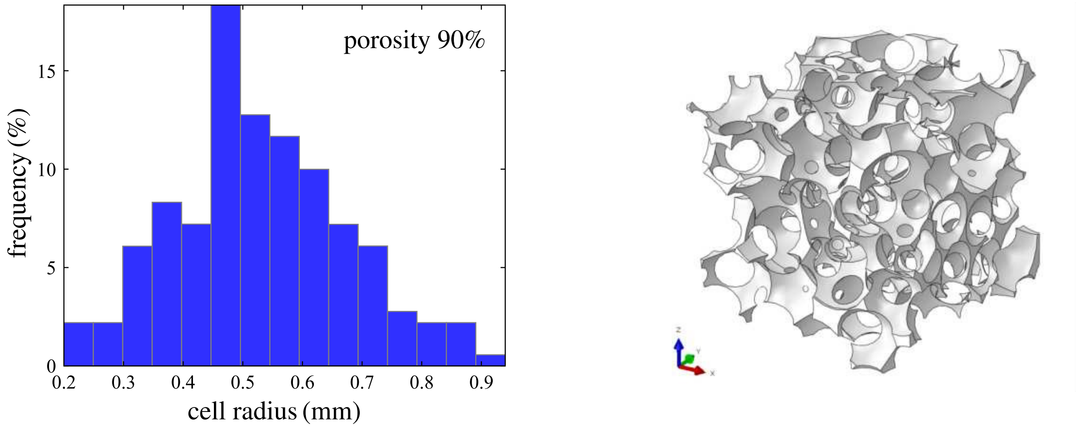

**Figure 10.** Generated foam structure with 90% porosity: histogram of cell size distribution (left); general view of the foam structure (right).

## 3. Numerical model of the geometry of open-cell foam with random cell structure

The structure of real foam is very complex, as shown in the 3D picture of alumina foam with 86% porosity obtained by the gelcasting method (Figure 1). The structure of the foam needs to be simplified. This is done by using a model with a random microstructure, which approximates the distribution and shape of the pores in real foam.

The process is based on impacted bubble simulations of random cells, which generates a foam skeleton that is representative of actual foam microstructures. The geometry of ceramic foams can be generated in three steps:

(1) Diameters of spherical bubbles and windows are estimated from microtomography, see Section 2B.

(2) Coordinates for the centers of the spherical bubbles are produced by Python scripts [Nowak 2014].

(3) The intersecting bubbles are subtracted from the bulk volume of any shape.

The initial data for generating the numerical model for the distribution (2-5), the assumed maximal and minimal values of cell diameter ($r_{\text{min}}$ and $r_{\text{max}}$) for foams of porosity 74%, 86% and 90%, and the assumed number of bubbles are presented in the paper [Nowak et al. 2013]. The particular generated foam structure is illustrated in Figure 10.

## 4. Finite element modeling

Simulating a compression test by means of a finite element modeling of the random alumina foam microstructure allows us to obtain valuable information about the kinematic state of its structure at some of the test stages. The challenge is then to mesh a representative foam volume to obtain a optimal number of degrees of freedom.

**4A.** ***Finite elements mesh generation.*** The volumetric meshing techniques based on tomographic images uses the natural discretization of the image by voxels. In this case, each voxel lying in the solid region would correspond to an 8-node cubic element. This technique, known as the voxel-element technique, can be easily implemented and does not require any additional expense for mesh generation.

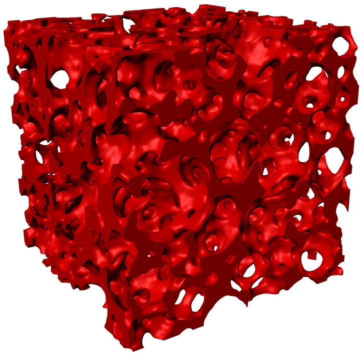

**Figure 11.** An example of rendered alumina 86% porosity foam.

| finite element mesh | (a) | (b) | (c) | (d) |
|---|---|---|---|---|
| average element size | $\frac{1}{5}r_c$ | $\frac{1}{10}r_c$ | $\frac{1}{20}r_c$ | $\frac{1}{40}r_c$ |
| number of elements | 24446 | 101857 | 203678 | 326599 |
| number of degrees of freedom | 53404 | 192294 | 363157 | 562621 |

**Table 2.** The parameters of four generated finite element mesh networks and the numbers of elements and variables in the alumina 90% porosity foam.

But using this method with a large volume would generate a mesh with millions of nodes, beyond the maximum required for convergence purposes. In another approach, an intermediate step is necessary before meshing the solid volume. In this work, the creation of surface models from the tomographic data is carried out automatically with the help of the advancing front method provided by the ScanIP software [ScanIP 2014], see Figure 11. The element type is C3D10 from the ABAQUS commercial code [ABAQUS 2013]. It is a 10-node tetrahedron with quadratic interpolation. Such second-order elements provide higher accuracy than first-order elements for problems that involve complex stress fields. They capture stress concentrations more effectively and are very effective in bending-dominated problems. The dependence of the number of elements and the number of degrees of freedom on the average size of tetrahedron elements in the case of alumina foam with 90% porosity is presented in Table 2.

Finally, the node coordinates information and element information are written into the INP file, and imported into the ABAQUS to calculate the result. The example of meshed alumina 90% porosity foam used in simulations is shown in Figure 12.

**4B.** ***Boundary conditions.*** The displacement boundary conditions are assumed. In a real uniaxial compression test, the bottom sample surface remains fixed and the top one is moved parallel to the z-axis while the lateral surfaces are free of constraint.

**4C.** ***Bulk material properties.*** It is assumed that the bulk material of alumina foam exhibits an elastic behavior. The model describing the behavior of the considered alumina foam is defined in the elastic range by Hooke's law. Two parameters are needed to calibrate this law: Young's modulus and Poisson's

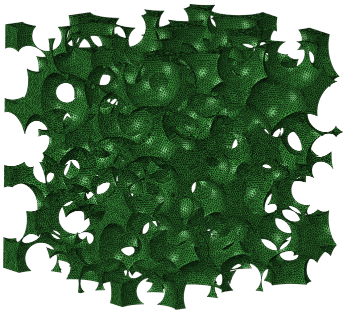

**Figure 12.** An example of meshed alumina 90% porosity foam used in simulations.

ratio. The following material data for $Al_2O_3$ have been retained: a Young's modulus $E_0$ = 370 GPa and a Poisson's ratio $\nu = 0.22$.

## 5. Size of the elementary representative volume

The elementary volume must be large enough to be representative of the overall behavior of the foam. In practice, the elementary volume size is limited, first by the tomography resolution and second by the hardware memory available to generate the mesh and to compute the model. To study the effect of the size of the elementary volume, we extracted regions with different sizes and meshed them with the same mesh density. Different cubic subvolumes with side lengths from 0.1 to 3.0 mm with step size of 0.05 mm were selected. All these volumes are meshed with a mean element volume $V_{el}$ = 0.08 $mm^3$ for the model with an average porosity of 90%.

Each marked point on the obtained curves corresponds to one numerical simulation of the uniaxial compression test. This leads to 150 independent numerical simulations (30 simulations per one sample). To manage this task, a special Python script for ABAQUS/CAE was developed. In the repetitive loop, the following steps are taken:

(a) The geometry based on tomographic images of a given size (from 0.6 mm to 2.5 mm) is created using ScanIP [2014].

(b) The finite element model is created using ABAQUS/CAE.

(c) The numerical simulation is performed using ABAQUS/Standard.

Completion of all numerical simulations took about two weeks. The reason for this is the nonlinearly increasing number of elements needed to discretize given sample during the repetitive loop. For example, with an Intel workstation 8x CPU, 2.66 MHz with 12Gb RAM, the numerical simulation of the 0.6 mm size sample took a few seconds and for the 2.5 mm size sample took a few hours.

The calculated Young's modulus normalized by the matrix Young's modulus ($E/E_0$, where $E_0$ is for alumina matrix) and the porosity of the different models are presented in Figure 13. The calculated normalized Young's modulus $E/E_0$ varies between 0.003–0.019. The sample standard deviation of the calculated normalized Young's modulus is bounded within the limits of 0.003 and 0.014, see Figure 14 (left). On the other hand, the sample standard deviation of the calculated porosity varies from

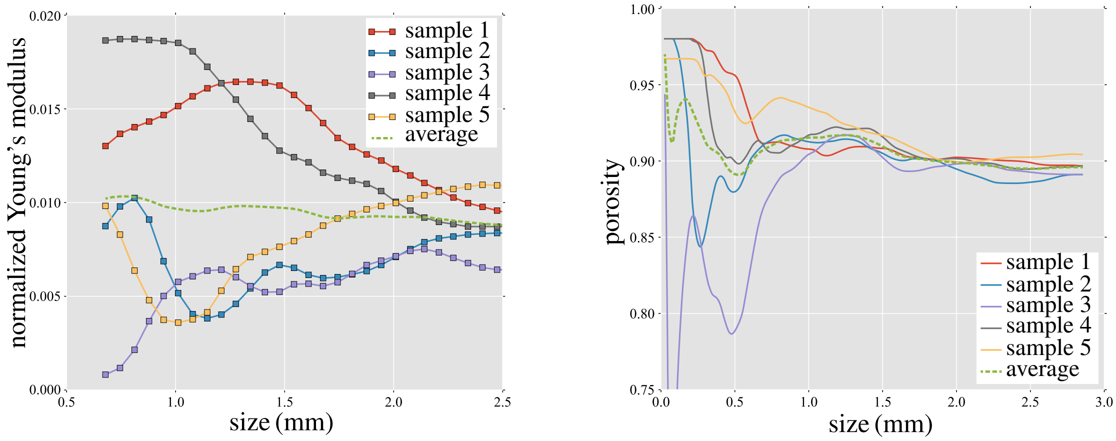

**Figure 13.** The calculated normalized Young's modulus and the porosity of the samples with average porosity 90% as a function of sample size: normalized Young's modulus (left) and porosity (right).

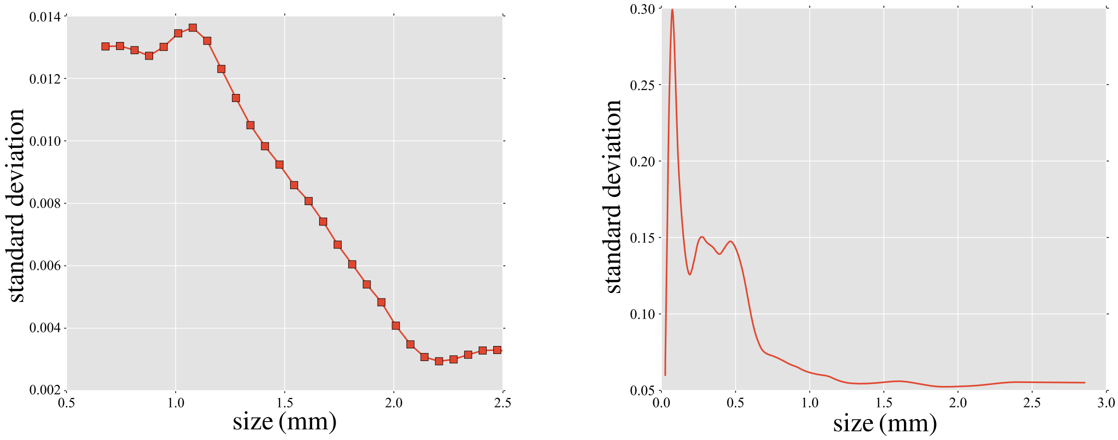

**Figure 14.** The sample standard deviation of the calculated normalized Young's modulus presented in Figure 13 (left) and of the porosity presented in Figure 13 (right) as a function of the sample size.

0.05 to 0.30, see Figure 14 (right). However, the model with a sample size less than 2.0 mm is several times faster than the FE computation and at least four times faster for the model creation. So, the sample with the size 2.0 mm seems to be a good trade-off between the representativity of the model and the macroscopic accuracy.

## 6. Summary and conclusions

The modeling methodology of the mechanical behavior of cellular materials was proposed and validated. The actual microstructure of ceramic foam was obtained by processing microtomographic data. The alumina skeleton of the foam was meshed with tetrahedral finite elements. Using the elastic constitutive

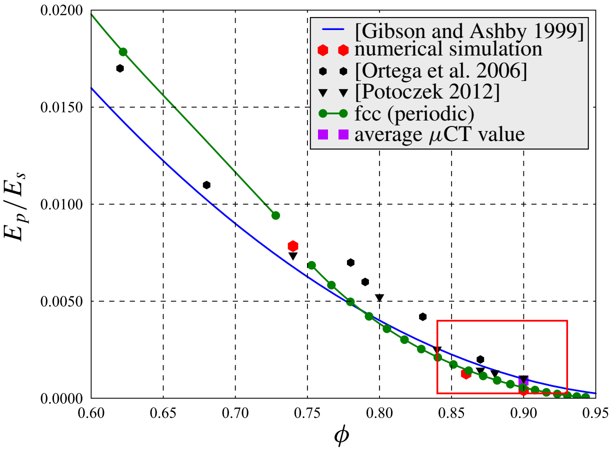

**Figure 15.** The comparison of numerically predicted normalized Young's modulus for the random model, the periodic model, and the analytical results of [Gibson and Ashby 1999] as a function of porosity with the experimental data from [Potoczek 2012; Ortega et al. 2006] and the numerical prediction of the $\mu$CT result presented as a slightly oscillating average-value line varying in sample size in Figure 13 (left).

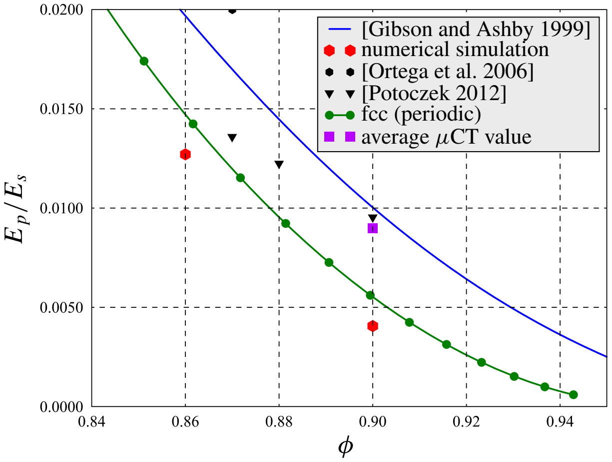

**Figure 16.** The magnified section of Figure 15.

equations of the bulk material, the uniaxial compression test was simulated. The ABAQUS FEM program [ABAQUS 2013] was used to numerically predict the response of alumina foam under compression. The macroscopic response and the local deformation mechanisms were correctly captured.

A comparison of the normalized Young's modulus and the numerical predictions for the random model with the experimental data for alumina foams obtained by gelcasting [Potoczek 2012; Ortega et al. 2006] is presented in Figure 15. A comparison with the results of the numerical simulation for the periodic structure of fcc type [Nowak 2014] is also displayed.

In Figure 16, the magnified section of Figure 15 containing the results of the numerical simulations for a periodic structure of the skeleton, the random structure of the skeleton, and for the average $\mu$CT

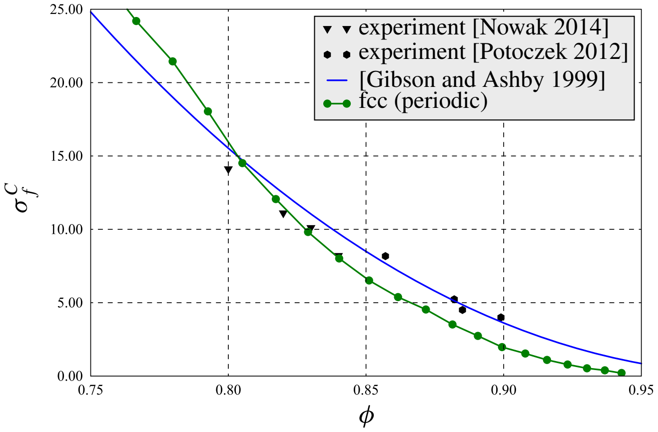

**Figure 17.** The comparison of numerically predicted failure strength for the periodic fcc model as a function of porosity with the experimental data from [Potoczek 2012; Nowak 2014] and the analytical results of [Gibson and Ashby 1999].

value versus the experimental data is shown. It is visible that the discrepancy between all displayed data is very large, what seems to be typical for low density ceramic foams. At the same time, it can be also observed that the difference between the predicted values of the relative Young's modulus for the periodic and random structure of the skeleton is not significant, e.g., for the porosity $\phi = 0.90$ it is equal to 27%. This observation leads to the hypothesis that the numerical simulation of the deformation process under compression until failure performed for the periodic structure of fcc type, see Figure 17, can provide an estimation of the failure strength $\sigma_f^C$ for real alumina foams considered in the paper. The advantage of such an approach lies in its low computational cost. The numerical simulations of the failure processes in alumina foam with random structure, which are much more time consuming, are under development.

The present approach can be very useful for designing cellular materials. Indeed, it enables the prediction of the elastic properties of the ceramic foam with feasible microstructures in order to obtain expected mechanical properties.

The novelty of the present paper is the application of the interconnected cell random model of the open-cell foam in simulations of elastic behavior at the microlevel. Using this model, it is possible to estimate the stress and strain fields in the ceramic phase. It is also possible to estimate the macroscopic properties of the foam as functions of the porosity.

The presented numerical approach in the modeling of open-cell foams has some limitations, which should be discussed. The influence of the mesh quality is not mentioned. It is clear that the accuracy of the finite element calculation depends on the quality of tetrahedral elements. Degenerated tetrahedrons with small volumes may lead to large local errors. For a complex foam structures, degenerated elements can not always be avoided. To eliminate them completely, each degenerated element has to be treated separately, which is a cumbersome operation.

In conclusion:

- The numerical model of real $Al_2O_3$ foam predicting the dependency of Young's modulus while varying porosity within the range from 60% to 95% is discussed and the comparison with experimental data is shown in Figure 15.

- The study of different unit cell sizes, Figure 13, shows that the random representative unit cell reflects the behavior of real $Al_2O_3$ foams produced by the gelcasting method in a satisfactory way.
- The results of the analysis of failure strength for the periodic skeleton structure of fcc type as a function of porosity are discussed, Figure 17.
- The new results obtained in the paper make a foundation for the development of an energy-based limit criterion of ceramic foams with the application of peridynamic states approach, proposed by Silling [2000] and, as a matter of fact, indicated much earlier by Gabrio Piola [dell'Isola et al. 2015]. The important role the criterion plays can be seen in the formula for critical energy density [Foster et al. 2011]:

  $$\omega_c = \frac{4G}{\pi \delta^4}, \tag{6-1}$$

  where $G$ denotes energy release rate and $\delta$ is a material parameter corresponding to the size of a horizon in peridynamic simulations. The analysis of foam microstructures presented in the paper appears to be useful in estimating the parameter $\delta$.

## References

[ABAQUS 2013] "ABAQUS/Standard user's manual", Dassault Systemes, Providence, RI, 2013. Version 6.13.

[Alié et al. 2006] C. Alié, F. Ferauche, A. Léonard, S. Lambert, N. Tcherkassova, B. Heinrichs, M. Crine, P. Marchot, E. Loukine, and J. Pirard, "Pd-Ag/$SiO_2$ xerogel catalyst forming by impregnation on alumina foams", *Chem. Eng. J.* **117**:1 (2006), 13–22.

[Atherton and Kerbyson 1999] T. J. Atherton and D. J. Kerbyson, "Size invariant circle detection", *Image Vision Comput.* **17**:11 (1999), 795–803.

[Bartholomew and Casagrande 1957] R. N. Bartholomew and R. M. Casagrande, "Measuring solids concentration in fluidized systems by gamma-ray absorption", *Ind. Eng. Chem.* **49**:3 (1957), 428–431.

[Baruchel et al. 2000] J. Baruchel, J.-Y. Buffiere, E. Maire, P. Merle, and G. Peix, *X-ray tomography in material sciences*, Hermes Science, Paris, 2000.

[dell'Isola et al. 2015] F. dell'Isola, U. Andreaus, and L. Placidi, "At the origins and in the vanguard of peridynamics, non-local and higher-gradient continuum mechanics: an underestimated and still topical contribution of Gabrio Piola", *Math. Mech. Solids* **20**:8 (2015), 887–928.

[Fey et al. 2015] T. Fey, B. Zierath, P. Greil, and M. Potoczek, "Microstructural, mechanical and thermal characterization of alumina gel-cast foams manufactured with the use of agarose as gelling agent", *J. Porous Mater.* **22**:5 (2015), 1305–1312.

[Foster et al. 2011] J. T. Foster, S. A. Silling, and W. Chen, "An energy based failure criterion for use with peridynamic states", *Int. J. Multiscale Comput. Eng.* **9**:6 (2011), 675–688.

[Gibson and Ashby 1999] L. J. Gibson and M. F. Ashby, *Cellular solids, structure and properties*, 2nd ed., Cambridge Universiy Press, 1999.

[Michalska and Pęcherski 2003] J. M. Michalska and R. B. Pęcherski, "Macroscopic properties of open-cell foams based on micromechanical modelling", *Technische Mechanik* **23**:24 (2003), 234–244.

[Nowak 2014] M. Nowak, *Analiza deformacji i zniszczenia struktur komórkowych w zastosowaniu do symulacji procesu infiltracji pianki korundowej ciekłym metalem*, Ph.D. thesis, 2014, available at http://www.ippt.pan.pl/_download/doktoraty/2014nowak_m_doktorat.pdf.

[Nowak et al. 2013] M. Nowak, Z. Nowak, R. B. Pęcherski, M. Potoczek, and R. Śliwa, "On the reconstruction method of ceramic foam structures and the methodology of Young modulus determination", *Arch. Metall. Mater.* **58** (2013), 1219–1222.

[Nowak et al. 2015] Z. Nowak, M. Nowak, R. B. Pęcherski, M. Potoczek, and R. E. Śliwa, "Mechanical properties of the ceramic open-cell foams of variable cel sizes", *Arch. Metall. Mater.* **60**:3 (2015), 1957–1963.

[Ortega et al. 2006] F. S. Ortega, J. A. Rodrigues, and V. C. Pandolfelli, "Elastic modulus of gelcast cellular ceramics at high temperatures", *Am. Ceram. Soc. Bull.* (2006), 9101–9106.

[Potoczek 2008] M. Potoczek, "Gelcasting of alumina foams using agarose solutions", *Ceram. Int.* **34**:3 (2008), 661–667.

[Potoczek 2012] M. Potoczek, "Kształtowanie mikrostruktury piankowych materiałów korundowych", Rzeszów University of Technology, Rzeszów, Poland, 2012.

[Roux et al. 2008] S. Roux, F. Hild, P. Viot, and D. Bernard, "Three-dimensional image correlation from X-ray computed tomography of solid foam", *Compos. A* **39**:8 (2008), 1253–1265.

[ScanIP 2014] "ScanIP", 2014, available at www.simpleware.com/software/scanip. Version 7.0.

[Silling 2000] S. A. Silling, "Reformulation of elasticity theory for discontinuities and long-range forces", *J. Mech. Phys. Solids* **48**:1 (2000), 175–209.

[Tulliani et al. 2013] J.-M. Tulliani, M. Lombardi, P. Palmero, M. Fornabaio, and L. J. Gibson, "Development and mechanical characterization of novel ceramic foams fabricated by gel-casting", *J. Eur. Ceram. Soc.* **33**:9 (2013), 1567–1576.

[Yang et al. 2011] J. Yang, J. Yu, and Y. Huang, "Recent developments in gelcasting of ceramics", *J. Eur. Ceram. Soc.* **31**:14 (2011), 2569–2591.

Received 29 Feb 2016. Revised 4 Aug 2016. Accepted 4 Nov 2016.

ZDZISŁAW NOWAK: `znowak@ippt.pan.pl`
*Institute of Fundamental Technological Research, Polish Academy of Sciences, Pawińskiego 5B, 02-106 Warsaw, Poland*

MARCIN NOWAK: `nowakm@ippt.pan.pl`
*Institute of Fundamental Technological Research, Polish Academy of Sciences, Pawińskiego 5B, 02-106 Warsaw, Poland*

RYSZARD PĘCHERSKI: `rpecher@ippt.pan.pl`
*Institute of Fundamental Technological Research, Polish Academy of Sciences, Pawińskiego 5B, 02-106 Warsaw, Poland*

MAREK POTOCZEK: `potoczek@prz.edu.pl`
*Faculty of Chemistry, Rzeszów University of Technology, 35-959 Rzeszów, Poland*

ROMANA ŚLIWA: `rsliwa@prz.edu.pl`
*Faculty of Mechanical Engineering and Aeronautics, Rzeszów University of Technology, 35-959 Rzeszów, Poland*

JOURNAL OF MECHANICS OF MATERIALS AND STRUCTURES
Vol. 12, No. 1, 2017
dx.doi.org/10.2140/jomms.2017.12.123

# GRADIENT-ENHANCED LARGE STRAIN THERMOPLASTICITY WITH AUTOMATIC LINEARIZATION AND LOCALIZATION SIMULATIONS

JERZY PAMIN, BALBINA WCISŁO AND KATARZYNA KOWALCZYK-GAJEWSKA

The paper deals with the thermomechanical extension of a large strain hyperelasto-plasticity model and focuses on algorithmic aspects and localization simulations. The formulation includes the degradation of the yield strength due to the increase of an averaged plastic strain measure and temperature, thus, three sources for loss of stability are included in the description. A gradient-enhancement of the model is incorporated through an additional differential equation, but localization is also influenced by heat conduction. The finite element analysis is performed for an elongated plate in plane strain conditions, using different finite elements and values of material parameters related to regularization (internal length scales are related to gradient averaging as well as heat conduction). In particular, the influence of the *F-bar* enrichment on the simulation results is studied. All computational tests are performed using self-programmed user subroutines prepared within a symbolic-numerical tool *AceGen* which is equipped with automatic differentiation options, allowing for automatic linearization of the governing equations.

## 1. Introduction

The research presented in this paper is focused on the development of a gradient-enhanced geometrically nonlinear thermoplasticity model and its numerical analysis with special attention paid to simulation of strain localization caused by material, thermal and geometrical softening. The phenomenon of localization, in which deformation concentrates in narrow bands whereas the rest of the material specimen experiences unloading, is closely connected with the notion of material instability, the theoretical basis of which goes back to [Hill 1958; Thomas 1961; Rice 1976]. Significant contributions to the subject of numerical simulation of instability and strain localization for isothermal conditions were offered among others in [Rudnicki and Rice 1975; Belytschko and Lasry 1989; de Borst et al. 1993; Sluys 1992; Vardoulakis and Sulem 1995; Tvergaard 1999; Menzel 2002; Forest and Lorentz 2004; Bigoni 2012; Benallal and Marigo 2007]. On the other hand, thermomechanical coupling in the instability analysis was discussed for instance in [Abeyaratne and Knowles 1999; Dunwoody and Ogden 2002; Rooney and Bechtel 2004] (thermoelasticity), [Duszek et al. 1992; Steinmann et al. 1999] (thermoplasticity under the assumption of adiabatic conditions) and [Benallal and Bigoni 2004] (analysis including heat conduction in thermoinelastic materials).

If a constitutive model includes material softening, regularization is required to prevent the governing equations from the loss of ellipticity and the simulation results from pathological mesh-dependence. In a thermomechanical problem a gradient-enhancement can be applied in two ways: as an introduction

---

Contract/grant sponsor: National Science Centre of Poland 2014/15/B/ST8/01322.
*Keywords:* thermoplasticity, softening, gradient averaging, strain localization, automatic linearization, AceGen package.

of higher order gradients of a mechanical field in the material description, e.g., gradient plasticity [Belytschko and Lasry 1989; Zbib and Aifantis 1988a; 1988b; 1988c; de Borst and Mühlhaus 1992] and gradient damage [Pijaudier-Cabot and Bažant 1987; Steinmann 1999], or as an incorporation of higher order gradients of a temperature field (spatial or time derivatives) in, for example, the energy balance or heat conduction equation [Müller and Ruggeri 1993; Aifantis 1992; Forest and Aifantis 2010]. It is worth mentioning at this point that already in book [Eringen 1967] incorporation of higher order gradients of displacements and temperature in the constitutive description is admitted. In the cited book, a material is called simple if mechanical and thermal grades are equal to one: G(1,1). The regularizing influence of heat conduction is discussed in several papers, e.g., in [LeMonds and Needleman 1986] or [Batra and Kim 1991] and publications cited therein.

The mechanical part of the presented model is based on the constitutive description introduced in [Geers 2004], which involves hyperelasto-plasticity with degradation of yield strength via a damage-like variable. However, combinations of plasticity with continuum damage can also be considered, e.g., [Areias et al. 2003] or [Wcisło et al. 2013]. The analysis of fully coupled thermoplasticity models in the thermodynamic framework, which is the basis for the development of the isothermal description, can be found in the papers by Simo, Miehe, Ortiz or Ristinmaa and their coworkers, e.g., [Simo and Miehe 1992; Simo and Hughes 1998; Yang et al. 2006; Ristinmaa et al. 2007].

The research presented in this paper is limited to phenomenological continuum modeling with the following assumptions: initial isotropy of the material, hyperelasticity, rate independent plasticity with associated flow rule, strain hardening with one internal variable and degradation of the yield strength due to the increase of temperature and of a plastic strain measure. However, the description can easily be extended to more elaborated models of plasticity. Furthermore, the thermal part of the formulation makes use of Fourier's law applied to the nonstationary heat flow. All models are developed in a three-dimensional space.

One of the aspects discussed in the paper is the application of symbolic-numerical *Ace* packages [Korelc 2011] as convenient tools for the implementation of complex models and their numerical verification. In this paper, all tests are performed using self-programmed user subroutines. In particular, an elongated plate in a plane strain regime is simulated using different finite elements and material parameters influencing the width of the localization band. The numerical analysis is focused on the following issues:

- the influence of finite element enhancement *F-bar* designed to prevent volumetric locking for volume preserving plasticity, and
- examination of the model response in the presence of gradient regularization, heat conduction and a combination of the two.

The research reveals that in the analyzed case of a tensioned plate, the application of the *F-bar* finite element enhancement influences the results only in the post peak regime, but the application of elements without enrichment can even preclude the shear band formation. Moreover, although both types of regularization (gradient-enhancement and heat conduction) influence the mesh-dependence of the results and the ductility of the material response, they produce different forms of deformation.

The paper is laid out as follows. The computational strategy adopted for the analyzed model is discussed in Section 2. In particular, the application of the symbolic-numerical tools is justified and a general procedure used for the finite element user subroutine is presented. Section 3 deals with the constitutive relations for the gradient-enhanced thermomechanical model of hyperelasto-plasticity. Moreover, three governing equations (the balance of linear momentum, the balance of energy and the averaging equation) are presented in their strong form alongside their corresponding weak formulations. Section 4 contains a concise discussion outlining the finite element implementation of the presented coupled model within *Ace* packages. The results of the numerical simulations are presented in Section 5. This section is divided into two parts: first, the isothermal model is tested for different finite elements and internal length scales, next the results for the thermomechanical model are included. The attention is focused mainly on the influence of the internal length scale introduced by gradient averaging and heat conduction as well as on the behavior of the sample when two localization limiters are applied simultaneously. The paper ends with concluding remarks.

## 2. Computational strategy

The constitutive model presented in the paper includes two sources of nonlinearities: geometrical (due to large deformations) and material (plasticity). Next, the gradient enhancement is adopted, which requires solving an additional averaging equation. Eventually, the full thermomechanical coupling is taken into account. All those issues produce a highly complex problem which can only be solved using user-programmed routines for a chosen solver.

To find the solution to the nonlinear problem within the Finite Element Method (FEM), the standard Newton–Raphson procedure is used, which guarantees the quadratic convergence of the computations. However, this method requires derivation of the consistent tangent matrix on the basis of linearized governing equations. For such an elaborated model, the analytical derivation of linearization is very difficult or even infeasible. Alternatively, numerical differentiation can be applied, e.g., in [Pérez-Foguet et al. 2000] that method is used for derivation of local (at the Gauss point level) and global tangent operators.

The next step after the formulation of the numerical algorithm is the implementation of the user supplied procedure within a programming language which is required by a selected FEM engine. Usually, standard FEM software relies on Fortran or C languages, which are not very convenient tools for matrix operations. This standard approach also has a significant inconvenience as any modification of the model or its development requires renewed linearization and, hence, reformulation of the numerical procedure.

To focus more on the developed model and on the behavior of the simulated material, instead of the analytical derivations and programming, symbolic-numerical tools have been chosen for the numerical analysis. The packages which are used in this paper are *AceGen* and *AceFEM* [Korelc 2011], which work in the *Wolfram Mathematica* environment. The *AceGen* package is a multilanguage numerical code generator which combines symbolic and algebraic capabilities of Mathematica, the automatic differentiation (AD) technique (particularly important if there is a need for linearization) and simultaneous optimization of expressions, which improves efficiency of the generated code. The work with the program includes preparation of the code in a specific symbolic metalanguage based on the notation of *Wolfram Mathematica* and generation of the code in a selected programming language (e.g., C, Fortran or Matlab). In particular, the package supports several FEM environments, such as *Abaqus* or *FEAP*, and also the *AceFEM* package which has been selected by the authors to perform the numerical tests. Although another FEM program could also be used, application of this engine guarantees full control of simulations due to close collaboration between packages. The important advantage of package *AceGen* is the possibility of relatively easy modifications of the model, as the tangent matrix is derived automatically by the code generator.

The general approach to the solution of a problem using *AceGen* and *AceFEM* package is as follows. First, the unknown field interpolation for a selected finite element has to be introduced,

$$\boldsymbol{d} = \boldsymbol{N}_I \cdot \boldsymbol{p}, \tag{2-1}$$

where $\boldsymbol{p}$ is a vector of nodal unknowns and $\boldsymbol{N}_I$ is a vector of interpolation (shape) functions.

According to the author of the packages, the most convenient and effective approach to *Ace* modeling is to define a potential $\Pi$ for the problem to be solved [Korelc 2008]. Upon the introduction of (2-1) into the functional definition, $\Pi$ becomes a function of the degrees of freedom $\boldsymbol{p}$. Then, for properly formulated dependencies and with the use of Gauss integration, the finite element residual vector for the Newton–Raphson iterative procedure can be automatically computed as

$$\boldsymbol{R}_{el} = \sum_{nG} w_G J_{XG} \boldsymbol{R}_G, \qquad \boldsymbol{R}_G = \frac{\partial \Pi}{\partial \boldsymbol{p}}, \tag{2-2}$$

where $nG$ is the number of Gauss points, $w_G$ is the weight of a Gauss point, $J_{XG}$ is the Jacobian of isoparametric mapping from the parent element to the element in the reference configuration. Accordingly, the tangent operator is derived through the formula

$$\boldsymbol{K}_{el} = \frac{\partial \boldsymbol{R}_{el}}{\partial \boldsymbol{p}}. \tag{2-3}$$

## 3. Presentation of model

The constitutive description of the isothermal model is based mainly on the paper by Geers [2004]. His model covers large strain hyperelasto-plasticity and can reproduce softening and failure of the material due to an isotropic plastic-damage variable. In this paper, the model is extended to encompass full thermomechanical coupling, including thermal expansion, dependence of the heat flux on deformation, production of heat due to the plastic process and thermal softening, which is understood as degradation of the yield strength with increasing temperature.

**3A.** ***Preliminaries and kinematics.*** To formulate the governing equations, let us introduce the following notation. The continuous body which deforms under an applied load is denoted by $\mathcal{B}$ and has a boundary labeled with $\partial\mathcal{B}$. The referential placement of the body particles at time $t_0$ is identified with vector $\boldsymbol{X}$, and the current position of particle $\boldsymbol{X}$ at time $t$ is denoted by vector $\boldsymbol{x}$.

The movement of the body is described by function $\boldsymbol{\varphi}$ such that $\boldsymbol{x} = \boldsymbol{\varphi}(\boldsymbol{X}, t)$. The deformation gradient and its determinant are defined as usual,

$$\boldsymbol{F} = \frac{\partial \boldsymbol{\varphi}(\boldsymbol{X}, t)}{\partial \boldsymbol{X}}, \qquad J = \det(\boldsymbol{F}). \tag{3-1}$$

Next, multiplicative decomposition of the deformation gradient is adopted basing on the concepts applied to elastoplasticity [Lee and Liu 1967; Lee 1969; Mandel 1974; Kroner and Teodosiu 1974] and thermomechanics [Stojanovitch et al. 1964; Lu and Pister 1975; Holzapfel 2000]. Here, the deformation gradient consists of three parts in order to separate the plastic and elastic deformation as well as thermal expansion,

$$\boldsymbol{F} = \boldsymbol{F}^\theta \boldsymbol{F}^e \boldsymbol{F}^p. \tag{3-2}$$

Note that the deformation gradient (3-2) can also be formulated in the forms

$$\begin{aligned} \boldsymbol{F} &= \boldsymbol{F}^r \boldsymbol{F}^p = \boldsymbol{F}^\theta \boldsymbol{F}^m, \\ \boldsymbol{F}^r &= \boldsymbol{F}^\theta \boldsymbol{F}^e, \\ \boldsymbol{F}^m &= \boldsymbol{F}^e \boldsymbol{F}^p, \end{aligned} \tag{3-3}$$

where $\boldsymbol{F}^r$ denotes reversible deformation and $\boldsymbol{F}^m$ the mechanical one. The decomposition (3-2) has been formulated and discussed in [Wcisło and Pamin 2016].

The thermal contribution $\boldsymbol{F}^\theta$ is assumed to be purely volumetric, so for the isotropic material it can be represented as

$$\boldsymbol{F}^\theta = (J^\theta)^{1/3} \boldsymbol{I}, \quad J^\theta = \det(\boldsymbol{F}^\theta), \tag{3-4}$$

where $\boldsymbol{I}$ is the second order identity tensor, and the deformation caused by a temperature change is determined according to Lu and Pister [1975] as

$$J^\theta = e^{3\alpha_T (T - T_0)}. \tag{3-5}$$

In (3-5), quantities $T$ and $T_0$ are the absolute temperature and the reference temperature for the stress- and strain-free state, and $\alpha_T$ is the coefficient of linear thermal expansion.

On the basis of the decomposition (3-3) and assumption (3-5), the mechanical part of the deformation gradient can be derived directly as

$$\boldsymbol{F}^m = e^{-\alpha_T (T - T_0)} \boldsymbol{F}. \tag{3-6}$$

The consequence of the deformation gradient splitting into the mechanical and thermal parts is the possible application of the algorithm for isothermal elastoplasticity presented in [Wcisło et al. 2013]. Note that in this approach the assumption of volume preserving plasticity is disregarded, contrary to the thermoplastic models presented in [Wriggers et al. 1992; Simo and Miehe 1992; Simo and Hughes 1998].

**3B.** ***Free energy function.*** The state of a material is expressed by the Helmholtz free energy potential, which is assumed in a decoupled form,

$$\psi(\boldsymbol{b}^e, T, \gamma) = \psi^e(\boldsymbol{b}^e) + \psi^\theta(T) + \psi^p(\gamma), \tag{3-7}$$

where the first part $\psi^e(\boldsymbol{b}^e)$ is related to the elastic response and includes thermomechanical coupling responsible for thermal expansion, $\psi^\theta(T)$ is a purely thermal part and, finally, the term $\psi^p(\gamma)$ denotes the potential of isotropic strain hardening in plasticity. In (3-7) $\boldsymbol{b}^e = \boldsymbol{F}^e(\boldsymbol{F}^e)^{\mathrm{T}}$ denotes the elastic left Cauchy–Green tensor and $\gamma$ is a scalar plastic strain measure. Note that tensor $\boldsymbol{b}^e$, computed from $\boldsymbol{F}^e$, depends on thermal expansion as presented in the third equation of (3-3) and relation (3-6).

In particular, the parts of the free energy potential are assumed in the subsequent analysis as follows [Simo and Hughes 1998; Miehe 1995]:

$$\psi^e(\boldsymbol{b}^e) = \tfrac{1}{2}G(\operatorname{tr}(J_{be}^{-1/3}\boldsymbol{b}^e) - 3) + \tfrac{1}{2}K\big(\tfrac{1}{2}(J_{be} - 1) + \tfrac{1}{2}\log J_{be}\big), \tag{3-8}$$

$$\psi^\theta(T) = c((T - T_0) - T\ln(T/T_0)), \tag{3-9}$$

$$\psi^p(\gamma) = \tfrac{1}{2}\sqrt{2/3}h_i\gamma^2 + (\sigma_{y\infty} - \sigma_{y0})\left(\gamma + \frac{1}{\sqrt{2/3}\delta}e^{-\sqrt{2/3}\delta\gamma}\right), \tag{3-10}$$

where $G$ and $K$ are the shear and bulk moduli, $J_{be}$ is the determinant of tensor $\boldsymbol{b}^e$, $c$ is the heat capacity at constant deformation, $\sigma_{y0}$ is the initial yield stress, $\sigma_{y\infty}$ is the residual yield stress, $h_i$ is the linear hardening coefficient whereas $\delta$ denotes the saturation parameter which governs the rate of hardening converging to an asymptote: oblique if the linear hardening term is incorporated or horizontal for $h_i = 0$.

In compliance with the second law of thermodynamics (which provides restrictions for the constitutive relations) and the assumed form of the free energy potential (3-7), the state functions of the Kirchhoff stress tensor, entropy and hardening function, are obtained:

$$\boldsymbol{\tau} = 2\frac{\partial\psi}{\partial\boldsymbol{b}^e}\boldsymbol{b}^e, \quad \eta = -\frac{\partial\psi}{\partial T}, \quad h = \frac{\partial\psi}{\partial\gamma}. \tag{3-11}$$

The heat capacity can also be derived from the laws of thermodynamics (see [Simo and Miehe 1992] or [Miehe 1995]) as

$$c = -T\frac{\partial^2\psi}{\partial T^2}. \tag{3-12}$$

It can be noticed that for the adopted form of the free energy potential the heat capacity is constant. This is valid only if we assume moderate changes of temperature [Ristinmaa et al. 2007].

**3C.** ***Constitutive relations for plasticity.*** To complete the constitutive description of the problem, the yield condition which governs the plastic regime is defined:

$$F_p(\boldsymbol{\tau}, \gamma, T) = f(\boldsymbol{\tau}) - \sqrt{2/3}\sigma_y(\gamma, T)(1 - \omega_p) \le 0, \tag{3-13}$$

where $f(\boldsymbol{\tau})$ is a stress measure and $\sigma_y(\gamma, T)$ denotes the evolving yield stress. The analyzed model includes isotropic degradation of the plastic properties of material, so the yield stress is multiplied by factor $(1 - \omega_p)$ depending on the plastic-damage variable $\omega_p$, which varies from 0 for the intact material to 1 for completely damaged material. The evolution of the value of $\omega_p$ can be formulated in different ways depending on the analyzed material, see [Geers 2004]. In this paper the following form is applied:

$$\omega_p = 1 - \exp(-\beta\kappa), \tag{3-14}$$

where $\beta$ is a ductility parameter and $\kappa$ is a history variable which is assumed here to be equal to $\sqrt{2/3}\gamma$, see [Geers 2004]. Thus, in this model the damage is assumed to begin simultaneously with the plastic process (no damage threshold is considered).

The yield stress appearing in (3-13) is subsequently assumed to include saturation-type strain hardening and linear thermal softening,

$$\sigma_y(\gamma, T) = \big(\sigma_{y0} + \sqrt{2/3}h_i\gamma + (\sigma_{y\infty} - \sigma_{y0})(1 - e^{-\sqrt{2/3}\delta\gamma})\big)(1 - h_T(T - T_0)), \tag{3-15}$$

where $h_T$ is the thermal softening modulus. Assuming that $\kappa = \sqrt{2/3}\gamma$, the general formulation presented above can be rewritten as

$$F_p = f(\boldsymbol{\tau}) - \sqrt{2/3}\big(\sigma_{y0} + \sqrt{2/3}h_i\gamma + (\sigma_{y\infty} - \sigma_{y0})(1 - e^{-\sqrt{2/3}\delta\gamma})\big)(1 - h_T(T - T_0))e^{-\beta\gamma} \leq 0. \tag{3-16}$$

The adopted stress measure is standard Huber–Mises–Hencky (HMH) definition, which is commonly used to reproduce the behavior of metals, although other measures of stress can be applied as well. The HMH function depends only on the second invariant of the deviatoric Kirchhoff stress tensor, and thus it models volume preserving plasticity.

Following [Simo 1988], the associated flow rule is adopted through the Lie derivative of $\boldsymbol{b}^e$,

$$\tfrac{1}{2}\mathscr{L}_v\boldsymbol{b}^e = \dot{\lambda}\boldsymbol{N}\boldsymbol{b}^e, \tag{3-17}$$

where $\boldsymbol{N} = \partial F_p/\partial\boldsymbol{\tau}$ and $\dot{\lambda}$ is the plastic multiplier satisfying the standard Kuhn–Tucker conditions

$$\dot{\lambda} \geq 0, \quad F_p \leq 0, \quad \dot{\lambda}F_p = 0. \tag{3-18}$$

For simplicity we will now limit our considerations to the plastic flow theories for which $\lambda \equiv \gamma$.

**3D.** ***Gradient enhancement.*** As it was mentioned, the model represents a plastic material, which involves a softening behavior and poses difficulties with modeling since the governing equations lose their ellipticity and the boundary valued problem becomes ill-posed [Forest and Lorentz 2004]. To obtain a material model capable of reproducing damage properly, regularization should be applied (e.g., a nonlocal model or higher-order gradient theory selected here) or discontinuous modeling using cohesive elements should be introduced [Li and Chandra 2003].

From the numerical point of view, the loss of well-posedness of the boundary-value problem related to material softening causes mesh-dependence of the results, since the strains tend to localize in the smallest possible volume of the material, which in the finite element model is determined by the size of the element. For that reason, the outcomes for different discretizations do not converge when the mesh is densified and the prediction of the material behavior may be wrong.

To obtain the gradient model, the local variable governing damage, $\kappa$, is replaced with its nonlocal counterpart denoted by $z$, computed using the additional differential equation

$$z - l^2\nabla^2 z = \kappa, \tag{3-19}$$

with homogeneous Neumann boundary conditions [Peerlings et al. 2001].

The implicit gradient averaging formulation (3-19) was introduced first in [Peerlings et al. 1996] for a model related to quasibrittle materials. The parameter $l$ appearing in (3-19) is an internal length scale which depends on the adopted material and is often related to its microstructure [Geers et al. 1999].

When the gradient enhancement is applied in the model including large deformations, a configuration to perform the averaging has to be selected [Steinmann 1999]. Here, the material averaging has been adopted on the basis of the analysis included in [Wcisło et al. 2013]. Thus, the gradient in (3-19) is calculated with respect to Lagrangian coordinates and the internal length $l$ is related to the undeformed configuration.

**3E.** ***Heat conduction.*** The constitutive assumption for heat conduction is the classic Fourier law for isotropic materials which is formulated here for the Kirchhoff heat flux vector $\boldsymbol{q}$,

$$\boldsymbol{q} = -k\nabla T, \tag{3-20}$$

where $k$ is the heat conduction coefficient. The Kirchhoff heat flux is defined in the current configuration. However, it is referred to the area in the undeformed state (it is the Cauchy heat flux multiplied by $J$ [Haupt 2002]). Thus, the temperature gradient $\nabla T$ is calculated with respect to Euler coordinates and coefficient $k$ is specified in the reference configuration.

**3F.** ***Governing equations and weak forms.*** The formulation of the problem consists of three governing equations: the linear momentum balance, the internal energy balance and the averaging equation.

**3F.1.** *Linear momentum balance.* The linear momentum equation is formulated in the spatial description, yet it is related to the volume of the undeformed body. Following [Simo and Miehe 1992], the local form of the first governing equation is written as

$$\rho_0 \frac{\partial^2 \boldsymbol{\varphi}}{\partial t^2} = J \operatorname{div}(\boldsymbol{\tau}/J) + \rho_0 \boldsymbol{B}. \tag{3-21}$$

In (3-21), div(·) is the divergence computed with respect to Eulerian coordinates, $\rho_0$ is a reference density and $\boldsymbol{B}$ is a given gravity field. For the static analysis, which is presented in the paper, the left-hand side of (3-21) is equal to zero.

The balance of linear momentum is completed with boundary conditions for displacements $\boldsymbol{u}$ and tractions $\boldsymbol{t}$,

$$\begin{aligned} \boldsymbol{u} &= \hat{\boldsymbol{u}} \quad \text{on} \quad \partial\mathcal{B}_u, \\ \boldsymbol{t} = \boldsymbol{\tau}\cdot\boldsymbol{n} &= \hat{\boldsymbol{\tau}} \quad \text{on} \quad \boldsymbol{\varphi}(\partial\mathcal{B}_\tau), \end{aligned} \tag{3-22}$$

where

$$\partial\mathcal{B}_u \cup \partial\mathcal{B}_\tau = \partial\mathcal{B} \quad \text{and} \quad \partial\mathcal{B}_u \cap \partial\mathcal{B}_\tau = \varnothing. \tag{3-23}$$

The finite element implementation is based on weak forms of the governing equations. Applying the standard derivation (multiplication by test function $\delta\boldsymbol{u}$, integration over body $\mathcal{B}$, applying the divergence theorem and Neumann boundary conditions) the balance of linear momentum has the weak form

$$\int_{\mathcal{B}} (\nabla\delta\boldsymbol{u} : \boldsymbol{\tau} + \delta\boldsymbol{u}\cdot\boldsymbol{B})\,\mathrm{d}V + \int_{\boldsymbol{\varphi}(\partial\mathcal{B}_\tau)} \delta\boldsymbol{u}\cdot\hat{\boldsymbol{\tau}}\,\mathrm{d}a = 0. \tag{3-24}$$

In this paper $\boldsymbol{B}$, is assumed to be zero.

**3F.2.** *Energy balance.* The second governing equation is the energy balance, which is expressed in the temperature form [Simo and Miehe 1992]

$$c\frac{\partial T}{\partial t} = J \operatorname{div}(-\boldsymbol{q}/J) + \mathcal{R}, \tag{3-25}$$

where $c$ is the heat capacity at constant deformation and $\mathcal{R}$ is a heat source density which represents heating resulting from plastic dissipation [Wriggers et al. 1992] and is given as

$$\mathcal{R} = \sqrt{2/3}\,\chi\sigma_y\dot{\gamma}. \tag{3-26}$$

Parameter $\chi$ in (3-26) denotes a dissipation heat factor, see [Taylor and Quinney 1934]. The constant value of $\chi$ is adopted in this paper for simplicity although consistently with the thermodynamic laws for associated plasticity, its value depends on the initial yield strength and the advancement of the plastic process [Ristinmaa et al. 2007].

Equation (3-25) is completed with appropriate boundary conditions

$$\begin{aligned} T &= \hat{T} \quad \text{on} \quad \partial\mathcal{B}_T, \\ \boldsymbol{q}\cdot\boldsymbol{n} &= \hat{q} \quad \text{on} \quad \boldsymbol{\varphi}(\partial\mathcal{B}_q), \end{aligned} \tag{3-27}$$

where

$$\partial\mathcal{B}_T \cup \partial\mathcal{B}_q = \partial\mathcal{B} \quad \text{and} \quad \partial\mathcal{B}_T \cap \partial\mathcal{B}_q = \varnothing. \tag{3-28}$$

Note that the essential boundary conditions for $\boldsymbol{u}$ and $T$ are specified in the reference configuration whereas the natural ones for $\boldsymbol{t}$ and $\boldsymbol{q}$ are defined in the current configuration, see [Miehe 1995].

The weak form of the energy balance (3-25) is also obtained using the standard procedure and the backward Euler scheme for time integration. The integral equation valid for the current time is obtained as

$$\int_{\mathcal{B}} \left( \delta T \frac{c}{\Delta t}(T - T_n) + \nabla\delta T k \nabla T - \delta T \mathcal{R} \right) \mathrm{d}V + \int_{\boldsymbol{\varphi}(\partial\mathcal{B}_q)} \delta T \hat{q}\, \mathrm{d}a = 0, \tag{3-29}$$

where $T_n$ is the value of temperature at the previous time and $\Delta t$ is the time increment.

**3F.3.** *Averaging equation.* The strong form of the averaging equation has already been presented in Section 3D whereas the weak form of the last governing equation needed for FEM implementation after applying the standard procedure is

$$\int_{\mathcal{B}} [(z - \kappa)\delta z + l^2 \nabla_0 z \cdot \nabla_0 \delta z]\, \mathrm{d}V = 0. \tag{3-30}$$

Obviously, equations (3-24), (3-29) and (3-30) must be valid for any admissible weighting functions $\delta\mathbf{u}$, $\delta T$ and $\delta z$.

## 4. Finite element implementation

This section presents the aspects of implementing the formulated model with symbolic-numerical tools. However, only necessary information is included here as the more in-depth discussion of numerical treatment of other models within Mathematica-based packages *Ace* has already been presented in papers [Wcisło et al. 2013] (gradient damage coupled to plasticity in isothermal conditions), [Wcisło and Pamin 2014; 2016], (entropic thermoelasticity and thermoplasticity with temperature averaging, respectively).

To begin with, the vector of the nodal unknowns of the finite element for the problem includes displacements, temperature and nonlocal variables,

$$\boldsymbol{p} = [\boldsymbol{u}_I, \boldsymbol{T}_I, \boldsymbol{z}_I], \tag{4-1}$$

where subscript $I$ is related to nodal quantities. In this paper the interpolation of all three fields is performed using linear shape functions $\boldsymbol{N}_I$. The order of the interpolation in multifield problems is discussed in many papers. For example, different types of finite elements are tested in [Wcisło et al.

2013] for a case of isothermal hyperelasto-plasticity coupled with gradient damage. The research reveals that in the considered examples the elements with linear interpolation and *F-bar* enhancement preventing volumetric locking (which is discussed below) give the same results as elements with the quadratic interpolation of displacements and linear interpolation of the nonlocal variable. On the other hand, in [Simone et al. 2003] which concerns implicit gradient-enhanced continuum damage models, it is substantiated that linear interpolation of both fields (displacement vector and nonlocal variable) can be used. The provided reasoning can also be applied to the case of thermomechanical coupling, and therefore linear interpolation of temperature and displacements is admissible.

For the sake of consistency with the approach convenient for the *Ace* packages, a potential for each governing equation is introduced [Korelc 2008]. For the balance of linear momentum this potential is the Helmholtz free energy function $\psi$ (for detailed explanation see [Korelc 2009]), the specific form of which is presented in Section 3B.

The potential for the second governing equation is proposed in the form

$$\Pi_{en} = \tfrac{1}{2}\frac{c}{\Delta t}(T - T_n)^2 + \tfrac{1}{2}k\nabla T \cdot \nabla T - \mathcal{R}T. \tag{4-2}$$

It is derived from of the weak form of the energy balance (3-29) in such a way that the variation of $\Pi_{en}$ is zero. The potential presented in (4-2) is valid provided that quantities $k$, $c$ and $\mathcal{R}$ are independent of temperature. However, it can be observed that $\mathcal{R}$, which includes plastic heating in accordance with (3-26), does not satisfy this requirement. To solve this problem without changing potential (4-2), an AD exception has been applied, i.e., additional information has been introduced into the AD process to obtain the proper derivative code [Korelc 2009]. Justification of this approach can be found in [Wcisło and Pamin 2016].

Finally, the potential for the averaging equation derived from the weak form (3-30) is proposed,

$$\Pi_z = \tfrac{1}{2}((z - \kappa)^2 + l^2 \nabla_0 z \cdot \nabla_0 z). \tag{4-3}$$

Therefore, the Gauss point contribution to the element residual vector $\boldsymbol{R}_G$ appearing in (2-2) consists of three parts,

$$\boldsymbol{R}_G = \left[ \frac{\partial \psi}{\partial \boldsymbol{u}_I}, \frac{\partial \Pi_{en}}{\partial \boldsymbol{T}_I}, \frac{\partial \Pi_z}{\partial z_I} \right], \tag{4-4}$$

and the tangent matrix is computed using (2-3).

As has been mentioned in Section 3C, the Huber–Mises–Hencky yield function is a volume-preserving plasticity model and the numerical results can be affected by the spurious effect of volumetric locking. There are two solutions to prevent such misrepresentation of the results: the first is to increase the interpolation order (which can cause a significant growth in the number of degrees of freedom and computational costs) and the second is to upgrade the finite elements using a chosen methodology as the enhanced assumed strain [Simo and Rifai 1990], selective integration or the B-bar method [Hughes 1980]. In this work the *F-bar* approach of de Souza Neto et al. [2008] has been applied. The method involves replacing the deformation gradient $\boldsymbol{F}$ with its modified counterpart $\bar{\boldsymbol{F}}$. The formulation is based on a multiplicative split of the deformation gradient into its volumetric and isochoric parts,

$$\boldsymbol{F} = \boldsymbol{F}_{\text{iso}} \boldsymbol{F}_{\text{vol}}, \tag{4-5}$$

where

$$\boldsymbol{F}_{\text{iso}} = J^{-1/3}\boldsymbol{F}, \qquad \boldsymbol{F}_v = J^{1/3}\boldsymbol{I}. \tag{4-6}$$

The modified deformation gradient $\bar{\boldsymbol{F}}$ for a selected Gauss point is computed using the isochoric part calculated at this integration point and the volumetric one derived at the centroid of the finite element.

## 5. Numerical tests

In this section the numerical examples simulating the response of a modeled material are presented. All computations were performed for a rectangular plate in tension in the plane strain regime.

The isothermal model is tested first with attention focused on the influence of the internal length scale on the material response and on the significance of element enhancement preventing the locking phenomenon.

The fully coupled thermomechanical model is subsequently tested for different values of the heat conduction coefficient and compared with the outcome of the model without temperature-dependence.

**5A.** ***Test parameters.*** The tested specimen is a plate of dimensions: $L = 0.20\,\text{m}$, $W = 0.10\,\text{m}$, $H = 0.0025\,\text{m}$ with a square imperfection in the center ($L_{\text{imp}} = 0.005\,\text{m}$). Due to the symmetry of the specimen, only one quarter of the plate is considered. The imperfection is assumed as the decrease of the initial yield stress to $\sigma_{y0,\text{imp}} = 0.8\sigma_{y0}$, and for simulations with the thermomechanical coupling an increase of the thermal softening coefficient to $H_{T,\text{imp}} = 1.05H_T$ is adopted additionally, see [LeMonds and Needleman 1986]. The material parameters used in the simulations are presented in Table 1. The upper part of the table is related to the isothermal model whereas the lower part refers to the thermal coupling. Two nonzero values of the internal length parameter are assumed. The first one $l = 0.0025\,\text{m}$ is equal to the plate thickness and the dimension of the imperfection in the relevant quarter of the plate. The plane strain conditions are applied (displacements at all nodes in the thickness direction are restrained) and the remaining mechanical boundary conditions are assumed in such a way as to allow deformation in the width direction. The enforced maximum displacement in the length-direction is equal to $u_{\max} = 0.04\,\text{m}$ and it is achieved after 1 s in adaptive steps.

For the coupled thermomechanical model, the whole sample is insulated, i.e., homogeneous Neumann boundary conditions are adopted on the whole surface. However, convection on the surface can also be considered, as in [Wcisło and Pamin 2016], where aspects of the convection implementation within *AceGen* are discussed.

The numerical model is generated using hexahedral elements and two finite element discretizations are used: 800 (mesh 1) and 3200 (mesh 2) elements, see Figure 1.

**5B.** ***Plate in tension in isothermal conditions.*** At the beginning the simulations are performed for the isothermal conditions. In the first test, the response of the sample is modeled with standard hexahedral elements with linear interpolation of both fields (displacements and the nonlocal variable), and no enrichment preventing the locking phenomenon is used.

Figure 2 presents the sum of reactions vs. load multiplier. It can be observed that all results for mesh 1 and mesh 2 coincide to a certain point in the post peak regime. When the load multiplier reaches the value 0.48, the diagrams become distinct. The value of the internal length scale does not influence the results much, and the outcomes of simulation using mesh 2 are very close to one another for each value of $l$.

| property | symbol | value | unit |
|---|---|---|---|
| bulk modulus | $K$ | $164.2 \cdot 10^9$ | N/m$^2$ |
| shear modulus | $G$ | $80.19 \cdot 10^9$ | N/m$^2$ |
| initial yield stress | $\sigma_{y0}$ | $0.45 \cdot 10^9$ | N/m$^2$ |
| residual yield stress | $\sigma_{y\infty}$ | $0.6 \cdot 10^9$ | N/m$^2$ |
| hardening modulus | $H$ | $129 \cdot 10^6$ | N/m$^2$ |
| saturation coefficient | $\delta$ | 16.93 | – |
| ductility | $\beta$ | 1 | – |
| internal length | $l$ | 0, 0.0025, 0.005 | m |
| density | $\rho$ | $7.8 \cdot 10^3$ | kg/m$^3$ |
| conductivity | $k$ | 0, 100 | J/(sKm) |
| heat capacity | $c$ | 460 | J/(kgK) |
| thermal expansion coefficient | $\alpha_T$ | $12 \cdot 10^{-6}$ | 1/K |
| thermal softening modulus | $H_T$ | 0.002 | 1/K |
| dissipation heat factor | $\chi$ | 0.9 | - |
| reference temperature | $T_0$ | 273.15 | K |

**Table 1.** Material parameters.

**Figure 1.** Finite element discretizations of plate in tension: mesh 1 (left) and mesh 2 (right); imperfection marked with red color.

The deformed meshes with the plastic strain measure distribution at the end of the deformation process are depicted in Figures 3 and 4. It can be noticed that the localization has the form of a neck rather than shear band and can be described as a diffuse mode. This is also consistent with the force-displacement diagram in the post peak regime and the little influence of the internal length on the results. Also the differences in deformations and plastic strain distributions for two values of internal length $l = 0$ and $l = 0.005$ m for both discretizations are minor. As expected, a smoother distribution of plastic strain is observed for the nonzero value of $l$. In fact, the results are slightly different for the adopted meshes, the introduction of the internal length scale does not prevent the results from being discretization-sensitive.

The next simulations are performed for a model with the finite element enrichment *F-bar*. Figure 5 shows the load-displacement relation for the nonlocal model with different values of the internal length scale. Similarly to the previous test, the curves overlap up to the value of the displacement multiplier equal to $\lambda = 0.48$ and subsequently diverge. However, in this case the internal length has a strong influence on the mesh-insensitivity, for the higher value the diagrams for mesh 1 and mesh 2 are very close, and they suggest a localized mode of deformation, which is confirmed in Figures 6 and 7, where shear bands are shown. It can be noticed that for the local model the strains localize in two rows of elements. The application of the gradient enhancement (see Figure 7) prevents the results from being affected by this pathological mesh-dependency and has a regularizing effect. To sum up, it is interesting that the choice of the kind of finite element has such strong influence on the results, which differ not only quantitatively but qualitatively as well.

The results obtained with the use of the finite elements with and without *F-bar* enrichment are compared in Figure 8 for the zero internal length scale. It was expected that the response of the material simulated with elements without enhancement would be stiffer, yet this has been confirmed only in the post peak regime.

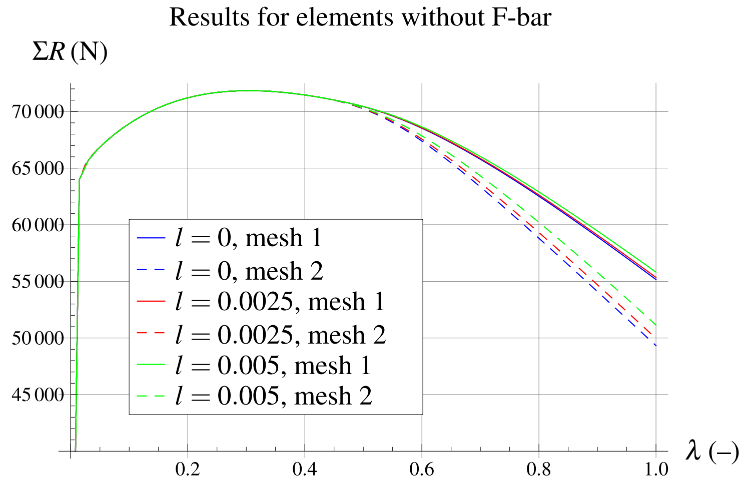

**Figure 2.** The sum of reactions vs. displacement multiplier (isothermal model).

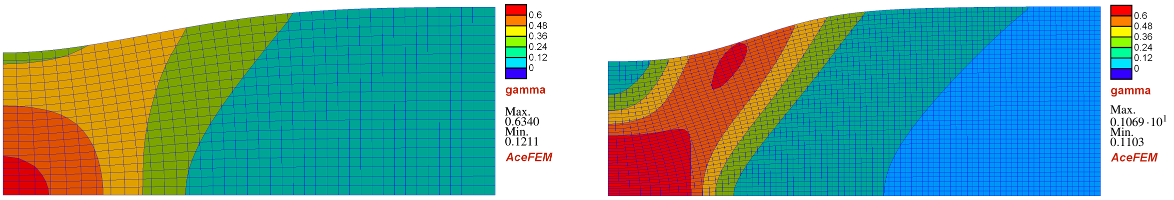

**Figure 3.** Deformed mesh with plastic strain measure $\gamma$ distribution for mesh 1 (left) and mesh 2 (right) — isothermal model, $l = 0$, elements without *F-bar* enhancement.

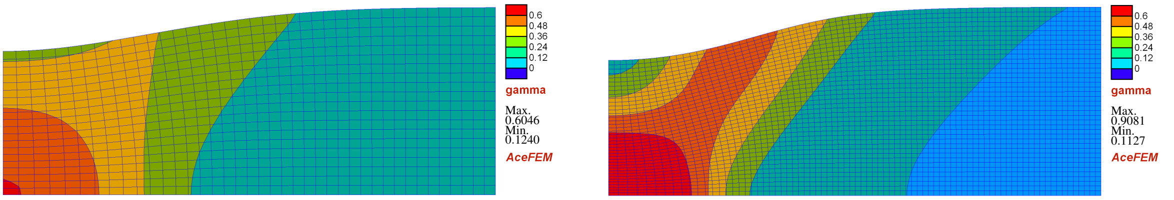

**Figure 4.** Deformed mesh with plastic strain measure $\gamma$ distribution for mesh 1 (left) and mesh 2 (right) — isothermal model, $l = 0.005$ m, elements without *F-bar* enhancement.

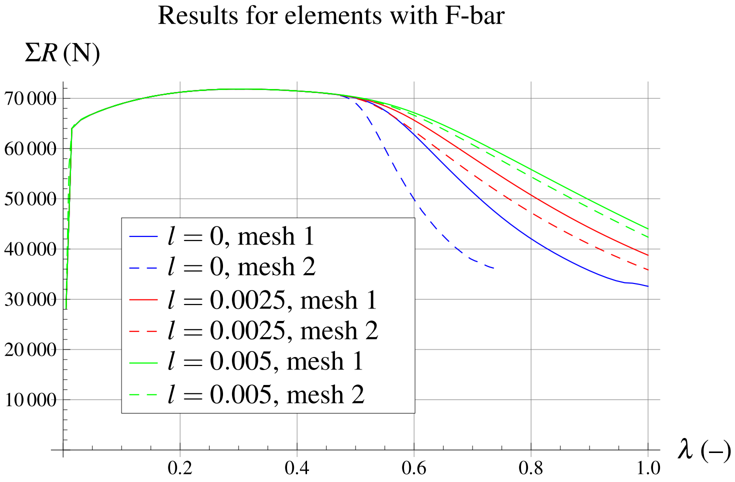

**Figure 5.** The sum of reactions vs. displacement multiplier (isothermal model).

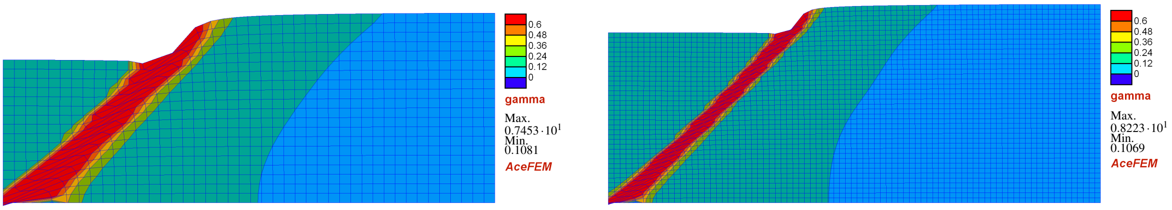

**Figure 6.** Deformed mesh with plastic strain measure $\gamma$ distribution for mesh 1 (left) and mesh 2 (right) — isothermal model, $l = 0$, elements with *F-bar* enhancement.

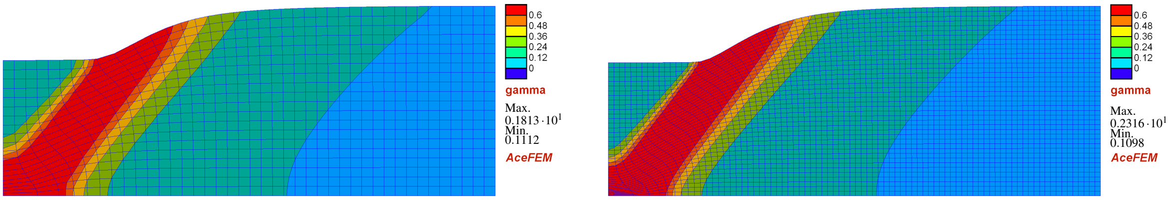

**Figure 7.** Deformed mesh with plastic strain measure $\gamma$ distribution for mesh 1 (left) and mesh 2 (right) — isothermal model, $l = 0.005$ m, elements with *F-bar* enhancement.

The application of elements with *F-bar* enhancement causes a softer response of the material and admits a jump into a localized shear band mode, while the standard finite elements seem to prevent localization and favor the diffuse deformation. To make the matters worse, the plane strain tests (e.g., 3D ones) are as a rule more sensitive to volumetric locking, but in our test the boundary conditions are imposed in such a way that globally the plate can extend while preserving the volume. Thus, it seems that the influence of the isochoric constraint manifests itself rather at the level of shear band formation.

**5C.** ***Plate in tension simulated with thermomechanical model.*** In this section the coupled model is tested using material properties presented in Table 1. Apart from the two sources of instabilities which have been discussed in the previous section (geometrical and material), now the third one is going to

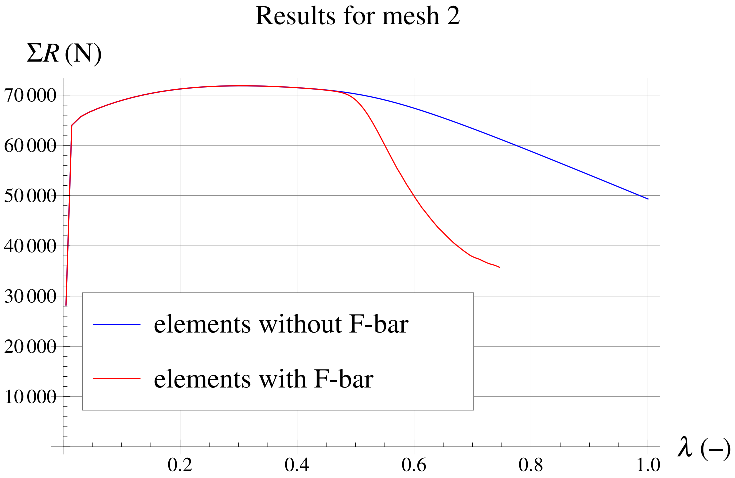

**Figure 8.** The sum of reactions vs. displacement multiplier (isothermal model).

be taken into account: reduction of the yield stress with temperature (thermal softening). Due to the complexity of the model and the large amount of tests which could be performed for different material and test parameters, the paper focuses solely on the presentation of the results of simulations carrier out for the following cases.

First, the finite elements without *F-bar* are used for simulations. The results for only one value of the internal length $l = 0.005\,\mathrm{m}$ are discussed due to the fact that in the isothermal case the diffuse mode was obtained and this deformation pattern (with very slight differences) is also observed for the thermomechanical coupling with different values of conductivity $k$ and internal length $l$ listed in Table 1.

In the next step, the elements with *F-bar* enrichment are tested. Here more cases are studied. We start by testing the influence of heat conduction for the model with zero internal length scale, next the effect of the internal parameter for the adiabatic case is examined, and finally the simulations are performed with two regularization effects (heat conduction and gradient averaging) included simultaneously.

**5C.1.** *Finite elements without F-bar.* The results obtained for the finite elements without enhancement are presented in Figures 9 and 10. The diagram presents the load-displacement paths for the nonlocal model with the internal length equal to $l = 0.005\,\mathrm{m}$ for three cases: isothermal, adiabatic $k = 0$ and heat conduction $k = 100\,\mathrm{J/(sKm)}$. The differences between these graphs are minor but the outcome for mesh 1 is stiffer than for the finer mesh. For the coarse mesh the diagrams for the adiabatic case and heat conduction coincide (red and green solid lines) so the influence of heat conduction is negligible. Also for the finer mesh that difference is insignificant.

Figure 10 presents the deformed mesh with temperature distribution. It can be observed that the presence of thermal softening does not influence significantly the shape of the deformation plot and the differences in temperature are also small.

**5C.2.** *Finite elements with F-bar.* First, the results are compared for zero internal length and different values of heat conduction to confirm that the conductivity has a regularizing effect in the absence of another length scale. The sum of reactions vs. enforced load multiplier for the analyzed cases are presented in Figure 11. It can be observed that the diagrams for heat conduction for the coarse mesh are

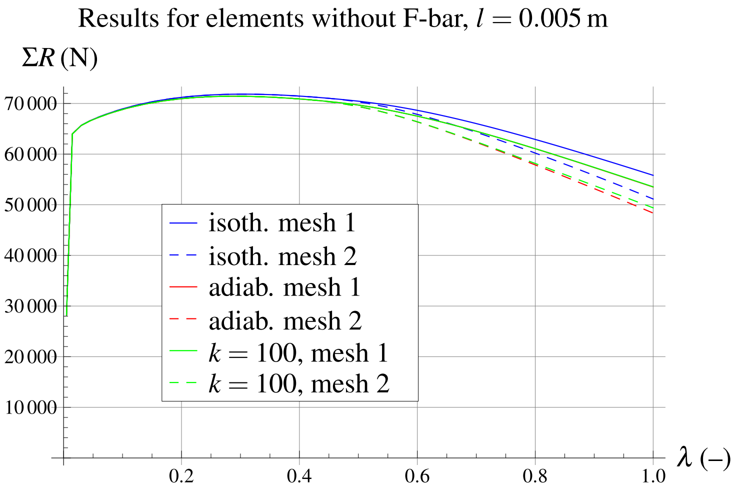

**Figure 9.** The sum of reactions vs. displacement multiplier.

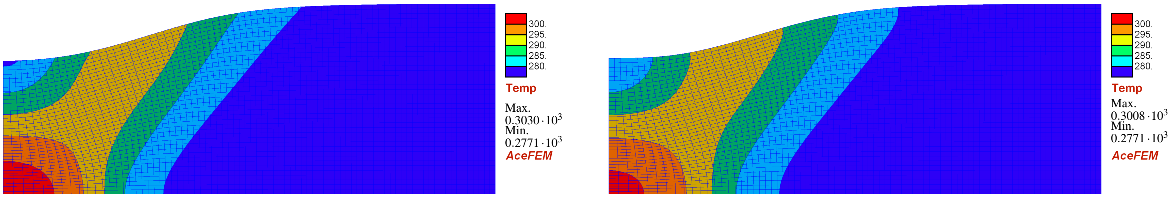

**Figure 10.** The deformed mesh with temperature distribution for $k = 0$ (left) and $k = 100$ J/(sKm) (right) — elements without *F-bar* enhancement, $l = 0.005$ m, mesh 2.

very close to the adiabatic case although some fluctuations are visible. Next, the response of the material with heat conduction was analyzed with the finer mesh. Just after the peak point the response is close to the adiabatic case, yet a convex curve is observed further on.

Figures 12 and 13 present the deformed meshes at the end of the elongation process with temperature distribution for the adiabatic case and heat conduction, respectively. As it was expected, in the first case, in which no regularization is applied the deformation is strongly localized and mesh-dependent. The generated shear band is inclined 45 degrees with respect to the longitudinal axis. However, if heat conduction is taken into account, see Figure 13, the shear band is wider but also bent. The deformations for the two considered discretizations differ slightly but have the same character — curved shear band with a narrowing in the middle of the modeled sample. When the evolution of the plastic strain measure is investigated along the line perpendicular to the initial shear band for mesh 2, it turns out that the shear band moves slightly, see Figure 14. In the first diagram we can observe a stationary shear band whereas for the case with heat conduction an evolving localization zone.

Now the influence of the internal length parameter will be examined for the adiabatic case to investigate the regularizing effect of the gradient averaging in the absence of heat conduction. Figure 15 presents the load-displacement diagrams for zero internal length scale and for $l = 0.005$ m. The results for isothermal variant are also presented for comparison. It can be noticed that the addition of thermal

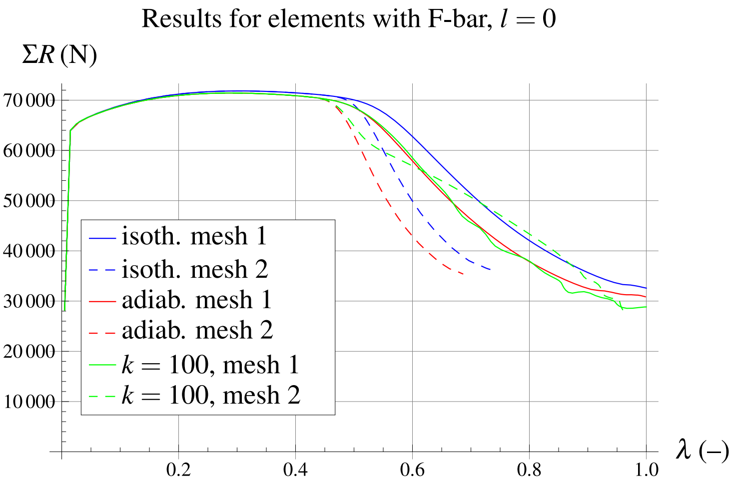

**Figure 11.** The sum of reactions vs. displacement multiplier.

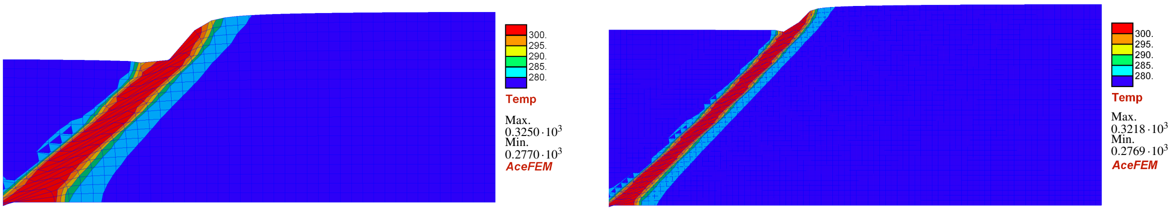

**Figure 12.** The deformed mesh with temperature distribution for mesh 1 (left) and mesh 2 (right) — elements with *F-bar* enhancement, $k = 0$, $l = 0$.

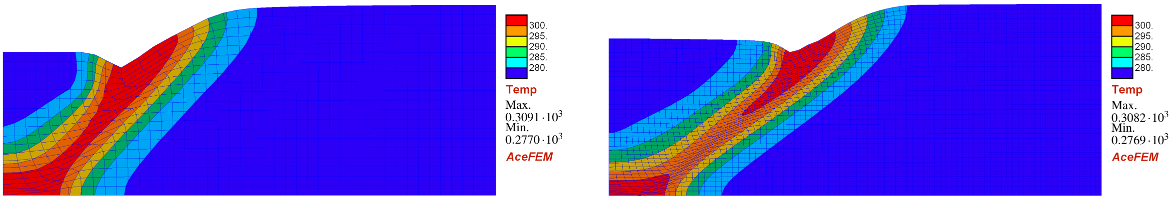

**Figure 13.** The deformed mesh with temperature distribution for mesh 1 (left) and mesh 2 (right) — elements with *F-bar* enhancement, $k = 100\,\mathrm{J/(sKm)}$, $l = 0$.

softening indeed causes a less stiff response of the material, though its influence is stronger for the model without averaging. The diagrams showing the model with regularization are very close for each mesh and exhibit higher ductility.

Figure 16 presents the deformed meshes with temperature distribution. The response of the sample is very similar to the isothermal deformation. The shear band has a constant width and the solution does not depend on the adopted mesh.

When the results for the two regularizations (heat conduction and gradient averaging) are compared, it must be stated that, although both of them influence the discretization-dependence and increase the

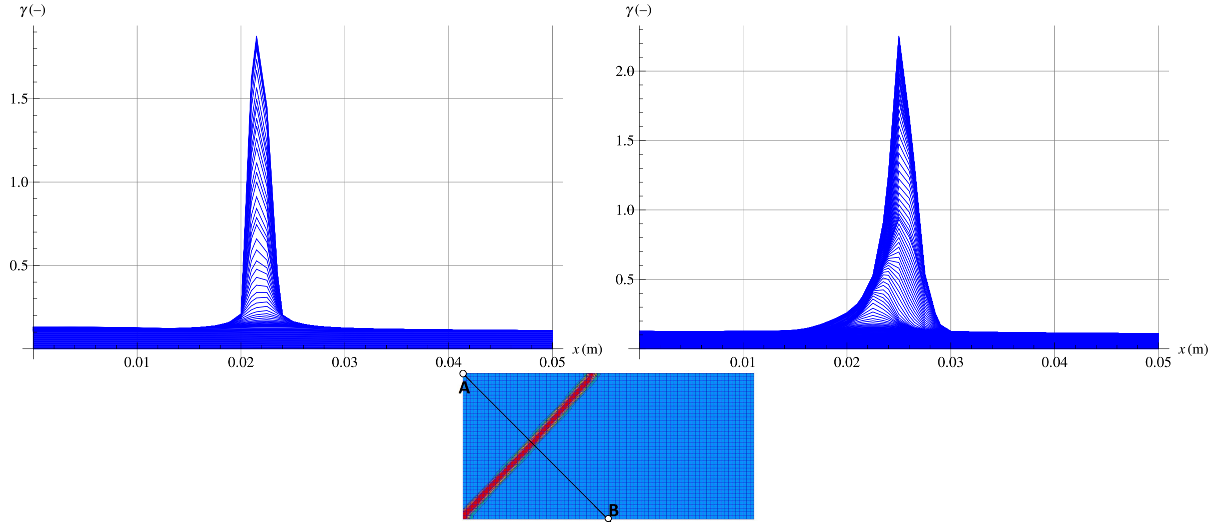

**Figure 14.** The evolution of the plastic strain measure $\gamma$ along line A-B in reference configuration for $k = 0$ (top left) and $k = 100\,\mathrm{J/(sKm)}$ (top right) — elements with *F-bar* enhancement, $l = 0$.

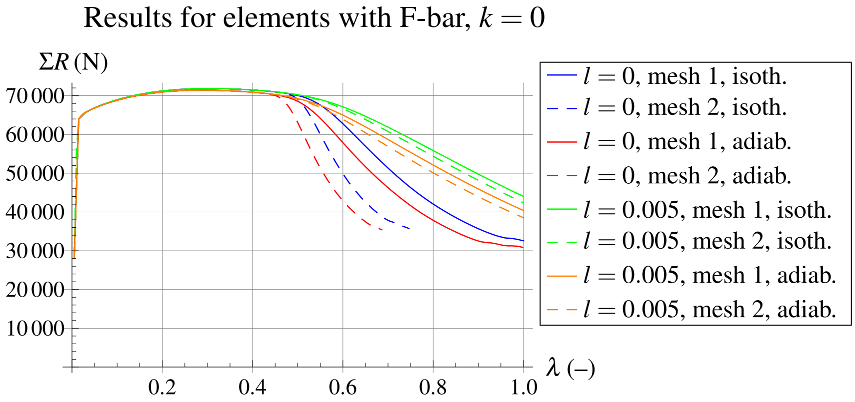

**Figure 15.** The sum of reactions vs. displacement multiplier.

ductility of the response, there are differences in the shape of the localization zone — the internal parameter $l$ visibly governs the width of the band whereas for heat conduction the band has no regular form (exhibits a variable width and a curved shape).

Finally, the results of the test performed for heat conduction and gradient averaging active simultaneously are discussed. Figure 17 presents the force-displacement diagram which compares the results for two regularizations acting together (red lines) with the previous results involving only one internal length. It can be observed that the outcome is close to the response obtained for the model with gradient averaging. In this test the results for mesh 1 also exhibit fluctuations which are apparently absent in the case of finer discretization. The material modeled with gradient averaging gives a slightly softer response for mesh 2 than the model with two nonzero scales (i.e., $l = 0.005\,\mathrm{m}$ and $k = 100\,\mathrm{J/(sKm)}$). Figure 18

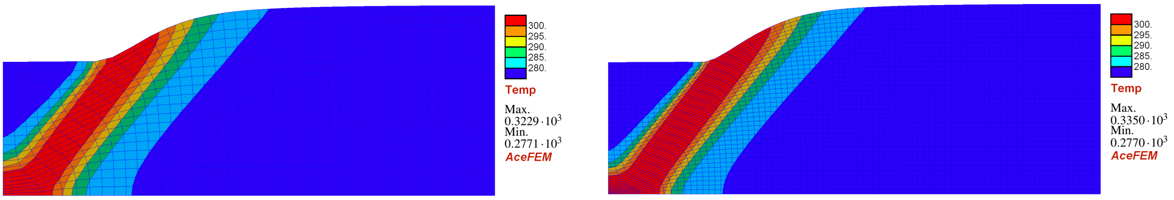

**Figure 16.** The deformed mesh with temperature distribution for mesh 1 (left) and mesh 2 (right) — elements with *F-bar* enhancement, $k = 0$, $l = 0.005\,\mathrm{m}$.

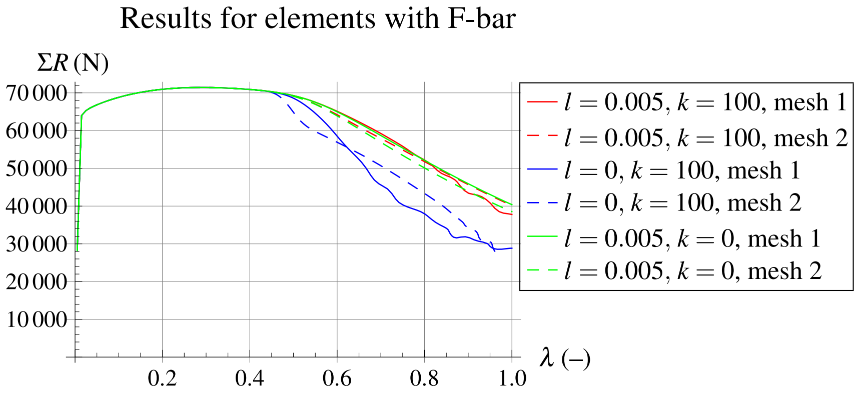

**Figure 17.** The sum of reactions vs. displacement multiplier for $k = 100\,\mathrm{J/(sKm)}$ and $l = 0.005\,\mathrm{m}$ — elements with *F-bar* enhancement.

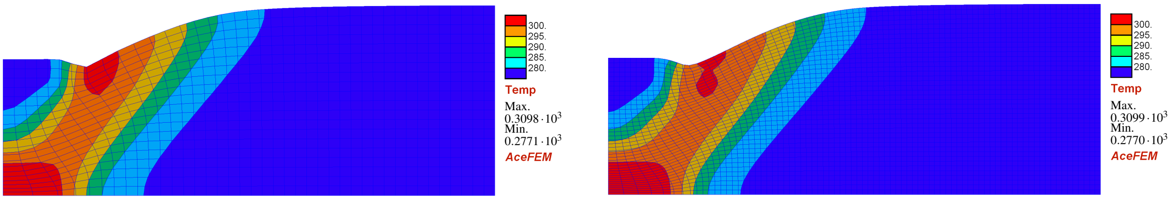

**Figure 18.** The deformed mesh with temperature distribution for mesh 1 (left) and mesh 2 (right) — elements with *F-bar* enhancement, $k = 100\,\mathrm{J/(sKm)}$, $l = 0.005\,\mathrm{m}$.

shows the deformations of the sample for two discretizations. The results are similar for both meshes and it can be observed that the shear band width is governed by the internal length scale, but it is also bent due to heat conduction. When the diagram presented in Figure 19 is analyzed, it turns out that the band also evolves.

## 6. Conclusion

This paper presents an analysis of the nonisothermal large-strain model of hyperelasto-plasticity with the yield strength degradation due to damage-like variable. The thermomechanical coupling, apart from thermal expansion and plastic heating, introduces also thermal softening as the third source of the loss of stability alongside the geometrical and material effects. To prevent the pathological mesh-dependence of simulation results, gradient averaging is incorporated in the model. This treatment introduces the internal

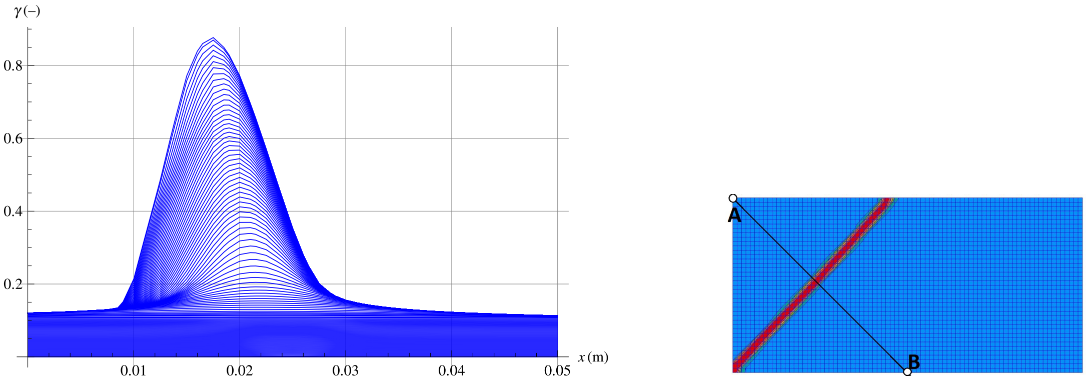

**Figure 19.** The evolution of the plastic strain measure $\gamma$ along line A-B in reference configuration for $k = 100\,\mathrm{J/(sKm)}$ and $l = 0.005\,\mathrm{m}$ — elements with *F-bar* enhancement.

length parameter which influences the ductility of the material response and the width of the shear band. Furthermore, heat conduction in the coupled thermomechanical model also has regularizing properties for the nonadiabatic case.

The developed model was implemented within FEM and tested using the elongated rectangular plate in plane strain conditions. The aims of the numerical analysis of the presented model were to examine the influence of the incorporated localization limiters on the simulation results and, moreover, to investigate how the results depend on the adopted finite elements with or without the enhancement preventing the locking phenomenon which might be encountered in a volume-preserving plasticity model.

Firstly, the isothermal analysis was carried out for different meshes, finite elements and values of the internal length scale. Although the volume-preserving plasticity was implemented, the results do not differ for the finite elements with and without *F-bar* enhancement in the plastic regime prior to strain localization. However, for the simulations with the finite elements without enrichment the response in the post peak zone has a diffuse form, practically independent of the internal length scale, whereas the *F-bar* enhancement involves strong localization of strains in a shear band. In the latter case the gradient averaging significantly influences the ductility of the material response removes the mesh-dependence.

The next step involved taking into account the nonisothermal model. The behavior of the sample simulated using the finite elements without *F-bar* enhancement is similar to the response in the isothermal conditions. What is more, the results actually do not depend on heat conduction. On the other hand, the response of the plate in tension modeled with the enriched elements exhibits the localized form of deformation, strongly discretization-dependent in the absence of regularization. The application of gradient averaging influences the shear band width whereas the heat conduction produces an irregular and evolving localization zone (with a changeable width and a curved shape). Thus, although both heat conduction and gradient averaging have regularizing properties, they affect deformation in different manners. The simultaneous presence of the heat conduction and gradient regularization results in the combination of these two effects.

Comparison of the obtained results shows that incorporation of finite element enhancement to avoid volumetric locking influences strongly the post peak behavior of the material, and its absence can prevent the localized form of deformation.

Finally, it is worth emphasizing the significance of the programming environment applied for the numerical treatment of the complex model. The package *AceGen*, involving automatic differentiation, is a convenient tool which allows researchers to focus on the model development and its examination instead of analytical derivation and programming.

To sum up, the new developments in the paper are:

- thermomechanical extension of the hyperelasto-plastic model including yield stress degradation and gradient averaging of the plastic strain measure,
- development of the algorithms for the symbolic-numerical tool *AceGen* based on the potential approach and performing automatic linearization of the governing equations,
- investigation of the influence of the finite element enhancement called *F-bar*, preventing the volumetric locking phenomenon on the results of the numerical simulations of shear band instability,
- analysis of the regularizing effects produced by heat conduction as well as gradient averaging and of the behavior of the sample in the simultaneous presence of two internal length parameters.

## References

[Abeyaratne and Knowles 1999] R. Abeyaratne and J. K. Knowles, "On the stability of thermoelastic materials", *J. Elasticity* **53**:3 (1999), 199–213.

[Aifantis 1992] E. C. Aifantis, "On the role of gradients in the localization of deformation and fracture", *Int. J. Eng. Sci.* **30**:10 (1992), 1279–1299.

[Areias et al. 2003] P. M. A. Areias, J. C. de Sá, and C. C. ao António, "A gradient model for finite strain elastoplasticity coupled with damage", *Finite Elements in Analysis and Design* **39**:13 (2003), 1191–1235.

[Batra and Kim 1991] R. C. Batra and C. H. Kim, "Effect of thermal conductivity on the initiation, growth and bandwidth of adiabatic shear bands", *Int. J. Eng. Sci.* **29**:8 (1991), 949–960.

[Belytschko and Lasry 1989] T. Belytschko and D. Lasry, "A study of localization limiters for strain-softening in statics and dynamics", *Comput. Struct.* **33**:3 (1989), 707–715.

[Benallal and Bigoni 2004] A. Benallal and D. Bigoni, "Effects of temperature and thermo-mechanical couplings on material instabilities and strain localization of inelastic materials", *J. Mech. Phys. Solids* **52**:3 (2004), 725–753.

[Benallal and Marigo 2007] A. Benallal and J.-J. Marigo, "Bifurcation and stability issues in gradient theories with softening", *Modelling Simul. Mater. Sci. Eng.* **15**:1 (2007), S283–S295.

[Bigoni 2012] D. Bigoni, *Nonlinear solid mechanics: bifurcation theory and material instability*, Cambridge University Press, 2012.

[de Borst and Mühlhaus 1992] R. de Borst and H.-B. Mühlhaus, "Gradient-dependent plasticity: formulation and algorithmic aspects", *Int. J. Num. Methods Eng.* **35**:3 (1992), 521–539.

[de Borst et al. 1993] R. de Borst, L. Sluys, H. Muhlhaus, and J. Pamin, "Fundamental issues in finite element analyses of localization of deformation", *Eng. Comput.* **10**:2 (1993), 99–121.

[de Souza Neto et al. 2008] E. de Souza Neto, D. Peric, and D. Owen, *Computational methods for plasticity: theory and applications*, Wiley, Chichester, England, 2008.

[Dunwoody and Ogden 2002] J. Dunwoody and R. Ogden, "On the thermodynamic stability of elastic heat-conducting solids subject to a deformation-temperature constraint", *Math. Mech. Solids* **7**:3 (2002), 285–306.

[Duszek et al. 1992] M. Duszek, P. Perzyna, and E. Stein, "Adiabatic shear band localization in elastic-plastic damaged solids", *Int. J. Plast.* **8**:4 (1992), 361–384.

[Eringen 1967] A. Eringen, *Mechanics of Continua*, Wiley, New York, 1967.

[Forest and Aifantis 2010] S. Forest and E. C. Aifantis, "Some links between recent gradient thermo-elasto-plasticity theories and the thermomechanics of generalized continua", *Int. J. Solids Struct.* **47**:25-26 (2010), 3367–3376.

[Forest and Lorentz 2004] S. Forest and E. Lorentz, "Localization phenomena and regularization methods", pp. 311–371 in *Local approach to fracture* (Paris), edited by J. Besson, Ecole d'été "Mécanique de l'endommagement et approche locale de la rupture" (MEALOR), Les presses de l'ecole des mines de paris, 2004.

[Geers 2004] M. Geers, "Finite strain logarithmic hyperelasto-plasticity with softening: a strongly non-local implicit gradient framework", *Computer Methods in Applied Mechanics and Engineering* **193**:30-32 (2004), 3377–3401.

[Geers et al. 1999] M. Geers, R. Borst, W. Brekelmans, and R. Peerlings, "Validation and internal length scale determination for a gradient damage model: application to short glass-fibre-reinforced polypropylene", *Int. J. Solids Struct.* **36**:17 (1999), 2557–2583.

[Haupt 2002] P. Haupt, *Continuum mechanics and theory of materials*, 2nd ed., Advanced Texts in Physics, Springer, Berlin, 2002.

[Hill 1958] R. Hill, "A general theory of uniqueness and stability in elastic-plastic solids", *J. Mech. Phys. Solids* **6** (1958), 236–249.

[Holzapfel 2000] G. A. Holzapfel, *Nonlinear solid mechanics: a continuum approach for engineering*, Wiley, Chichester, England, 2000.

[Hughes 1980] T. J. R. Hughes, "Generalization of selective integration procedures to anisotropic and nonlinear media", *Int. J. Num. Methods Eng.* **15**:9 (1980), 1413–1418.

[Korelc 2008] J. Korelc, *Nonlinear finite element methods*, Springer, Berlin, 2008.

[Korelc 2009] J. Korelc, "Automation of primal and sensitivity analysis of transient coupled problems", *Comput. Mech.* **44**:5 (2009), 631–649.

[Korelc 2011] J. Korelc, "AceGen and AceFEM user manual", Technical report, University of Ljubljana, 2011, Available at http://symech.fgg.uni-lj.si/.

[Kroner and Teodosiu 1974] E. Kroner and C. Teodosiu, "Lattice defect approach to plasticity and viscoplasticity", in *Problems of plasticity* (Leiden, Netherlands, 1972), edited by A. Sawczuk, 1974.

[Lee 1969] E. Lee, "Elastic-plastic deformation at finite strains", *J. Appl. Mech.* (*ASME*) **36**:1 (1969), 1–6.

[Lee and Liu 1967] E. Lee and D. Liu, "Finite-strain elastic-plastic theory with application to plane-wave analysis", *J. Appl. Phys.* **38**:1 (1967), 19–27.

[LeMonds and Needleman 1986] J. LeMonds and A. Needleman, "Finite element analyses of shear localization in rate and temperature dependent solids", *Mech. Mater.* **5**:4 (1986), 339–361.

[Li and Chandra 2003] H. Li and N. Chandra, "Analysis of crack growth and crack-tip plasticity in ductile materials using cohesive zone models", *Int. J. Plasticity* **19**:6 (2003), 849–882.

[Lu and Pister 1975] S. Lu and K. Pister, "Decomposition of deformation and representation of the free energy function for isotropic thermoelastic solids", *Int. J. Solids and Struct.* **11**:7 (1975), 927–934.

[Mandel 1974] J. Mandel, "Thermodynamics and plasticity", in *Foundations of continuum thermodynamics*, edited by J. Delgado, Macmillan, New York, 1974.

[Menzel 2002] A. Menzel, *Modelling and computation of geometrically nonlinear anisotropic inelasticity*, Ph.D. Thesis, University of Kaiserslautern, 2002, Available at https://kluedo.ub.uni-kl.de/frontdoor/index/index/year/2002/docId/1334.

[Miehe 1995] C. Miehe, "Entropic thermoelasticity at finite strains: aspects of the formulation and numerical implementation", *Comput. Methods Appl. Mech. Eng.* **120**:3-4 (1995), 243–269.

[Müller and Ruggeri 1993] I. Müller and T. Ruggeri, *Extended thermodynamics*, Springer Tracts in Natural Philosophy **37**, Springer, New York, 1993.

[Peerlings et al. 1996] R. Peerlings, R. de Borst, W. Brekelmans, and J. de Vree, "Gradient enhanced damage for quasi-brittle materials", *Int. J. Num. Methods Eng.* **39**:19 (1996), 3391–3403.

[Peerlings et al. 2001] R. Peerlings, M. Geers, R. de Borst, and W. Brekelmans, "A critical comparison of nonlocal and gradient-enhanced softening continua", *Int. J. Solids Struct.* **38**:44-45 (2001), 7723–7746.

[Pérez-Foguet et al. 2000] A. Pérez-Foguet, A. Rodríguez-Ferran, and A. Huerta, "Numerical differentiation for local and global tangent operators in computational plasticity", *Comput. Methods in Appl. Mech. Eng.* **189**:1 (2000), 277–296.

[Pijaudier-Cabot and Bažant 1987] G. Pijaudier-Cabot and Z. P. Bažant, "Nonlocal damage theory", *J. Eng. Mech.* **113**:10 (1987), 1512–1533.

[Rice 1976] J. R. Rice, "The localization of plastic deformation", pp. 207–220 in *Proceedings of the 14th international congress on theoretical and applied mechanics*, edited by W. Koiter, North-Holland Publishing Company, 1976.

[Ristinmaa et al. 2007] M. Ristinmaa, M. Wallin, and N. Ottosen, "Thermodynamic format and heat generation of isotropic hardening plasticity", *Acta Mech.* **194**:1-4 (2007), 103–121.

[Rooney and Bechtel 2004] F. J. Rooney and S. E. Bechtel, "Constraints, constitutive limits, and instability in finite thermoelasticity", *J. Elasticity* **74**:2 (2004), 109–133.

[Rudnicki and Rice 1975] J. Rudnicki and J. Rice, "Conditions for the localization of deformation in pressure-sensitive dilatant materials", *J. Mech. Phys. Solids* **23**:6 (1975), 371–394.

[Simo 1988] J. C. Simo, "A framework for finite strain elastoplasticity based on maximum plastic dissipation and the multiplicative decomposition, I: Continuum formulation", *Comput. Methods Appl. Mech. Eng.* **66**:2 (1988), 199–219.

[Simo and Hughes 1998] J. C. Simo and T. J. R. Hughes, *Computational inelasticity*, Interdisciplinary Applied Mathematics **7**, Springer, New York, 1998.

[Simo and Miehe 1992] J. Simo and C. Miehe, "Associative coupled thermoplasticity at finite strains: formulation, numerical analysis and implementation", *Comput. Methods Appl. Mech. Eng.* **98**:1 (1992), 41–104.

[Simo and Rifai 1990] J. C. Simo and M. S. Rifai, "A class of mixed assumed strain methods and the method of incompatible modes", *Int.. J. Num. Methods Eng.* **29**:8 (1990), 1595–1638.

[Simone et al. 2003] A. Simone, H. Askes, R. Peerlings, and L. Sluys, "Interpolation requirements for implicit gradient-enhanced continuum damage models", *Commun. Numer. Methods Eng.* **19**:7 (2003), 563–572. Errata in **20** (2004), 163–165.

[Sluys 1992] L. Sluys, *Wave propagation, localization and dispersion in softening solids*, Ph.D. Thesis, Delft University of Technology, 1992, Available at http://resolver.tudelft.nl/uuid:0f9b3de9-e0ec-4d9b-a42f-fcf704d5d40e.

[Steinmann 1999] P. Steinmann, "Formulation and computation of geometrically non-linear gradient damage", *Int. J. Num. Methods Eng.* **46**:5 (1999), 757–779.

[Steinmann et al. 1999] P. Steinmann, K. Runesson, and R. Larsson, "On the analysis of adiabatic strong discontinuities within thermoplastic solids", in *IUTAM Symposium on micro- and macrostructural aspects of thermoplasticity* (Dordrecht, Netherlands), edited by O. Bruhns and E. Stein, Kluwer Academic Publisher, 1999.

[Stojanovitch et al. 1964] R. Stojanovitch, S. Djuritch, and L. Vujoshevitch, "On finite thermal deformations", *Arch. Mech.* **16** (1964), 103–108.

[Taylor and Quinney 1934] G. I. Taylor and H. Quinney, "The latent energy remaining in a metal after cold working", *Proc. Royal Soc. Lond. A* **143**:849 (1934), 307–326.

[Thomas 1961] Y. Thomas, *Plastic flow and fracture of solids*, Academic Press, New York, 1961.

[Tvergaard 1999] V. Tvergaard, "Studies of elastic-plastic instabilities", *J. Appl. Mech.* (*ASME*) **66**:1 (1999), 3–9.

[Vardoulakis and Sulem 1995] I. Vardoulakis and J. Sulem, *Bifurcation Analysis in Geomechanics*, Blackie Academic & Professional, 1995.

[Wcisło and Pamin 2014] B. Wcisło and J. Pamin, "Entropic thermoelasticity for large deformations and its AceGen implementation", pp. 319–326 in *Recent advances in computational mechanics*, edited by T. Łodygowski et al., CRC Press, London, 2014.

[Wcisło and Pamin 2016] B. Wcisło and J. Pamin, "Local and non-local thermomechanical modeling of elastic-plastic materials undergoing large strains", *Int. J. Num. Methods Eng.* **65** (2016).

[Wcisło et al. 2013] B. Wcisło, J. Pamin, and K. Kowalczyk-Gajewska, "Gradient-enhanced damage model for large deformations of elastic-plastic materials", *Arch. Mech.* **65**:5 (2013), 407–428.

[Wriggers et al. 1992] P. Wriggers, C. Miehe, M. Kleiber, and J. C. Simo, "On the coupled thermomechanical treatment of necking problems via finite element methods", *Int. J. Num. Meth. Eng.* **33**:4 (1992), 869–883.

[Yang et al. 2006] Q. Yang, L. Stainier, and M. Ortiz, "A variational formulation of the coupled thermo-mechanical boundary-value problem for general dissipative solids", *J. Mech. Phys. Solids* **54**:2 (2006), 401–424.

[Zbib and Aifantis 1988a] H. M. Zbib and E. C. Aifantis, "On the localization and postlocalization behavior of plastic deformation, I: On the initiation of shear bands", *Res Mechanica* **23** (1988), 261–277.

[Zbib and Aifantis 1988b] H. M. Zbib and E. C. Aifantis, "On the localization and postlocalization behavior of plastic deformation, II: On the evolution and thickness of shear bands", *Res Mechanica* **23** (1988), 279–292.

[Zbib and Aifantis 1988c] H. M. Zbib and E. C. Aifantis, "On the localization and postlocalization behavior of plastic deformation, III: On the structure and velocity of Portevin – Le Chatelier bands", *Res Mechanica* **23** (1988), 293–305.

Received 1 Mar 2016. Revised 17 Jun 2016. Accepted 31 Aug 2016.

JERZY PAMIN: `jpamin@L5.pk.edu.pl`
*Institute for Computational Civil Engineering, Cracow University of Technology, ul. Warszawska 24, 31-155 Kraków, Poland*

BALBINA WCISŁO: `bwcislo@L5.pk.edu.pl`
*Institute for Computational Civil Engineering, Cracow University of Technology, ul. Warszawska 24, 31-155 Kraków, Poland*

KATARZYNA KOWALCZYK-GAJEWSKA: `kkowalcz@ippt.pan.pl`
*Institute of Fundamental Technological Research, Polish Academy of Sciences, ul. Pawińskiego 5B, 02-106 Warszawa, Poland*

# SUBMISSION GUIDELINES

## ORIGINALITY

Authors may submit manuscripts in PDF format online at the Submissions page. Submission of a manuscript acknowledges that the manuscript is original and has neither previously, nor simultaneously, in whole or in part, been submitted elsewhere. Information regarding the preparation of manuscripts is provided below. Correspondence by email is requested for convenience and speed. For further information, write to contact@msp.org.

## LANGUAGE

Manuscripts must be in English. A brief abstract of about 150 words or less must be included. The abstract should be self-contained and not make any reference to the bibliography. Also required are keywords and subject classification for the article, and, for each author, postal address, affiliation (if appropriate), and email address if available. A home-page URL is optional.

## FORMAT

Authors can use their preferred manuscript-preparation software, including for example Microsoft Word or any variant of TEX. The journal itself is produced in LATEX, so accepted articles prepared using other software will be converted to LATEX at production time. Authors wishing to prepare their document in LATEX can follow the example file at www.jomms.net (but the use of other class files is acceptable). At submission time only a PDF file is required. After acceptance, authors must submit all source material (see especially Figures below).

## REFERENCES

Bibliographical references should be complete, including article titles and page ranges. All references in the bibliography should be cited in the text. The use of BibTEX is preferred but not required. Tags will be converted to the house format (see a current issue for examples); however, for submission you may use the format of your choice. Links will be provided to all literature with known web locations; authors can supply their own links in addition to those provided by the editorial process.

## FIGURES

Figures must be of publication quality. After acceptance, you will need to submit the original source files in vector format for all diagrams and graphs in your manuscript: vector EPS or vector PDF files are the most useful. (EPS stands for Encapsulated PostScript.)

Most drawing and graphing packages—Mathematica, Adobe Illustrator, Corel Draw, MATLAB, etc.—allow the user to save files in one of these formats. Make sure that what you're saving is vector graphics and not a bitmap. If you need help, please write to graphics@msp.org with as many details as you can about how your graphics were generated.

Please also include the original data for any plots. This is particularly important if you are unable to save Excel-generated plots in vector format. Saving them as bitmaps is not useful; please send the Excel (.xls) spreadsheets instead. Bundle your figure files into a single archive (using zip, tar, rar or other format of your choice) and upload on the link you been given at acceptance time.

Each figure should be captioned and numbered so that it can float. Small figures occupying no more than three lines of vertical space can be kept in the text ("the curve looks like this:"). It is acceptable to submit a manuscript with all figures at the end, if their placement is specified in the text by means of comments such as "Place Figure 1 here". The same considerations apply to tables.

## WHITE SPACE

Forced line breaks or page breaks should not be inserted in the document. There is no point in your trying to optimize line and page breaks in the original manuscript. The manuscript will be reformatted to use the journal's preferred fonts and layout.

## PROOFS

Page proofs will be made available to authors (or to the designated corresponding author) at a Web site in PDF format. Failure to acknowledge the receipt of proofs or to return corrections within the requested deadline may cause publication to be postponed.

Journal of Mechanics of Materials and Structures

Volume 12, No. 1 January 2017

# Special issue on

# Coupled Field Problems and Multiphase Materials

**Preface** Corina S. Drapaca, Stefan Hartmann, Jacek Leszczyński, Sivabal Sivaloganathan and Wojciech Sumelka 1

**Variational methods for the solution of fractional discrete/continuous Sturm–Liouville problems** Ricardo Almeida, Agnieszka B. Malinowska, M. Luísa Morgado and Tatiana Odzijewicz 3

**Analytical and numerical solution of the fractional Euler–Bernoulli beam equation** Tomasz Blaszczyk 23

**Fractional calculus in neuronal electromechanics** Corina S. Drapaca 35

**Time-adaptive finite element simulations of dynamical problems for temperature-dependent materials** Matthias Grafenhorst, Joachim Rang and Stefan Hartmann 57

**Computer simulation of the effective viscosity in Brinkman filtration equation using the Trefftz method** Jan Adam Kołodziej, Magdalena Mierzwiczak and Jakub Krzysztof Grabski 93

**Numerical simulations of mechanical properties of alumina foams based on computed tomography** Zdzisław Nowak, Marcin Nowak, Ryszard Pęcherski, Marek Potoczek and Romana Śliwa 107

**Gradient-enhanced large strain thermoplasticity with automatic linearization and localization simulations** Jerzy Pamin, Balbina Wcisło and Katarzyna Kowalczyk-Gajewska 123

1559-3959(2017)12:1;1-4