Archivum mathematicum

Proceedings of Equadiff 15

Section
Numerical analysis and applications

FOREWORD TO PROCEEDINGS OF EQUADIFF 15

The Conference on Differential Equations and Their Applications – abbreviated as Equadiff – is one of the oldest active series of mathematical conferences in the world. The tradition of the Czechoslovak Equadiff dates back to 1962 when Equadiff 1 took place in Prague. The subsequent Czechoslovak Equadiff conferences are held since then periodically in Prague, Bratislava, and Brno every four years (with few exceptions). The Western Equadiff conferences are organized in various cities in Western Europe, starting in Marseille in 1970 and with the last meeting in Leiden in 2019.

The last Equadiff was held in Brno in summer 2022 as the 15th conference within the Czechoslovak Equadiff series, and hence it bears the name Equadiff 15. The conference was rescheduled to the year 2022 from the original date in July 2021 due to an unstable pandemic situation in the world. The proceedings from all previous Czechoslovak Equadiff conferences are available via the Czech Digital Mathematics Library at

`https://dml.cz/handle/10338.dmlcz/700001`.

The conference Equadiff 15 was organized by joint efforts of the Faculty of Science of Masaryk University (and its Department of Mathematics and Statistics) with the Faculty of Civil Engineering of Brno University of Technology, the Institute of Mathematics of the Czech Academy of Sciences, and the Brno branch of The Union of Czech Mathematicians and Physicists. The conference took place at the campus of the Faculty of Economics and Administration of Masaryk University from July 11 till July 15, 2022. More than 250 participants from 37 countries from all over the world attended the 241 talks of the conference, including 6 plenary talks, 17 invited talks, 124 talks in 33 organized minisymposia, 75 contributed talks, and 19 posters.

The proceedings of Equadiff 15 cover the theory of differential equations in a broad sense, including their theoretical aspects, numerical methods, and applications. The proceedings contain 29 scientific articles written by participants of Equadiff 15. The papers are divided into three sections according to the program of the conference:
  – ordinary differential equations (15 papers),
  – partial differential equations (9 papers),
  – numerical analysis and applications (5 papers).
Each manuscript underwent a rigorous refereeing process to ensure its scientific quality. This issue contains the contributions from section Numerical analysis and applications.

We would like to take this opportunity to express our special thanks to all the participants for their active contributions to the success of the Equadiff 15 conference. Our gratitude and appreciation belong to the members of the Scientific Committee who ensured the high standards of the scientific activities of the conference, to the organizers and supporting PhD students for their efforts towards the realization of the conference, to the administration of the Faculty of Economics and Administration of Masaryk University for providing the venue for the conference and for their organizational support, to the management and employees of the Accommodation and Catering Services of Masaryk University for their help with the organization and realization of the catering during the conference, to the workers of the Botanical Garden of the Faculty of Science of Masaryk University for providing the flower decoration, and to the director of the Department of Mathematics and Statistics of the Faculty of Science of Masaryk University for financial support. We also thank to Ilona Lukešová from the Editorial Office of Archivum Mathematicum for her extensive editorial work on these proceedings.

Brno – Prague, Czech Republic
January, 2023

Zuzana Došlá
Jan Chleboun
Pavel Krejčí
Martin Kružík
Šárka Nečasová
Roman Šimon Hilscher
(The editors)

The guest editors

Zuzana Došlá: Masaryk University, Faculty of Science, Kotlářská 2, 611 37 Brno, Czech Republic, `dosla@math.muni.cz`

Jan Chleboun: Czech Technical University in Prague, Faculty of Civil Engineering, Thákurova 7, 166 29 Praha 6, Czech Republic, `jan.chleboun@cvut.cz`

Pavel Krejčí: Institute of Mathematics of the Czech Academy of Sciences, Žitná 25, 115 67 Praha 1, Czech Republic, `pavel.krejci@cvut.cz`

Martin Kružík: Institute of Information Theory and Automation of the Czech Academy of Sciences, Pod vodárenskou věží 4, 182 00 Praha 8, Czech Republic, `kruzik@utia.cas.cz`

Šárka Nečasová: Institute of Mathematics of the Czech Academy of Sciences, Žitná 25, 115 67 Praha 1, Czech Republic, `matus@math.cas.cz`

Roman Šimon Hilscher: Masaryk University, Faculty of Science, Kotlářská 2, 611 37 Brno, Czech Republic, `hilscher@math.muni.cz`

DEEP LEARNING FOR GRADIENT FLOWS USING THE BREZIS–EKELAND PRINCIPLE

Laura Carini, Max Jensen, and Robert Nürnberg

Abstract. We propose a deep learning method for the numerical solution of partial differential equations that arise as gradient flows. The method relies on the Brezis–Ekeland principle, which naturally defines an objective function to be minimized, and so is ideally suited for a machine learning approach using deep neural networks. We describe our approach in a general framework and illustrate the method with the help of an example implementation for the heat equation in space dimensions two to seven.

1. Introduction

In this paper we advocate a deep learning approach for solving parabolic partial differential equations (PDEs)

$$u_t + \partial\phi(u) = f \,,$$

that arise as evolution equations for gradient flows. We exploit the variational principle of the seminal papers by Brezis and Ekeland, [2,3], now commonly known as the Brezis–Ekeland principle, [17,19]. We also refer to [21] for related work, as well as to [7] for alternative variational formulations for large classes of PDEs.

Using neural networks for the numerical solution of PDEs has become increasingly popular over the last decade. An advantage of neural network-based approaches is their suitability for high-dimensional problems. For a comprehensive review of current developments we refer to [1,12]. Among currently popular techniques are methods based on residual minimization, e.g. see [15,16,18], and on the reformulation as a backward stochastic differential equation, e.g. [8,9,10,14].

Most relevant for this work are variational approaches. For an elliptic problem, E and Yu, [6], proposed a deep learning method based on a variational principle that leads to a natural optimization framework. An approach connecting variational principles with convex duality for stationary equations was taken in [11], which mirrors some aspects of our work for gradient flows.

The aim of our approach is three-fold:

(1) Machine learning approaches have been criticized for their less developed methodology to bound, or at least estimate, the approximation error. It is therefore interesting that the minimum of the Brezis–Ekeland functional, which is minimized during the learning process, is guaranteed to be zero for the exact solution, thus providing an error measure that is known at the point of computing the neural net approximation.

(2) Adversarial networks have very successfully been applied across multiple problem classes of machine learning. We were intrigued by the question whether duality can be a context in which the concept of adversarial networks is translated to partial differential equations as well as to convex analysis, by introducing a neural network for the primal and another one for the dual problem. In a resulting min-max formulation the training stages of the respective networks take opposing, or adversarial, roles in finding the value of the joint loss functional.

(3) Finally, we wish to construct a method which takes advantage of the specific structural properties of gradient flows, based on the relevance of these properties in the literature for the construction of finite element methods for time-dependent PDEs.

The outline of the remainder of the paper is as follows: in section 2 we introduce gradient flows and the Brezis–Ekeland principle; in section 3 we formulate our deep learning approach; in section 4 we discuss the computer implementation of the method; in section 5 we present numerical experiments, followed by conclusions.

2. The Brezis–Ekeland principle for gradient flows

Let $V \subset H \subset V^*$ be a Gelfand triple and $T > 0$ a fixed time. In addition, let $\phi \in C^1(V)$ be convex, $f \in L^2(0, T; V^*)$ and $u_0 \in V$. We consider the gradient flow

$$(2.1) \qquad u_t + \partial\phi(u) = f \quad \text{a.e. in } (0, T), \quad u(0) = u_0 \,.$$

The Brezis–Ekeland principle asserts that solutions $u \in Y$ to (2.1), with

$$Y := \{ w \in L^2(0, T; V) \cap H^1(0, T; V^*) \ : \ w(0) = u_0 \} \,,$$

are the global minimizers of the functional $\Phi : Y \to [0, \infty]$ defined by

$$\Phi(w) = \tfrac{1}{2}\|w(T)\|_H^2 - \tfrac{1}{2}\|w(0)\|_H^2 + \int_0^T \phi(w) + \phi^*(f - w_t) - \langle f, w \rangle \, \mathrm{d}t$$

$$(2.2) \qquad = \int_0^T \phi(w) + \phi^*(f - w_t) + \langle w_t - f, w \rangle \, \mathrm{d}t \,,$$

using $\int_0^T \langle w_t, w \rangle \, \mathrm{d}t = \tfrac{1}{2}\|w(T)\|_H^2 - \tfrac{1}{2}\|w(0)\|_H^2$. Here $\langle \cdot, \cdot \rangle$ is the duality pairing between V and V^*, and $\phi^*(w) = \sup_{v \in V} \langle w, v \rangle - \phi(v)$ is the conjugate of ϕ. In fact, owing to [17, Theorem 8.99], u solves (2.1) if and only if

$$(2.3) \qquad \Phi(u) = \min\{\Phi(w) \ : \ w \in Y\} = 0 \,.$$

Inspired by the work in [20], we now consider a time-discrete variant of (2.2) and the associated minimization problem. For that purpose, we divide $[0, T]$ into N sub-intervals with end points $t_0 = 0 < t_1 < \cdots < t_N = T$.

The parabolic nature of (2.1) ensures that $u(t)$ only depends on $u(s)$ if $s \leq t$, but not if $s > t$. Together with the Brezis–Ekeland principle (2.3) this guarantees that minimization and summation may be interchanged:

$$(2.4) \quad \min_{w \in Y} \Phi(w)$$

$$= \min_{w \in Y} \sum_{n=1}^{N} \int_{t_{n-1}}^{t_n} \phi(w) + \phi^*(f - w_t) + \langle w_t - f, w \rangle \, \mathrm{d}t$$

$$= \sum_{n=1}^{N} \min_{\substack{w_n \in Y \\ w_n(t_{n-1}) = w_{n-1}(t_{n-1})}} \int_{t_{n-1}}^{t_n} \phi(w_n) + \phi^*(f - (w_n)_t) + \langle (w_n)_t - f, w_n \rangle \, \mathrm{d}t,$$

where we have defined $w_0(0) = u_0$. The representation (2.4) of the minimization problem suggests a clear strategy for our deep learning method: we will sequentially solve N optimization problems for the PDE (2.1) on the time intervals $[t_{n-1}, t_n]$, where the initial data is either given by u_0 at the first step, or by the previously computed solution at time t_{n-1}.

The heat equation. The canonical example of a gradient flow is the heat equation. Given a Lipschitz domain $\Omega \subset \mathbb{R}^d$, $d \geq 1$, we consider the PDE:

$$(2.5) \quad \begin{cases} u_t - \kappa \Delta u = f, & \text{in} \quad (0, T) \times \Omega, \\ u(0, \cdot) = u_0, & \text{in} \quad \Omega, \\ u = 0, & \text{on} \quad (0, T) \times \partial \Omega, \end{cases}$$

where $\kappa > 0$. Upon defining $H = L^2(\Omega)$, $V = H_0^1(\Omega)$ and $\phi(u) = \frac{\kappa}{2} \|\nabla u\|_{L^2}^2$, the problem (2.5) is a special case of (2.1), and the Brezis–Ekeland functional (2.2) in this case reduces to:

$$(2.6) \quad \Phi(w) = \int_0^T \frac{\kappa}{2} \|\nabla w\|_{L^2}^2 + \frac{1}{2\kappa} \left[\sup_{v \neq 0} \frac{\langle f - w_t, v \rangle}{\|\nabla v\|_{L^2}} \right]^2 + \langle w_t - f, w \rangle \, \mathrm{d}t,$$

where we have used that $\phi^*(w) = \frac{1}{2\kappa} \left[\sup_{v \in H_0^1(\Omega) \setminus \{0\}} \frac{\langle w, v \rangle}{\|\nabla v\|_{L^2}} \right]^2$, e.g. see [17]. We remark that ϕ^* defines a norm on $H^{-1}(\Omega) = (H_0^1(\Omega))^*$. Combining (2.4) and (2.6) we find that $\min_{w \in Y} \Phi(w)$ equals

$$\sum_{n=1}^{N} \min_{\substack{w_n(t_{n-1}) \\ = w_{n-1}(t_{n-1})}} \int_{t_{n-1}}^{t_n} \frac{\kappa}{2} \|\nabla w_n\|_{L^2}^2 + \frac{1}{2\kappa} \left[\sup_{v \neq 0} \frac{\langle f - (w_n)_t, v \rangle}{\|\nabla v\|_{L^2}} \right]^2$$

$$(2.7) \qquad\qquad + \langle (w_n)_t - f, w_n \rangle \, \mathrm{d}t,$$

with the solution u of (2.5) being a minimizer in the sense that $\Phi(u) = 0$ and that the choice $w_n = u$, $n \in \{1, \ldots, N\}$, yields a minimizer of (2.7) over Y^N.

3. The deep learning approach

In this section we discuss a deep learning algorithm to find approximations of the solution to (2.1). For simplicity we restrict our attention to the heat equation, so that (2.7) is our starting point.

We wish to find the approximations $u_h^n \approx u(t_n, \cdot)$, $n = 1, \ldots, N$, where u_h^n is given by a neural network. More generally, in this paper we view a neural network $\hat{u}_h$ as a function determined through its weights θ. Given θ the neural network takes a position $x \in \Omega$ as input and returns $\hat{u}_h(x; \theta)$ as output. The approximation set containing the functions u_h^n is thus given by

$$\mathbb{U}_h := \{\hat{u}_h(\cdot; \theta) \, : \, \theta \in \Theta\},$$

where Θ is the set of possible weights. In this notation $u_h^n = \hat{u}_h(\cdot; \theta)$ is the neural network with a choice of weights θ determined through the method described in this section.

In order to define a discrete version of the Brezis–Ekeland functional (2.7) and perform its minimization over $\mathbb{U}_h$, we turn our attention to the interpretation of v in $\phi^*(w) = \frac{1}{2\kappa} \left[\sup_{v \in H_0^1(\Omega) \setminus \{0\}} \langle w, v \rangle / \|\nabla v\|_{L^2} \right]^2$. Also the v are approximated by neural networks. Since their architecture may be different compared to $\hat{u}_h$, we introduce $\hat{v}_h(x; \eta)$ and

$$\mathbb{V}_h := \{\hat{v}_h(\cdot; \eta) \, : \, \eta \in \mathrm{H}\},$$

where H is the set of possible weights in $\hat{v}_h$.

While we assume $\mathbb{U}_h, \mathbb{V}_h \subset H^1(\Omega)$ throughout, elements of $\mathbb{U}_h$ and $\mathbb{V}_h$ will in general not belong to $H_0^1(\Omega)$ because the Dirichlet boundary conditions may not be satisfied homogeneously. Therefore, similarly to [6], we introduce a penalty term into $\phi^*(w)$ to obtain the functional

$$(3.1) \qquad \phi_h^*(w_h) = \frac{1}{2\kappa} \left[\sup_{v_h \in \mathbb{V}_h \setminus \{0\}} \frac{(w_h, v_h)}{\left(\|\nabla v_h\|_{L^2}^2 + \lambda \|v_h\|_{L^2(\partial\Omega)}^2 \right)^{\frac{1}{2}}} \right]^2,$$

where $\lambda > 0$ is a penalty parameter depending on the structure of the neural net. Then the denominator in (3.1) cannot vanish for $v_h \neq 0$ due to a Poincaré-Friedrichs inequality and the penalization weakly imposes homogeneous boundary conditions on any maximising v_h as $\lambda \to \infty$. We note that such penalization strategies to enforce Dirichlet boundary conditions are common in deep learning approaches, see e.g. [6], and have a long tradition in discontinuous Galerkin approximation methods, [5, §4.2].

It remains to discretize the time derivatives in the Brezis–Ekeland functional. To this end we substitute $(w_n)_t$ in (2.7) by backward time differences. Let $\Delta t_n = t_n - t_{n-1}$ and let $(\cdot, \cdot)$ denote the L^2-inner product over Ω. Inspired by (2.7), we then define the solution of the deep learning method through the following sequence of optimization problems: Given $u_h^{n-1} \in \mathbb{U}_h$, for $n = 1, \ldots, N$, find a minimizer

$u_h^n \in \mathbb{U}_h$ to

$$\Phi_n(w_h) = \frac{\kappa \Delta t_n}{2} \|\nabla w_h\|_{L^2}^2 + \Delta t_n \phi_h^* \Big(f - \frac{w_h - u_h^{n-1}}{\Delta t_n} \Big)$$

(3.2)
$$+ (w_h - u_h^{n-1}, w_h) + \lambda \|w_h\|_{L^2(\partial\Omega)}^2 \,,$$

where we have once again added a penalization term; this time to weakly impose homogeneous Dirichlet boundary conditions on w_h.

Obtaining the minimizer of (3.2) requires a maximization to evaluate $\phi_h^*(f - (w_h - u_h^{n-1})/\Delta t_n)$, see (3.1). For the remainder of this section we focus on an algorithm for solving this min-max problem. In the subsequent text it will be convenient to refer to

$$\widetilde{\phi}_h^*(w_h; v_h) = \frac{1}{2\kappa} \left[\frac{(w_h, v_h)}{(\|\nabla v_h\|_{L^2}^2 + \lambda \|v_h\|_{L^2(\partial\Omega)}^2)^{\frac{1}{2}}} \right]^2 \,,$$

which is equal to $\phi_h^*(w_h)$ if v_h is a maximizer. Similarly, we write

(3.3) $\widetilde{\Phi}_n(w_h; p_h) = \frac{\kappa \Delta t_n}{2} \|\nabla w_h\|_{L^2}^2 + \Delta t_n\, p_h + (w_h - u_h^{n-1}, w_h) + \lambda \|w_h\|_{L^2(\partial\Omega)}^2 \,,$

which equals $\Phi_n(w_h)$ upon choosing $p_h = \phi_h^*\big(f - \frac{w_h - u_h^{n-1}}{\Delta t_n}\big)$.

Algorithm 1

1: Compute an approximation $u_h^0 \in \mathbb{U}_h$ to u_0
2: **for** $n = 1, \dots, N$ **do**
3:     $u_h^{n,0} \leftarrow u_h^{n-1}$, $k \leftarrow 0$
4:     **while** termination_criterion($u_h^{n,k}$, k) = FALSE **do**
5:         $k \leftarrow k + 1$
6:         $p_h^{n,k} = \max_{v_h \in \mathbb{V}_h \setminus \{0\}} \widetilde{\phi}_h^*(f - (u_h^{n,k-1} - u_h^{n-1})/\Delta t_n; v_h)$
7:         $u_h^{n,k} \in \arg\min_{w_h \in \mathbb{U}_h} \widetilde{\Phi}_n(w_h; p_h^{n,k})$
8:     **end while**
9:     $u_h^n \leftarrow u_h^{n,k}$
10: **end for**

We shall base the minimization of (3.2) on Algorithm 1. The approximation of the initial conditions in line 1 of the algorithm is a supervised learning problem. Lines 2 and 10 frame the iteration over the time steps. Lines 4 and 8 implement a loop where the optimization of $\widetilde{\phi}_h^*$ (line 6) and $\widetilde{\Phi}_n$ (line 7) are alternated.

Remark 3.1. Alternatively to the above, one could use $\phi^*(w) = \frac{1}{2\kappa} \|\nabla \Delta^{-1} w\|_{L^2}^2$ for the formulation of the method, see [17, Example 8.104]. In this scenario we envisage $\Delta^{-1} u$ being approximated by a neural net, using an existing methodology for solving the Laplace problem.

4. BENNO: Brezis–Ekeland Neural Network Optimizer

We make a full Python implementation [4] of our deep learning approach available on Github, which is called *Brezis–Ekeland Neural Network Optimizer*, in short BENNO.

Neural Network Structure. We describe the internal structure of the neural networks $\hat{u}_h$ and $\hat{v}_h$, which were introduced in the previous section. Both these neural networks use five densely-connected layers with a linear activation function for the input and output layers and a (leaky) rectified linear unit activation function $\sigma(s) = \max\{0, s\} + \mu \min\{0, s\}$, $\mu \geq 0$, for the inner layers. Each layer is made of m nodes, except for the output layer that presents a single node. This means that the neural network $\hat{u}_h$ has the following architecture:

$$S^1(x) = W^1 x + b^1, \; S^2(x) = \sigma(W^2 S^1(x) + b^2), \; S^3(x) = \sigma(W^3 S^2(x) + b^3),$$

$$S^4(x) = \sigma(W^4 S^3(x) + b^4), \; S^5(x) = W^5 S^4(x) + b^5, \quad \text{and} \quad \hat{u}_h(x; \theta) = S^5(x),$$

where the set of parameters of the neural network $\hat{u}_h(\cdot; \theta)$ are given by

$$\theta = \{W^1, b^1, W^2, b^2, W^3, b^3, W^4, b^4, W^5, b^5\},$$

with $W^1 \in \mathbb{R}^{m \times d}$, $W^i \in \mathbb{R}^{m \times m}$ for $i = 2, 3, 4$, $W^5 \in \mathbb{R}^{1 \times m}$, $b^j \in \mathbb{R}^m$ for $j = 1, \dots, 4$ and $b^5 \in \mathbb{R}$. With a slight abuse of notation, the application of the activation function is understood elementwise: $\sigma(z)$ is the vector $(\sigma(z_1), \dots, \sigma(z_m))$ for $z = (z_1, \dots, z_m) \in \mathbb{R}^m$. By default, we set $\mu = 0.03$ in the definition of σ.

Also the network $\hat{v}_h$ has an architecture of this type; however, generally with a different parameter m. As indicated in the previous section, we denote the weights of $\hat{v}_h$ by η.

Adam Optimizer, Loss Functions and Algorithm. We implemented the neural networks with the help of the Tensorflow Sequential API. Both neural networks were trained with the Adam Optimizer, a variant of the stochastic gradient descent method based on an adaptive estimation of first-order and second-order moments that improves the speed of convergence [13]. For the optimization parameters, we use the standard values $\beta_1 = 0.9$, $\beta_2 = 0.999$ and $\varepsilon = 10^{-8}$ in the notation of [13].

There are three distinct optimization scenarios with their respective loss functions:

(1) The approximation of u_0 by $u_h^0 \in \mathbb{U}_h$ in line 1 of Algorithm 1: It is a supervised learning problem with the loss function $\mathcal{L} = \|u_0 - w_h\|_{L^2}^2$. We use the constant learning rate $\alpha = 10^{-3}$.

(2) The maximization of $\widetilde{\phi}_h^*$ in line 6 of Algorithm 1: We use the constant learning rate $\alpha = 10^{-5}$. By default the training extends over 500 epochs.

(3) The minimization of $\widetilde{\Phi}_n$ in line 7 of Algorithm 1: We employ the k-dependent decaying learning rate

$$\alpha(k) = 10^{-5} \mathbf{1}_{\{k \leq 5\}} + 10^{-6} \mathbf{1}_{\{5 < k \leq 50\}} + 10^{-7} \mathbf{1}_{\{50 < k \leq 120\}}$$
$$+ 10^{-8} \mathbf{1}_{\{120 < k \leq 140\}} + 10^{-9} \mathbf{1}_{\{140 < k \leq 180\}} + 10^{-10} \mathbf{1}_{\{180 < k\}}.$$

By default the training extends over 50 epochs.

The integrals appearing in these loss functions are evaluated with the help of a Monte-Carlo integration method, using the sampling points $\{x_i : i = 1, \ldots, N_s\} \subset \overline{\Omega}$. Here $N_s = N_i + N_b$, with N_i points drawn from a uniform distribution in Ω and N_b points drawn from a uniform distribution on $\partial\Omega$.

Finally, the default termination criterion in line 4 of Algorithm 1 is

$$\text{termination\_criterion}(u_h^{n,k}, k) = \begin{cases} \text{TRUE} & : \ k > 200\,, \\ \text{FALSE} & : \ k \le 200\,, \end{cases}$$

which is employed in all numerical experiments of the forthcoming section.

5. NUMERICAL RESULTS

We consider problem (2.5), with $f = 0$, on the domain $\Omega = (0, \pi)^d$, for $d = 2, 3, 5, 7$. Given the initial condition $u_0(x) = \prod_{i=1}^d \sin(a_i x_i)$, for $a \in \mathbb{N}^d$ and $x \in \overline{\Omega}$, the exact solution to (2.7) with $\kappa = [\sum_{i=1}^d a_i^2]^{-1}$ is $u(t, x) = e^{-t} \prod_{i=1}^d \sin(a_i x_i)$.

We investigate the following types of approximation errors:

$$\text{MSE} = \frac{1}{N_s} \sum_{i=1}^{N_s} (u(t_n, x_i) - u_h^n(x_i))^2,$$

$$\varepsilon_{\text{abs},L^\infty} = \max_{i=1,\ldots,N_s} |u(t_n, x_i) - u_h^n(x_i)|, \ \ \varepsilon_{\text{rel},L^2} = \left[\frac{\sum_{i=1}^{N_s} (u(t_n, x_i) - u_h^n(x_i))^2}{\sum_{i=1}^{N_s} (u(t_n, x_i))^2} \right]^{\frac{1}{2}},$$

$$\varepsilon_{\text{rel},H^1} = \left[\frac{\sum_{i=1}^{N_s} (u(t_n, x_i) - u_h^n(x_i))^2 + |\nabla u(t_n, x_i) - \nabla u_h^n(x_i)|^2}{\sum_{i=1}^{N_s} (u(t_n, x_i))^2 + |\nabla u(t_n, x_i)|^2} \right]^{\frac{1}{2}},$$

where MSE stands for mean square error, and the other quantities define approximations of the L^∞-norm error and of the relative L^2- and H^1-norm errors, respectively.

In addition the Brezis–Ekeland functional itself represents a measure of the accuracy of the deep learning algorithm since we look for u such that $\Phi(u) = \min \Phi = 0$. It follows that values of the loss function $\widetilde{\Phi}_n$ give us information about the quality of the training and the approximate solution u_h^n.

Unless otherwise stated, we use the layer width $m = m_{v_h} = 30$ for the neural networks $\hat{v}_h$, while the the layer width $m = m_{u_h}$ will be varied for $\hat{u}_h$ depending on the dimension d. For the time discretization we use uniform time steps $\Delta t_n = \Delta t$, $n = 1, \ldots, N$, where we always choose $\Delta t = 10^{-4}$. Finally, for the boundary value penalty parameter we always use $\lambda = 100$.

Energy landscapes for a 5D problem. Let $d = 5$ and $a = (2, 2, 1, 2, 3)^\intercal$. We choose $N_i = 10^5$ inner and $N_b = 10^3$ boundary sampling points. The optimization to obtain the initial value approximation u_h^0 is done over $5 \cdot 10^4$ epochs. We use $m_{u_h} = 60$ for $\hat{u}_h$.

We are interested in the shape of the graphs of the two objective functions $\widetilde{\Phi}_n$ and $\widetilde{\phi}_h^*$ as functions of the neural network weights θ and η, respectively. This will allow us to gain insight into how challenging the training of the neural nets is.
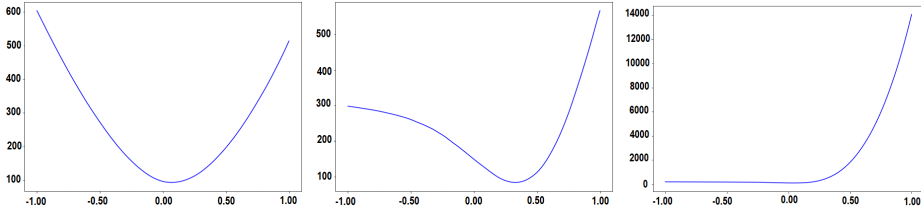
FIG. 1. Plots of $\widetilde{\Phi}_n(\hat{u}_h(\theta); p_h^{1,200})$ against different components of θ: b_{60}^1 (left), $W_{46,60}^3$ (middle) and $W_{7,45}^3$ (right).
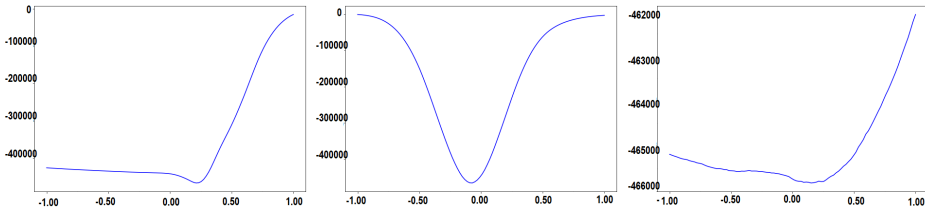

FIG. 2. Plots of $-\widetilde{\phi}_h^*((u_h^{1,199} - u_h^0)/\Delta t; \hat{v}_h(\eta))$ against different components of η: b_{30}^2 (left), $W_{2,1}^3$ (middle) and $W_{21,3}^4$ (right).

Having computed $u_h^1 = u_h^{1,K} = \hat{u}_h(\hat{\theta})$, in Figure 1 we plot the loss function $\widetilde{\Phi}_n(\hat{u}_h(\theta); p_h^{n,K})$, for $n = 1$ and $K = 200$, against selected entries of θ. In particular, for each plot we keep all the weights in $\theta = \hat{\theta}$ fixed, apart from a single entry of θ, that we continuously vary from -1 to 1. In this way it is possible to visualize how the Brezis–Ekeland functional varies depending on certain parameters of the neural network $\hat{u}_h$. While generally smooth, we note that the right plot in Figure 1 shows that $\widetilde{\Phi}_n$ has a nearly vanishing gradient when the parameter $W_{7,45}^3$ varies in $[-1,0]$, which may require attention during the optimization process.

Similarly, in Figure 2 we show the loss function $-\widetilde{\phi}_h^*((u_h^{n,k-1} - u_h^{n-1})/\Delta t; \hat{v}_h(\eta))$ plotted against selected entries of the neural network weights η. Once again we observe nearly flat parts in the graph, but now in addition we see also some non-convex and non-smooth regions, which may pose challenges during the optimization.

Error quantities for a 5D problem. We use the previous example to compute error quantities for the trained neural networks. Table 1 shows values of the Brezis–Ekeland loss function ($\widetilde{\Phi}_n$), the mean square error (MSE), the absolute error ($\varepsilon_{\mathrm{abs},L^\infty}$) and the relative errors ($\varepsilon_{\mathrm{rel},L^2}$, $\varepsilon_{\mathrm{rel},H^1}$) for the neural networks u_h^n for every t_n, $n = 0, \ldots, N$. The reported values of $\widetilde{\Phi}_n$ are large, possibly caused by the fact that the measure of Ω is considerable with $|\Omega| = \pi^5 \approx 306$, as is the scaling of the boundary term $\lambda \| \cdot \|_{L^2(\partial\Omega)}^2$ with $\lambda = 100$ and $|\partial\Omega| = 2 \cdot 5 \cdot \pi^4 \approx 974$. We observe that all the other error quantities slowly increase with time, which is typical for approximations of parabolic PDEs.

	$\widetilde{\Phi}_n$	MSE	$\varepsilon_{abs,L^\infty}$	ε_{rel,L^2}	ε_{rel,H^1}
t_0	—	5.894e-04	0.142	0.138	0.420
t_1	112.762	6.195e-04	0.134	0.142	0.438
t_2	67.164	6.516e-04	0.146	0.140	0.439
t_3	72.938	6.785e-04	0.145	0.149	0.439
t_4	62.323	7.214e-04	0.150	0.153	0.441
t_5	51.332	7.466e-04	0.150	0.156	0.442
t_6	85.143	7.804e-04	0.158	0.160	0.444
t_7	45.920	8.211e-04	0.161	0.164	0.445
t_8	48.652	8.196e-04	0.162	0.164	0.446
t_9	58.798	8.686e-04	0.167	0.169	0.447
t_{10}	37.395	8.780e-04	0.166	0.170	0.448

TAB. 1. Error quantities at times t_n, $n = 0, \ldots, N$, for the 5D test problem.

Apart from the global error properties, we are also interested in how these quantities change during the training process. In Figure 3 we plot the loss function $\widetilde{\Phi}_n(u_h^{n,k}; p_h^{n,k})$ and the four contributions to it against k, for $k = 1, \ldots, K = 200$, during the training for the time $t_4 = 4 \cdot 10^{-4}$, i.e. $n = 4$. We observe a significant decrease of the Brezis–Ekeland functional $\widetilde{\Phi}_n$ during the training, from about 400 for $k = 1$ to about 60 for $k = 200$, when the weights seem to have converged. Observe also that the decrease is non-monotone, with a global maximum of about 1000, and that the graph is rather oscillatory. In addition, we note that after an initial increase, the functional decays rapidly at first and then slower as the iteration proceeds. The plot of the four contributions reveals that the term $\lambda \|u_h^{n,k}\|^2_{L^2(\partial\Omega)}$ is the dominant contribution in the Brezis–Ekeland functional (3.3) once the iterative scheme settles down.

The analogous plot for the mean square error (MSE) is shown on the left of Figure 4, where we again notice an oscillatory decrease until convergence is reached. In addition, on the right of Figure 4 we show the concatenated plots of the MSE against k, for every time t_n, $n = 1, \ldots, N$. To help differentiate the different time steps, we indicate the start of the training for a new time step with vertical lines. The figure demonstrates that overall the MSE increases in time, but that each training procedure decreases the MSE until convergence can be observed.

Dependence on the dimension d. Here we let $d = 2, 3$ or 5, and set $a = (2, 2)^\intercal$, $a = (2, 2, 3)^\intercal$ and $a = (2, 2, 1, 2, 3)^\intercal$, respectively. The number of sampling points in Ω is $N_i = 10^4$ for $d = 2$ and $N_i = 10^5$ for $d = 3, 5$, while $N_b = 400, 600, 1000$ for $d = 2, 3, 5$, respectively. Moreover, for the training of the initial conditions u_h^0 we use $5 \cdot 10^3$ epochs in the cases $d = 2, 3$, and $5 \cdot 10^4$ for $d = 5$. We set $m_{u_h} = 60$ throughout.

Table 2 shows the values of the Brezis–Ekeland loss function $\widetilde{\Phi}_n$, the MSE, the absolute error $\varepsilon_{\mathrm{abs},L^\infty}$, the relative error $\varepsilon_{\mathrm{rel},L^2}$ and the GPU time used for the training of the neural networks $\hat{u}_h$ and $\hat{v}_h$, for the tenth time step, $t_{10} = 0.001$, for
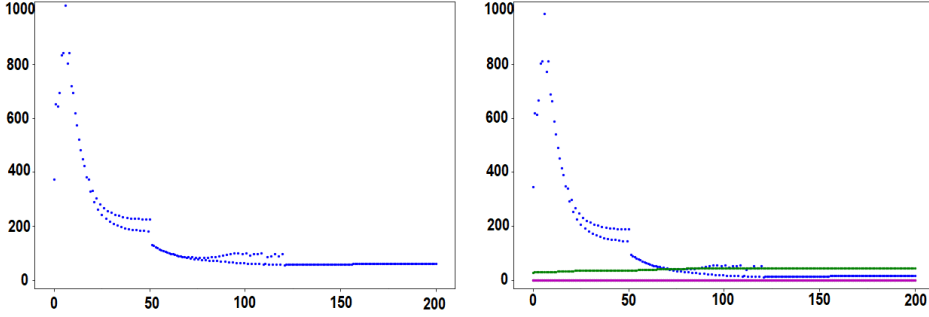
FIG. 3. Plots of $\widetilde{\Phi}_n(u_h^{n,k}; p_h^{n,k})$ (left) and of the 4 terms contributing to it (right) against k, for $n = 4$, for the 5D test problem. For the right plot the chosen colours are red for $\frac{\kappa\Delta t}{2}\|\nabla u_h^{n,k}\|_{L^2}^2$, blue for $\Delta t\, p_h^{n,k}$, magenta for $(u_h^{n,k} - u_h^{n-1}, u_h^{n,k})$ and green for $\lambda\|u_h^{n,k}\|_{L^2(\partial\Omega)}^2$.
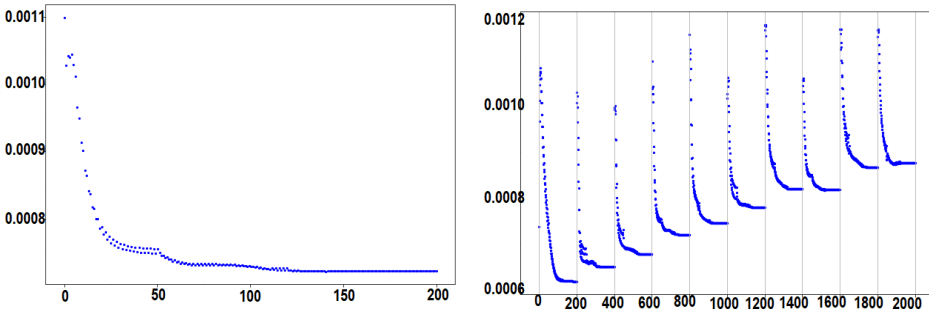

FIG. 4. Plots of the MSE against k at time $t_4 = 4 \cdot 10^{-4}$ (left), and of the MSE against $(n-1)K + k$, for $n = 1, \ldots, N$. (right)

the three problems $d = 2, 3, 5$. It can be observed that an increase in the dimension leads to a decrease in the accuracy of the algorithm. Indeed, reading the table from left to right, we notice that all the measures increase monotonically with the dimension, with the only exception being the absolute error $\varepsilon_{\mathrm{abs}, L^\infty}$, which actually reduces slightly going from the 3D to the 5D problem. While a growth of the error is to be expected with an increase of the dimension, the figures here indicate that further improvements of the neural net architecture and the training methodology should be investigated.

Effect of the number of nodes m for a 7D problem. Let $d = 7$ and $a = (2, 2, 1, 3, 2, 2, 3)^\intercal$. We choose $N_i = 10^5$ and $N_b = 1400$ sample points. The initial conditions are training over $5 \cdot 10^4$ epochs. The layer width m_{u_h} of the networks $\hat{u}_h$ varies between 60 and 100.

	2D	3D	5D
$\widetilde{\Phi}_n$	4.54e-02	3.97	39.93
MSE	1.74e-04	1.00e-03	7.67e-03
$\varepsilon_{abs,L^\infty}$	6.25e-03	0.14	0.13
ε_{rel,L^2}	2.71e-02	8.98e-02	0.15
GPU time [s]	58195	73595	100641

TAB. 2. Comparison of error quantities and the GPU time for the three test problems with $d = 2, 3, 5$.

m_{u_h}	60	100
$\widetilde{\Phi}_n$	469.11	60.44
MSE	3.80e-03	1.21e-03
$\varepsilon_{abs,L^\infty}$	0.59	0.31
ε_{rel,L^2}	0.70	0.38
GPU time [s]	119972	117657

TAB. 3. Comparison of error quantities and the GPU time for the 7D test problem with either 60 or 100 nodes per layer.

From the error quantities reported in Table 3 we can immediately see that using more nodes in the network architecture is beneficial in higher dimensions. In fact, all the reported quantities are lower when considering the case with $m_{u_h} = 100$.

Finally, in Figure 5 we present plots of the MSE against $k = 1, \ldots, K = 200$ for the training of the third time step, $n = 3$, comparing the performance of the two different network structures. We note that the MSE is on average increasing for the network with 60 nodes per layer, possibly indicating that the neural network does not have enough expressivity to make the learning effective. In contrast, when training with wider layers one is able to decrease the MSE after the typical initial increase. These results strongly indicate that for higher dimensional problems, more elaborate networks perform better in practice. Unfortunately, memory and GPU time limitations mean that at present we are not able to investigate this trend for even higher dimensional problems.

6. CONCLUSIONS

We introduced a novel deep learning approach for the numerical solution of PDEs using the Brezis–Ekeland principle. As a proof of concept we implemented a practical algorithm for the heat equation and presented results for experiments up to dimension 7. Higher dimensional problems are particularly computationally challenging, and more research into the optimal design for the employed neural networks is needed. Similarly to other neural network based approaches, we expect our method to be superior to classical algorithms for very high dimensional problems, whereas in lower space dimensions classical algorithms are likely more efficient. In
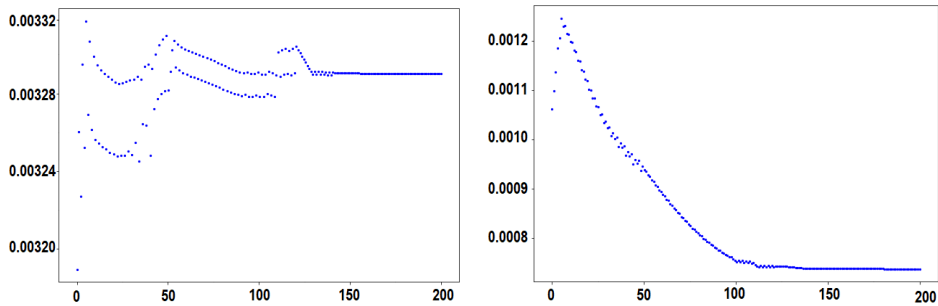
FIG. 5. Plots of the MSE against k at time $t_3 = 3 \cdot 10^{-4}$ for the 7D problem with neural networks using 60 (left) and 100 (right) nodes per layer.

addition, an extension of the implemented method to nonlinear problems is part of future research.

REFERENCES

[1] Blechschmidt, J., Ernst, O.G., *Three ways to solve partial differential equations with neural network - a review*, GAMM-Mitt. **44** (2021), no. 2, Paper No. e202100006, 29.

[2] Brézis, H., Ekeland, I., *Un principe variationnel associé à certaines équations paraboliques. Le cas dépendant du temps*, C. R. Acad. Sci. Paris Sér. A-B **282** (1976), no. 20, Ai, A1197–A1198.

[3] Brezis, H., Ekeland, I., *Un principe variationnel associé à certaines équations paraboliques. Le cas indépendant du temps*, C. R. Acad. Sci. Paris Sér. A-B **282** (1976), no. 17, Aii, A971–A974.

[4] Carini, L., *BENNO*, https://github.com/LauraCarini/BENNO (2022).

[5] Di Pietro, D.A., Ern, A., *Mathematical aspects of discontinuous Galerkin methods*, Mathématiques & Applications (Berlin), vol. 69, Springer, Heidelberg, 2012.

[6] E, W., Yu, B., *The deep Ritz method: a deep learning-based numerical algorithm for solving variational problems*, Commun. Math. Stat. **6** (2018), no. 1, 1–12.

[7] Ghoussoub, N., *Self-dual partial differential systems and their variational principles*, Springer Monographs in Mathematics, Springer, New York, 2009.

[8] Han, J., E, W., *Deep Learning Approximation for Stochastic Control Problems*, arXiv **cs.LG** (2016).

[9] Han, J., Jentzen, A., E, W., *Solving high-dimensional partial differential equations using deep learning*, Proceedings of the National Academy of Sciences of the United States of America **115** (2018), no. 34, 8505–8510.

[10] Henry-Labordère, P., *Deep Primal-Dual Algorithm for BSDEs: Applications of Machine Learning to CVA and IM*, SSRN Electronic Journal (2017).

[11] Kaltenbach, A., Zeinhofer, M., *The Deep Ritz Method for Parametric p-Dirichlet Problems*, arXiv:2207.01894, 2022.

[12] Karniadakis, G.E., Kevrekidis, I.G., Lu, L., Perdikaris, P., Wang, S., Yang, L., *Physics-informed machine learning*, Nature Reviews Physics **3** (2021), no. 6, 422–440.

[13] Kingma, D.P., Ba, J., *Adam: A Method for Stochastic Optimization*, arXiv:1412.6980, 2014.

[14] Raissi, M., *Forward-Backward Stochastic Neural Networks: Deep Learning of High-dimensional Partial Differential Equations*, arXiv **stat.ML** (2018).

[15] Raissi, M., Karniadakis, G.E., *Hidden physics models: machine learning of nonlinear partial differential equations*, J. Comput. Phys. **357** (2018), 125–141. DOI: 10.1016/j.jcp.2017.11.039

[16] Raissi, M., Perdikaris, P., Karniadakis, G.E., *Physics-informed neural networks: a deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations*, J. Comput. Phys. **378** (2019), 686–707. DOI: 10.1016/j.jcp.2018.10.045

[17] Roubíček, T., *Nonlinear partial differential equations with applications*, second ed., International Series of Numerical Mathematics, vol. 153, Birkhäuser/Springer Basel AG, Basel, 2013.

[18] Sirignano, J., Spiliopoulos, K., *DGM: a deep learning algorithm for solving partial differential equations*, J. Comput. Phys. **375** (2018), 1339–1364. DOI: 10.1016/j.jcp.2018.08.029

[19] Stefanelli, U., *The Brezis-Ekeland principle for doubly nonlinear equations*, SIAM J. Control Optim. **47** (2008), no. 3, 1615–1642.

[20] Stefanelli, U., *The discrete Brezis-Ekeland principle*, J. Convex Anal. **16** (2009), no. 1, 71–87.

[21] Visintin, A., *Extension of the Brezis–Ekeland–Nayroles principle to monotone operators*, Adv. Math. Sci. Appl. **18** (2008), no. 2, 633–650.

DIPARTIMENTO DI MATHEMATICA, UNIVERSITÀ DI TRENTO,
38123 TRENTO, ITALY
E-mail: `laura.carini@studenti.unitn.it`

MATHEMATICS DEPARTMENT, UNIVERSITY COLLEGE LONDON,
25 GORDON STREET, LONDON, WC1H 0AY, UNITED KINGDOM
E-mail: `max.jensen@ucl.ac.uk`

DIPARTIMENTO DI MATHEMATICA, UNIVERSITÀ DI TRENTO,
38123 TRENTO, ITALY
E-mail: `robert.nurnberg@unitn.it`

AN UNCONDITIONALLY STABLE FINITE ELEMENT SCHEME FOR ANISOTROPIC CURVE SHORTENING FLOW

Klaus Deckelnick and Robert Nürnberg

Abstract. Based on a recent novel formulation of parametric anisotropic curve shortening flow, we analyse a fully discrete numerical method of this geometric evolution equation. The method uses piecewise linear finite elements in space and a backward Euler approximation in time. We establish existence and uniqueness of a discrete solution, as well as an unconditional stability property. Some numerical computations confirm the theoretical results and demonstrate the practicality of our method.

1. Introduction

In this paper we study a fully discrete numerical scheme for parametric anisotropic curve shortening flow. This evolution law arises as a natural gradient flow for the energy

$$(1.1) \qquad \mathcal{E}(\Gamma) = \int_\Gamma a(z)\gamma(z,\nu)\, \mathrm{d}\mathcal{H}^1(z) = \int_\Gamma a\, \gamma(\cdot,\nu)\, \mathrm{d}\mathcal{H}^1 \,,$$

where Γ is a closed curve with unit normal ν contained in a given convex domain $\Omega \subset \mathbb{R}^2$. Furthermore, $a \in C^1(\Omega, \mathbb{R}_{>0})$ is a weight function and $\gamma \in C^0(\Omega \times \mathbb{R}^2, \mathbb{R}_{\geq 0}) \cap C^2(\Omega \times (\mathbb{R}^2 \setminus \{0\}), \mathbb{R}_{>0})$ denotes an anisotropy function satisfying

$$(1.2) \qquad \gamma(z, \lambda p) = |\lambda| \gamma(z, p) \quad \text{for all } p \in \mathbb{R}^2,\ \lambda \in \mathbb{R},\ z \in \Omega \,.$$

In addition, we assume that γ is strictly convex in the sense that for every compact $K \subset \Omega$ there exists $c_K > 0$ such that

$$\gamma_{pp}(z,p)q \cdot q \geq c_K |q|^2 \quad \text{for all } z \in K,\ p, q \in \mathbb{R}^2 \text{ with } p \cdot q = 0, |p| = 1.$$

Here, and in what follows, γ_p and γ_{pp} denote gradient and Hessian of the function $p \mapsto \gamma(\cdot, p)$. Anisotropic energies of the form (1.1) play a role in applications, such as materials science, crystal growth, phase transitions and image processing, and in differential geometry, and we refer to e.g. [1, 9, 10, 12, 13, 23, 24, 25, 32] for examples

and further details. It can be shown, see [16, Appendix A], that the first variation of $\mathcal{E}$ in the direction of a smooth vector field V is given by

$$(1.3) \qquad d\mathcal{E}(\Gamma; V) = - \int_\Gamma a\,\gamma(\cdot, \nu)\,\varkappa_\gamma\,V \cdot \nu_\gamma\,\mathrm{d}\mathcal{H}^1\,,$$

where

$$\nu_\gamma = \frac{\nu}{\gamma(\cdot, \nu)} \quad \text{and} \quad \varkappa_\gamma = \varkappa\gamma_{pp}(\cdot, \nu)\tau \cdot \tau - \sum_{i=1}^{2} \gamma_{p_i z_i}(\cdot, \nu) - \frac{\nabla a}{a} \cdot \gamma_p(\cdot, \nu) \qquad \text{on } \Gamma$$

denote the anisotropic normal and the anisotropic curvature of Γ, respectively, with τ and $\varkappa$ the tangent and curvature of Γ.

In view of (1.3), a natural gradient flow for the energy $\mathcal{E}$ evolves a family of closed curves $(\Gamma(t))_{t \in [0,T]}$ according to

$$(1.4) \qquad \mathcal{V}_\gamma = \varkappa_\gamma \qquad \text{on } \Gamma(t)\,,$$

where $\mathcal{V}_\gamma = \frac{1}{\gamma(\cdot, \nu)}\mathcal{V}$ and $\mathcal{V}$ is the normal velocity of $\Gamma(t)$. We remark that solutions of (1.4) satisfy the energy relation

$$\frac{\mathrm{d}}{\mathrm{d}t} \int_{\Gamma(t)} a\,\gamma(\cdot, \nu)\,\mathrm{d}\mathcal{H}^1 + \int_{\Gamma(t)} |\mathcal{V}_\gamma|^2\,a\,\gamma(\cdot, \nu)\,\mathrm{d}\mathcal{H}^1 = 0\,.$$

Note that in the isotropic case, $a(z) = 1$ and $\gamma(z, p) = |p|$, the flow (1.4) collapses to the well–known curve shortening flow $\mathcal{V} = \varkappa$. The isotropic curve shortening flow and its higher dimensional analogue, the mean curvature flow, have been studied extensively both analytically and numerically over the last few decades, and we refer to the works [8, 15, 20, 27] for more details.

In the spatially homogeneous case, $a(z) = 1$ and $\gamma(z, p) = \gamma_0(p)$, the flow (1.4) reduces to the classical anisotropic curve shortening flow

$$(1.5) \qquad \frac{1}{\gamma_0(\nu)}\mathcal{V} = \varkappa_{\gamma_0}\,,$$

where $\varkappa_{\gamma_0} = \varkappa\gamma_0''(\nu)\tau \cdot \tau$ denotes the usual anisotropic curvature. An example for a nonconstant function a is given by the geodesic curvature flow in a Riemannian manifold, see §3.2 and [16, Appendix B] for details.

In this paper we focus on a parametric description of the evolving curves, i.e. $\Gamma(t) = x(I, t)$ with $x \colon I \times [0, T] \ni (\rho, t) \mapsto x(\rho, t) \in \mathbb{R}^2$ and $I = \mathbb{R}/\mathbb{Z}$. Hence the evolution law (1.4) translates into

$$(1.6) \qquad \frac{1}{\gamma(x, \nu)}x_t \cdot \nu = \varkappa_\gamma\,,$$

where $\nu = \tau^\perp = (\frac{x_\rho}{|x_\rho|})^\perp$ and $p^\perp = \binom{p_1}{p_2}^\perp = \binom{-p_2}{p_1}$ denotes an anti-clockwise rotation of p by $\frac{\pi}{2}$. Note that here, and from now on, we think of τ, ν, $\varkappa$ and $\varkappa_\gamma$ as being defined on $I \times [0, T]$.

In order to obtain solutions of (1.6), frequently the partial differential equation (PDE) given by

$$(1.7) \qquad \frac{1}{\gamma(x, \nu)}x_t = \varkappa_\gamma \nu$$

is solved, with the initial condition $x(\cdot, 0) = x_0$, where x_0 is a parameterization of the initial curve Γ_0. Since the right hand side of (1.7) is a geometric invariant, the above PDE appears to be a natural choice. Let us focus for a moment on the isotropic case $a(z) = 1$ and $\gamma(z, p) = |p|$. Then the system (1.7) takes the form

$$(1.8) \qquad x_t = \varkappa\nu = \frac{1}{|x_\rho|}\left(\frac{x_\rho}{|x_\rho|}\right)_\rho = \frac{1}{|x_\rho|^2}\left[x_{\rho\rho} - (x_{\rho\rho} \cdot \tau)\,\tau\right],$$

so that the underlying PDE is only weakly parabolic, causing difficulties for the numerical analysis. We refer to Dziuk's seminal paper [18] for the details. A simple remedy is to apply the so-called DeTurck trick, and to consider, in place of (1.8), the strictly parabolic PDE

$$(1.9) \qquad x_t = \frac{x_{\rho\rho}}{|x_\rho|^2},$$

whose solutions clearly still satisfy $x_t \cdot \nu = \varkappa$. This formulation was proposed and analysed for the first time in [14], see also [21] for a possible generalization.

Extending the DeTurck trick (1.9) for the isotropic flow to the anisotropic evolution equation (1.6) is highly nontrivial. However, the main idea is the same: derive a strictly parabolic PDE whose solutions satisfy (1.6). In this way, a uniquely defined tangential velocity is prescribed together with the normal velocity (1.6), yielding a unique parameterization of the evolving curve. In fact, in the recent paper [16], the authors proved that solutions to the strictly parabolic PDE

$$(1.10) \qquad H(x, x_\rho)x_t = [\Phi_p(x, x_\rho)]_\rho - \Phi_z(x, x_\rho)$$

also satisfy (1.6). Here

$$(1.11) \qquad \Phi(z, p) = \tfrac{1}{2}a^2(z)\gamma^2(z, p^\perp),$$

with Φ_z denoting the gradient of $z \mapsto \Phi(z, \cdot)$, and the matrix

$$H(z, p) = \frac{a^2(z)\gamma(z, p^\perp)}{|\gamma_p(z, p^\perp)|^2}\begin{pmatrix} \gamma(z, p^\perp) & \gamma_p(z, p^\perp) \cdot p \\ -\gamma_p(z, p^\perp) \cdot p & \gamma(z, p^\perp) \end{pmatrix} \quad \forall\, z \in \Omega,\ p \in \mathbb{R}^2 \setminus \{0\}$$

is positive definite in $\Omega \times (\mathbb{R}^2 \setminus \{0\})$ with

$$(1.12) \qquad H(z, p)\xi \cdot \xi = \frac{a^2(z)\gamma^2(z, p^\perp)}{|\gamma_p(z, p^\perp)|^2}|\xi|^2 \quad \forall\, z \in \Omega,\ p \in \mathbb{R}^2 \setminus \{0\},\ \xi \in \mathbb{R}^2\,.$$

The weak formulation of (1.10) is obtained by multiplying it with a test function, integrating over I and performing one integration by parts. It reads as follows. Given $x_0 : I \to \Omega$, find $x : I \times [0, T] \to \Omega$ such that $x(\cdot, 0) = x_0$ and, for $t \in (0, T]$,

$$(1.13)$$
$$\int_I H(x, x_\rho)x_t \cdot \eta \, \mathrm{d}\rho + \int_I \Phi_p(x, x_\rho) \cdot \eta_\rho \, \mathrm{d}\rho + \int_I \Phi_z(x, x_\rho) \cdot \eta \, \mathrm{d}\rho = 0 \quad \forall\, \eta \in [H^1(I)]^2\,.$$

For a continuous-in-time semidiscrete finite element approximation of (1.13) using piecewise linear elements the authors were then able to prove an optimal H^1–error bound, see [16, Theorem 4.1].

In this paper we propose and analyse a fully discrete finite element approximation of (1.13). The scheme, which will be introduced in Section 2, is nonlinear and uses both explicit and implicit approximations in $\Phi_p(x, x_\rho)$ and $\Phi_z(x, x_\rho)$ that are

chosen in such a way as to yield unconditional stability. Here the discrete stability bound will mimic the natural estimate

$$\frac{\mathrm{d}}{\mathrm{d}t} \int_I \Phi(x, x_\rho) \, \mathrm{d}\rho = - \int_I H(x, x_\rho) x_t \cdot x_t \, \mathrm{d}\rho \leq 0 \,,$$

which follows from choosing $\eta = x_t$ in (1.13). Furthermore, we prove the existence of a unique solution under a suitable CFL condition. Then in Section 3 we present some numerical simulations, demonstrating the practicality of the method, as well as the good properties with respect to stability and the distribution of vertices. Let us finally mention that alternative numerical approximations of anisotropic variants of curve shortening flow, which are based on a parametric description of the moving curve, have also been considered in [3, 4, 5, 7, 11, 19, 26, 28, 29, 30, 31].

2. Finite element approximation

Let $[0, 1] = \bigcup_{j=1}^J I_j$, $J \geq 3$, be a decomposition of $[0, 1]$ into the intervals $I_j = [q_{j-1}, q_j]$, where, for simplicity, $q_j = jh$, $j = 0, \ldots, J$, with $h = \frac{1}{J}$. Within I we identify $q_J = 1$ with $q_0 = 0$ and define the finite element space $\underline{V}^h = \{\chi \in C^0(I, \mathbb{R}^2) : \chi|_{I_j} \text{ is affine, } j = 1, \ldots, J\}$. For two piecewise continuous functions, with possible jumps at the nodes $\{q_j\}_{j=1}^J$, we define the mass lumped L^2–inner product

$$(2.1) \qquad (u, v)^h = \tfrac{1}{2} \sum_{j=1}^J h_j \left[(u \cdot v)(q_j^-) + (u \cdot v)(q_{j-1}^+) \right] \,,$$

where $(u \cdot v)(q_j^\pm) = \lim_{\delta \searrow 0} (u \cdot v)(q_j \pm \delta)$. We define the associated norm on $\underline{V}^h$ via $\|u\|_h^2 = (u, u)^h$.

In order to discretize in time, let $t_m = m\Delta t$, $m = 0, \ldots, M$, with the uniform time step $\Delta t = \frac{T}{M} > 0$. On recalling (1.11), we assume that there exists a splitting $\Phi = \Phi^+ + \Phi^-$ such that $\Phi^\pm \in C^1(\Omega \times \mathbb{R}^2)$ and $z \mapsto \pm\Phi^\pm(z, p)$ are convex in Ω for all $p \in \mathbb{R}^2$. Furthermore, we assume that for every compact set $K \subset \Omega$ there exists $L_K \geq 0$ such that

$$(2.2) \quad |\Phi_z^\pm(z, p) - \Phi_z^\pm(z, q)| \leq L_K(|p| + |q|)|p - q| \quad \text{ for all } z \in K, \; p, q \in \mathbb{R}^2 \,.$$

Then our finite element scheme is defined as follows. Given $x_h^m \in \underline{V}^h$ with $\Gamma_h^m := x_h^m(I) \subset \Omega$, for $m = 0, \ldots, M-1$, find $x_h^{m+1} \in \underline{V}^h$ such that $\Gamma_h^{m+1} \subset \Omega$ and

$$\frac{1}{\Delta t} \left(H(x_h^m, x_{h,\rho}^m)(x_h^{m+1} - x_h^m), \eta_h \right)^h + \left(\Phi_p(x_h^m, x_{h,\rho}^{m+1}), \eta_{h,\rho} \right)^h$$

$$(2.3) \qquad + \left(\Phi_z^+(x_h^{m+1}, x_{h,\rho}^{m+1}) + \Phi_z^-(x_h^m, x_{h,\rho}^{m+1}), \eta_h \right)^h = 0 \qquad \forall \, \eta_h \in \underline{V}^h \,.$$

The convex/concave splitting employed for the implicit/explicit approximation of $\Phi_z(\cdot, x_{h,\rho}^{m+1})$ in (2.3), is by now standard practice in the numerical analysis community. This technique goes back to [22], see also [2, 6, 7, 16] for subsequent applications of such splittings. It leads to an unconditionally stable approximation, as we show in our first result.

Theorem 2.1. *Any solution of* (2.3) *satisfies the energy estimate*

$$\left(\Phi(x_h^{m+1}, x_{h,\rho}^{m+1}), 1\right)^h + \frac{1}{\Delta t}\left(H(x_h^m, x_{h,\rho}^m)(x_h^{m+1} - x_h^m), x_h^{m+1} - x_h^m\right)^h$$

(2.4)
$$\leq \left(\Phi(x_h^m, x_{h,\rho}^m), 1\right)^h.$$

Proof. From the convexity properties of Φ and $\pm\Phi^{\pm}$ we infer that

$$\left(\Phi_p(x_h^m, x_{h,\rho}^m + \eta_{h,\rho}), \eta_{h,\rho}\right)^h \geq \left(\Phi(x_h^m, x_{h,\rho}^m + \eta_{h,\rho}), 1\right)^h - \left(\Phi(x_h^m, x_{h,\rho}^m), 1\right)^h,$$

$$\left(\Phi_z^+(x_h^m + \eta_h, x_{h,\rho}^m + \eta_{h,\rho}), \eta_h\right)^h \geq \left(\Phi^+(x_h^m + \eta_h, x_{h,\rho}^m + \eta_{h,\rho}), 1\right)^h$$
$$- \left(\Phi^+(x_h^m, x_{h,\rho}^m + \eta_{h,\rho}), 1\right)^h,$$

$$\left(\Phi_z^-(x_h^m, x_{h,\rho}^m + \eta_{h,\rho}), \eta_h\right)^h \geq \left(\Phi^-(x_h^m + \eta_h, x_{h,\rho}^m + \eta_{h,\rho}), 1\right)^h$$

(2.5)
$$- \left(\Phi^-(x_h^m, x_{h,\rho}^m + \eta_{h,\rho}), 1\right)^h,$$

for all $\eta_h \in \underline{V}^h$. Choosing $\eta_h = x_h^{m+1} - x_h^m$ in (2.3) and applying (2.5) yields the bound (2.4). $\square$

We note that the fully discrete finite element approximation (2.3) can be seen as a generalization of two fully discrete schemes introduced by the authors in [16]. In particular, in the special case of a spatially homogeneous anisotropy, recall (1.5), the scheme (2.3) reduces to [16, (5.4)]. Similarly, in the case when (1.6) models geodesic flow in a Riemannian manifold, the approximation [16, (5.10)] is a special case of the scheme (2.3). Moreover, we remark that the nonlinear systems of equations arising from (2.3) can be solved with a Newton method or with a Picard-type iteration. In our experience, in general, in practice these solution methods converge within a few iterations.

Let us next address the existence and uniqueness for the nonlinear system (2.3). We assume that $x_h^m \in \underline{V}^h$ is given with $x_{h,\rho}^m \neq 0$ in I and $\Gamma_h^m \subset \Omega$. There exists $R > 0$ such that $K_0 := \{z \in \mathbb{R}^2 : \text{dist}(z, \Gamma_h^m) \leq R\} \subset \Omega$. Before we present our main theorem, we collect the following auxiliary results.

Lemma 2.2. *There exists a constant $C_0 > 0$ depending on x_h^m such that*

(2.6a)
$$\Phi(z, p) \geq C_0|p|^2 \qquad \forall\, p \in \mathbb{R}^2, z \in K_0,$$

(2.6b)
$$\left(\Phi_p(z, q) - \Phi_p(z, p)\right) \cdot (q - p) \geq C_0|q - p|^2 \qquad \forall\, p, q \in \mathbb{R}^2, z \in K_0,$$

(2.6c)
$$H(x_h^m, x_{h,\rho}^m)\xi \cdot \xi \geq C_0|\xi|^2 \qquad \forall\, \xi \in \mathbb{R}^2 \quad in\ I.$$

Proof. The bound (2.6a) follows from (1.2) and $\Phi(z, p) > 0$ for $z \in K_0$ and $|p| = 1$. Similarly, we have from (1.12) and $\min_I |x_{h,\rho}^m| > 0$ that (2.6c) holds.

It remains to show (2.6b). Since $a \geq a_0 > 0$ in K_0, it is sufficient to carry out the proof for $\Phi(z, p) = \frac{1}{2}\gamma^2(z, p)$. Note that in this case $\Phi \in C^1(\Omega \times \mathbb{R}^2, \mathbb{R}_{\geq 0}) \cap C^2(\Omega \times (\mathbb{R}^2 \setminus \{0\}), \mathbb{R}_{>0})$. Furthermore, according to [24, Remark 1.7.5], $\gamma_{pp}^2(z, p) := [\gamma^2]_{pp}(z, p)$ is positive definite for $p \neq 0$. In particular, there exists $c_0 > 0$ such that

$$\gamma_{pp}^2(z, p)q \cdot q \geq c_0|q|^2 \qquad \text{for all } z \in K_0,\ p, q \in \mathbb{R}^2, |p| = 1.$$

Observing that $\Phi_{pp}(z,p) = A\gamma_{pp}^2(z,p^\perp)A^T$ with $A = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$, we infer that

$$(2.7) \qquad \Phi_{pp}(z,p)q \cdot q = \gamma_{pp}^2(z,p^\perp)A^Tq \cdot A^Tq \geq c_0 |A^Tq|^2 = c_0|q|^2$$

for all $z \in K_0$, p, $q \in \mathbb{R}^2$, $|p| = 1$. Let us fix $z \in K_0$ and $p, q \in \mathbb{R}^2$. We distinguish two cases:

Case 1: $sq + (1-s)p \neq 0$ for all $s \in [0,1]$. Then

$$\big(\Phi_p(z,q) - \Phi_p(z,p)\big) \cdot (q-p) = \int_0^1 \Phi_{pp}(z, sq+(1-s)p)(q-p) \cdot (q-p) \, \mathrm{d}s \geq c_0|q-p|^2,$$

using the fact that $p \mapsto \Phi_{pp}(z,p)$ is 0-homogeneous and (2.7).

Case 2: There exists $s \in [0,1]$ with $sq + (1-s)p = 0$. We may assume that $s \in [0,1)$, so that $p = -\frac{s}{1-s}q$. As $\Phi(z,\lambda q) = \lambda^2 \Phi(z,q)$ for $\lambda \in \mathbb{R}$, recall (1.2), we have that $\Phi_p(z,\lambda q) = \lambda \Phi_p(z,q)$ and $\Phi_p(z,q) \cdot q = 2\Phi(z,q)$. Hence we obtain that

$$\begin{aligned}
\big(\Phi_p(z,q) - \Phi_p(z,p)\big) \cdot (q-p) &= \big(\Phi_p(z,q) - \Phi_p(z, -\tfrac{s}{1-s}q)\big) \cdot \big(q + \tfrac{s}{1-s}q\big) \\
&= \big(1 + \tfrac{s}{1-s}\big)^2 \Phi_p(z,q) \cdot q = 2\big(1 + \tfrac{s}{1-s}\big)^2 \Phi(z,q) \\
&\geq 2C_0\big(1 + \tfrac{s}{1-s}\big)^2 |q|^2 = 2C_0|q-p|^2,
\end{aligned}$$

on noting (2.6a). $\square$

We are now in a position to prove our main result.

Theorem 2.3. *There exists $\delta > 0$ such that for $\Delta t \leq \delta h$ there is a unique element $x_h^{m+1} \in \underline{V}^h$ with $\Gamma_h^{m+1} \subset K_0$ which solves (2.3). The constant δ depends on R, C_0, L_{K_0} and $\Phi(x_h^m, x_{h,\rho}^m)$.*

Proof. We denote by $\{\varphi_j\}_{j=1}^{2J}$ the basis of $\underline{V}^h$ satisfying $\varphi_j(q_k) = \delta_{jk}e_1$ and $\varphi_{j+J}(q_k) = \delta_{jk}e_2$ for $1 \leq j, k \leq J$, and associate with every $\alpha \in \mathbb{R}^{2J}$ the element $v_\alpha = \sum_{j=1}^{2J} \alpha_j \varphi_j \in \underline{V}^h$, so that $v_\alpha(q_j) = \begin{pmatrix} \alpha_j \\ \alpha_{j+J} \end{pmatrix}$. We have for $|\alpha| \leq R$ and $\rho \in I$ that

$$\mathrm{dist}((x_h^m + v_\alpha)(\rho), \Gamma_h^m) \leq \|v_\alpha\|_\infty \leq |\alpha| \leq R,$$

so that $(x_h^m + v_\alpha)(\rho) \in K_0$. Let us next define the continuous map $F \colon \overline{B_R(0)} \times [0,1] \to \mathbb{R}^{2J}$ via

$$\begin{aligned}
[F(\alpha,\lambda)]_i = {}& \frac{1}{\Delta t}\big(H(x_h^m, x_{h,\rho}^m)v_\alpha, \varphi_i\big)^h + \lambda\big(\Phi_p(x_h^m, x_{h,\rho}^m + v_{\alpha,\rho}), \varphi_{i,\rho}\big)^h \\
&+ \lambda\big(\Phi_z^+(x_h^m + v_\alpha, x_{h,\rho}^m + v_{\alpha,\rho}) + \Phi_z^-(x_h^m, x_{h,\rho}^m + v_{\alpha,\rho}), \varphi_i\big)^h.
\end{aligned}$$

In what follows we make use of standard results for the Brouwer degree $d(f, B_R(0), 0)$ of a continuous function $f \colon \overline{B_R(0)} \to \mathbb{R}^{2J}$, if $0 \notin f(\partial B_R(0))$, see [17, Chapter 1]. Clearly, the mapping $F(\cdot, 0) =: A$ is linear with

$$[A\alpha]_i = \frac{1}{\Delta t}\sum_{j=1}^{2J} \alpha_j \big(H(x_h^m, x_{h,\rho}^m)\varphi_j, \varphi_i\big)^h, \quad i = 1, \ldots, 2J,$$

and invertible in view of (2.6c). Hence it follows from [17, Theorem 1.1] that

$$(2.8) \qquad d(F(\cdot, 0), B_R(0), 0) = d(A, B_R(0), 0) = \mathrm{sgn}\det A \neq 0.$$

Next, it holds for $\lambda \in [0, 1]$ that

$$F(\alpha, \lambda) \cdot \alpha = \frac{1}{\Delta t} \left(H(x_h^m, x_{h,\rho}^m) v_\alpha, v_\alpha \right)^h + \lambda \left(\Phi_p(x_h^m, x_{h,\rho}^m + v_{\alpha,\rho}), v_{\alpha,\rho} \right)^h$$

$$\text{(2.9)} \qquad + \lambda \left(\Phi_z^+(x_h^m + v_\alpha, x_{h,\rho}^m + v_{\alpha,\rho}) + \Phi_z^-(x_h^m, x_{h,\rho}^m + v_{\alpha,\rho}), v_\alpha \right)^h .$$

Inserting (2.5) with $\eta_h = v_\alpha$ into (2.9) yields

$$F(\alpha, \lambda) \cdot \alpha \geq \frac{1}{\Delta t} \left(H(x_h^m, x_{h,\rho}^m) v_\alpha, v_\alpha \right)^h$$

$$+ \lambda \left(\Phi(x_h^m + v_\alpha, x_{h,\rho}^m + v_{\alpha,\rho}), 1 \right)^h - \lambda \left(\Phi(x_h^m, x_{h,\rho}^m), 1 \right)^h .$$

If we combine this estimate with (2.6c) we finally obtain for $|\alpha| = R$ that

$$F(\alpha, \lambda) \cdot \alpha \geq \frac{C_0}{\Delta t} \|v_\alpha\|_h^2 - \left(\Phi(x_h^m, x_{h,\rho}^m), 1 \right)^h$$

$$= \frac{C_0}{\Delta t} h |\alpha|^2 - \left(\Phi(x_h^m, x_{h,\rho}^m), 1 \right)^h \geq \frac{C_0 R^2}{\delta} - \left(\Phi(x_h^m, x_{h,\rho}^m), 1 \right)^h$$

using (2.1) and the relation $\Delta t \leq \delta h$. By choosing δ sufficiently small we obtain $F(\alpha, \lambda) \cdot \alpha > 0$ and therefore $F(\alpha, \lambda) \neq 0$ for all $(\alpha, \lambda) \in \partial B_R(0) \times [0, 1]$. Using the homotopy invariance of the Brouwer degree together with (2.8) we deduce that $d(F(\cdot, 1), B_R(0), 0) = d(F(\cdot, 0), B_R(0), 0) \neq 0$, so that the existence property of the degree shows that there is $\alpha \in B_R(0)$ such that $F(\alpha, 1) = 0$. Clearly, $x_h^{m+1} := x_h^m + v_\alpha$ is then a solution of (2.3).

In order to prove uniqueness, suppose that $x_h^{m+1}, \tilde{x}_h^{m+1} \in \underline{V}^h$ are two solutions of (2.3) with $\Gamma_h^{m+1}, \tilde{\Gamma}_h^{m+1} \subset K_0$. To begin, we infer from (2.4) and (2.6a) that

$$\text{(2.10)} \qquad C_0 \|x_{h,\rho}^{m+1}\|_h^2 \leq \left(\Phi(x_h^m, x_{h,\rho}^m), 1 \right)^h, \quad C_0 \|\tilde{x}_{h,\rho}^{m+1}\|_h^2 \leq \left(\Phi(x_h^m, x_{h,\rho}^m), 1 \right)^h .$$

We have that

$$\frac{1}{\Delta t} \left(H(x_h^m, x_{h,\rho}^m)(x_h^{m+1} - \tilde{x}_h^{m+1}), \eta_h \right)^h + \left(\Phi_p(x_h^m, x_{h,\rho}^{m+1}) - \Phi_p(x_h^m, \tilde{x}_{h,\rho}^{m+1}), \eta_{h,\rho} \right)^h$$

$$+ \left(\Phi_z^+(x_h^{m+1}, x_{h,\rho}^{m+1}) - \Phi_z^+(\tilde{x}_h^{m+1}, x_{h,\rho}^{m+1}), \eta_h \right)^h$$

$$= \left(\Phi_z^+(\tilde{x}_h^{m+1}, \tilde{x}_{h,\rho}^{m+1}) - \Phi_z^+(\tilde{x}_h^{m+1}, x_{h,\rho}^{m+1}), \eta_h \right)^h$$

$$+ \left(\Phi_z^-(x_h^m, \tilde{x}_{h,\rho}^{m+1}) - \Phi_z^-(x_h^m, x_{h,\rho}^{m+1}), \eta_h \right)^h$$

for all $\eta_h \in \underline{V}^h$. Choosing $\eta = x_h^{m+1} - \tilde{x}_h^{m+1}$ we deduce with the help of (2.6c), (2.6b), (2.2), (2.10) and the convexity of $z \mapsto \Phi^+(z, p)$ that

$$\frac{C_0}{\Delta t} \|x_h^{m+1} - \tilde{x}_h^{m+1}\|_h^2 + C_0 \|x_{h,\rho}^{m+1} - \tilde{x}_{h,\rho}^{m+1}\|_h^2$$

$$\leq 2L \left(\|x_{h,\rho}^{m+1}\|_h + \|\tilde{x}_{h,\rho}^{m+1}\|_h \right) \|x_{h,\rho}^{m+1} - \tilde{x}_{h,\rho}^{m+1}\|_h \|x_h^{m+1} - \tilde{x}_h^{m+1}\|_\infty$$

$$\leq 4L C_0^{-\frac{1}{2}} h^{-\frac{1}{2}} \sqrt{\left(\Phi(x_h^m, x_{h,\rho}^m), 1 \right)^h} \|x_{h,\rho}^{m+1} - \tilde{x}_{h,\rho}^{m+1}\|_h \|x_h^{m+1} - \tilde{x}_h^{m+1}\|_h$$

$$\leq C_0 \|x_{h,\rho}^{m+1} - \tilde{x}_{h,\rho}^{m+1}\|_h^2 + 4L^2 \left(\Phi(x_h^m, x_{h,\rho}^m), 1 \right)^h C_0^{-2} h^{-1} \|x_h^{m+1} - \tilde{x}_h^{m+1}\|_h^2 ,$$

with $L = L_{K_0}$. By choosing $\delta > 0$ so small that $4L^2\big(\Phi(x_h^m, x_{h,\rho}^m), 1\big)^h \delta < C_0^3$ we deduce that $x_h^{m+1} = \tilde{x}_h^{m+1}$. $\square$

3. NUMERICAL RESULTS

For all our numerical simulations we use $J = 256$ and $\Delta t = 10^{-4}$. On recalling (1.1), for $\chi_h \in \underline{V}^h$ we define the discrete energy

$$\mathcal{E}^h(\chi_h) = \big(\gamma(\chi_h, \chi_{h,\rho}^\perp), a(\chi_h)\big)^h .$$

We also consider the ratio

$$\mathfrak{r}^m = \frac{\max_{j=1,\ldots,J} |x_h^m(q_j) - x_h^m(q_{j-1})|}{\min_{j=1,\ldots,J} |x_h^m(q_j) - x_h^m(q_{j-1})|}$$

between the longest and shortest element of $\Gamma_h^m = x_h^m(I)$, and are often interested in the evolution of this ratio over time. We stress that no redistribution of vertices was necessary during any of our numerical simulations. We remark that a convergence experiment for (2.3), for an anisotropy of the form $\gamma(z, p) = \sqrt{p_1^2 + \delta^2 p_2^2}$ with $\delta > 0$, which confirms the theoretically obtained optimal H^1–error bound, can be found in [16, §6.1].

3.1. **The spatially homogeneous case.** In the case that

$$\gamma(z, p) = \gamma_0(p) \quad \text{and} \quad a(z) = 1 \qquad \forall\, z \in \Omega = \mathbb{R}^2 ,$$

for an anisotropy function $\gamma_0 \in C^2(\mathbb{R}^2 \setminus \{0\}, \mathbb{R}_{>0})$, the flow (1.6) reduces to classical anisotropic curvature flow, (1.5). Most existing approaches for the numerical approximation of anisotropic curve shortening flow deal with this simpler case, see e.g. [3, 5, 19, 31].

For our first experiment we choose the anisotropy from [19, (7.1)], with

$$(3.1) \quad \gamma_0(p) = |p|(1 + \delta \cos(k\theta(p))), \quad p = |p|\begin{pmatrix} \cos\theta(p) \\ \sin\theta(p) \end{pmatrix}, \quad k = 6, \ \delta = 0.028 ,$$

and as initial curve use the one given in [28, p. 1494], i.e. we let

$$(3.2) \quad x(\rho, 0) = \begin{pmatrix} \cos u(\rho) \\ \frac{1}{2}\sin u(\rho) + \sin(\cos u(\rho)) + \sin u(\rho)\,[\frac{1}{5} + \sin u(\rho)\,\sin^2 u(3\rho)] \end{pmatrix} ,$$

where $u(\rho) = 2\pi\rho$. The evolution is shown in Figure 1, where we can observe that the shrinking curve becomes convex, with its form soon approaching a scaled Wulff shape of the six-fold symmetric anisotropy (3.1). In addition, we once again note that the discrete energy $\mathcal{E}^h$ is monotonically decreasing, while the tangential motion induced by (1.10) leads to only a moderate initial increase in $\mathfrak{r}^m$, before it decreases towards the end.

With our next experiment we would like to demonstrate that our scheme can easily be extended to situations where a forcing term appears in the flow, e.g.

$$(3.3) \qquad\qquad \mathcal{V}_\gamma = \varkappa_\gamma + f(\cdot, \nu) \quad \text{on } \Gamma(t) ,$$

in place of (1.4), where $f : \Omega \times \mathbb{R}^2 \to \mathbb{R}$. This leads to the additional term $\int_I f(x, \nu) H(x, x_\rho) \nu \cdot \eta \, d\rho$ on the right hand side of (1.13), and analogously to
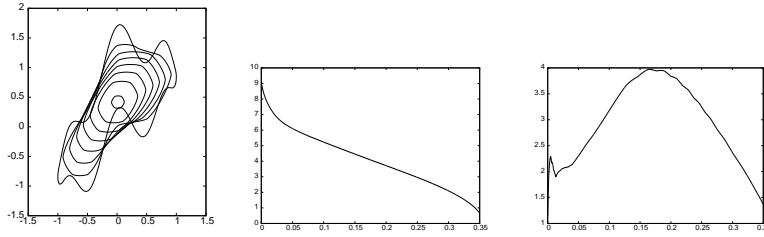
FIG. 1: Solution at times $t = 0, 0.05, \ldots, 0.35$. We also show
a plot of the discrete energy $\mathcal{E}^h(x_h^m)$ (middle) and of the ratio $\mathfrak{r}^m$
over time (right).

the additional term $(f(x_h^m, \frac{(x_{h,\rho}^m)^\perp}{|x_{h,\rho}^m|})H(x_\rho^m, x_{h,\rho}^m)\frac{(x_{h,\rho}^m)^\perp}{|x_{h,\rho}^m|}, \eta_h)^h$ on the right hand side
of (2.3). In our numerical experiments we choose

$$f(z, \nu) = f_0 \in \mathbb{R},$$

so that (3.3) overall reduces to $\frac{1}{\gamma_0(\nu)}\mathcal{V} = \varkappa_{\gamma_0} + f_0$, compare with (1.5). Starting
this flow from the same initial data (3.2), but now with the forcing $f_0 = 1.15$,
leads to an expanding curve that, upon an appropriate rescaling, approaches the
boundary of the Wulff shape, see Figure 2. What is particularly interesting in the
observed evolution is that the ratio $\mathfrak{r}^m$ appears to tend towards unity, indicating
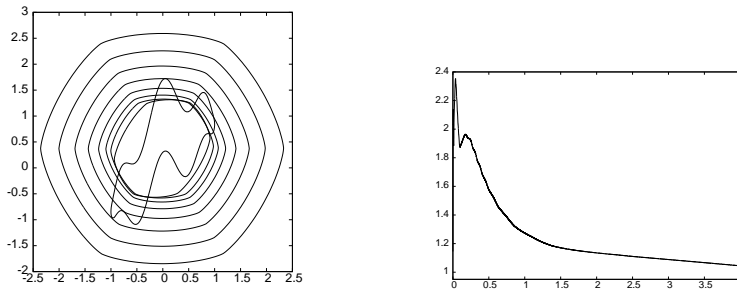an asymptotic equidistribution of the vertices on the polygonal curve.

FIG. 2: Solution at times $t = 0, 0.5, \ldots, 4$. We also show a plot of
the ratio $\mathfrak{r}^m$ over time.

3.2. **Geodesic curvature flow in Riemannian manifolds.** Let $(\mathcal{M}, g)$ be a
two-dimensional Riemannian manifold, with local parameterization $F : \Omega \to \mathcal{M}$
and corresponding basis $\{\partial_1, \partial_2\}$ of the tangent space. We define

$$\gamma(z, p) = \sqrt{G^{-1}(z)p \cdot p} \quad \text{and} \quad a(z) = \sqrt{\det G(z)}.$$

Here, $G(z) = (g_{ij}(z))^2_{i,j=1}$, with $g_{ij}(z) = g_{F(z)}(\partial_i, \partial_j)$, $z \in \Omega$. Then (1.1) reduces to $\mathcal{E}(\Gamma) = \int_\Gamma \sqrt{G\tau \cdot \tau} \, d\mathcal{H}^1$, the Riemannian length of the curve $\tilde{\Gamma} = F(\Gamma) \subset \mathcal{M}$, and it can be shown that (1.4) is now equivalent to geodesic curvature flow in $\mathcal{M}$, see [16, Appendix B] for details. Furthermore, in [16, §5.2] a possible construction of the splitting $\Phi_z^\pm$ is given, with the help of which the scheme (2.3) reduces to (5.10) in [16].

As an example we consider the case when $(\mathcal{M}, g)$ is a hypersurface in the Euclidean space $\mathbb{R}^3$. Assuming that $\mathcal{M}$ can be written as a graph, we let

$$F(z) = (z_1, z_2, \varphi(z))^T, \quad \varphi \in C^3(\Omega).$$

The induced matrix G is then given by $G(z) = \text{Id} + \nabla\varphi(z) \otimes \nabla\varphi(z)$, and we have that $\Phi(z, p) = \frac{1}{2}G(z)p \cdot p$. For the splitting $\Phi = \Phi^+ + \Phi^-$ it is natural to let $\Phi^+(z, p) = \frac{1}{2}G_+(z)p \cdot p$, where $G_+(z) = G(z) + c_\varphi |z|^2 \text{Id}$ and $c_\varphi \in \mathbb{R}_{\geq 0}$ is chosen sufficiently large. In our computation we observe a monotonically decreasing discrete energy when choosing $c_\varphi = 0$, and so we let $\Phi^+ = \Phi$.
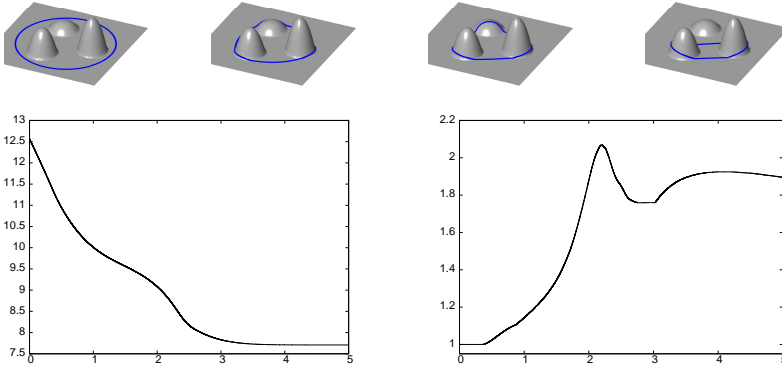

FIG. 3: Geodesic curvature flow on the graph defined by (3.4). We show the evolution of $F(x_h^m)$ on $\mathcal{M}$ at times $t = 0, 1, 2, 4$. Below we show a plot of the discrete energy $\mathcal{E}^h(x_h^m)$ (left) and of the ratio $\mathfrak{r}^m$ over time (right).

For our numerical simulation, following [33], we define a surface with three "mountains" via

$$(3.4) \quad \varphi(z) = \sum_{i=1}^3 \lambda_i \psi(2|z - \mu_i|^2) \text{ with } \Omega = \mathbb{R}^2, \quad \text{where} \quad \psi(s) = \begin{cases} e^{-\frac{1}{1-s}} & s < 1, \\ 0 & s \geq 1, \end{cases}$$

and where $\mu_1 = 0$, $\mu_2 = \binom{2}{0}$, $\mu_3 = \binom{1}{\sqrt{3}}$ and $(\lambda_1, \lambda_2, \lambda_3) = (1, 3, 4)$. On letting the initial polygonal curve be defined by an equidistributed approximation of a circle of radius 2 in Ω, centred at $\frac{1}{3}\sum_{i=1}^3 \mu_i$, we show the evolution for geodesic curvature flow on the defined hypersurface in Figure 3. During the flow the curve tries to decrease its (Euclidean) length, while remaining on the manifold. As the initial

circle begins to shrink, the curve is able to pass over the smallest of the three "mountains", but then reaches a steady state solution encompassing the two taller mountains. Here the curve cannot reduce its length further, because to rise higher up would yield an increase in its length, since it needs to remain attached to the flat part of the surface between the two mountains. We note that as soon as one of the larger "mountains" is not enclosed by the initial curve, then the evolution is going to lead to extinction in finite time. We show this in Figure 4, where the initial circle is now centred at $\left(\begin{smallmatrix}0\\\frac{1}{2}\end{smallmatrix}\right)$. Here the curve can continually decrease its length, until it reaches the peak of the tallest "mountain", where it shrinks to a point.

FIG. 4: Geodesic curvature flow on the graph defined by (3.4). We show the evolution of $F(x_h^m)$ on $\mathcal{M}$ at times $t = 0, 1, 2, 4$.

References

[1] Alfaro, M., Garcke, H., Hilhorst, D., Matano, H., Schätzle, R., *Motion by anisotropic mean curvature as sharp interface limit of an inhomogeneous and anisotropic Allen-Cahn equation*, Proc. Roy. Soc. Edinburgh Sect. A **140** (4) (2010), 673–706.

[2] Barrett, J.W., Blowey, J.F., *Finite element approximation of a model for phase separation of a multi-component alloy with non-smooth free energy*, Numer. Math. **77** (1) (1997), 1–34.

[3] Barrett, J.W., Garcke, H., Nürnberg, R., *Numerical approximation of anisotropic geometric evolution equations in the plane*, IMA J. Numer. Anal. **28** (2) (2008), 292–330.

[4] Barrett, J.W., Garcke, H., Nürnberg, R., *Numerical approximation of gradient flows for closed curves in $\mathbb{R}^d$*, IMA J. Numer. Anal. **30** (1) (2010), 4–60.

[5] Barrett, J.W., Garcke, H., Nürnberg, R., *The approximation of planar curve evolutions by stable fully implicit finite element schemes that equidistribute*, Numer. Methods Partial Differential Equations **27** (1) (2011), 1–30.

[6] Barrett, J.W., Garcke, H., Nürnberg, R., *Stable phase field approximations of anisotropic solidification*, IMA J. Numer. Anal. **34** (4) (2014), 1289–1327.

[7] Barrett, J.W., Garcke, H., Nürnberg, R., *Numerical approximation of curve evolutions in Riemannian manifolds*, IMA J. Numer. Anal. **40** (3) (2020), 1601–1651.

[8] Barrett, J.W., Garcke, H., Nürnberg, R., *Parametric finite element approximations of curvature driven interface evolutions*, Handb. Numer. Anal. (Bonito, A., Nochetto, R.H., eds.), vol. 21, Elsevier, Amsterdam, 2020, pp. 275–423.

[9] Bellettini, G., *Anisotropic and crystalline mean curvature flow*, A sampler of Riemann-Finsler geometry, Math. Sci. Res. Inst. Publ., vol. 50, Cambridge Univ. Press, 2004, pp. 49–82.

[10] Bellettini, G., Paolini, M., *Anisotropic motion by mean curvature in the context of Finsler geometry*, Hokkaido Math. J. **25** (3) (1996), 537–566.

[11] Beneš, M., Mikula, K., *Simulation of anisotropic motion by mean curvature-comparison of phase field and sharp interface approaches*, Acta Math. Univ. Comenian. (N.S.) **67** (1) (1998), 17–42.

[12] Caselles, V., Kimmel, R., Sapiro, G., *Geodesic active contours*, Int. J. Comput. Vis. **22** (1) (1997), 61–79.

[13] Clarenz, U., Dziuk, G., Rumpf, M., *On generalized mean curvature flow in surface processing*, Geometric Analysis and Nonlinear Partial Differential Equations, Springer-Verlag, Berlin, 2003, pp. 217–248.

[14] Deckelnick, K., Dziuk, G., *On the approximation of the curve shortening flow*, Calculus of Variations, Applications and Computations (Pont-à-Mousson, 1994) (Bandle, C., Bemelmans, J., Chipot, M., Paulin, J.S.J., Shafrir, I., eds.), Pitman Res. Notes Math. Ser., Longman Sci. Tech., Harlow, 1995, pp. 100–108.

[15] Deckelnick, K., Dziuk, G., Elliott, C.M., *Computation of geometric partial differential equations and mean curvature flow*, Acta Numer. **14** (2005), 139–232.

[16] Deckelnick, K., Nürnberg, R., *A novel finite element approximation of anisotropic curve shortening flow*, arXiv:2110.04605 (2021).

[17] Deimling, K., *Nonlinear functional analysis*, Springer-Verlag, Berlin, 1985.

[18] Dziuk, G., *Convergence of a semi-discrete scheme for the curve shortening flow*, Math. Models Methods Appl. Sci. **4** (4) (1994), 589–606.

[19] Dziuk, G., *Discrete anisotropic curve shortening flow*, SIAM J. Numer. Anal. **36** (6) (1999), 1808–1830.

[20] Ecker, K., *Regularity Theory for Mean Curvature Flow*, Birkhäuser, Boston, 2004.

[21] Elliott, C.M., Fritz, H., *On approximations of the curve shortening flow and of the mean curvature flow based on the DeTurck trick*, IMA J. Numer. Anal. **37** (2) (2017), 543–603.

[22] Elliott, C.M., Stuart, A.M., *The global dynamics of discrete semilinear parabolic equations*, SIAM J. Numer. Anal. **30** (6) (1993), 1622–1663.

[23] Garcke, H., Lam, K.F., Nürnberg, R., Signori, A., *Overhang penalization in additive manufacturing via phase field structural topology optimization with anisotropic energies*, arXiv:2111.14070 (2021).

[24] Giga, Y., *Surface evolution equations*, vol. 99, Birkhäuser, Basel, 2006.

[25] Gurtin, M.E., *Thermomechanics of Evolving Phase Boundaries in the Plane*, Oxford Mathematical Monographs, The Clarendon Press Oxford University Press, New York, 1993.

[26] Haußer, F., Voigt, A., *A numerical scheme for regularized anisotropic curve shortening flow*, Appl. Math. Lett. **19** (8) (2006), 691–698.

[27] Mantegazza, C., *Lecture notes on mean curvature flow*, Progress in Mathematics, vol. 290, Birkhäuser/Springer Basel AG, Basel, 2011.

[28] Mikula, K., Ševčovič, D., *Evolution of plane curves driven by a nonlinear function of curvature and anisotropy*, SIAM J. Appl. Math. **61** (5) (2001), 1473–1501.

[29] Mikula, K., Ševčovič, D., *A direct method for solving an anisotropic mean curvature flow of plane curves with an external force*, Math. Methods Appl. Sci. **27** (13) (2004), 1545–1565.

[30] Mikula, K., Ševčovič, D., *Computational and qualitative aspects of evolution of curves driven by curvature and external force*, Computing Vis. Sci. **6** (4) (2004), 21–225.

[31] Pozzi, P., *Anisotropic curve shortening flow in higher codimension*, Math. Methods Appl. Sci. **30** (11) (2007), 1243–1281.

[32] Taylor, J.E., Cahn, J.W., Handwerker, C.A., *Geometric models of crystal growth*, Acta Metall. Mater. **40** (7) (1992), 1443–1474.

[33] Wu, C., Tai, X., *A level set formulation of geodesic curvature flow on simplicial surfaces*, IEEE Trans. Vis. Comput. Graph. **16** (4) (2010), 647–662.

Institut für Analysis und Numerik,
Otto-von-Guericke-Universität Magdeburg,
39106 Magdeburg, Germany
E-mail: `klaus.deckelnick@ovgu.de`

Dipartimento di Mathematica, Università di Trento,
38123 Trento, Italy
E-mail: `robert.nurnberg@unitn.it`

STRESS-CONTROLLED HYSTERESIS AND LONG-TIME DYNAMICS OF IMPLICIT DIFFERENTIAL EQUATIONS ARISING IN HYPOPLASTICITY

Victor A. Kovtunenko, Ján Eliaš, Pavel Krejčí, Giselle A. Monteiro, and Judita Runcziková

Abstract. A long-time dynamic for granular materials arising in the hypo-plastic theory of Kolymbas type is investigated. It is assumed that the granular hardness allows exponential degradation, which leads to the densification of material states. The governing system for a rate-independent strain under stress control is described by implicit differential equations. Its analytical solution for arbitrary inhomogeneous coefficients is constructed in closed form. Under cyclic loading by periodic pressure, finite ratcheting for the void ratio is derived in explicit form, which converges to a limiting periodic process (attractor) when the number of cycles tends to infinity.

1. Introduction

In this paper we study long-time dynamics of the constitutive stress-strain relation for granular materials like cohesionless soils or broken rocks. The constitutive law is based on the hypoplastic concept proposed by Kolymbas [17], the model is of the rate type and incrementally nonlinear. Compared to hyper- and hypoelastic material laws, the hypoplastic responses are different for loading and unloading, that is typical for inelastic materials. In contrast to the classical elastoplastic concept, the strain is not decomposed into elastic and plastic parts. Physical aspects of hypoplastic models can be found in [26, 27]. For other representatives of incrementally nonlinear constitutive equations, see the models by Armstrong–Frederick [2],

endochronic [30], octolinear [10], and CLoE [9]. For mathematical modelling granular and multiphase media we cite [1, 11, 12, 13, 22, 23, 24], while for well-posedness analysis we refer to [8, 16, 25].

In an earlier work [7], we have considered the strain-stress law as a nonlinear differential equation for the stress under a given proportional strain (the strain-control). The model therein is a simplified version of the hypoplastic model by Bauer [3] and Gudehus [14], in which the pressure and density dependent properties of granular materials were omitted. In this way we also make a close link to barodesy models [18]. The existence of an exact solution made it possible to describe analytically various scenarios of the behavior of stress paths obtained from monotonic compression, extension, and isochoric deformations [7, 19, 20, 21].

Our ultimate goal in the current work is to study the phenomenon of ratcheting, that is, the shift of the hysteresis loops under periodic loading and unloading cycles. The theoretical ratcheting is infinite when the influence of the void ratio of the granular material is neglected, which is not consistent with experimental observations. In the present paper, we consider a stress-controlled constitutive law within the hypoplastic theory, which allows degradation of the granular hardness. This results in inhomogeneous material parameters, and the granular body becomes asymptotically rigid when the number of cycles tends to infinity. Since we are in the stress-controlled case, the strain rate is the unknown of the problem and has to be found as a solution of an implicit system of differential equations. Like in the strain-controlled case, the solution is found in closed form. Moreover, we prove that the void ratio subjected to periodic loading-unloading pressure cycles converges to an equilibrium, independently of whether the proportional stress paths are isotropic or not.

Within the nonlinear theory of rate-independent materials we consider a constitutive response between the Cauchy stress $\boldsymbol{\sigma}$, linearized strain $\boldsymbol{\varepsilon}$, and its rates $\dot{\boldsymbol{\sigma}}$ and $\dot{\boldsymbol{\varepsilon}}$, which is expressed by an implicit function [15, 28]:

$$\mathfrak{f}(\boldsymbol{\sigma}, \dot{\boldsymbol{\sigma}}, \boldsymbol{\varepsilon}, \dot{\boldsymbol{\varepsilon}}) = 0 \,.$$

For the function $\mathfrak{f}$ positively homogeneous of degree one with respect to rates:

$$(1.1) \qquad \mathfrak{f}(\boldsymbol{\sigma}, s\dot{\boldsymbol{\sigma}}, \boldsymbol{\varepsilon}, s\dot{\boldsymbol{\varepsilon}}) = s\mathfrak{f}(\boldsymbol{\sigma}, \dot{\boldsymbol{\sigma}}, \boldsymbol{\varepsilon}, \dot{\boldsymbol{\varepsilon}}) \quad \text{for } s > 0,$$

the constitutive relation is rate-independent. As a special case of $\mathfrak{f}$, the hypoelastic law linear with respect to both rates constitutes [29]:

$$(1.2) \qquad \dot{\boldsymbol{\sigma}} - \mathbf{L}_4(\boldsymbol{\sigma})\dot{\boldsymbol{\varepsilon}} = 0 \,,$$

where $\mathbf{L}_4$ is a fourth-order symmetric tensor. To extend (1.2) for an inelastic behavior such that

$$\mathfrak{f}(\boldsymbol{\sigma}, \dot{\boldsymbol{\sigma}}, \boldsymbol{\varepsilon}, -\dot{\boldsymbol{\varepsilon}}) \neq -\mathfrak{f}(\boldsymbol{\sigma}, \dot{\boldsymbol{\sigma}}, \boldsymbol{\varepsilon}, \dot{\boldsymbol{\varepsilon}}) \,,$$

the nonlinearity in $\dot{\boldsymbol{\varepsilon}}$ in the function $\mathfrak{f}$ can be expressed as

$$(1.3) \qquad \dot{\boldsymbol{\sigma}} - \mathbf{L}_4(\boldsymbol{\sigma})\dot{\boldsymbol{\varepsilon}} - \mathbf{N}(\boldsymbol{\sigma})\|\dot{\boldsymbol{\varepsilon}}\| = 0 \,,$$

with a second-order symmetric tensor $\mathbf{N}$ and the Frobenius norm $\|\dot{\boldsymbol{\varepsilon}}\| = \sqrt{\dot{\boldsymbol{\varepsilon}} : \dot{\boldsymbol{\varepsilon}}}$.

The hypoplastic law (1.3) satisfying (1.1) arises in engineering by description of granular materials. For cohesionless grains only non-positive principal stresses

$$\sigma_1 \leq 0, \quad \sigma_2 \leq 0, \quad \sigma_3 \leq 0$$

are physically relevant, in that case, the non-negative mechanical pressure reads

(1.4) $$p = -\frac{1}{3}\mathrm{tr}(\boldsymbol{\sigma}) \geq 0, \quad \mathrm{tr}(\boldsymbol{\sigma}) = \sigma_1 + \sigma_2 + \sigma_3 \leq 0.$$

A deformable granular body consists of solid particles of volume V_s assumed constant, and empty voids of variable volume V_v characterized by a void ratio

$$e = \frac{V_v}{V_s} \in (e_d, e_i),$$

whose bounds e_d and e_i depend on the pressure proportionally [4]. More precisely:

(1.5) $$e_d = e_{\min}f_p, \quad e_i = e_{\max}f_p, \quad 0 < e_{\min} < e_{\max}$$

by means of the unified factor

(1.6) $$f_p(p/h_s) = \exp\left(-\left(\frac{3p}{h_s}\right)^n\right) \in (0, \exp(1)], \quad n \geq 0.$$

with $h_s > 0$ denoting the granular hardness. According to [5], h_s allows degradation:

$$\dot{h}_s = \frac{1}{c}\left(h_s^\infty - h_s\right), \quad c \geq 0,$$

which is expressed for prescribed $0 < h_s^\infty < h_s^0$ by an exponential function

(1.7) $$h_s(t) = h_s^\infty + \left(h_s^0 - h_s^\infty\right)\exp\left(\frac{t_0 - t}{c}\right) \quad \text{for } t > t_0 \geq 0.$$

We assume that the void ratio fulfills the following equation for rates [5]:

(1.8) $$\dot{e} = ne\left(\frac{3p}{h_s}\right)^n\left(\frac{\dot{h}_s}{h_s} - \frac{\dot{p}}{p}\right).$$

For prescribed $p_0 \geq 0$ and $e_0 \in (e_{\min}, e_{\max})f_p(p_0/h_s^0)$, the solution to (1.8) is expressed in the form akin to (1.5) and (1.6):

(1.9) $$e(p/h_s) = \frac{e_0}{f_p(p_0/h_s^0)}\exp\left(-\left(\frac{3p}{h_s}\right)^n\right) \in (e_d(p/h_s), e_i(p/h_s)).$$

The degradation of granular hardness $h_s(t)$ is shown in the left plot of Figure 1, for the example parameters $h_s^0 = 120$ (MPa), $h_s^\infty = 78.5$ (MPa) and $c = 4$ (h). In the right plot of Figure 1, the void ratio $e_d < e < e_i$ versus p/h_s is drawn, for the example parameters $e_{\min} = 0.1$, $e_0 = 0.2$, $e_{\max} = 0.3$ and $n = 0.82$.

The response relation (1.3) depends on the void ratio e as described next. We will consider the specific model (1.3) due to Bauer [3] and Gudehus [14]:

(1.10) $$\mathbf{L}_4(\boldsymbol{\sigma}) = f_s\left(a^2\mathrm{tr}(\boldsymbol{\sigma})\mathbf{I}_4 + \frac{\boldsymbol{\sigma} \otimes \boldsymbol{\sigma}}{\mathrm{tr}(\boldsymbol{\sigma})}\right), \quad \mathbf{N}(\boldsymbol{\sigma}) = f_s f_d a\left(2\boldsymbol{\sigma} - \frac{1}{3}\mathrm{tr}(\boldsymbol{\sigma})\mathbf{I}\right)$$

using the dyadic product of the stress tensor, where $a > 0$ is the yield strength, while $\mathbf{I}$ and $\mathbf{I}_4$ stand for the second-order and the forth-order identity tensors,
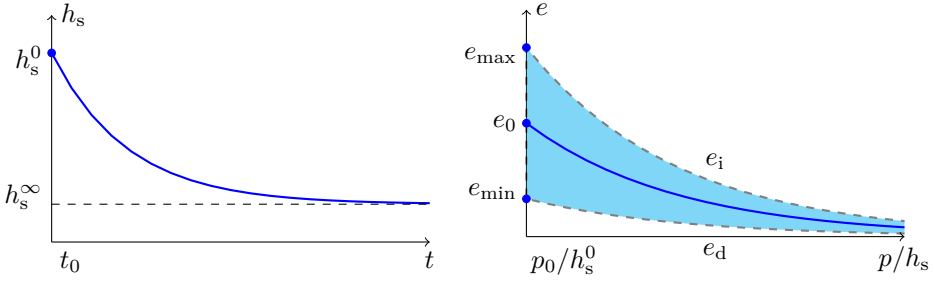
FIG. 1: The granular hardness $h_\mathrm{s}(t)$ degradation (left plot); admissible area for void ratio $e_\mathrm{d} < e < e_\mathrm{i}$ versus p/h_s (right plot).

respectively. The density $f_\mathrm{d} > 0$ and stiffness $f_\mathrm{s} < 0$ factors depend on the void ratio as follows [4]:

$$(1.11) \qquad f_\mathrm{d}(e) = \left(\frac{e - e_\mathrm{d}}{e_\mathrm{c} - e_\mathrm{d}} \right)^\alpha, \quad f_\mathrm{s}(e) = -b \left(\frac{e_\mathrm{i}}{e} \right)^\beta$$

for the parameters $\alpha \in (0, 0.5)$, $\beta > 1$, $b > 0$, and the critical void ratio

$$(1.12) \qquad e_\mathrm{c} = e_\mathrm{crt} f_\mathrm{p}(p/h_\mathrm{s}), \quad e_\mathrm{min} < e_\mathrm{crt} < e_\mathrm{max}.$$

Thanks to the identity (1.9) and using the definitions of $e_\mathrm{d}, e_\mathrm{i}$ in (1.5), one can suppress the void ratio dependence of both functions defined in (1.11), more precisely, the density as well as the stiffness are constants given by

$$(1.13) \qquad f_\mathrm{d} = \left(\frac{\frac{e_0}{f_\mathrm{p}(p_0/h_\mathrm{s}^0)} - e_\mathrm{min}}{e_\mathrm{crt} - e_\mathrm{min}} \right)^\alpha, \quad f_\mathrm{s} = -b \left(\frac{e_\mathrm{max} f_\mathrm{p}(p_0/h_\mathrm{s}^0)}{e_0} \right)^\beta.$$

Inserting (1.10) into (1.3) we write explicitly the differential equation

$$(1.14) \qquad \dot{\boldsymbol{\sigma}} = f_\mathrm{s} \left\{ a^2 \mathrm{tr}(\boldsymbol{\sigma}) \dot{\varepsilon} + \frac{\dot{\boldsymbol{\varepsilon}} : \boldsymbol{\sigma}}{\mathrm{tr}(\boldsymbol{\sigma})} \boldsymbol{\sigma} + a f_\mathrm{d} \left(2\boldsymbol{\sigma} - \frac{1}{3} \mathrm{tr}(\boldsymbol{\sigma}) \mathbf{I} \right) \|\dot{\varepsilon}\| \right\}.$$

Spatial dependence is omitted from the consideration such that (1.14) implies a coupled system of nonlinear dynamic equations.

2. ANALYSIS OF THE MODEL

In this study we focus on a modeling setup when the strain $\boldsymbol{\varepsilon}$ is fully controlled by the stress $\boldsymbol{\sigma}$. The stress-controlled proportional loading consists in choosing a fixed second-order symmetric tensor $\mathbf{T}$ such that

$$(2.1) \qquad \boldsymbol{\sigma} = s\mathbf{T}, \quad \dot{\boldsymbol{\sigma}} = \dot{s}\,\mathbf{T},$$

where a loading parameter $s(t) > 0$ is a strictly monotone differentiable function with respect to time $t \in (t_0, t_1)$ for some $0 \le t_0 < t_1$. This is what we call a *proportional loading* if s is increasing by the mean of $\dot{s} > 0$, and *proportional unloading* if s is decreasing, i.e., $\dot{s} < 0$.

After the substitution of (2.1) into (1.14) we get the implicit system of the first-order ordinary differential equations with respect to the strain rate $\dot{\boldsymbol{\varepsilon}}(t)$ as

$$(2.2) \qquad \dot{s}\mathbf{T} = s f_{\mathrm{s}} \Big\{ a^2 \mathrm{tr}(\mathbf{T})\dot{\boldsymbol{\varepsilon}} + \frac{\dot{\boldsymbol{\varepsilon}} : \mathbf{T}}{\mathrm{tr}(\mathbf{T})}\mathbf{T} + a f_{\mathrm{d}}\Big(2\mathbf{T} - \frac{1}{3}\mathrm{tr}(\mathbf{T})\mathbf{I}\Big)\|\dot{\boldsymbol{\varepsilon}}\| \Big\},$$

which is under our study. Using for brevity the notation

$$(2.3) \qquad \mathbf{X}(t) := \frac{s(t)}{\dot{s}(t)}f_{\mathrm{s}}\dot{\boldsymbol{\varepsilon}}, \quad \hat{\mathbf{T}} := \frac{\mathbf{T}}{\mathrm{tr}(\mathbf{T})}, \quad \mathrm{tr}(\hat{\mathbf{T}}) = 1,$$

equation (2.2) after division by $\mathrm{tr}(\mathbf{T})$ turns into

$$(2.4) \qquad \hat{\mathbf{T}} = a^2\mathbf{X} + (\mathbf{X} : \hat{\mathbf{T}})\hat{\mathbf{T}} \mp a f_{\mathrm{d}}\Big(2\hat{\mathbf{T}} - \frac{1}{3}\mathbf{I}\Big)\|\mathbf{X}\| \quad \text{for } \pm\dot{s} > 0,$$

with two signs $\mp\|\mathbf{X}\|$ corresponding to $\pm\dot{s} > 0$ due to $f_{\mathrm{s}} < 0$.

Taking the scalar product of (2.4) with $\mathbf{X}$ and gathering like terms we calculate

$$\mathbf{X} : \hat{\mathbf{T}} = \frac{\|\hat{\mathbf{T}}\|^2}{a^2 + \|\hat{\mathbf{T}}\|^2} \pm a f_{\mathrm{d}}\frac{2\|\hat{\mathbf{T}}\|^2 - \frac{1}{3}}{a^2 + \|\hat{\mathbf{T}}\|^2}\|\mathbf{X}\| \quad \text{for } \pm\dot{s} > 0.$$

Its substitution into (2.4) yields

$$\frac{a^2}{a^2 + \|\hat{\mathbf{T}}\|^2}\hat{\mathbf{T}} = a^2\mathbf{X} \mp a f_{\mathrm{d}}\Big(\frac{2a^2 + \frac{1}{3}}{a^2 + \|\hat{\mathbf{T}}\|^2}\hat{\mathbf{T}} - \frac{1}{3}\mathbf{I}\Big)\|\mathbf{X}\| \quad \text{for } \pm\dot{s} > 0.$$

Or, after division by a^2 and using for brevity the notation

$$(2.5) \qquad \mathbf{A} := \frac{\hat{\mathbf{T}}}{a^2 + \|\hat{\mathbf{T}}\|^2}, \quad \mathbf{B} := \frac{1}{a^2}\Big[\Big(2a^2 + \frac{1}{3}\Big)\mathbf{A} - \frac{1}{3}\mathbf{I}\Big],$$

an equivalent equation with respect to $\mathbf{X}$ follows:

$$(2.6) \qquad \mathbf{X} = \mathbf{A} \pm a f_{\mathrm{d}}\|\mathbf{X}\|\mathbf{B} \quad \text{for } \pm\dot{s} > 0.$$

Theorem 2.1 (Analytical solution). *Under a solvability condition*

$$(2.7) \qquad f_{\mathrm{d}}^2 \leq \frac{1}{a^2\big(\|\boldsymbol{B}\|^2 - \big(\frac{\boldsymbol{A}:\boldsymbol{B}}{\|\boldsymbol{A}\|}\big)^2\big)} = \frac{3a^2\|\hat{\boldsymbol{T}}\|^2}{\|\hat{\boldsymbol{T}}\|^2 - \frac{1}{3}} =: f_{\max},$$

a solution to the nonlinear system (2.6) *is given in the closed form:*

$$(2.8) \qquad \boldsymbol{X} = \boldsymbol{A} + \frac{a f_{\mathrm{d}}\|\boldsymbol{A}\|\boldsymbol{B}}{\pm\sqrt{D} - a f_{\mathrm{d}}\frac{\boldsymbol{A}:\boldsymbol{B}}{\|\boldsymbol{A}\|}} \quad \text{for } \pm\dot{s} > 0,$$

or, explicitly in terms of $\hat{\boldsymbol{T}}$:

$$(2.9) \qquad \boldsymbol{X} = \frac{\pm\|\hat{\boldsymbol{T}}\|\sqrt{D}\,\hat{\boldsymbol{T}} - \frac{1}{3a}f_{\mathrm{d}}\big(\|\hat{\boldsymbol{T}}\|^2\boldsymbol{I} - \hat{\boldsymbol{T}}\big)}{\pm\|\hat{\boldsymbol{T}}\|(a^2 + \|\hat{\boldsymbol{T}}\|^2)\sqrt{D} - a f_{\mathrm{d}}(2\|\hat{\boldsymbol{T}}\|^2 - \frac{1}{3})}, \quad D := 1 - \frac{f_{\mathrm{d}}^2}{f_{\max}}.$$

Proof. Taking the norm of the expression (2.6) we get

$$\|\mathbf{X}\|^2 = \big\|\mathbf{A} \pm a f_{\mathrm{d}}\|\mathbf{X}\|\mathbf{B}\big\|^2 = \|\mathbf{A}\|^2 \pm 2a f_{\mathrm{d}}(\mathbf{A} : \mathbf{B})\|\mathbf{X}\| + a^2 f_{\mathrm{d}}^2\|\mathbf{B}\|^2\|\mathbf{X}\|^2,$$

which turns into the quadratic equation with respect to $\lambda = \|\mathbf{X}\|$:

$$(2.10) \qquad \big(1 - a^2 f_{\mathrm{d}}^2\|\mathbf{B}\|^2\big)\lambda^2 \mp 2a f_{\mathrm{d}}(\mathbf{A} : \mathbf{B})\lambda - \|\mathbf{A}\|^2 = 0 \quad \text{for } \pm\dot{s} > 0,$$

where according to (2.5)

$$(2.11) \qquad \|\mathbf{A}\| = \frac{\|\hat{\mathbf{T}}\|}{a^2 + \|\hat{\mathbf{T}}\|^2}, \quad \mathrm{tr}(\mathbf{A}) = \frac{1}{a^2 + \|\hat{\mathbf{T}}\|^2},$$

and the coefficients are

$$(2.12) \quad \mathbf{A} : \mathbf{B} = \left(2 + \frac{1}{3a^2}\right)\|\mathbf{A}\|^2 - \frac{\mathrm{tr}(\mathbf{A})}{3a^2} = \frac{2\|\hat{\mathbf{T}}\|^2 - \frac{1}{3}}{(a^2 + \|\hat{\mathbf{T}}\|^2)^2},$$

$$\|\mathbf{B}\|^2 = \left(2 + \frac{1}{3a^2}\right)\left(\mathbf{A} : \mathbf{B} - \frac{\mathrm{tr}(\mathbf{A})}{3a^2}\right) + \frac{1}{3a^4} = \left(2 + \frac{1}{3a^2}\right)\frac{(2 - \frac{1}{3a^2})\|\hat{\mathbf{T}}\|^2 - \frac{2}{3}}{(a^2 + \|\hat{\mathbf{T}}\|^2)^2} + \frac{1}{3a^4}.$$

The discriminant for this equation

$$(2.13) \quad \mathrm{Disc} = \|\mathbf{A}\|^2 \left(1 - a^2 f_{\mathrm{d}}^2 \left[\|\mathbf{B}\|^2 - \left(\frac{\mathbf{A} : \mathbf{B}}{\|\mathbf{A}\|}\right)^2\right]\right) = \frac{\|\hat{\mathbf{T}}\|^2 - \frac{1}{3a^2} f_{\mathrm{d}}^2 (\|\hat{\mathbf{T}}\|^2 - \frac{1}{3})}{(a^2 + \|\hat{\mathbf{T}}\|^2)^2}$$

is non-negative when the solvability condition (2.7) holds, provided by the lower bound $\|\hat{\mathbf{T}}\|^2 \geq 1/3$ since $\mathrm{tr}(\hat{\mathbf{T}}) = 1$. Then the roots of (2.10) are

$$(2.14) \qquad \lambda = \frac{\pm a f_{\mathrm{d}}(\mathbf{A} : \mathbf{B}) + \sqrt{\mathrm{Disc}}}{1 - a^2 f_{\mathrm{d}}^2 \|\mathbf{B}\|^2}, \quad \lambda = \frac{\pm a f_{\mathrm{d}}(\mathbf{A} : \mathbf{B}) - \sqrt{\mathrm{Disc}}}{1 - a^2 f_{\mathrm{d}}^2 \|\mathbf{B}\|^2}$$

for $\pm \dot{s} > 0$. Since $\|\mathbf{X}\|$ has a sense only for positive values of λ, noting that $\mathbf{A} : \mathbf{B} > 0$ in (2.12) and using (2.13) to write

$$1 - a^2 f_{\mathrm{d}}^2 \|\mathbf{B}\|^2 = \frac{(\sqrt{\mathrm{Disc}})^2 - (a f_{\mathrm{d}}(\mathbf{A} : \mathbf{B}))^2}{\|\mathbf{A}\|^2},$$

from (2.14) we deduce

$$(2.15) \qquad \|\mathbf{X}\| = \frac{\|\mathbf{A}\|^2}{\sqrt{\mathrm{Disc}} \mp a f_{\mathrm{d}}(\mathbf{A} : \mathbf{B})} > 0 \quad \text{for } \pm \dot{s} > 0.$$

The substitution of (2.15) into (2.6) gives the analytical formulas (2.8) and (2.9) for the solution, where $D = \mathrm{Disc}/\|\mathbf{A}\|^2$ from (2.13). This finishes the proof. $\quad\square$

Note that the expressions obtained in Theorem 2.1 show no dependence in time resulting in a constant value for X implicitly related to the loading parameter. The formulas above together with (2.3) allow us to derive the strain rate in terms of $\mathbf{X}$ provided $f_{\mathrm{d}}^2 \leq f_{\max}$, in other words, equation (2.3) gives

$$(2.16) \qquad \dot{\varepsilon} = \frac{\dot{s}}{s f_{\mathrm{s}}} \mathbf{X},$$

and from (2.9) we infer the scalar expression

$$(2.17) \qquad \mathrm{tr}(\dot{\varepsilon}) = \left(\frac{\dot{s}}{s f_{\mathrm{s}}}\right) \frac{\pm \|\hat{\mathbf{T}}\| \sqrt{1 - \frac{f_{\mathrm{d}}^2}{f_{\max}}} - \frac{1}{a} f_{\mathrm{d}}(\|\hat{\mathbf{T}}\|^2 - \frac{1}{3})}{\pm \|\hat{\mathbf{T}}\|(a^2 + \|\hat{\mathbf{T}}\|^2) \sqrt{1 - \frac{f_{\mathrm{d}}^2}{f_{\max}}} - a f_{\mathrm{d}}(2\|\hat{\mathbf{T}}\|^2 - \frac{1}{3})}$$

for $\pm \dot{s} > 0$, which we use for numerical simulation tests below.

2.1. Isotropic loading. In particular, for the isotropic case

$$(2.18) \qquad \mathbf{T} = -\mathbf{I}, \quad \text{tr}(\mathbf{T}) = -3, \quad \hat{\mathbf{T}} = \frac{1}{3}\mathbf{I}, \quad \|\hat{\mathbf{T}}\| = \frac{1}{\sqrt{3}},$$

then $f_{\max} = \infty$ and $D = 1$, thus (2.7) always holds. From (2.9) we have

$$(2.19) \qquad \mathbf{X} = \frac{1}{3a^2 + 1 \mp \sqrt{3}af_{\mathrm{d}}}\mathbf{I} \quad \text{for } \pm\dot{s} > 0,$$

and from (2.17) get respectively

$$(2.20) \qquad \text{tr}(\mathbf{X}) = \frac{3}{3a^2 + 1 \mp \sqrt{3}af_{\mathrm{d}}} \quad \text{for } \pm\dot{s} > 0.$$

2.2. Example of shear stress. Let us consider the shear stress matrix

$$(2.21) \qquad \mathbf{T} = \begin{bmatrix} -0.5 & 0.5 & 0 \\ 0.5 & -0.5 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \quad \text{tr}(\mathbf{T}) = -1, \quad \hat{\mathbf{T}} = -\mathbf{T}, \quad \|\hat{\mathbf{T}}\| = 1$$

such that the principal stresses $\sigma_1 = -1$, $\sigma_2 = \sigma_3 = 0$, and $f_{\max} = 4.5a^2$ in (2.7).
Let us consider the functions in (1.11) rescaled as follows

$$(2.22) \qquad f_{\mathrm{d}}(e/f_{\mathrm{p}}) = \left(\frac{e/f_{\mathrm{p}} - e_{\min}}{e_{\mathrm{crt}} - e_{\min}}\right)^{\alpha}, \quad f_{\mathrm{s}}(e/f_{\mathrm{p}}) = -b\left(\frac{e_{\max}}{e/f_{\mathrm{p}}}\right)^{\beta}.$$

From (2.17) and (2.22) we find

$$(2.23) \qquad \frac{1}{f_{\mathrm{s}}}\text{tr}(\mathbf{X}) = \frac{1}{f_{\mathrm{s}}}\frac{\pm\sqrt{1 - \frac{2f_{\mathrm{d}}^2}{9a^2}} - \frac{2}{3a}f_{\mathrm{d}}}{\pm(a^2 + 1)\sqrt{1 - \frac{2f_{\mathrm{d}}^2}{9a^2}} - \frac{5a}{3}f_{\mathrm{d}}} \quad \text{for } \pm\dot{s} > 0$$

as a function of two variables a and $e/f_{\mathrm{p}} \in [e_{\min}, e_{\max}]$.
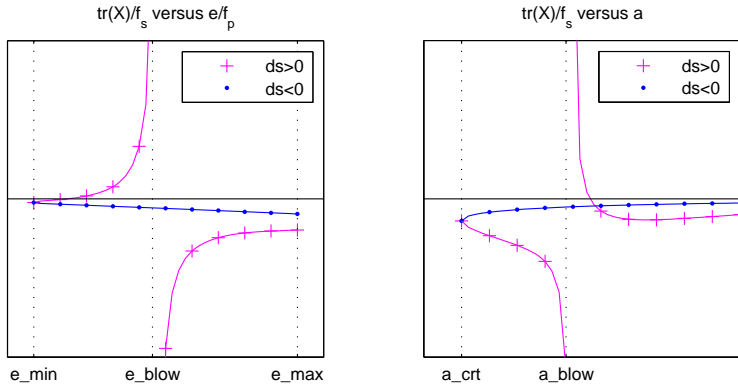
FIG. 2: Example $\text{tr}(\mathbf{X})/f_{\mathrm{s}}$ for $\pm\dot{s} > 0$ versus parameter e/f_{p}
entering f_{d} and f_{s} (left plot); versus yield strength a (right plot).

For the parameters used in Figure 1 and $\alpha = 0.18$, $\beta = 1.05$, $b = 1$, two graphs of $\mathrm{tr}(\mathbf{X})/f_{\mathrm{s}}$ for $\dot{s} > 0$ and $\dot{s} < 0$ versus e/f_{p} and fixed $a = 0.6$ are depicted in the window $|\mathrm{tr}(\mathbf{X})/f_{\mathrm{s}}| \leq 10$ in the left plot of Figure 2. They coincide at $e/f_{\mathrm{p}} = e_{\min}$ because $f_{\mathrm{d}} = 0$ and $\mathbf{X} = \mathbf{A}$ in (2.8) is unique in this case. The graph as $\dot{s} < 0$ is continuous, whereas as $\dot{s} > 0$ it blows up when the denominator in (2.23), that is,

$$(2.24) \qquad \|\hat{\mathbf{T}}\|(a^2 + \|\hat{\mathbf{T}}\|^2)\sqrt{1 - \frac{f_{\mathrm{d}}^2}{f_{\max}}} - af_{\mathrm{d}}\left(2\|\hat{\mathbf{T}}\|^2 - \frac{1}{3}\right),$$

tends to zero at $e_{\mathrm{blow}} \approx 0.19$.

In the right plot of Figure 2, two graphs of $\mathrm{tr}(\mathbf{X})/f_{\mathrm{s}}$ for $\pm\dot{s} > 0$ are depicted versus the yield strength $a \geq a_{\mathrm{crt}}$ in the window $|\mathrm{tr}(\mathbf{X})/f_{\mathrm{s}}| \leq 10$ and $a \leq 1.2$ for fixed $e/f_{\mathrm{p}} = e_0/f_{\mathrm{p}}(p_0/h_{\mathrm{s}}^0) \approx 0.2278$ used in Figure 1. There $a_{\mathrm{crt}} \approx 0.4638$ can be found from the solvability condition (2.7) such that

$$(2.25) \qquad a_{\mathrm{crt}} = \frac{\sqrt{\|\hat{\mathbf{T}}\|^2 - \frac{1}{3}}}{\sqrt{3}\|\hat{\mathbf{T}}\|} f_{\mathrm{d}}\,.$$

When (2.25) holds, the corresponding discriminant $D = 0$ and the solution $\mathbf{X} = \mathbf{A} + \|\mathbf{A}\|^2\mathbf{B}/(\mathbf{A} : \mathbf{B})$ in (2.8) is unique for both $\pm\dot{s} > 0$. As $\dot{s} > 0$ the denominator in (2.24) tends to zero and causes the blow up at $a_{\mathrm{blow}} \approx 0.7398$.

3. HYSTERESIS UNDER CYCLIC LOADING

Let us consider a time discretization based on equidistant points $t_k = k\tau$ for $k = 0, 1, 2, \ldots$ and a fixed period $\tau > 0$. For prescribed $0 < s_{\mathrm{even}} \leq s(t_0) < s_{\mathrm{odd}}$, we introduce cyclic loading by a continuous periodic function $s(t)$ in (2.1) such that

$$(3.1) \qquad \begin{cases} \dot{s} > 0 & \text{for } t \in (t_{2j}, t_{2j+1}), \\ \dot{s} < 0 & \text{for } t \in (t_{2j+1}, t_{2j+2}), \end{cases} \quad j = 0, 1, 2, \ldots$$

where the node values for the loading parameter are set

$$(3.2) \qquad s(t_2) = s(t_4) = \ldots = s_{\mathrm{even}}, \quad s(t_1) = s(t_3) = \ldots = s_{\mathrm{odd}}.$$

For example, solving the equation $\dot{s} = s$ with $s(t_0) = s_{\mathrm{even}}$ we get

$$(3.3) \qquad \begin{cases} s(t) = s(t_{2j})\exp(t - t_{2j}) & \text{for } t \in (t_{2j}, t_{2j+1}), \\ s(t) = s(t_{2j+1})\exp(t_{2j+1} - t) & \text{for } t \in (t_{2j+1}, t_{2j+2}), \end{cases}$$

which is continuous when $s_{\mathrm{odd}} = s_{\mathrm{even}}\exp(\tau)$ and is illustrated in Figure 3. Having in mind the identities in (2.1) with the cyclic loading described by (3.1) and (3.2), the corresponding equation for the mechanical pressure (1.4) becomes

$$(3.4) \qquad p(t) = -s(t)\frac{\mathrm{tr}(\mathbf{T})}{3} \quad \text{for } t \in (t_{2j}, t_{2j+1}) \cup (t_{2j+1}, t_{2j+2})\,,$$

therefore, it is continuous and periodic with

$$(3.5) \qquad p(t_{2j+2}) = -s_{\mathrm{even}}\frac{\mathrm{tr}(\mathbf{T})}{3} =: p_{\mathrm{even}}, \quad p(t_{2j+1}) = -s_{\mathrm{odd}}\frac{\mathrm{tr}(\mathbf{T})}{3} =: p_{\mathrm{odd}},$$
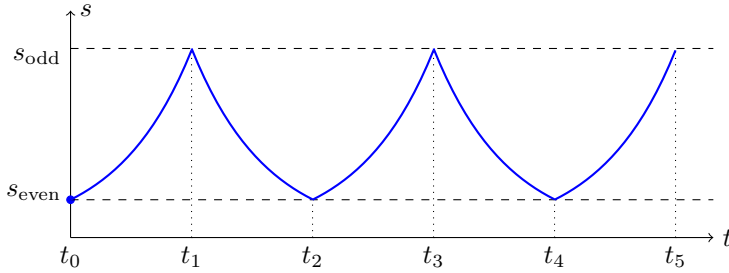
FIG. 3: Example cyclic loading $s(t)$.

where $0 < p_{\mathrm{even}} \leq p(t_0) < p_{\mathrm{odd}}$ because of $\mathrm{tr}(\mathbf{T}) \leq 0$ in (1.4). The granular hardness in (1.7) after discretization implies a monotonically decaying sequence

$$(3.6) \qquad h_{\mathrm{s}}(t_k) = h_{\mathrm{s}}^{\infty} + \left(h_{\mathrm{s}}^0 - h_{\mathrm{s}}^{\infty}\right) \exp\!\left(\frac{t_0 - t_k}{c}\right) \searrow h_{\mathrm{s}}^{\infty} \quad \text{as } k \to \infty.$$

For the pressure and granular hardness-dependent function $f_{\mathrm{p}}(p/h_{\mathrm{s}})$ defined in (1.6), let the initial void ratio be prescribed at $t = t_0$:

$$e(t_0) \in (e_{\min}, e_{\max}) f_{\mathrm{p}}\big(p(t_0)/h_{\mathrm{s}}(t_0)\big).$$

The void ratio in (1.9) yields a time-dependent relation as $k = 0, 1, 2, \dots$

$$(3.7) \qquad e(t) = \frac{e(t_k)}{f_{\mathrm{p}}(p(t_k)/h_{\mathrm{s}}(t_k))} f_{\mathrm{p}}(p(t)/h_{\mathrm{s}}(t)) \quad \text{for } t \in (t_k, t_{k+1}).$$

The function in (3.7) is continuous at $t = t_{k+1}$. The factors in (1.13) are determined by $e_0 = e(t_0)$, $p_0 = p(t_0)$, and $h_{\mathrm{s}}^0 = h_{\mathrm{s}}(t_0)$.

Assuming that $f_{\mathrm{d}}^2 \leq f_{\max}$ according to (2.7) in Theorem 2.1, from (2.8) and (2.16) we get the strain rate for $j = 0, 1, 2, \dots$

$$(3.8) \quad \begin{cases} \dot{\boldsymbol{\varepsilon}}(t) = \dfrac{\dot{s}(t)}{s(t) f_{\mathrm{s}}}\left(\mathbf{A} + \dfrac{a f_{\mathrm{d}} \|\mathbf{A}\| \mathbf{B}}{\sqrt{1 - \frac{f_{\mathrm{d}}^2}{f_{\max}}} - a f_{\mathrm{d}} \frac{\mathbf{A}:\mathbf{B}}{\|\mathbf{A}\|}}\right) & \text{for } t \in (t_{2j}, t_{2j+1}), \\[6mm] \dot{\boldsymbol{\varepsilon}}(t) = \dfrac{\dot{s}(t)}{s(t) f_{\mathrm{s}}}\left(-\mathbf{A} + \dfrac{a f_{\mathrm{d}} \|\mathbf{A}\| \mathbf{B}}{\sqrt{1 - \frac{f_{\mathrm{d}}^2}{f_{\max}}} + a f_{\mathrm{d}} \frac{\mathbf{A}:\mathbf{B}}{\|\mathbf{A}\|}}\right) & \text{for } t \in (t_{2j+1}, t_{2j+2}). \end{cases}$$

Theorem 3.1 (Attractor). *Under the stress control* (2.1) *by cyclic loading* (3.1) *and* (3.2) *the void ratio is found in the closed form:*

$$(3.9) \qquad e(t) = e(t_0) \exp\!\left((-\mathrm{tr}(\boldsymbol{T}))^n \left[\left(\frac{s(t_0)}{h_{\mathrm{s}}(t_0)}\right)^n - \left(\frac{s(t)}{h_{\mathrm{s}}(t)}\right)^n\right]\right) \quad \text{for} \quad t \geq t_0.$$

As $t \to \infty$ it tends exponentially to an attractor with end-points

$$(3.10) \quad e_i^{\infty} = e(t_0) \exp\!\left((-\mathrm{tr}(\boldsymbol{T}))^n \left[\left(\frac{s(t_0)}{h_{\mathrm{s}}(t_0)}\right)^n - \left(\frac{s_i}{h_{\mathrm{s}}^{\infty}}\right)^n\right]\right) \quad \text{for } i \in \{\mathrm{even}, \mathrm{odd}\}.$$

Proof. From (3.7), the following formula can be justified by induction:

$$(3.11) \qquad e(t_k) = \frac{e(t_0)}{f_{\mathrm{p}}(p(t_0)/h_{\mathrm{s}}(t_0))} f_{\mathrm{p}}\big(p(t_k)/h_{\mathrm{s}}(t_k)\big) \quad \text{for } k = 0, 1, 2, \ldots$$

Inserting here (1.6), (3.5) and (3.6) we calculate the void ratio (3.9). Based on the periodicity (3.2) and convergence (3.6), the assertion follows. $\qquad\square$

3.1. **Example of densification.** Under the cyclic loading given by (3.3) for $s_{\mathrm{even}} = 1$, $\tau = 1$, and $s(t_0) = s_{\mathrm{odd}} = \exp(1) \approx 2.7183$, we put $\mathrm{tr}(\mathbf{T}) = -1$ from the example stress (2.21) such that the pressure bounds $p_{\mathrm{even}} \approx 0.3333$ and $p(t_0) = p_{\mathrm{odd}} \approx 0.9061$ in (3.5). The graph of the void ratio $e(t)$ computed by formula (3.9) is drawn versus $p(t)$ in Figure 4 after 10 cycles implying $j = 0, 1, \ldots, 9$ in (3.1) with $t_0, t_1, \ldots, t_{19}$ time points. The consequential cycles of $e(t)$ are visually indistinguishable and approach the attractor according to Theorem 3.1.
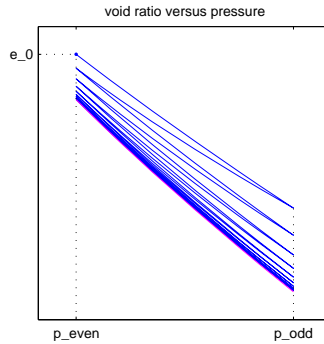
FIG. 4: Void ratio $e(t)$ versus pressure $p(t)$ under cyclic loading.

4. CONCLUSION

We conclude with some principal findings of our theoretical and numerical study.
- An implicit system of 1st-order ordinary differential equations is studied to model a granular media within the hypoplastic theory, which allows degradation of
- Under stress control, an analytical formula for the strain rate is constructed for arbitrary inhomogeneous coefficients in the governing equations.
- Under loading-unloading cycles, the void ratio exhibits ratcheting of hysteresis loops towards densification of the granular media.
- The ratcheting phenomenon is finite, pressure-void ratio states generate an attractor with respect to time.

References

[1] Annin, B.D., Kovtunenko, V.A., Sadovskii, V.M., *Variational and hemivariational inequalities in mechanics of elastoplastic, granular media, and quasibrittle cracks*, Analysis, Modelling, Optimization, and Numerical Techniques (Tost, G.O., Vasilieva, O., eds.), vol. 121, Springer Proc. Math. Stat., 2015, pp. 49–56.

[2] Armstrong, P.J., Frederick, C.O., *A mathematical representation of the multiaxial Bauschinger effect*, C.E.G.B. Report RD/B/N, 1966.

[3] Bauer, E., *Calibration of a comprehensive hypoplastic model for granular materials*, Soils Found. **36** (1996), 13–26.

[4] Bauer, E., *Analysis of shear band bifurcation with a hypoplastic model for a pressure and density sensitive granular material*, Mech. Mater. **31** (1999), 597–609.

[5] Bauer, E., *Long-term behavior of coarse-grained rockfill material and their constitutive modeling*, Dam Engineering - Recent Advances in Design and Analysis (Fu, Z., Bauer, E., eds.), IntechOpen, 2021.

[6] Bauer, E., Kovtunenko, V.A., Krejčí, P., Krenn, N., Siváková, L., Zubkova, A.V., *Modified model for proportional loading and unloading of hypoplastic materials*, Extended Abstracts Spring 2018. Singularly Perturbed Systems, Multiscale Phenomena and Hysteresis: Theory and Applications (Korobeinikov, A., Caubergh, M., Lázaro, T., Sardanyés, J., eds.), Trends in Mathematics, vol. 11, Birkhäuser, Hamburk, 2019, pp. 201–210.

[7] Bauer, E., Kovtunenko, V.A., Krejčí, P., Krenn, N., Siváková, L., Zubkova, A.V., *On proportional deformation paths in hypoplasticity*, Acta Mechanica **231** (2020), 1603–1619.

[8] Brokate, M., Krejčí, P., *Wellposedness of kinematic hardening models in elastoplasticity*, RAIRO Modél. Math. Anal. Numér. **32** (1998), 177–209.

[9] Chambon, R., Desrues, J., Hammad, W., Charlier, R., *CLoE, a new rate-type constitutive model for geomaterials, theoretical basis and implementation*, Int. J. Num. Anal. Methods Geomech. **18** (1994), 253–278.

[10] Darve, F., *Incrementally non-linear constitutive relationships*, Geomaterials, Constitutive Equations and Modelling (Darve, F., ed.), Elsevier, Horton, Greece, 1990, pp. 213–238.

[11] Fellner, K., Kovtunenko, V.A., *A singularly perturbed nonlinear Poisson–Boltzmann equation: uniform and super-asymptotic expansions*, Math. Methods Appl. Sci. **38** (2015), 3575–3586.

[12] Fellner, K., Kovtunenko, V.A., *A discontinuous Poisson–Boltzmann equation with interfacial transfer: homogenisation and residual error estimate*, Appl. Anal. **95** (2016), 2661–2682.

[13] González Granada, J.R., Kovtunenko, V.A., *Entropy method for generalized Poisson–Nernst–Planck equations*, Anal. Math. Phys. **8** (2018), 603–619.

[14] Gudehus, G., *A comprehensive constitutive equation for granular materials*, Soils Found. **36** (1996), 1–12.

[15] Hron, J., Málek, J., Rajagopal, K.R., *Simple flows of fluids with pressure dependent viscosities*, Proc. Roy. Soc. A **457** (2001), 1603–1622.

[16] Khludnev, A.M., Kovtunenko, V.A., *Analysis of Cracks in Solids*, WIT-Press, Southampton, Boston, 2000.

[17] Kolymbas, D., *Introduction to Hypoplasticity*, A.A. Balkema, Rotterdam, 2000.

[18] Kolymbas, D., Medicus, G., *Genealogy of hypoplasticity and barodesy*, Int. J. Numer. Anal. Methods Geomech. **40** (2016), 2530–2550.

[19] Kovtunenko, V.A., Bauer, E., Eliaš, J., Krejčí, P., Monteiro, G.A., Straková (Siváková), L., *Cyclic behavior of simple models in hypoplasticity and plasticity with nonlinear kinematic hardening*, J. Sib. Fed. Univ. - Math. Phys. **14** (2021), 756–767.

[20] Kovtunenko, V.A., Krejčí, P., Bauer, E., Siváková, L., Zubkova, A.V., *On Lyapunov stability in hypoplasticity*, Proc. Equadiff 2017 Conference (Mikula, K., Ševčovič, D., Urbán, J., eds.), Slovak University of Technology, Bratislava, 2017, pp. 107–116.

[21] Kovtunenko, V.A., Krejčí, P., Krenn, N., Bauer, E., Siváková, L., Zubkova, A.V., *On feasibility of rate-independent stress paths under proportional deformations within hypoplastic constitutive model for granular materials*, Math. Models Eng. **5** (2019), 119–126.

[22] Kovtunenko, V.A., Zubkova, A.V., *Mathematical modeling of a discontinuous solution of the generalized Poisson–Nernst–Planck problem in a two-phase medium*, Kinet. Relat. Models **11** (2018), 119–135.

[23] Kovtunenko, V.A., Zubkova, A.V., *Homogenization of the generalized Poisson–Nernst–Planck problem in a two-phase medium: correctors and estimates*, Appl. Anal. **100** (2021 a), 253–274.

[24] Kovtunenko, V.A., Zubkova, A.V., *Existence and two-scale convergence of the generalised Poisson–Nernst–Planck problem with non-linear interface conditions*, Eur. J. Appl. Math. **32** (2021 b), 683–710.

[25] Krejčí, P., *Hysteresis, Convexity and Dissipation in Hyperbolic Equations*, Gakkotosho, Tokyo, 1996.

[26] Mašín, D., *Modelling of Soil Behaviour with Hypoplasticity: Another Approach to Soil Constitutive Modelling*, Springer Nature, Switzerland, 2019.

[27] Niemunis, A., Herle, I., *Hypoplastic model for cohesionless soils with elastic strain range*, Mech. Cohes.-Frict. Mat. **2** (1997), 279–299.

[28] Rajagopal, K.R., Srinivasa, A.R., *On a class of non-dissipative materials that are not hyperelastic*, Proc. Roy. Soc. A **465** (2009), 493–500.

[29] Truesdell, C., *Remarks on hypo-elasticity*, J. Res. Natl. Bur. Stand. B **67B** (1963), 141–143.

[30] Valanis, K.C., *A theory of viscoplasticity without a yield surface*, Arch. Mech. **23** (1971), 517–533.

Institute for Mathematics and Scientific Computing,
University of Graz, NAWI Graz,
Heinrichstr. 36, 8010 Graz, Austria
and
Lavrent'ev Institute of Hydrodynamics,
Siberian Division of the Russian Academy of Sciences,
630090 Novosibirsk, Russia
E-mail: `victor.kovtunenko@uni-graz.at`

Institute for Mathematics and Scientific Computing,
University of Graz, NAWI Graz,
Heinrichstr. 36, 8010 Graz, Austria
E-mail: `jan.elias3@gmail.com`

Faculty of Civil Engineering,
Czech Technical University in Prague,
Thákurova 7, 166 29 Praha 6, Czech Republic
E-mail: `krejci@math.cas.cz`

Institute of Mathematics, Czech Academy of Sciences,
Žitná 25, 115 67 Praha 1, Czech Republic
E-mail: `gam@math.cas.cz`

Faculty of Civil Engineering,
Czech Technical University in Prague,
Thákurova 7, 166 29 Praha 6, Czech Republic
E-mail: `judita.runczikova@fsv.cvut.cz`

ON EULER METHODS FOR CAPUTO FRACTIONAL DIFFERENTIAL EQUATIONS

Petr Tomášek

ABSTRACT. Numerical methods for fractional differential equations have specific properties with respect to the ones for ordinary differential equations. The paper discusses Euler methods for Caputo differential equation initial value problem. The common properties of the methods are stated and demonstrated by several numerical experiments. Python codes are available to researchers for numerical simulations.

1. Introduction

Fractional differential equations have become an area of mathematics being widely investigated in recent years. Several approaches to fractional derivative were formulated, e.g. Grünwald–Letnikov, Riemann–Liouville, Atangana–Baleanu, Riesz, Caputo and others. The Caputo approach has become popular in connection with formulation of fractional differential equations initial value problems. It was due to the form of initial conditions where integer order derivatives are employed. Thence physical meaning of such conditions is more clear and understandable. For basics of fractional calculus and its applications we refer to [1], [7]. The numerical analysis of these tasks has been developed following the fractional calculus advancement. A survey of some numerical approaches can be found in monograph [6] and the references therein. In the paper we introduce the basic Euler numerical schemes to the Caputo type differential equations. We restrict our discussions to a scalar problem with fractional order derivative $\alpha \in (0, 1)$ but it is possible to generalize the following considerations to $\alpha > 1$ and also to a vector counterpart of the problem. The aim of the paper is to introduce the methods, to mention their properties and mainly demonstrate them by numerical experiments. Moreover, we present the Python codes of the methods to be utilized by researchers.

We consider the initial value problem

$$(1.1) \qquad \begin{cases} D^\alpha y(t) = f\big(t, y(t)\big) \,, \\ y(0) = y_0 \,, \end{cases}$$

where $\alpha \in (0,1)$ and $t \in [0, t_f]$. Caputo differential operator is introduced as

$$D^\alpha f(t) = \frac{1}{\Gamma(1-\alpha)} \int_0^t (t-s)^{-\alpha} f'(s) \, ds \,, \qquad \alpha \in (0,1) \,.$$

First of all we mention conditions, which locally ensure existence and uniqueness of a solution of initial value problem (1.1) (see [6, p. 97]).

Theorem 1.1. *Let $M := [0, \chi^\star] \times [y_0 - \delta, y_0 + \delta]$ with $\chi^\star > 0$ and $\delta > 0$. Let $f\colon M \to \mathbb{R}$ be a bounded function on M satisfying the Lipschitz condition with respect to the second variable $|f(t,u) - f(t,v)| \le L|u - v|$ with $L > 0$ independent of t, u, v. Let $\chi = \min\{\chi^\star, (\delta\Gamma(1+\alpha)/|f|)^{1/\alpha}\}$. Then there exists a unique solution $y : [0, \chi] \to \mathbb{R}$ of (1.1).*

We often need to employ numerical methods to obtain approximation of the solution since an analytical form of the solution can be found in a very narrow class of initial value problems.

2. Numerical methods

There are two common approaches to obtain a numerical formula for problem (1.1): direct discretization of the Caputo derivative and transformation of a fractional differential equation to a fractional integral equation with a subsequent discretization of the fractional integral. The transformation is based on application of α Caputo integral to (1.1), which gives

$$(2.1) \qquad y(t) = y_0 + D^{-\alpha} f\big(t, y(t)\big) = y_0 + \frac{1}{\Gamma(\alpha)} \int_0^t (t-s)^{\alpha-1} f\big(s, y(s)\big) \, ds \,.$$

We use the second approach. In the following, we consider an equidistant mesh with stepsize $h > 0$ ($h = t_f/N$), i.e. $t_0 = 0$, $t_n = nh, n = 0, 1, 2, \ldots, N$. Then y_n denotes a numerical approximation of the exact solution value $y(t_n)$.

2.1. **Fractional forward Euler method.** The fractional forward Euler formula (FFEM) solving (1.1) is introduced as follows:

$$y_{n+1} = y_0 + h^\alpha \sum_{j=0}^n b_{j,n+1} f(t_j, y_j) \,, \quad n = 0, 1, \ldots, N-1 \,,$$

where

$$(2.2) \quad b_{j,n+1} = \frac{(n-j+1)^\alpha - (n-j)^\alpha}{\Gamma(1+\alpha)} \,, \quad n = 0, 1, \ldots, N-1, \; j = 0, 1, \ldots, n \,.$$

2.2. **Fractional backward Euler method.** The fractional backward Euler formula (FBEM) solving (1.1) is introduced as follows:

$$y_{n+1} = y_0 + h^\alpha \sum_{j=0}^n b_{j,n+1} f(t_{j+1}, y_{j+1}), \quad n = 0, 1, \ldots, N-1,$$

where $b_{j,n+1}$ is given by (2.2).

As we can see, the integral (2.1) is approximated by the left-point and right-point rule in the case of FFEM and FBEM, respectively. In the following we mention some crucial properties of these methods.

2.3. **Properties of FFEM and FBEM.** First we introduce an assertion which classifies FFEM and FBEM as first order methods (see [6, p. 103]).

Theorem 2.1. *Let $y(t)$ be a solution of (1.1), $f(t, y)$ satisfy the Lipschitz condition with respect to the second argument and $f(t, y(t))$, $y(t) \in C^1[0, t_f]$. Let $y_n, n = 0, 1, 2, \ldots, N$ be the approximations of $y(t_n)$ by FFEM (resp. FBEM). Then*

$$|y(t_n) - y_n)| \leq Kh, \qquad n = 0, 1, 2, \ldots, N,$$

where $K > 0$ is a constant independent of h and n.

Both the methods are convergent and stable considering the assumptions of Theorem 2.1 are fulfilled. The stability means that small perturbations in the initial conditions would not lead to large errors in the numerical solution (see [6, p. 102]).

3. Numerical experiments

Example 3.1. Consider the following initial value problem

$$(3.1) \qquad \begin{cases} D^\alpha y(t) = -y(t), & t \in [0, 10] \\ y(0) = 1, \end{cases}$$

which may serve as a test initial value problem for A-stability. Its analytical solution is well known: $y(t) = E_\alpha(-t^\alpha)$, where $E_\alpha(z) = \sum_{k=0}^{\infty} \frac{z^k}{\Gamma(\alpha k + 1)}$ is the Mittag-Leffler function. The numerical solution obtained by both the Euler's schemes are shown in Figures 1-5 with fractional derivative order α equal to 0.05, 0.25, 0.5, 0.75, 0.95, in sequence. For these values of α there are shown solutions considered on meshes with various stepsizes h (0.01, 0.5, 1 and 2). We can observe that, in analogy with Euler's methods for ordinary differential equations, the backward fractional Euler method has much better asymptotic stability behavior then the forward one. All the presented parameter settings lead to a positive decreasing solution in the case of FBEM. On the contrary, in the case of FFEM, larger stepsizes cause a more prominent oscillation of the numerical solution.

Example 3.2. Consider the following initial value problem

$$(3.2) \qquad \begin{cases} D^{1/3} y(t) = \frac{9t^{5/3}}{10\Gamma(2/3)}, & t \in [0, 10]. \\ y(0) = 1. \end{cases}$$

The exact solution of this problem is $y(t) = t^2/2 + 1$. The numerical solution for various stepsizes h (0.01, 0.5, 1 and 2) is shown in Figure 6. Propagation of absolute error by FFEM and FBEM is shown in Table 1 and 2, respectively. The computed absolute errors coincide with the estimate introduced in Theorem 2.1.
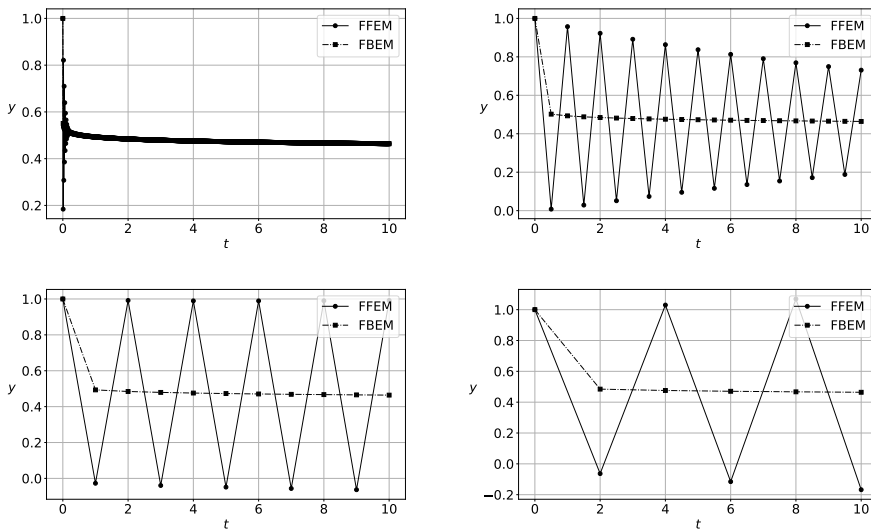
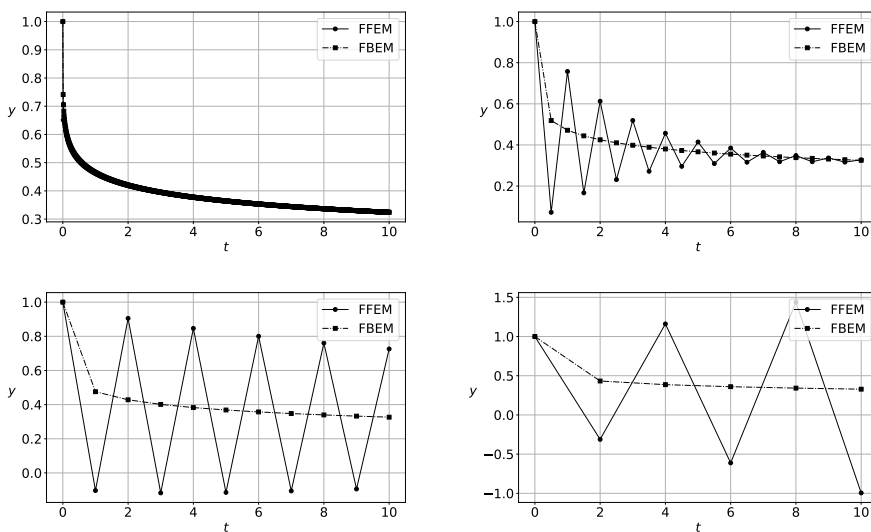FIG. 1: Numerical solution of (3.1) with $\alpha = 0.05$ and $h = 0.01, 0.5, 1, 2$.

FIG. 2: Numerical solution of (3.1) with $\alpha = 0.25$ and $h = 0.01, 0.5, 1, 2$.

Remark 3.3. One can easily check, that $y(t) = t^2/2 + 1$ is also the solution of the problem

$$(3.3) \qquad \begin{cases} D^{1/3}y(t) = \frac{9t(2y(t)-2)^{2/3}}{10\Gamma(2/3)}, & t \in [0,10]. \\ y(0) = 1. \end{cases}$$
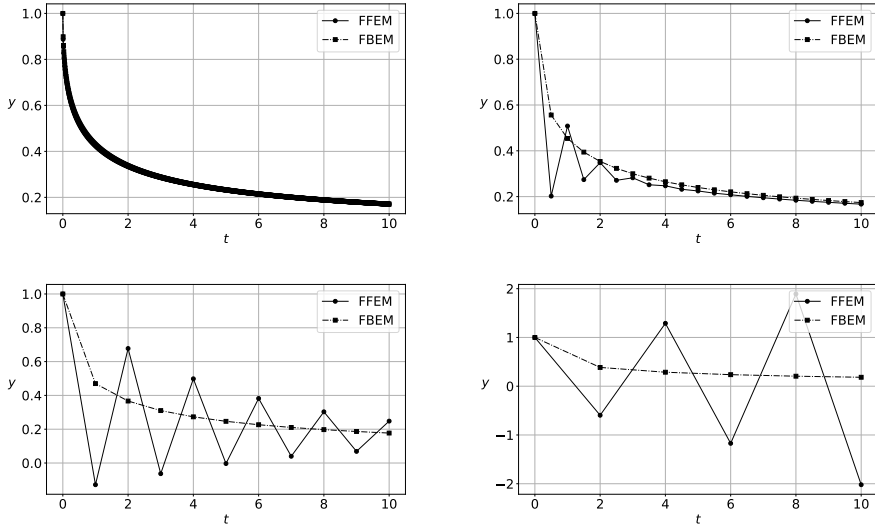
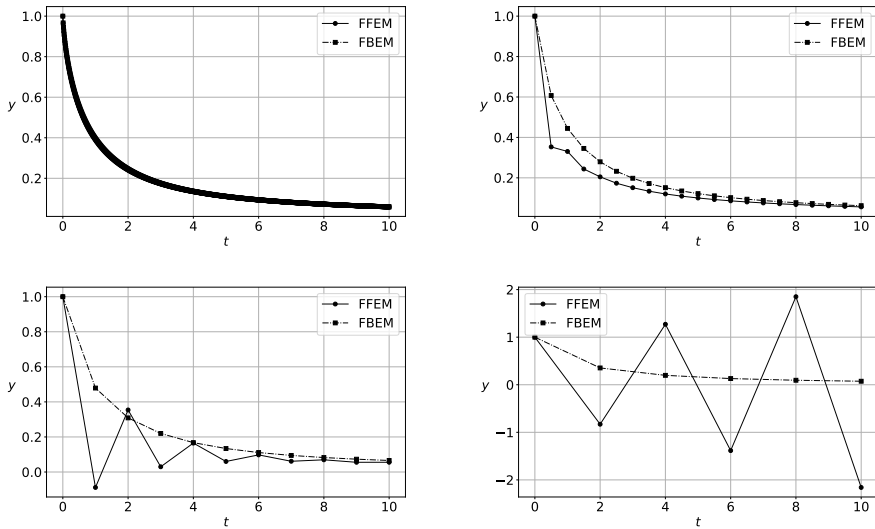FIG. 3: Numerical solution of (3.1) with $\alpha = 0.5$ and $h = 0.01, 0.5, 1, 2$.

FIG. 4: Numerical solution of (3.1) with $\alpha = 0.75$ and $h = 0.01, 0.5, 1, 2$.

Nevertheless, both the fractional Euler's methods give the constant numerical solution $y_n = 1$, $n = 0, 1, 2, \ldots$, which corresponds to the trivial constant solution
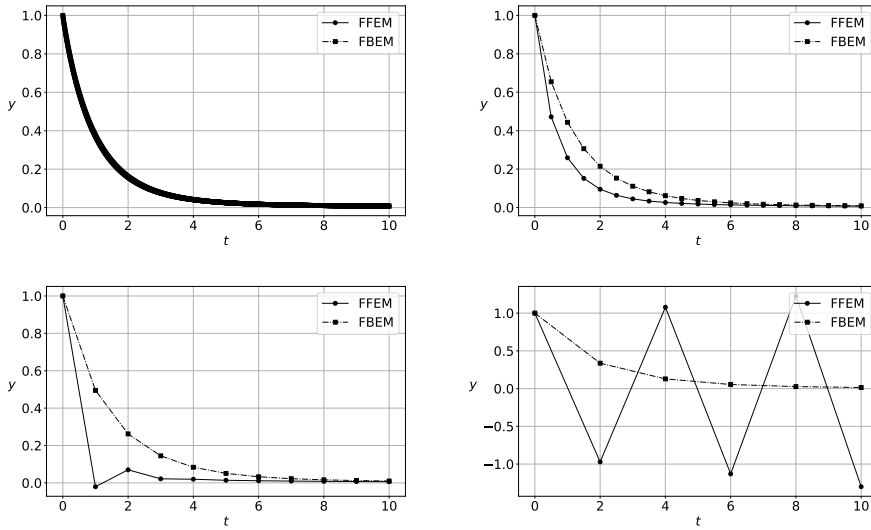
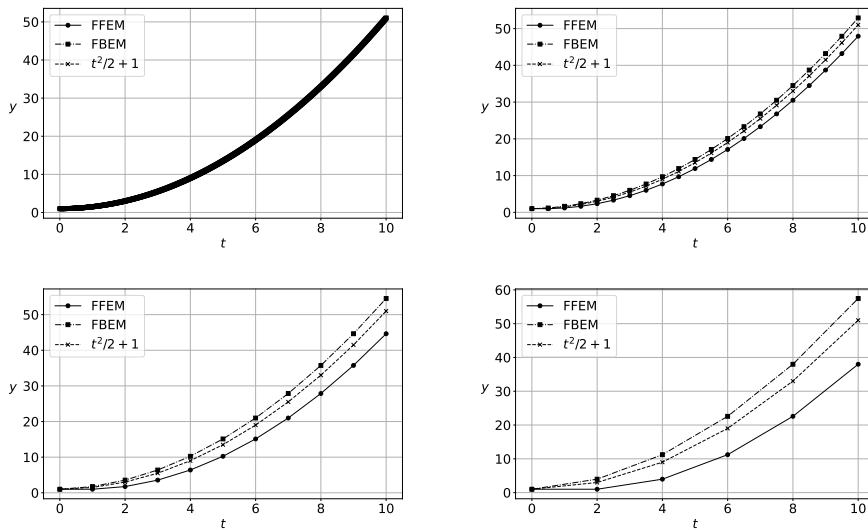FIG. 5: Numerical solution of (3.1) with $\alpha = 0.95$ and $h = 0.01, 0.5, 1, 2$.

FIG. 6: Numerical solution of (3.2) for $h = 0.01, 0.5, 1, 2$.

$y(t) = 1$ of (3.3). The uniqueness of solution is not ensured for this initial value problem.

t	$h = 0.1$	$h = 0.01$	$h = 0.001$
0	0.0000e-00	0.0000e-00	0.0000e-00
1	6.3616e-02	5.7214e-03	5.3422e-04
2	1.2314e-01	1.1157e-02	1.0544e-03
3	1.8108e-01	1.6523e-02	1.5714e-03
4	2.3814e-01	2.1849e-02	2.0865e-03
5	2.9463e-01	2.7149e-02	2.6004e-03
6	3.5069e-01	3.2429e-02	3.1134e-03
7	4.0641e-01	3.7694e-02	3.6257e-03
8	4.6187e-01	4.2947e-02	4.1374e-03
9	5.1710e-01	4.8189e-02	4.6487e-03
10	5.7214e-01	5.3422e-02	5.1595e-03

TAB. 1: Absolute errors of FFEM for (3.2).

t	$h = 0.1$	$h = 0.01$	$h = 0.001$
0	0.0000e-00	0.0000e-00	0.0000e-00
1	3.5315e-02	4.2737e-03	4.6575e-04
2	7.6014e-02	8.8387e-03	9.4555e-04
3	1.1818e-01	1.3474e-02	1.4286e-03
4	1.6119e-01	1.8148e-02	1.9135e-03
5	2.0475e-01	2.2848e-02	2.3996e-03
6	2.4873e-01	2.7568e-02	2.8866e-03
7	2.9303e-01	3.2303e-02	3.3743e-03
8	3.3760e-01	3.7051e-02	3.8625e-03
9	3.8239e-01	4.1809e-02	4.3513e-03
10	4.2737e-01	4.6575e-02	4.8405e-03

TAB. 2: Absolute errors of FBEM for (3.2).

4. CONCLUDING REMARKS

In the previous sections we have introduced forward and backward fractional Euler methods. We have compared numerical solutions obtained by these methods under various parameter settings. The computational costs significantly increase during the process since the computation of value y_{n+1} depends on approximation of the definite integral over $[0, t_n]$ in the case of FFEM and $[0, t_{n+1}]$ in the case of FBEM. The whole history of the solution must be taken into account. This phenomenon is common for all reasonable numerical schemes due to the Caputo fractional derivative nature. On this account there were considered parallel algorithms for efficient employment of available processors to deal with the vast number of computations to get the solution in reasonable time, see, e.g. [2], [8].

In the literature, we can also find some simplifying numerical approaches which neglect a part of the process' history. However, such techniques appeared to be unreliable in general, see [5] and references therein.

Python codes by the author related to the above discussed problems are available at: *https://math.fme.vutbr.cz/Home/kundrat/software*. With respect to the above formulated problems it is enough to modify the input part of the code to get numerical solution of own initial value problems.

As another valuable source of numerical programs, the paper [4] may serve. It introduces various numerical schemes and offers link to their Matlab codes. The paper also serves as a manual to these programs.

Finally, in connection with numerical methods for fractional differential equations, we mention recent paper [3] where open problems in fractional calculus are stated. Particularly error analysis of numerical methods for fractional differential equations is found unsatisfactory and it offers many research challenges.

References

[1] Diethelm, K., *The analysis of fractional differential equations*, Springer, Heidelberg, 2010.

[2] Diethelm, K., *An efficient parallel algorithm for the numerical solution of fractional differential equations*, Fract. Calc. Appl. Anal. **84** (3) (2011), 475–490.

[3] Diethelm, K., Kiryakova, V., Luchko, Y., Machado, J.A.T., Tarasov, V.E., *Trends, directions for further research, and some open problems of fractional calculus*, Nonlinear Dyn. **107** (2022), 3245–3270.

[4] Garrappa, R., *Numerical solution of fractional differential equations: A survey and software tutorial*, Mathematics **2018** (6) (2018), 1–23.

[5] Garrappa, R., *Neglecting nonlocality leads to unreliable numerical methods for fractional differential equations*, Commun. Nonlinear Sci. Numer. Simul. **70** (2019), 302–306.

[6] Li, C.P., Zeng, F.H., *Numerical methods for fractional calculus*, Chapman & Hall/CRC, Boca Raton, FL, 2015.

[7] Podlubny, I., *Fractional differential equations*, Academic Press, San Diego, CA, 1999.

[8] Rosu, F., *Parallel algorithm for numerical methods applied to fractional-order system*, Parallel algorithm for numerical methods applied to fractional-order system **21** (4) (2020), 701–707.

Institute of Mathematics, Faculty of Mechanical Engineering,
Brno University of Technology,
Technická 2, 616 69 Brno, Czech Republic
E-mail: Petr.Tomasek@vutbr.cz

NUMERICAL APPROACHES TO THE MODELLING OF QUASI-BRITTLE CRACK PROPAGATION

Jiří Vala

Abstract. Computational analysis of quasi-brittle fracture in cement-based and similar composites, supplied by various types of rod, fibre, etc. reinforcement, is crucial for the prediction of their load bearing ability and durability, but rather difficult because of the risk of initiation of zones of microscopic defects, followed by formation and propagation of a large number of macroscopic cracks. A reasonable and complete deterministic description of relevant physical processes is rarely available. Thus, due to significance of such materials in the design and construction of buildings, semi-heuristic computational models must be taken into consideration. These models generate mathematical problems, whose solvability is not transparent frequently, which limits the credibility of all results of ad hoc designed numerical simulations. In this short paper such phenomena are demonstrated on a simple model problem, covering both micro- and macro-cracking, with references to needful generalizations and more realistic computational settings.

1. Introduction

Cement-based composites, supplied by various type of fibre, rod, etc. reinforcement, are the most frequently used materials in building structures thorough the world. Their load bearing ability and durability is conditioned by the minimization of the risk of initiation and propagation of fracture. Due its rather complicated structure, the so-called quasi-brittle fracture can be expected here, using the nomenclature of [27], unlike simple fracture models as the brittle or ductile ones. In the rough qualitative classification, under mechanical, thermal, etc. loads 4 deformation stages can be distinguished: i) reversible elastic deformation, ii) initiation of zones of microscopic defects, iii) formation and propagation of systems of macroscopic cracks, iv) destruction of material structure, from local to total one. The development of advanced materials, structure and technologies can rarely come from the experience with classical ones, moreover the significant size effect limits

the replacement of the expensive results from long-time observations in situ by those from simplified laboratory experiment, thus some deeper both physical and mathematical analysis for the design of tools for computational modelling and simulation is required. This can be based on the principles of classical thermomechanics by [3], working with conservation of mass, (linear and angular) momentum and energy, supplied by appropriate constitutive relations. Unfortunately, their complete quantitative formulation, coming from exact microstructural considerations like [20], is not available because of a complicated material structure; consequently some semi-heuristic approach could be useful, as proposed by [13], although its well-possedness might be doubtful and further generalizations are needed.

Such time-delayed development of mathematical theory is typical for many algorithms needed in engineering problems: e.g. most engineering journals celebrated 70 years of the finite element method (FEM), connected with the design of wings of Boeing YB52 prototype by [30] (tested 15th April 1952, published 1956), whereas the history of the mathematical theory of FEM is about 15 years shorter, as evident from [36] and its references. Moreover, very detailed deterministic models suffer from the complicated (or quite impossible) identification of reasonable material characteristics for constitutive relations, generating non-trivial ill-posed inverse problems; thus all practical computational tools can be seen as certain compromises between the general physical theory, its simplified mathematical presentation and the design of effective and robust numerical algorithms, up to their software and hardware implementations, limited by the amount of time and money for such complete analysis, covering both the formal verification and the practical validation in the sense of [29]. Thus in this contributed conference paper we shall introduce a model problem based on a linear Neumann and Dirichlet boundary value and Cauchy initial value problem for one partial differential equation of evolution for i), coming from the conservation of momentum, modified by certain nonlinear terms covering ii) and iii) (Section 2), followed by its existence and convergence analysis (Section 3), with some references to useful generalizations (Section 4); this approach is not able to handle later stages of both ii) and iii) tending to iv) without substantial improvements, containing numerous open questions, not discussed here.

The principal idea for ii) is the implementation of certain damage factor $\mathfrak{D}$, following [13], relying on the nonlocal approach suggested by [7].The ill-possedness of such approach for most engineering formulations, if applied to the evaluation of strains or stresses in general, criticized by [8], unlike the preliminary existence analysis of [1] (valid for a pure Dirichlet boundary value problem), can be fortunately avoided here thanks to a careful (rather complicated) choice of $\mathfrak{D}$. The long tradition of an intuitive need of such regularization, motivated by micromechanical considerations, can be documented on [2]. For iii) the model of cohesive interfaces by [23] for activation and development of cracks for pre-defined potential crack positions can be used, applying the classical (extrinsic) formulation of the extended finite element method (XFEM) by [18], working with additional degrees of freedom (i. e. with new parameters for the evaluation of unknown function(s)) for adaptive local enrichment of standard FEM bases on all tips of macroscopic cracks and along them; alternatively the intrinsic version of XFEM by [10], modifying such bases in

a more complicated way without increase of degrees of freedom, can be utilized. The method of discretization of time, based on the properties of Rothe seqences, will be considered as the first choice for any time t from a time interval $\mathcal{I} = [0, T]$ with a prescribed finite positive time T; for the discretization of a deformable body Ω, including all boundary and interface conditions, in the 3-dimensional Euclidean space $\mathcal{R}^3$, supplied by some fixed Cartesian coordinate system $x = (x_1, x_2, x_3)$, XFEM is then available.

2. A MODEL PROBLEM

For simplicity (to avoid technical difficulties in proofs), let a deformable body Ω occupy a unit of a finite number of domains with Lipschitz continuous boundaries, whose unit consists of disjoint parts Θ (for Dirichlet boundary conditions), Γ (for Neumann boundary conditions) and Λ (for interface boundary conditions, required by iii) only); the pair of Cauchy initial conditions will be prescribed for $t = 0$. The standard notation of Lebesgue, Sobolev, Bochner–Sobolev, etc. (abstract) function spaces by [24], *Parts* 1 and 7, will be utilized. For the brevity we shall also introduce $H = L^2(\Omega)^3$, $M = L^\infty(\Omega)$, $X = L^2(\Gamma)^3$, $Z = L^2(\Lambda)$ and $V = \{v \in W^{1,2}(\Omega)^3 \colon v = o \text{ on } \Theta\}$ where o means the zero vector in $\mathcal{R}^3$, together with the notation of scalar products $(.,.)$ in H, $\langle .,. \rangle$ in X, $\langle .,. \rangle_*$ in Z and $((.,.))$ in $L^2(\Omega)^{3 \times 3}$; the symbol $[\,.\,]$ will be reserved for the jumps in normal components of values from V on Λ in the sense of traces, using some predefined orientation of unit normals $n = (n_1, n_2, n_3)$ to Λ, i.e. $[v] = v^+ - v^- \in Z$, v^+ and v^- understood as $v_1 n_1 + v_2 n_2 + v_3 n_3$ from both sides of Λ for any $v \in V$. We shall also use the notation $\varepsilon(v) \in L^2(\Omega)^{3 \times 3}_{\mathrm{sym}}$ for the tensor of small strains, well-known in the linear theory of elasticity, assuming $\varepsilon_{ij}(v) = (\partial v_i / \partial x_j + \partial v_j / \partial x_i)/2$ for all $i, j \in \{1, 2, 3\}$; all upper dots will be applied as abbreviations of $\partial / \partial t$.

The weak formulation of the linear momentum conservation reads

$$(2.1) \qquad (v, \rho(\ddot{u} + \alpha \dot{u})) + ((\varepsilon(v), \sigma)) = (v, f) + \langle v, g \rangle - \langle [v], \tau \rangle_*$$

for any virtual displacement $v \in V$, related to the reference configuration of Ω, Θ, Γ and Λ; here $\rho \in M$ denotes the material density, not lesser, everywhere on Ω, than some prescribed positive constant, and $\alpha \in M$ the always non-negative mass damping factor (forcing certain energy dissipation even for i), to respect the reality of a physically open system), $f \in L^2(\mathcal{I}, H)$ refers to the applied volume forces and $g \in C_L(\mathcal{I}, X)$ to the applied surface forces (where C_L denotes the Lipschitz continuity). Moreover (2.1) contains an unknown time-dependent actual displacement $u \in W^{2,2,2,2}(\mathcal{I}, V, V, V^*)$ (using V^* dual to V: for more details see [24], *Part* 7.1, namely $u, \dot{u} \in L^2(\mathcal{I}, V)$, $\ddot{u} \in L^2(\mathcal{I}, V^*)$, cf. the Gelfand triple $V \subset H \subset V^*$) and some still undefined stresses $\sigma \in L^2(\mathcal{I}, L^2(\Omega)^{3 \times 3}_{\mathrm{sym}})$ and interface tractions $\tau \in L^2(\mathcal{I}, Z)$, always normal to Λ, which must be evaluated from appropriate constitutive relations.

The linearized theory of elasticity, applicable to i), works with the empiric Hooke law $\sigma = C\varepsilon(u)$ with $C \in L^\infty(\Omega)^{(3 \times 3) \times (3 \times 3)}_{\mathrm{sym}}$ where C contains 21 different material characteristics in general (for any Boltzmann continuum the symmetry can be derived from the conservation of angular momentum), reducible up to the pair of

the well-known Lamé factors (or to the Young modulus and the Poisson coefficient in most engineering applications) in the isotropic case. In general, it is natural to assume $a \cdot Ca \geq ca \cdot a$ everywhere on Ω for any $a \in \mathcal{R}^{3 \times 3}_{\mathrm{sym}}$ and a positive constant c independent of $x \in \Omega$ and a; the central dot denotes the standard scalar product in $\mathcal{R}^{3 \times 3}$ here. In addition to α, the structural damping factor $\beta \in M$, not lesser, everywhere on Ω, than some prescribed positive constant, is required in the Kelvin parallel viscoelastic model, assuming $\sigma = C\varepsilon(u + \beta\dot{u})$. For ii) we shall use the seemingly slight modification of this relation in the form $\sigma = (1 - \mathfrak{D})C\varepsilon(u + \beta\dot{u})$ where the most delicate step is the calculation of the nonlocal factor of irreversible damage $\mathfrak{D}$ with values in M, always between 0 (for no damage) and some positive constant ς lesser than 1 (to avoid the total damage iv)), and depending on u in a rather complicated way, which will be sketched later. In particular, we are allowed to set $\mathfrak{D} = 0$ for $t = 0$, as required by (2.2) below, as well as for $t < 0$ formally in the difference schemes (3.1) and (3.2). For iii) the assessment of the traction-separation law, activated by reaching sufficient level of deformation energy on Λ locally, is needed; for all details see [16]. Here, for simplicity, we introduce only some continuous real function $\mathfrak{T}$ and insert $\tau = \mathfrak{T}([u])$ into (2.1). Thus we obtain

$$(2.2) \quad (v, \rho(\ddot{u} + \alpha\dot{u})) + ((\varepsilon(v), (1 - \mathfrak{D})C\varepsilon(u + \beta\dot{u}))) = (v, f) + \langle v, g \rangle - \langle [v], \mathfrak{T}([u]) \rangle_*$$

for any $v \in V$ again.

Such approach has been applied in [32] to the quasi-static version of (2.2) and in [33] to certain dynamical simplification of (2.2), relying on [13] in both cases; its limitation, excluding any realistic description of iv), has been discussed by [31] in details. However, this limitation might be not crucial for engineering applications because most practical numerical simulations are expected to detect the risk of material destruction in advance, whereas the detailed quantitative description of its subsequent disintegration is less interesting. Thus the more significant step in the upgrade of these formulations is some reasonable incorporation of different material behaviour under tension and compression, typical just for quasi-brittle cracking in cement-based composites. We shall adopt the access, coming from [22], modified by [11], referred as the "Mazars' model" in software packages, for simplicity; some significant later improvements will be mentioned in *Section* **??**. The final aim will be the evaluation of $\mathfrak{D}$ for (2.2).

As the 1st step, to preserve the objectivity of our analysis, let us evaluate (at certain fixed time step $t \in \mathcal{I}$ in practice, as will be evident from Section 3)the scalar principal values ε_i with $i \in \{1, 2, 3\}$ of $\varepsilon(u)$, i.e. the triple of eigenvalues from the condition $\det(\varepsilon(u) - \varepsilon_i I) = 0$, I being the unit matrix in $\mathcal{R}^3$. As the 2nd step, we can evaluate an equivalent strain $\bar{\varepsilon}$, using some bounded continuous functions ω of 6 real non-negative (at most 3 non-zero) arguments, as $\tilde{\varepsilon} = \omega(-\varepsilon_{1-}, \varepsilon_{1+}, -\varepsilon_{2-}, \varepsilon_{2+}, -\varepsilon_{3-}, \varepsilon_{3+})$ where ε_{i+} and ε_{i-} refer to the positive and negative parts of ε_i (for each $x \in \Omega$ locally); for an example of such admissible function ω see [5]. As the 3rd step, the nonlocal form of $\tilde{\varepsilon}$ reads

$$(2.3) \qquad \bar{\varepsilon}(x, t) = \int_\Omega \mathcal{K}(x, \tilde{x})\, \tilde{\varepsilon}(\tilde{x}, t)\, \mathrm{d}\tilde{x}, \qquad \int_\Omega \mathcal{K}(x, \tilde{x})\, \mathrm{d}\tilde{x} = 1$$

for any $x \in \Omega$, still at a fixed $t \in \mathcal{I}$, using some regularizing kernel $\mathcal{K} \in L^2(\Omega \times \Omega)$ in the sense of [9] (the choice $\mathcal{K}(x, \tilde{x}) = \delta(x - \tilde{x})$, δ being the Dirac measure, forcing $\bar{\varepsilon} = \tilde{\varepsilon}$, is not allowed intentionally), e.g. the Gaussian one, recommended by [11]; for its numerical approximation using various radial basis functions cf. [25]. As the 4th step, we have to evaluate the trial value $\mathfrak{D}_\star$, using some non-decreasing continuous functions ϖ of 1 real argument, returning values between 0 and ς, as $\mathfrak{D}_\star = \varpi(\bar{\varepsilon})$; for an example of such admissible function ϖ see [21]. The final 5th step, forcing the irreversibility of damage, can be then written in the form $\mathfrak{D}(., t) = \max_{0 \le \xi \le t} \mathfrak{D}_\star(., \xi)$.

The practical design of ω, $\mathcal{K}$ and ϖ is typically not easy, being conditioned, beyond the scope of this paper, by the careful design of material parameters, by the analysis of related sensitivity and inverse problems and by the extensive experimental research. Here we remind only the following result, derived (in 2 different ways) by [6], *Part* 2.2: an arbitrary operator introduced by (2.3) is compact as an operator from $L^2(\Omega)$ to itself. In particular, for each $t \in \mathcal{I}$ a sequence $\{\tilde{\varepsilon}^m(., t)\}_{m=1}^\infty$ with a weak limit $\tilde{\varepsilon}(., t)$ is converted to a sequence $\{\bar{\varepsilon}^m(., t)\}_{m=1}^\infty$ with a strong limit $\bar{\varepsilon}(., t)$. We shall see in Section 3 that such property for the approximation of $\mathfrak{D}$ by a sequence of simple functions on $\mathcal{I}$ will be required.

3. Existence and convergence analysis

To be able to evaluate the time development of u, we need to set the Cauchy initial conditions $u(., 0) = o$ (no displacements occur in the reference configuration) and $\dot{u}(., 0) = w$ where the initial displacement rates $w \in V$ must be prescribed. Then the existence of solution of (2.2) can be verified in the constructive way, using the limit passage for a positive integer $m \to \infty$ from

$$(v, \rho \tilde{u}^m) + (v, \alpha \rho \dot{u}^m) + ((\varepsilon(v), (1 - \bar{\mathfrak{D}}_\times^m) C \varepsilon(\bar{u}^m)))$$

$$(3.1) \qquad + ((\varepsilon(v), (1 - \bar{\mathfrak{D}}_\times^m) \beta C \varepsilon(\dot{u}^m))) = (v, f^m) + \langle v, g^m \rangle - \langle [v], \mathfrak{T}([u_\times^m]) \rangle_*$$

on $\mathcal{I}$ where $h = T/m$ for brevity. The following approximations are used in (3.1): $u^m(t) = u_{s-1}^m + (u_s^m - u_{s-1}^m)(t - (s-1)h)/h$ (linear Lagrange splines) and $\bar{u}^m(t) = u_s^m$, assuming $(s-1)h < t \le sh$ (simple functions) with $s \in \{1, \ldots, m\}$, whereas $u_\times^m(t) = u^m(t - h)$ and $\bar{u}_\times^m(t) = \bar{u}^m(t - h)$ (retarded versions of preceding types of functions), using $u_0^m = o$ and $u_{-1}^m = -hw$; $\bar{\mathfrak{D}}_\times^m$ must be considered as $\mathfrak{D}$ evaluated for $\bar{u}_\times^m$. Such functions generate 4 different types of Rothe sequences $\{u^m\}_{m=1}^\infty$, $\{\bar{u}^m\}_{m=1}^\infty$, $\{u_\times^m\}_{m=1}^\infty$ and $\{\bar{u}_\times^m\}_{m=1}^\infty$; we shall need $\{\tilde{u}^m\}_{m=1}^\infty$ defined by $\tilde{u}^m = (\dot{u}^m - \dot{u}_\times^m)/h$, too. Using the values u_{-1}^m, u_0^m and u_s^m with $s \in \{1, \ldots, m\}$ for a fixed m, we can omit all upper indices m for brevity and rewrite (3.1) to its more transparent form for step-by-step evaluations of u_s from formally linear elliptic equations (whose solvability can be verified applying the Lax‑Milgram theorem) at all times $t = sh$

$$(v, \rho \mathcal{D}^2 u_s) + (v, \alpha \rho \mathcal{D} u_s) + ((\varepsilon(v), (1 - \mathfrak{D}_{s-1}) C \varepsilon(u_s)))$$

$$(3.2) \qquad + ((\varepsilon(v), (1 - \mathfrak{D}_{s-1}) C \varepsilon(\mathcal{D} u_s))) = (v, f_s) + \langle v, g_s \rangle - \langle [v], \mathfrak{T}([u_{s-1}]) \rangle_*,$$

taking f_s and g_s as the mean values of $f(., t)$, $g(., t)$ over t between $(s-1)h$ and sh (which is the Clément quasi-interpolation by [24], *Part* 8.2) and using the

obvious notation of the 1st and 2nd relative differences $\mathcal{D}u_s = (u_s - u_{s-1})/h$ and $\mathcal{D}^2 u_s = (\mathcal{D}u_s - \mathcal{D}u_{s-1})/h$.

In particular, for $v = 2h\mathcal{D}u_s$ in (3.2) we can derive some useful a priori estimates. Namely, using $s \in \{1, \ldots, r\}$ as the Einstein summation index for an arbitrary fixed $r \in \{1, \ldots, m\}$, for the left-hand side of (3.2) we have

$$(3.3) \quad 2h(\mathcal{D}u_s, \rho\mathcal{D}^2 u_s) = (u_r - u_{r-1}, \rho(u_r - u_{r-1}))/h^2 - (w, \rho w)$$
$$+ (u_s - 2u_{s-1} + u_{s-2}, \rho(u_s - 2u_{s-1} + u_{s-2}))/h^2\,,$$

$$2h(\mathcal{D}u_s, \alpha\rho\mathcal{D}u_s) = 2(u_s - u_{s-1}, \alpha\rho(u_s - u_{s-1}))/h\,,$$

$$2h((\varepsilon(\mathcal{D}u_s), (1-\mathfrak{D}_{s-1})\beta C\varepsilon(\mathcal{D}u_s))) = ((\varepsilon(u_s - u_{s-1}), (1-\mathfrak{D}_{s-1})\beta C\varepsilon(u_s - u_{s-1})))/h\,,$$

$$2h((\varepsilon(\mathcal{D}u_s), (1-\mathfrak{D}_{s-1})C\varepsilon(u_s))) = ((\varepsilon(u_r), (1 - \mathfrak{D}_{r-1})C\varepsilon(u_r)))$$
$$+ ((\varepsilon(u_s - u_{s-1})(1-\mathfrak{D}_{s-1})C\varepsilon(u_s - u_{s-1}))) - ((\varepsilon(u_{s-1})(\mathfrak{D}_{s-1}-\mathfrak{D}_{s-2})C\varepsilon(u_{s-1})))\,.$$

To avoid very long formulae, for the right-hand side of (3.2) it is sufficient, applying the Cauchy–Schwarz and Young inequalities with any positive ϵ, using the standard norms in H and V, to present the estimates

$$(3.4)\ 2h(\mathcal{D}u_s, f_s) \leq \frac{\epsilon}{h}\sum_{s=1}^{r}\|u_s - u_{s-1}\|_H^2 + \frac{h}{\epsilon}\sum_{s=1}^{r}\|f_s\|_H^2\,,$$

$$2h\langle \mathcal{D}u_s, g_s\rangle = 2\langle u_r, g_r\rangle - 2\langle u_{s-1}, g_s - g_{s-1}\rangle$$
$$\leq \gamma\epsilon\|u_r\|_V^2 + \frac{1}{\epsilon}\|g_r\|_X^2 + \gamma\epsilon h\sum_{s=1}^{r}\|u_{s-1}\|_V^2 + \frac{1}{\epsilon h}\sum_{s=1}^{r}\|g_s - g_{s-1}\|_X^2\,,$$

$$-2h\langle [\mathcal{D}u_s], \mathfrak{T}([u_{s-1}])\rangle_* \leq 4\gamma\epsilon h\sum_{s=1}^{r}\|u_s\|_V^2 + \frac{4\gamma}{\epsilon h}\|u_s - u_{s-1}\|_V^2$$

where the positive constant γ stems from the trace theorem (duplicated on Λ, unlike Γ). Using the upper estimates (3.4) and the similar lower ones generated by (3.3) (all details must be left to the curious reader), the discrete Gronwall lemma, applied to (3.2), gives

$$(3.5) \qquad \|\mathcal{D}u_r\|_H^2 + h^2\sum_{s=1}^{r}\|\mathcal{D}^2 u_s\|_H^2 + h\sum_{s=1}^{r}\|\mathcal{D}u_s\|_V^2 + \|u_r\|_V^2 \leq \mathcal{C}$$

where $\mathcal{C}$ is some fixed positive constant (sufficiently large, depending on T).

In terms of (3.1), the a priori estimate (3.5) can be interpreted as follows: $h\{\tilde{u}^m\}_{m=1}^{\infty}$ is bounded in $L^2(I, H)$ and $\{\dot{u}^m\}_{m=1}^{\infty}$ is bounded in $L^2(I, V)$, whereas $\{\bar{u}^m\}_{m=1}^{\infty}$ and $\{\bar{u}_\times^m\}_{m=1}^{\infty}$ are bounded in $L^\infty(I, V)$. The Eberlein‑Shmul'yan theorem then guarantees that, up to subsequences, $\{\dot{u}^m\}_{m=1}^{\infty}$ has a weak limit $u' \in L^2(I, V)$, whereas $\{\bar{u}^m\}_{m=1}^{\infty}$ and $\{\bar{u}_\times^m\}_{m=1}^{\infty}$ have their weak limits $\bar{u}, \bar{u}_\times \in L^\infty(I, V)$. Finally, using these results together with (3.1), we can see that $\{\tilde{u}^m\}_{m=1}^{\infty}$ has a weak limit $u'' \in L^2(I, V^*)$. For any $t \in \mathcal{I}$ let us now define

$$(3.6)u(.,t) = \int_0^t u'(.,\xi)\,\mathrm{d}\xi\,, \quad \hat{u}(.,t) = w + \int_0^t u''(.,\xi)\,\mathrm{d}\xi\,, \quad \hat{u}^m(.,t) = w + \int_0^t \tilde{u}^m(.,\xi)\,\mathrm{d}\xi\,,$$

valid for any positive integer m in the last case. Since the Aubin‑Lions lemma yields also the strong convergence of $\{\bar{u}^m\}_{m=1}^{\infty}$ to $\bar{u}$ and of $\{\bar{u}_\times^m\}_{m=1}^{\infty}$ to $\bar{u}_\times$ in $L^2(I, H)$ (in both cases), it is not difficult to identify u with both $\bar{u}$ and $\bar{u}_\times$ and $\dot{u}$ with both u' and $\hat{u}$: namely the square of the norm of $u^m - \bar{u}^m$ in $L^2(I, H)$ can be estimated,

using (3.5), from above as the sum of all $\|u_s - u_{s-1}\|_H^2 = h^2\|\mathcal{D}u_s\|_H^2 \leq \mathcal{C}h^2$ over $s \in \{1, \ldots, m\}$, lesser or equal to $\mathcal{C}Th$, vanishing with $h \to 0$, thus $u = \bar{u}$, etc. Even the (seemingly strange) 2nd left-hand-side additive term of (3.5) is required here for the identification of $\hat{u}$ with $\dot{u}$: considering $\{\hat{u}^m\}_{m=1}^{\infty}$ by the last relation of (3.6), we obtain $\hat{u}^m(t) - \dot{u}^m(t) = (t - sh)\mathcal{D}^2 u_s$, thus the square of the norm of $\hat{u}^m - \dot{u}^m$ in $L^2(I, H)$ is just the sum of all $\|(t - sh)\mathcal{D}^2 u_s\|_H^2$, whose upper bound is $\mathcal{C}h/3$, tending to zero with $h \to 0$. This enables us to perform the limit passage from (3.1) to (2.2) where $u, \dot{u} \in L^2(\mathcal{I}, V)$ and $\ddot{u} \in L^2(\mathcal{I}, V^*)$, which can be expressed as $u \in W^{2,2,2,2}(\mathcal{I}, V, V, V^*)$.

4. CONCLUSIONS AND GENERALIZATIONS

The formulation and analysis of a model problem, supplied by the main ideas of proofs, in this paper was intended as the demonstration of numerical considerations for a class of initial and boundary value problems for partial differential equations of evolution, connected with extraordinarily significant tasks of modelling and simulation of behaviour of advanced materials, building end structures. Some preliminary computational results, referring namely to [32] and [33], for i), ii) and iii) have been presented (with numerous illustrative figures) and discussed in [34]. The much more detailed analysis is under development, including e. g. such processes as dynamics of multiple contacts / impacts of deformable bodies with potential micro- and macro-cracking by [26], requiring an explicit time integration scheme like [4], as well as an appropriate parallel / distributed computing platform.

The approach of [26] demonstrates also the need of incorporation of a complete set of conservation laws by the 1st principle of classical thermodynamics in formulations covering several physical processes (known as "multiphysics" in the unofficial language of scientific computing), together with the compatibility of constitutive relations with the 2nd and 3rd ones, respecting the finite (not only linearized small) strains and related stresses. Numerous inspirations can be found in the reviews of crack branching [28] and of XFEM-based simulations [17], opening the possibility of computational modelling of physically realistic development of all active interfaces Λ. Since only the linear elastic and viscous components have been combined in our model problem, one natural generalization could be the proper analysis of plastic zones, observed namely on crack tips in practice, combining the intuitive engineering approach of [12] with the deep mathematical analysis by [19]. Also the thermodynamical study of initiation and propagation of anisotropic damage, introduced by [14], revised by [35] and [15], taking $\mathfrak{D}$ as matrix characteristics, should belong to the research priorities for the next years.

REFERENCES

[1] Altan, S., *Existence in nonlocal elasticity*, Arch. Mech. **47** (1989), 25–36.

[2] Bažant, Z.P., *Why continuum damage is nonlocal: micromechanics arguments*, J. Eng. Mech. **117** (1991), 1070–1089.

[3] Bermúdez de Castro, A., *Continuum Thermomechanics*, Birkhäuser, Basel, 2005.

[4] Bybordiani, M., Dias da Costa, D., *A consistent finite element approach for dynamic crack propagation with explicit time integration*, Comput. Methods Appl. Mech. Eng. **376** (2021), 1–32, 113652.

[5] de Vree, J.H.P., Brekelmans, W.A.M., van Gils, M.A.J., *Comparison of nonlocal approaches in continuum damage mechanics*, Comput. Struct. **55** (1995), 581–588.

[6] Drábek, P., Milota, I., *Methods of Nonlinear Analysis*, Birkhäuser, Basel, 2013.

[7] Eringen, A.C., *Theory of Nonlocal Elasticity and Some Applications*, Tech. report, Princeton University, Princeton, 1984.

[8] Evgrafov, A., Bellido, J.-C., *From nonlocal Eringen's model to fractional elasticity*, Math. Mech. Solids **24** (2019), 1935–1953.

[9] Fasshauer, G.E., Ye, Q., *Reproducing kernels of generalized Sobolev spaces via a Green function approach with distributional operators*, Numer. Math. **119** (2011), 585–611.

[10] Fries, T.P., Belytschko, T., *The intrinsic XFEM: a method for arbitrary discontinuities without additional unknowns*, Int. J. Numer. Methods Eng. **68** (2006), 1358–1385.

[11] Giry, C., Dufour, F., Mazars, J., *Stress-based nonlocal damage model*, Int. J. Solids Struct. **48** (2011), 3431–3443.

[12] Hashiguchi, K., *Elastoplasticity Theory*, Springer Berlin, 2014.

[13] Havlásek, P., Grassl, P., Jirásek, M., *Analysis of size effect on strength of quasi-brittle materials using integral-type nonlocal models*, Eng. Fract. Mech. **157** (2016), 72–85.

[14] Ju, J.W., *Isotropic and anisotropic damage variables in continuum damage*, J. Eng. Mech. **116** (1990), 2764–2770.

[15] Kamińska, I., Szwed, A., *A thermodynamically consistent model of quasibrittle elastic damaged materials based on a novel Helmholtz potential and dissipation function*, MDPI Materials **14** (2021), 1–30, 6323.

[16] Kozák, V., Chlup, Z., Padělek, P., Dlouhá, I., *Prediction of the traction separation law of ceramics using iterative finite element modelling*, Solid State Phenomena **258** (2017), 186–189.

[17] Li, H., Li, J., Yuan, H., *A review of the extended finite element method on macrocrack and microcrack growth simulations*, Theor. Appl. Fract. Mech. **97** (2018), 236–249.

[18] Mariani, S., Perego, U., *Extended finite element method for quasi-brittle fracture*, Int. J. Numer. Meth. Engn. **58** (2003), 103–126.

[19] Mielke, A., Roubíček, T., *Rate-Independent Systems*, Springer, New York, 2015.

[20] Mousavi, S.M., *Dislocation-based fracture mechanics within nonlocal and gradient elasticity of bi-Helmholtz type*, Int. J. Solids Struct. **87** (2016), 92–93, 105–120.

[21] Peerlings, R.H.J., R.de Borst, Brekelmans, W.A.M., Geers, M., *Gradient enhanced damage modelling of concrete fracture*, Int. J. Numer. Anal. Methods Geomech. **3** (1998), 323–342.

[22] Pijaudier-Cabot, G., Mazars, J., *Damage models for concrete*, Handbook of Materials Behavior Models (Lemaitre, J., ed.), Academic Press, Cambridge (Massachusetts, USA), 2001, pp. 500–512.

[23] Pike, M.G., Oskay, C., *XFEM modeling of short microfiber reinforced composites with cohesive interfaces*, Finite Elem. Anal. Des. **106** (2015), 16–31.

[24] Roubíček, T., *Nonlinear Partial Differential Equations with Applications*, Birkhäuser, Basel, 2005.

[25] Skala, V., *A practical use of radial basis functions interpolation and approximation*, Investigación Operacional **37** (2016), 137–144.

[26] Štekbauer, H., Němec, I., Lang, R., Burkart, D., ValaSte22, J., *On a new computational algorithm for impacts of elastic bodies*, Appl. Math. **67** (2022), 28 pp., in print.

[27] Sumi, Y., *Mathematical and Computational Analyses of Cracking Formation*, Springer, Tokyo, 2014.

[28] Sun, Y., Edwards, M.G., Chen, B., Li, C., *A state-of-the-art review of crack branching*, Eng. Fract. Mech. **257** (2021), 1–33, 108036.

[29] Szabó, B., Babuška, I., *Finite Element Analysis: Method, Verification and Validation*, J. Wiley & Sons, Hoboken, 2021.

[30] Turner, M.J., Clough, R.W., Martin, H.C., Top, L.J., *Stiffness and deflection analysis of complex structures*, Journal of the Aeronautical Sciences **23** (1956), 805–823.

[31] Vala, J., *On a computational smeared damage approach to the analysis of strength of quasi-brittle materials*, WSEAS Trans. Appl. Theor. Mech. **16** (2021), 283–292.

[32] Vala, J., Kozák, V., *Computational analysis of quasi-brittle fracture in fibre reinforced cementitious composites*, Theor. Appl. Fract. Mech.. **107** (2020), 1–8, 102486.

[33] Vala, J., Kozák, V., *Nonlocal damage modelling of quasi-brittle composites*, Appl. Math. **66** (2021), 701–721.

[34] Vala, J., Kozák, V., Jedlička, M., *Scale bridging in computational modelling of quasi-brittle fracture of cementitious composites*, Solid State Phenomena **325** (2021), 56–64.

[35] Vilppo, J., Kouhia, R., Hartikainen, J., Kolari, K., Fedoroff, A., Calonius, K., *Anisotropic damage model for concrete and other quasi-brittle materials*, Int. J. Solids Struct. **225** (2021), 1–13, 111048.

[36] Zlámal, M., *On the finite element method*, Numer. Math. **12** (1968), 394–409.

INSTITUTE OF MATHEMATICS AND DESCRIPTIVE GEOMETRY,
FACULTY OF CIVIL ENGINEERING, BRNO UNIVERSITY OF TECHNOLOGY,
VEVEŘÍ 331/95, 602 00 BRNO, CZECH REPUBLIC
E-mail: `vala.j@fce.vutbr.cz`

CONTENTS

INSTRUCTIONS TO AUTHORS

Archivum Mathematicum welcomes the submission of high quality research papers that are of interest to a substantial numbers of its readers.

The submission of a paper implies the author's assurance that it has not been published or submitted for publication elsewhere.

Papers to be considered for publication should be sent electronically (in the form of a pdf file) to the Editorial Office. The authors are invited to suggest the Editor(s) to be associated with the manuscript during the submission process.

Manuscripts should contain a short abstract, 2020 Mathematics Subject Classification, key words, and author's address including the e-mail address. If the title is long, a shortened version should be supplied for the running heads. Joint authors should indicate on the manuscript which of them should receive the proof.

Authors are encouraged to use TEX typesetting software to prepare their papers. The LATEX macro package is especially recommended. Once an article has reached its final form and is accepted for publication, authors will be asked to supply the appropriate electronic file. This should be a single file, with any relevant macros included in the preamble, and named after the article number (as assigned by the Editorial Office, e.g. AM1234.tex). Figures in PostScript form are also welcome, and should be submitted at the same time as the TEX file for the paper.

One proof is supplied, which should be corrected and returned promptly.

MUNI
PRESS

MUNI
SCI