Nonlinear Differ. Equ. Appl. 20 (2013), 1243–1275

1021-9722/13/031243-33
published online November 28, 2012
DOI 10.1007/s00030-012-0207-9

Nonlinear Differential Equations and Applications NoDEA

Nonlinearly coupled thermo-visco-elasticity

Tomáš Roubíček

Abstract. The d-dimensional thermo-visco-elasticity system for Kelvin–Voigt-type materials at small strains with a general nonlinear coupling is considered. Thermodynamical consistency leads to a heat capacity dependent both on temperature and on the strain. Using higher-gradient theory, namely the concept of so-called second-grade non-simple materials (or of hyper-stresses), existence of a weak solution to a system arising after an enthalpy-type transformation is proved by a suitably regularized Rothe method, fine a-priori estimates for the temperature gradient performed for the coupled system, and a subsequent limit passage.

Mathematics Subject Classification (2010). Primary 35K55; Secondary 74A30, 74F05, 74H20, 80A17.

Keywords. Kelvin–Voigt rheology, Small strains, Nonsimple materials, Nonlinear thermodynamics, Weak solutions.

1. Introduction

Thermodynamical systems governed by specific free energy are treated in mathematical literature with only few exception only under partly linearized thermo-coupling. More specifically, considering this free energy as $\psi = \psi(z, \theta)$ with z some mechanical-like variable and θ temperature, it means the simplifying assumption $\psi'''_{z\theta\theta} = 0$. This admits only a partly linearized ansatz of the type $\psi(z, \theta) = \phi_0(z) + \theta\phi_1(z) + \phi_2(\theta)$, cf. Remark 2.1 and Example 5.4 below, and leads to the heat capacity $-\theta\psi''_{\theta\theta}$ that is independent of z. Such partial linearization is well acceptable because also a lot of physical studies use such a linearization, usually meant around some reference temperature. Anyhow, this special form has surely some limitations in application and it is certainly an ultimate challenge to allow for $\psi'''_{z\theta\theta} \neq 0$. This is just the goal of this article, which is demonstrated on a quite canonical example from

Partially supported from the grants A 100750802 (GA AV ČR), LC 06052 (MŠMT ČR), 201/09/0917 and 201/10/0357 (GA ČR), and the institutional support RVO: 61388998 (ČR).

solid mechanics, namely the thermo-visco-elastic system for materials without internal parameters in Kelvin–Voigt rheology at small strains. Then $z = e$ with e the small-strain tensor, yet the results apply more generally for cases where z is not necessarily e and even not necessarily a mechanical variable at all, cf. Remark 5.7 below.

It seems to be an ultimate need to use a gradient theory for the mechanical part, namely the concept of so-called second-grade nonsimple materials (also called multipolar solids), cf. e.g. [18,20,33,46], alternatively also referred as the concept of hyper-stresses [35] or also of microforces or capillarity. This concept has already been used even for the linearly coupled problems (where some results are available even without higher-order terms, cf. Remark 2.1 below) rather for the reason to allow for possibly nonconvex free energies $\psi(\cdot, \theta)$, see [7, Chap.5] or [40, Sect. 3.4] for a survey of broad references in the context of smart materials. Energies which are nonconvex in terms of e are also broadly used also in rock mechanics, accompanied by experimental data but not accompanied by any mathematical analysis even in isothermal case, cf. e.g. [21,27,28].

A general nonlinear coupling leading to the heat capacity depending on both θ and the mechanical variables occurred in mathematical literature in a particular context: Colli and Sprekels [11,12] handled a special mechanical system with internal parameters and, of course, involving also gradients of this parameters, similarly as also in [10,22], then Krejčí and Sprekels [23,24] investigated coupling through a certain hysteresis operator in one dimension. Besides, Krejčí and Stefanelli [25] have investigated a special problem with thermodynamical coupling through an internal parameter in one dimension.

Recently, Pawłow and Zajaçkowski [36,37] have used regularity in the 2- or 3-dimensional situations under the assumption of a smooth domain with zero Dirichlet boundary conditions, a regular external loading, and a special form of the higher-gradient term, i.e. a special $\mathbb{H}$, allowing however for zero higher viscosity, i.e. for $\mathbb{G} = 0$ in our notation below.

The novelty of the presented article is in using the enthalpy transformation suitably modified to cope with the heat capacity dependent also on the mechanical variables, which creates additional nonlinear terms in the transformed system; this is in some sense similar to the transformation devised by Niezgódka and Pawłow [32] where dependence on space/time variables has been treated in an analogous way. Another novelty is also the usage of the fully implicit time discretisation (= Rothe's method) combined with a certain monotone-type regularization which guarantees existence of the discrete solution but requires a particularly careful limit passage (just because of this regularization). The definite advantage (comparing e.g. to a Galerkin procedure) is that the comparison principle (yielding here the ultimately needed non-negativity of temperature) is at disposal already for the discrete solution and the limit passage thus may depend just on one parameter. Also, careful and repeated usage of Gagliardo–Nirenberg's interpolation gives estimates for a natural qualification of the external loading without any regularity. This seems to be quite universally applicable approach, cf. [44,45] for applications to some problems with internal parameters.

The paper is organised as follows: Sect. 2 formulates the problem and relates it to the standard thermodynamical scheme. In Sect. 3, the problem is transformed by introducing another variable replacing temperature and the definition of a weak solution is devised. In Sect. 4 the existence of such weak solutions is proved by the Rothe method, exploiting also fine a-priori estimates of the (transformed) temperature gradient. Eventually, some aspects of the used data qualification and the hyper-stress concept are discussed in Sect. 5.

2. Problem formulation and underlying thermodynamics

Let $\Omega \subset \mathbb{R}^d$ be a bounded domain with a boundary $\Gamma := \partial\Omega$ which is assumed smooth for the purpose of the derivation of the model in this section. We consider a fixed time horizon $T > 0$, and abbreviate $I := (0, T)$, $Q := I \times \Omega$, and $\Sigma := I \times \Gamma$. Let us list the basic notation we will use:

$u : \Omega \to \mathbb{R}^d$ the displacement,
$\theta : \Omega \to \mathbb{R}$ temperature,
$e = e(u) = \frac{1}{2}(\nabla u)^\top + \frac{1}{2}\nabla u$ the small-strain tensor,
ψ specific free energy,
ξ specific dissipation rate,
σ stress, $\sigma = \sigma(e(u), e(\frac{\partial u}{\partial t}), \theta)$,
$\mathfrak{h}$ hyper-stress, $\mathfrak{h} = \mathfrak{h}(\nabla e(u), \nabla e(\frac{\partial u}{\partial t}))$,
ϱ mass density,
$\mathfrak{c}$ is heat capacity, $\mathfrak{c} = \mathfrak{c}(e, \theta)$,
$\mathbb{K}$ a 2nd-order heat-conduction tensor, $\mathbb{K} = \mathbb{K}(e, \theta)$,
$\mathbb{D}$ a 4th-order viscous-like dissipation tensor,
$\mathbb{H}$ a 6th-order tensor of "hyper" elastic response,
$\mathbb{G}$ a 6th-order tensor of "hyper" viscous response,
g the prescribed specific bulk force,
h the prescribed surface force,
f the external prescribed heat flux through the boundary Γ.

The fully nonlinearly-coupled Helmholtz free energy ψ will need also to use a gradient theory for e, cf. (3.5) below, which leads us to postulating an enhanced free energy $\hat{\psi}$ in the form

$$\hat{\psi}(e, \theta, \nabla e) := \psi(e, \theta) + \frac{1}{2}\mathbb{H}(\nabla e)^2. \tag{2.1}$$

Further we postulate the dissipation rate

$$\xi(\dot{e}, \nabla\dot{e}) := \mathbb{D}(\dot{e})^2 + \mathbb{G}(\nabla\dot{e})^2. \tag{2.2}$$

In (2.1)–(2.2) we abbreviated

$$\mathbb{H}(\nabla e)^2 = \sum_{i,j,k,l,m,n=1}^{d} \mathbb{H}_{ijklmn} \frac{\partial e_{ij}}{\partial x_k} \frac{\partial e_{lm}}{\partial x_n}, \tag{2.3a}$$

$$\mathbb{D}(\dot{e})^2 = \sum_{i,j,k,l=1}^{d} \mathbb{D}_{ijkl}\dot{e}_{ij}\dot{e}_{kl}\,, \qquad \text{and} \qquad \mathbb{G}(\nabla\dot{e})^2 = \sum_{i,j,k,l,m,n=1}^{d} \mathbb{G}_{ijklmn}\frac{\partial \dot{e}_{ij}}{\partial x_k}\frac{\partial \dot{e}_{lm}}{\partial x_n}. \tag{2.3b}$$

The gradient term $\frac{1}{2}\mathbb{H}(\nabla e)^2$ is often used in the static variational problems (either generalized non-quadratic as e.g. in [2, Sect. 6] or even in a simplified variant as $\mathbb{H}_{ijklmn} = \epsilon\delta_{kl}\delta_{mn}$ or $\epsilon\delta_{km}\delta_{ln}$ with δ the Kronecker symbol to yield respectively terms of the type $\epsilon|\Delta u|^2$ or $\epsilon|\nabla^2 u|^2$, cf. e.g. [26,29]) without scrutinizing explicitly any natural boundary conditions as (2.7b,c) and (3.10a,b) below. The quadratic form of (2.2) is related to linear viscosity, which is relevant for relatively slow (and not activated) mechanical processes. Up to the factor $\frac{1}{2}$, this form is simultaneously a (pseudo)potential of the dissipative forces, cf. (2.4b,c) below. One can think about $\mathbb{G} = \tau_{\rm relax}\mathbb{H}$ with a (presumably only small) relaxation time $\tau_{\rm relax} > 0$ in the Kelvin–Voigt rheological model.

We define the *entropy* classically by *Gibbs' formula*

$$s = -\hat{\psi}'_\theta(e,\theta,\nabla e) = -\psi'_\theta(e,\theta), \tag{2.4a}$$

and, postulating the Kelvin–Voigt rheology, we further define the stress σ by

$$\sigma = \sigma(e,\dot{e},\nabla e,\nabla\dot{e},\theta) := \hat{\psi}'_e(e,\theta,\nabla e) + \frac{1}{2}\xi'_{\dot{e}}(\dot{e},\nabla\dot{e}) = \psi'_e(e,\theta) + \mathbb{D}\dot{e}, \tag{2.4b}$$

and the so-called hyperstress $\mathfrak{h}$ by

$$\mathfrak{h} = \mathfrak{h}(e,\dot{e},\nabla e,\nabla\dot{e},\theta) := \hat{\psi}'_{\nabla e}(e,\theta,\nabla e) + \frac{1}{2}\xi'_{\nabla\dot{e}}(\dot{e},\nabla\dot{e}) = \mathbb{H}\nabla e + \mathbb{G}\nabla\dot{e}, \tag{2.4c}$$

where the ansatz (2.1)–(2.2) has been taken into account. Standardly, the *entropy equation* reads as

$$\theta\frac{\partial s}{\partial t} = \xi_1\Big(e\Big(\frac{\partial u}{\partial t}\Big), \nabla e\Big(\frac{\partial u}{\partial t}\Big)\Big) + \operatorname{div} j, \tag{2.5a}$$

where j is the heat flux, and the *force balance* reads as

$$\varrho\frac{\partial^2 u}{\partial t^2} - \operatorname{div}(\sigma - \operatorname{div}\mathfrak{h}) = g. \tag{2.5b}$$

Thus, in view of (2.4b,c), the equation (2.5b) takes the form

$$\varrho\frac{\partial^2 u}{\partial t^2} - \operatorname{div}\Big(\mathbb{D}e\Big(\frac{\partial u}{\partial t}\Big) + \psi'_e(e(u),\theta) - \operatorname{div}\Big(\mathbb{G}\nabla e\Big(\frac{\partial u}{\partial t}\Big) + \mathbb{H}\nabla e(u)\Big)\Big) = g. \tag{2.6a}$$

Considering an anisotropic nonlinear *Fourier law* for the heat flux $j = -\mathbb{K}(e(u),\theta)\nabla\theta$, from (2.5a) we obtain the *heat equation*

$$\mathfrak{c}(e(u),\theta)\frac{\partial\theta}{\partial t} - \operatorname{div}(\mathbb{K}(e(u),\theta)\nabla\theta) = \mathbb{D}\Big(e\Big(\frac{\partial u}{\partial t}\Big)\Big)^2 + \mathbb{G}\Big(\nabla e\Big(\frac{\partial u}{\partial t}\Big)\Big)^2 \\ + \theta\psi''_{e\theta}(e(u),\theta) : e\Big(\frac{\partial u}{\partial t}\Big) \qquad \text{with} \quad \mathfrak{c}(e,\theta) = -\theta\psi''_{\theta\theta}(e,\theta). \tag{2.6b}$$

We complete the system (2.6) by the initial conditions

$$u(0,\cdot) = u_0, \qquad \frac{\partial u}{\partial t}(0,\cdot) = v_0, \qquad \theta(0,\cdot) = \theta_0 \quad \text{on } \Omega, \tag{2.7a}$$

and by the natural boundary conditions of Robin (or sometimes called Newton) type, i.e. we prescribe the traction stress h and the heat flux f, which means here

$$\Big(\mathbb{D}e\Big(\frac{\partial u}{\partial t}\Big)+\psi_e'(e(u),\theta) - \operatorname{div}\mathfrak{h}\Big)\cdot\nu - \operatorname{div}_{\mathrm{S}}\Big(\mathfrak{h}\cdot\nu\Big) = h \quad \text{and} \tag{2.7b}$$

$$\mathfrak{h}{:}(\nu\otimes\nu) = 0 \text{with} \quad \mathfrak{h} = \mathbb{G}\nabla e\Big(\frac{\partial u}{\partial t}\Big) + \mathbb{H}\nabla e(u), \quad \text{and} \tag{2.7c}$$

$$\mathbb{K}(e(u),\theta)\frac{\partial\theta}{\partial\nu} = f \quad \text{on } \Sigma. \tag{2.7d}$$

In (2.7b), $\operatorname{div}_{\mathrm{S}}$ denotes the $(d{-}1)$-dimensional "surface divergence" defined by $\operatorname{div}_{\mathrm{S}} := \operatorname{tr}(\nabla_{\mathrm{S}})$ with "tr" as the trace (here of a $(d{-}1)\times(d{-}1)$-matrix) and with ∇_{S} denoting the tangential derivative, i.e. being defined by $\nabla_{\mathrm{S}} v = \nabla v - \frac{\partial v}{\partial \nu}\nu$. For justification of (2.7b,c) (in the form (3.10a,b) below) as indeed natural boundary conditions see (3.13)–(3.15), cf. also e.g. [35].

To see the energetics of this model, let us define the *internal energy* as

$$w = \hat{\psi} + \theta s. \tag{2.8}$$

Using (2.4a) and (2.5a), we have

$$\begin{aligned}\frac{\partial w}{\partial t} &= \psi_e'(e(u),\theta){:}e\Big(\frac{\partial u}{\partial t}\Big) + \mathbb{H}\nabla e(u)\,\vdots\,\nabla e\Big(\frac{\partial u}{\partial t}\Big) + \psi_\theta'(e(u),\theta)\frac{\partial\theta}{\partial t} + \theta\frac{\partial s}{\partial t} + \frac{\partial\theta}{\partial t}s\\ &= \psi_e'(e(u),\theta){:}e\Big(\frac{\partial u}{\partial t}\Big) + \mathbb{H}\nabla e(u)\,\vdots\,\nabla e\Big(\frac{\partial u}{\partial t}\Big) + \theta\frac{\partial s}{\partial t}\\ &= \psi_e'(e(u),\theta){:}e\Big(\frac{\partial u}{\partial t}\Big) + \mathbb{H}\nabla e(u)\,\vdots\,\nabla e\Big(\frac{\partial u}{\partial t}\Big)\\ &\quad + \xi_1\Big(e\Big(\frac{\partial u}{\partial t}\Big), \nabla e\Big(\frac{\partial u}{\partial t}\Big)\Big) + \operatorname{div} j.\end{aligned} \tag{2.9}$$

To complete the energy balance, we test (2.6a) together with the boundary conditions (2.7b,c) by $\frac{\partial u}{\partial t}$ and use (2.9) to obtain:

$$\underbrace{\frac{\mathrm{d}}{\mathrm{d}t}\int_\Omega\Big(\frac{\varrho}{2}\Big|\frac{\partial u}{\partial t}\Big|^2 + w\Big)\mathrm{d}x}_{\substack{\text{kinetic + internal energy}\\ \text{=total energy}}} = \underbrace{\int_\Omega g\cdot\frac{\partial u}{\partial t}\,\mathrm{d}x + \int_\Gamma h\cdot\frac{\partial u}{\partial t} + f\,\mathrm{d}S}_{\text{power of external load and heat}}. \tag{2.10}$$

Integrating (2.5a) divided by θ and using Green's formula and the boundary condition (2.7d) yield formally the *Clausius–Duhem inequality* (the 2nd law of thermodynamics):

$$\frac{\mathrm{d}}{\mathrm{d}t}\int_\Omega s\,\mathrm{d}x = \int_\Omega \frac{\xi_1}{\theta} - \frac{\mathrm{div}(\mathbb{K}\nabla\theta)}{\theta}\,\mathrm{d}x = \underbrace{\int_\Omega \frac{\xi_1}{\theta} + \frac{\mathbb{K}(\nabla\theta)^2}{\theta^2}\,\mathrm{d}x}_{\text{entropy production rate}} + \underbrace{\int_\Gamma \frac{f}{\theta}\,\mathrm{d}S}_{\text{entropy flux}} \geq 0 \tag{2.11}$$

provided $f \geq 0$, $\mathbb{K} \geq 0$, and $\theta > 0$. The last requirement about the positivity of temperature is indeed guaranteed if temperature is positive at $t = 0$, cf. [19, Sect. 4.2.1].

Remark 2.1 (Conventional model). As already mentioned in Sect. 1, most of literature about thermo-visco-elasticity considers the partly linearized ansatz

$$\psi(e,\theta) = \varphi(e) + \theta\phi_1(e) + \phi_2(\theta), \tag{2.12}$$

thus $\mathfrak{c}(e,\theta) = \mathfrak{c}(\theta) = -\theta\phi_2''(\theta)$. In this case, often also $\mathbb{H} = 0$ and $\mathbb{G} = 0$ is considered, i.e. no hyperstresses: $\mathfrak{h} = 0$. The system (2.6) then reduces to

$$\varrho\frac{\partial^2 u}{\partial t^2} - \mathrm{div}\Big(\mathbb{D}e\Big(\frac{\partial u}{\partial t}\Big) + \varphi'\big(e(u)\big) + \theta\phi_1'\big(e(u)\big)\Big) = g, \tag{2.13a}$$

$$\mathfrak{c}(\theta)\frac{\partial\theta}{\partial t} - \mathrm{div}\big(\mathbb{K}(e(u),\theta)\nabla\theta\big) = \mathbb{D}\Big(e\Big(\frac{\partial u}{\partial t}\Big)\Big)^2 + \theta\phi_1'\big(e(u)\big){:}e\Big(\frac{\partial u}{\partial t}\Big). \tag{2.13b}$$

A special case φ quadratic and ϕ_1 linear represents a classical problem of thermodynamically consistent linear thermo-visco-elasticity. Existence of its solution was proved already in [13,14] in 1982 for $d = 1$ and $\mathfrak{c}$ and $\mathbb{K}$ constant, while the multidimensional case remained open many decades in spite of great effort, and it was recognised that some natural modifications help. One option involving a non-constant heat capacity $\mathfrak{c}$ depending on temperature θ with a growth at least $\theta^{1/2+\varepsilon}$, $\varepsilon > 0$, was considered in [3] where existence of a solution has been proved by a Schauder fixed point, cf. also [42] for the proof by the Galerkin method and [43] for $\mathfrak{c}$ growing only like $\theta^{6/5+\varepsilon}$. Temperature-dependent heat conductivity $\mathbb{K}$ was also considered in [16], requiring a growth bigger than $1 - 2/d$ if $\mathfrak{c}$ is constant, see also [17, Sect. 5.4.2.1]. Alternatively, also dependence of $\mathbb{K}$ on $\nabla\theta$ pronouncing entropy production for bigger temperature gradient may help, cf. [17, Sect. 5.4.2.2] and [30,31]. Yet, making heat flux depending nonlinearly on $\nabla\theta$ does not seem relevant in most applications, which is why we consider only a linear Fourier law thorough the whole article.

3. Enthalpy-like transformation and weak solution

The peculiarity of the system (2.6) is that the heat capacity $\mathfrak{c}$ depends also on the mechanical variable e so that one cannot perform the conventional enthalpy transformation by using just a primitive function of $\mathfrak{c}$. Anyhow, one can make an *enhanced enthalpy transformation*, which is to some extent similar as in [32] where $\mathfrak{c}$ depends on (t,x) instead of e. This procedure is based on the following elementary calculus:

$$\mathfrak{c}(e,\theta)\frac{\partial\theta}{\partial t} = \frac{\partial\mathfrak{C}(e,\theta)}{\partial t} - \mathfrak{C}_e'(e,\theta){:}\frac{\partial e}{\partial t} \tag{3.1}$$

where

$$\mathfrak{C}(e,\theta) := \int_0^\theta \mathfrak{c}(e,\hat\theta)\,\mathrm{d}\hat\theta\,, \qquad \text{and thus} \qquad \mathfrak{C}_e'(e,\theta) = \int_0^\theta \mathfrak{c}_e'(e,\hat\theta)\,\mathrm{d}\hat\theta. \tag{3.2}$$

We introduce the substitution

$$\vartheta = \mathfrak{C}(e(u),\theta). \tag{3.3}$$

It is physically natural to assume the heat capacity $\mathfrak{c}$ positive, which makes $\mathfrak{C}(e,\cdot)$ increasing and thus also invertible. Thus allows us to define

$$\Theta(e,\vartheta) := \big[\mathfrak{C}(e,\cdot)\big]^{-1}(\vartheta), \tag{3.4a}$$
$$\mathcal{K}_0(e,\vartheta) := \mathbb{K}(e,\Theta(e,\vartheta))\Theta_\vartheta'(e,\vartheta), \tag{3.4b}$$
$$\mathcal{K}_1(e,\vartheta) := \mathbb{K}(e,\Theta(e,\vartheta))\Theta_e'(e,\vartheta), \tag{3.4c}$$
$$\mathcal{A}(e,\vartheta) := \Theta(e,\vartheta)\psi_{\theta e}''(e,\Theta(e,\vartheta)) + \mathfrak{C}_e'(e,\Theta(e,\vartheta)). \tag{3.4d}$$

Since $\theta = \Theta(e(u),\vartheta)$, one can express the heat flux in terms of ϑ as

$$\begin{aligned}\mathbb{K}(e(u),\theta)\nabla\theta &= \mathbb{K}(e(u),\Theta(e(u),\vartheta))\nabla\Theta(e(u),\vartheta)\\ &= \mathcal{K}_0(e(u),\vartheta)\nabla\vartheta + \mathcal{K}_1(e(u),\vartheta)\nabla e(u).\end{aligned} \tag{3.5}$$

Let us notice that the last term in (3.5) involves $\nabla e(u)$, from which the necessity of using a gradient theory for $e(u)$ clearly arises. Also, we have $\mathcal{A}(e,\vartheta) = \mathbb{A}(e,\theta) := \theta\psi_{e\theta}''(e,\theta) + \int_0^\theta \mathfrak{c}_e'(e,\hat\theta)\,\mathrm{d}\hat\theta$ for $\theta = \Theta(e,\vartheta)$ so that, since $\mathfrak{c}_e'(e,\theta) = -\theta\psi_{e\theta\theta}'''(e,\theta)$, we have

$$\big[\mathcal{A}{-}\psi_e'{\circ}\Theta\big](e,\vartheta) = \big[\mathbb{A}{-}\psi_e'\big](e,\theta) = \theta\psi_{e\theta}''(e,\theta) - \int_0^\theta t\psi_{e\theta\theta}'''(e,t)\,\mathrm{d}t - \psi_e'(e,\theta). \tag{3.6}$$

Furthermore,

$$\big[\mathbb{A}{-}\psi_e'\big]_\theta'(e,\theta) = \theta\psi_{e\theta\theta}'''(e,\theta) + \psi_{e\theta}''(e,\theta) - \theta\psi_{e\theta\theta}'''(e,\theta) - \psi_{e\theta}''(e,\theta) = 0,$$

so that $\mathcal{A}{-}\psi_e'{\circ}\Theta$ is independent of ϑ. Moreover, from (3.6), one can see that $[\mathbb{A}{-}\psi_e'](e,0) = -\psi_e'(e,0)$. Therefore, abbreviating

$$\varphi(\cdot) := \psi(\cdot,0), \tag{3.7}$$

we can deduce that

$$\big[\psi_e'{\circ}\Theta - \mathcal{A}\big](e,\vartheta) = \psi_e'(e,0) = \varphi'(e). \tag{3.8}$$

This transforms the system (2.6) into the form

$$\varrho\frac{\partial^2 u}{\partial t^2} - \mathrm{div}\Big(\mathbb{D}e\Big(\frac{\partial u}{\partial t}\Big) + \varphi'(e(u)) + \mathcal{A}(e(u),\vartheta) - \mathrm{div}\Big(\mathbb{G}\nabla e\Big(\frac{\partial u}{\partial t}\Big) + \mathbb{H}\nabla e(u)\Big)\Big) = g, \tag{3.9a}$$

$$\frac{\partial\vartheta}{\partial t} - \mathrm{div}\big(\mathcal{K}_0(e(u),\vartheta)\nabla\vartheta\big) = \mathbb{D}\Big(e\Big(\frac{\partial u}{\partial t}\Big)\Big)^2 + \mathbb{G}\Big(\nabla e\Big(\frac{\partial u}{\partial t}\Big)\Big)^2 + \mathcal{A}\big(e(u),\vartheta\big){:}e\Big(\frac{\partial u}{\partial t}\Big) + \mathrm{div}\big(\mathcal{K}_1(e(u),\vartheta)\nabla e(u)\big), \tag{3.9b}$$

with the boundary conditions

$$\Big(\mathbb{D}e\Big(\frac{\partial u}{\partial t}\Big) + \varphi'(e(u)) + \mathcal{A}(e(u),\vartheta) - \mathrm{div}\,\mathfrak{h}\Big)\cdot\nu - \mathrm{div}_{\mathrm{S}}\big(\mathfrak{h}\cdot\nu\big) = h, \tag{3.10a}$$

$$\mathfrak{h}{:}(\nu\otimes\nu) = 0 \qquad \text{with } \mathfrak{h} \text{ again from (2.7c), and} \tag{3.10b}$$

$$\mathcal{K}_0(e(u),\vartheta)\frac{\partial\vartheta}{\partial\nu} + \mathcal{K}_1(e(u),\vartheta)\frac{\partial e(u)}{\partial\nu} = f \quad \text{on } \Sigma, \tag{3.10c}$$

and the initial conditions

$$u(0,\cdot) = u_0, \quad \frac{\partial u}{\partial t}(0,\cdot) = v_0, \quad \vartheta(0,\cdot) = \mathfrak{C}(e(u_0),\theta_0) \qquad \text{in } \Omega. \tag{3.10d}$$

One can observe that the internal energy (2.8) now takes the form

$$w = \hat\psi + \theta s = \varphi(e) + \vartheta + \frac{1}{2}\mathbb{H}(\nabla e)^2. \tag{3.11}$$

Indeed, denoting the residuum in (3.11) by $r(e,\theta) := \psi(e,\theta) + \theta s - \varphi(e) - \vartheta$ with $s = -\psi'_\theta(e,\theta)$ and $\vartheta = \mathfrak{C}(e,\theta)$, we can see that $r(e,0) = \psi(e,0) - \varphi(e) - \mathfrak{C}(e,0) = 0$ and that $r'_\theta(e,\theta) = \psi'_\theta(e,\theta) - \theta\psi''_{\theta\theta}(e,\theta) - \psi'_\theta(e,\theta) - \mathfrak{c}(e,\theta) = 0$, hence $r \equiv 0$. Thus, in view of (3.11), the energy balance (2.10) written in terms of ϑ takes the form

$$\frac{\mathrm{d}}{\mathrm{d}t}\int_\Omega \frac{\varrho}{2}\Big|\frac{\partial u}{\partial t}\Big|^2 + \varphi\big(e(u)\big) + \vartheta + \frac{1}{2}\mathbb{H}\big(\nabla e(u)\big)^2\,\mathrm{d}x = \int_\Omega g\cdot\frac{\partial u}{\partial t}\,\mathrm{d}x + \int_\Gamma h\cdot\frac{\partial u}{\partial t} + f\,\mathrm{d}S. \tag{3.12}$$

For the balance of mere mechanical energy see (4.54) below.

The weak formulation of the initial-boundary-value problem (3.9)–(3.10) can be designed by using by-part integration for both equations in (3.9), and Green formula twice for (3.9a) and once for (3.9b), and by incorporating both the initial and the boundary conditions (3.10). For this, in case of (3.9a), one must use still another Green formula along the boundary, namely, written for a matrix-valued field w and a scalar function v,

$$\int_\Gamma w{:}\big((\nabla_{\mathrm{S}}v)\otimes\nu\big)\,\mathrm{d}S = \int_\Gamma \Big((\mathrm{div}_{\mathrm{S}}\nu)\big(w{:}(\nu\otimes\nu)\big) - \mathrm{div}_{\mathrm{S}}(w\cdot\nu)\Big)v\,\mathrm{d}S, \tag{3.13}$$

where $\mathrm{div}_{\mathrm{S}}\nu$ is (up to a factor $-\frac{1}{2}$) the mean curvature of the surface Γ; cf. [20,34,35,47]. Just for illustration, (3.13) is a compact form for what can be written more in detail as

$$\int_\Gamma \sum_{i,j=1}^d w_{ij}\Big(\frac{\partial v}{\partial x_i} - \sum_{k=1}^d \frac{\partial v}{\partial x_k}\nu_k\nu_i\Big)\nu_j\,\mathrm{d}S$$
$$= \int_\Gamma \Big(\sum_{l=1}^d\Big(\frac{\partial \nu_l}{\partial x_l} - \sum_{k=1}^d \frac{\partial \nu_k}{\partial x_k}\nu_k\nu_l\Big)\sum_{i,j=1}^d w_{ij}\nu_i\nu_j$$
$$-\sum_{l=1}^d\Big(\frac{\partial}{\partial x_l}\sum_{m=1}^d w_{lm}\nu_m - \sum_{k=1}^d \nu_k\nu_l\frac{\partial}{\partial x_k}\sum_{m=1}^d w_{km}\nu_m\Big)\Big)v\,\mathrm{d}S; \qquad (3.14)$$

in fact, the expressions for ∇_{S} and $\mathrm{div}_{\mathrm{S}}$ has a component corresponding to the normal direction zero, so the summation can use only $d-1$ indices if expressed in suitable local coordinates.

Then, using the vectorial variant of (3.13) for $\mathfrak{h}$ and z in place of w and v, and using also the decomposition $\nabla z = \frac{\partial z}{\partial\nu}\nu + \nabla_{\mathrm{S}} z$, we have

$$\int_\Omega \big(\mathrm{div}^2\mathfrak{h}\big)\cdot z\,\mathrm{d}x = \int_\Gamma \big(\mathrm{div}\,\mathfrak{h}\big):(z\otimes\nu)\,\mathrm{d}S - \int_\Omega \big(\mathrm{div}\,\mathfrak{h}\big):\nabla z\,\mathrm{d}x$$
$$= \int_\Omega \mathfrak{h}\,\vdots\,\nabla^2 z\,\mathrm{d}x + \int_\Gamma \big(\mathrm{div}\,\mathfrak{h}\big):(z\otimes\nu) - \mathfrak{h}\,\vdots\,(\nabla z\otimes\nu)\,\mathrm{d}S$$
$$= \int_\Omega \mathfrak{h}\,\vdots\,\nabla^2 z\,\mathrm{d}x + \int_\Gamma \Big(\mathrm{div}\,\mathfrak{h}\Big):(z\otimes\nu) - \mathfrak{h}\,\vdots\Big(\Big(\frac{\partial z}{\partial\nu}\nu + \nabla_{\mathrm{S}} z\Big)\otimes\nu\Big)\,\mathrm{d}S$$
$$= \int_\Omega \mathfrak{h}\,\vdots\,\nabla^2 z\,\mathrm{d}x + \int_\Gamma \Big(\big(\mathrm{div}\,\mathfrak{h}\big)\cdot\nu + \mathrm{div}_{\mathrm{S}}(\mathfrak{h}\cdot\nu) - (\mathrm{div}_{\mathrm{S}}\nu)\big(\mathfrak{h}:(\nu\otimes\nu)\big)\Big)\cdot z$$
$$-\big(\mathfrak{h}:(\nu\otimes\nu)\big)\cdot\frac{\partial z}{\partial\nu}\,\mathrm{d}S \qquad (3.15)$$

where "$\vdots$" abbreviates summation over 3 indices, consistently with "$:$" and "$\cdot$" meaning summation over 2 or 1 indices, respectively. It reveals that one of the natural boundary conditions is to prescribe $\mathfrak{h}:(\nu\otimes\nu)$. Our (quite natural) choice (2.7c) simplifies the situation because the term containing the mean curvature $\mathrm{div}_{\mathrm{S}}\nu$ vanishes and we can see that h in (2.7b) indeed means the true (prescribed) boundary force.

If both $\mathbb{G}$ and $\mathbb{H}$ are symmetric, so is $\mathfrak{h}$, and we have $\mathfrak{h}\,\vdots\,\nabla^2 z = \mathfrak{h}\,\vdots\,\nabla e(z)$. This leads to the weak formulation of the mechanical part (3.9a)-(3.10a,b) in the form (3.16a) below:

Definition 3.1 (*Weak solution.*) We call the pair $(u,\vartheta) \in W^{1,2}(I;W^{2,2}(\Omega;\mathbb{R}^d)) \times L^1(I;W^{1,1}(\Omega))$ a weak solution to the initial-boundary-value problem (3.9)–(3.10) if

$$\int_Q \Big(\mathbb{D}e\Big(\frac{\partial u}{\partial t}\Big) + \varphi'(e(u)) + \mathcal{A}(e(u),\vartheta)\Big){:}e(v)$$
$$+\Big(\mathbb{G}\nabla e\Big(\frac{\partial u}{\partial t}\Big) + \mathbb{H}\nabla e(u)\Big)\vdots\nabla e(v) - \varrho\frac{\partial u}{\partial t}\cdot\frac{\partial v}{\partial t}\,\mathrm{d}x\mathrm{d}t + \int_\Omega \varrho\frac{\partial u}{\partial t}(T)\cdot v(T)\,\mathrm{d}x$$
$$= \int_\Omega \varrho v_0\cdot v(T)\,\mathrm{d}x + \int_Q g\cdot v\,\mathrm{d}x\mathrm{d}t + \int_\Sigma h\cdot v\,\mathrm{d}S\mathrm{d}t \tag{3.16a}$$

for any $v \in W^{1,2}(I;L^2(\Omega;\mathbb{R}^d)) \cap L^2(I;W^{2,2}(\Omega;\mathbb{R}^d))$, and

$$\int_Q \Big(\mathcal{K}_0(e(u),\vartheta)\nabla\vartheta + \mathcal{K}_1(e(u),\vartheta)\nabla e(u)\Big)\cdot\nabla z - \vartheta\frac{\partial z}{\partial t}$$
$$-\Big(\mathbb{D}\Big(e\Big(\frac{\partial u}{\partial t}\Big)\Big)^2 + \mathbb{G}\Big(\nabla e\Big(\frac{\partial u}{\partial t}\Big)\Big)^2 + \mathcal{A}\big(e(u),\vartheta\big){:}e\left(\frac{\partial u}{\partial t}\right)\Big)v\,\mathrm{d}x\mathrm{d}t$$
$$= \int_\Omega \mathfrak{C}(e(u_0),\theta_0)z(0)\,\mathrm{d}x + \int_\Sigma fz\,\mathrm{d}S\mathrm{d}t \tag{3.16b}$$

for any $z \in C^1(\bar{Q})$ with $z(T,\cdot) = 0$, and if the remaining initial condition not involved in (3.16a,b) also holds:

$$u(0) = u_0. \tag{3.16c}$$

We first state our main existence result. For this, we first summarize the assumptions. We use the notation of p' being the conjugate exponent $p/(p{-}1)$ and of p^* being the critical Sobolev exponent defined here (just with a slight loss of generality) as $dp/(d{-}p)$ if $p < d$, or we choose arbitrary large $p^* < \infty$ but fixed if $p \geq d$. Thus, e.g., we have the compact embedding $W^{2,2}(\Omega) \subset W^{1,2^*-\epsilon}(\Omega)$ with ϵ arbitrarily small positive number. We then impose the following assumptions formulated in terms of the transformed data:

$\varphi(\cdot) = \psi(\cdot,0) : \mathbb{R}^{d\times d}_{\mathrm{sym}} \to \mathbb{R}^+$ is continuously differentiable, and
$\exists\, C_\varphi \in \mathbb{R}\ \epsilon > 0\ \forall e \in \mathbb{R}^{d\times d}_{\mathrm{sym}}$:

$$e \mapsto \varphi(e) + C_\varphi|e|^2 \text{ is convex,} \quad \varphi(e) \leq C_\varphi(1{+}|e|^{2^*}), \tag{3.17a}$$

$$\varphi'(e){:}e + C_\phi \geq 0, \qquad \big|\varphi'(e)\big| \leq C_\varphi\big(1{+}|e|^{2^*-1-\epsilon}\big), \tag{3.17b}$$

$\exists\, C_{\mathcal{K}_0}, C_{\mathcal{K}_1} \in \mathbb{R},\ \kappa_0, \epsilon > 0\ \ \forall\, (e,\vartheta) \in \mathbb{R}^{d\times d}_{\mathrm{sym}}{\times}\mathbb{R},\ \xi \in \mathbb{R}^d$:

$$\mathcal{K}_0(e,\vartheta)\xi\cdot\xi \geq \kappa_0|\xi|^2, \tag{3.17c}$$

$$\big|\mathcal{K}_0(e,\vartheta)\big| \leq C_{\mathcal{K}_0}\big(1{+}|e|^{2^*/(d+2)-\epsilon}{+}|\vartheta|^{1/d-\epsilon}\big), \tag{3.17d}$$

$$\big|\mathcal{K}_1(e,\vartheta)\big| \leq C_{\mathcal{K}_1}\sqrt{1{+}|\vartheta|}, \tag{3.17e}$$

$\exists C_{\mathcal{A}} < \infty,\ q > (2d{+}8)/(d{+}6),\ s < q:$

$$\forall\,(e,\vartheta) \in \mathbb{R}^{d\times d}_{\rm sym}{\times}\mathbb{R}: \quad \big|\mathcal{A}(e,\vartheta)\big| \le C_{\mathcal{A}}\big(1{+}|e|^{2^*/q}{+}|\vartheta|^{s/q}\big), \tag{3.17f}$$

$$\varrho > 0, \quad \mathbb{D},\mathbb{G},\mathbb{H} \text{ symmetric positive definite, } \mathbb{H} \text{ having a potential}, \tag{3.17g}$$

$$g \in L^1(I;L^2(\Omega;\mathbb{R}^d)), \quad h \in W^{1,1}(I;L^{p_h}(\Gamma;\mathbb{R}^d)), \tag{3.17h}$$

$$u_0 \in W^{2,2}(\Omega;\mathbb{R}^d), \qquad v_0 \in L^2(\Omega;\mathbb{R}^d), \tag{3.17i}$$

$$\theta_0 \ge 0, \quad f \ge 0, \quad \mathfrak{C}(\theta_0,e(u_0)) \in L^1(\Omega), \quad f \in L^1(\Sigma), \tag{3.17j}$$

with $p_h = 2{-}2/d$ if $d > 2$, or $p_h > 1$ for $d \le 2$ (or even $p_h = 1$ if $d = 1$). In fact, we will need (3.17) only for $\vartheta \ge 0$ in what follows. Note that (3.17d) ensures integrability of $\mathcal{K}_0(e(u),\vartheta)\nabla\vartheta$ if $e(u) \in L^{2^*}(Q;\mathbb{R}^{d\times d})$, $\vartheta \in L^{(d+2)/d-\epsilon}(Q)$, and $\nabla\vartheta \in L^{(d+2)/(d+1)-\epsilon}(Q;\mathbb{R}^d)$. The weakened convexity of φ in (3.17a) is called semiconvexity. The symmetry in (3.17g) means that $\mathbb{D}:\mathbb{R}^{d\times d}_{\rm sym} \to \mathbb{R}^{d\times d}_{\rm sym}$, i.e. $\mathbb{D}_{ijkl} = \mathbb{D}_{jikl}$, and similarly $\mathbb{G}_{ijklmn} = \mathbb{G}_{ikjlmn}$, while for $\mathbb{H}$ we need $\mathbb{H}_{ijklmn} = \mathbb{H}_{ikjlmn} = \mathbb{H}_{lmnijk}$. The assumption $\varphi'(e){:}e + C_\phi \ge 0$ in (3.17b) is not restrictive in applications and simplifies some considerations below, but, in fact, can be relaxed.

The following main result is a consequence of Lemma 4.1 together with Proposition 4.5 below when taking into account that, due to the qualification (3.17i) of u_0, the regularization $u_{0\tau}$ with all properties required in Proposition 4.5 always exist.

Theorem 3.2. (Existence of a weak solution). *Let* (3.17) *hold and* $d \le 3$. *Then the transformed system* (3.9) *with the initial/boundary conditions* (3.10) *possesses a weak solution* (u,ϑ) *according to the Definition* 3.1 *such that, in addition,*

$$\frac{\partial^2 u}{\partial t^2} \in L^2(I;W^{2,2}(\Omega;\mathbb{R}^d)^*), \tag{3.18a}$$

$$\vartheta \in L^r(I;W^{1,r}(\Omega)) \cap L^\infty(I;L^1(\Omega)) \cap W^{1,1}(I;W^{3,2}(\Omega)^*), \qquad r < \frac{d{+}2}{d{+}1}. \tag{3.18b}$$

Remark 3.1 (Internal energy balance). The formula (3.11) says that $\vartheta = w - \varphi(e) - \frac{1}{2}\mathbb{H}(\nabla e)^2$, which suggests to subtract from the internal-energy balance (2.9) the balance of the stored-energy rate versus the power of conservative parts of (hyper)stresses, i.e. $\frac{\partial}{\partial t}(\varphi(e(u)) + \frac{1}{2}\mathbb{H}(\nabla e(u))^2 = \varphi'(e(u)){:}e\big(\frac{\partial u}{\partial t}\big) + \mathbb{H}\nabla e(u)\,\vdots\,\nabla e\big(\frac{\partial u}{\partial t}\big)$. In this way, we obtain $\frac{\partial\vartheta}{\partial t} + {\rm div}\, j = (\sigma{-}\varphi'(e(u))){:}e\big(\frac{\partial u}{\partial t}\big) + (\mathfrak{h}{-}\mathbb{H}\nabla e(u))\,\vdots\,\nabla e\big(\frac{\partial u}{\partial t}\big)$ which, after eliminating temperature as we did by the substitution (3.3), would again result to the transformed heat equation (3.9b). This reveals the physical character of the transformed system (3.9) and a certain conceptual similarity with thermodynamics of fluid where internal energy is sometimes used instead of temperature for analysis, cf. e.g. [8,9].

4. Analysis by time discretisation

We analyse the system (3.9) by Rothe's method with a suitable regularization to compensate the growth of the non-monotone terms on the right-hand side of the heat equation (3.9b). More specifically, for $\tau \le 4$ and a suitable $\eta > 0$ specified later, we consider

$$\varrho\frac{u_\tau^k-2u_\tau^{k-1}+u_\tau^{k-2}}{\tau^2}-\mathrm{div}\Big(\mathbb{D}e\Big(\frac{u_\tau^k-u_\tau^{k-1}}{\tau}\Big)+\varphi'\big(e(u_\tau^k)\big)\\ +\mathcal{A}(e(u_\tau^k),\vartheta_\tau^k)+\tau\big|e(u_\tau^k)\big|^{\eta-2}e(u_\tau^k)-\mathrm{div}\,\mathfrak{h}_\tau^k\Big)=g_\tau^k\\ \text{with}\ \ \mathfrak{h}_\tau^k=\mathbb{G}\nabla e\Big(\frac{u_\tau^k-u_\tau^{k-1}}{\tau}\Big)+\mathbb{H}\nabla e(u_\tau^k)+\tau\big|\nabla e(u_\tau^k)\big|^{\eta-2}\nabla e(u_\tau^k), \tag{4.1a}$$

$$\frac{\vartheta_\tau^k-\vartheta_\tau^{k-1}}{\tau}-\mathrm{div}\big(\mathcal{K}_0(e(u_\tau^k),\vartheta_\tau^k)\nabla\vartheta_\tau^k\big)\\ =\Big(1-\frac{\sqrt{\tau}}{2}\Big)\mathbb{D}\Big(e\Big(\frac{u_\tau^k-u_\tau^{k-1}}{\tau}\Big)\Big)^2+\mathbb{G}\Big(\nabla e\Big(\frac{u_\tau^k-u_\tau^{k-1}}{\tau}\Big)\Big)^2\\ +\mathcal{A}\Big(e(u_\tau^k),\vartheta_\tau^k\Big){:}e\Big(\frac{u_\tau^k-u_\tau^{k-1}}{\tau}\Big)+\mathrm{div}\big(\mathcal{K}_1(e(u_\tau^k),\vartheta_\tau^k)\nabla e(u_\tau^k)\big) \tag{4.1b}$$

completed with the boundary conditions

$$\Big(\mathbb{D}e\Big(\frac{u_\tau^k-u_\tau^{k-1}}{\tau}\Big)+\varphi'(e(u_\tau^k))+\mathcal{A}(e(u_\tau^k),\vartheta_\tau^k))\\ -\mathrm{div}\,\mathfrak{h}_\tau^k+\tau|e(u_\tau^k)|^{\eta-2}e(u_\tau^k)\Big)\cdot\nu-\mathrm{div}_{\mathrm{S}}\big(\mathfrak{h}_\tau^k\cdot\nu\big)=h_\tau^k, \tag{4.2a}$$

$$\mathfrak{h}_\tau^k{:}(\nu\otimes\nu)=0 \qquad \text{and} \tag{4.2b}$$

$$\mathcal{K}_0(e(u_\tau^k),\vartheta_\tau^k)\frac{\partial\vartheta_\tau^k}{\partial\nu}+\mathcal{K}_1(e(u_\tau^k),\vartheta_\tau^k)\frac{\partial e(u_\tau^k)}{\partial\nu}=f_\tau^k \quad \text{on }\Gamma. \tag{4.2c}$$

This recursive scheme for $k=1$ is started by considering the initial conditions

$$u_\tau^0=u_{0\tau},\quad u_\tau^{-1}=u_\tau^0-\tau v_0,\quad \vartheta_\tau^0=\mathfrak{C}(e(u_{0\tau}),\theta_0)\qquad \text{in }\Omega \tag{4.2d}$$

involving a suitably regularized initial displacement $u_{0\tau}$. In (4.1)–(4.2), we used a suitable approximation of the right-hand sides; to be more specific, for $k=1,\dots,T/\tau$, we consider

$$g_\tau^k=\frac{1}{\tau}\int_{(k-1)\tau}^{k\tau}g(t,\cdot)\,\mathrm{d}t,\qquad f_\tau^k=\frac{1}{\tau}\int_{(k-1)\tau}^{k\tau}f(t,\cdot)\,\mathrm{d}t,\qquad h_\tau^k=\frac{1}{\tau}\int_{(k-1)\tau}^{k\tau}h(t,\cdot)\,\mathrm{d}t. \tag{4.3}$$

Lemma 4.1 (Existence of Rothe's solution). *Let* (3.17) *hold, let* $u_{0\tau}\in W^{2,\eta}(\Omega;\mathbb{R}^d)$ *with* η *sufficiently large, namely* $\eta>\max(4,2(2^*+q)/q,2q/(q-s))$. *Then the recursive boundary-value problem* (4.1)–(4.2) *has a weak solution* $(u_\tau^k,\vartheta_\tau^k)\in W^{2,\eta}(\Omega;\mathbb{R}^d)\times W^{1,2}(\Omega)$ *such that* $\vartheta_\tau^k\ge 0$ *for any* $k=1,\dots,T/\tau$.

Proof. In contrast to isothermal situation, the system (4.1)–(4.2) does not seem to have a structure of a variational problem and thus existence of its solution is to be proved only in a nonconstructive way by using the very classical Brézis theorem [4] based on the concept of coercive pseudomonotone operators and using Brouwer fixed-point argument. More specifically, here we apply this theorem to an abstract equation $A(u,\vartheta) = F$ with the nonlinear mapping $A = (A_1, A_2) : W^{2,\eta}(\Omega;\mathbb{R}^d)\times W^{1,2}(\Omega) \to W^{2,\eta}(\Omega;\mathbb{R}^d)^*\times W^{1,2}(\Omega)^*$ given, at the k level, by

$$\langle A_1(u,\vartheta), v\rangle := \int_\Omega \varrho\frac{u}{\tau^2}\cdot v + \Big(\mathbb{D}e\Big(\frac{u}{\tau}\Big) + \varphi'(e(u)) + \mathcal{A}(e(u),\vartheta) + \tau\big|e(u)\big|^{\eta-2}e(u)\Big):e(v)$$
$$+ \Big(\mathbb{G}\nabla e\Big(\frac{u}{\tau}\Big) + \mathbb{H}\nabla e(u) + \tau\big|\nabla e(u)\big|^{\eta-2}\nabla e(u)\Big)\vdots\nabla e(v)\,\mathrm{d}x \tag{4.4a}$$

$$\langle A_2(u,\vartheta), z\rangle := \int_\Omega \Big(\frac{\vartheta}{\tau} + \Big(1-\frac{\sqrt{\tau}}{2}\Big)\Big(\mathbb{D}\Big(e\Big(\frac{u}{\tau}\Big)\Big)^2 + 2\mathbb{D}e\Big(\frac{u}{\tau}\Big):e\Big(\frac{u_\tau^{k-1}}{\tau}\Big)\Big)$$
$$+ \mathbb{G}\Big(\nabla e\Big(\frac{u}{\tau}\Big)\Big)^2 + 2\mathbb{G}\nabla e\Big(\frac{u}{\tau}\Big)\vdots\nabla e\Big(\frac{u_\tau^{k-1}}{\tau}\Big) + \mathcal{A}(e(u),\vartheta):e\Big(\frac{u-u_\tau^{k-1}}{\tau}\Big)\Big)z$$
$$+ \big(\mathcal{K}_0(e(u),\vartheta)\nabla\vartheta\big) + \mathcal{K}_1(e(u),\vartheta)\nabla e(u)\cdot\nabla z\,\mathrm{d}x \tag{4.4b}$$

and the functional $F = (F_1, F_2) \in W^{2,\eta}(\Omega;\mathbb{R}^d)^*\times W^{1,2}(\Omega)^*$ given by

$$\langle F_1, v\rangle := \int_\Omega \Big(\varrho\frac{2u_\tau^{k-1}-u_\tau^{k-2}}{\tau^2} + g_\tau^k\Big)\cdot v + \mathbb{D}e\Big(\frac{u_\tau^{k-1}}{\tau}\Big):e(v)$$
$$+ \mathbb{G}\nabla e\Big(\frac{u_\tau^{k-1}}{\tau}\Big)\vdots\nabla e(v)\,\mathrm{d}x + \int_\Gamma h_\tau^k\cdot v\,\mathrm{d}S, \tag{4.5a}$$

$$\langle F_2, z\rangle := \int_\Omega \Big(\frac{\vartheta_\tau^{k-1}}{\tau} + \Big(1-\frac{\sqrt{\tau}}{2}\Big)\mathbb{D}\Big(e\Big(\frac{u_\tau^{k-1}}{\tau}\Big)\Big)^2$$
$$+ \mathbb{G}\Big(\nabla e\Big(\frac{u_\tau^{k-1}}{\tau}\Big)\Big)^2\Big)z\,\mathrm{d}z + \int_\Gamma f_\tau^k z\,\mathrm{d}S. \tag{4.5b}$$

To prove the coercivity of A, we plug $v = u$ and $= \vartheta$ into (4.4). This results to

$$\langle A_1(u,\vartheta), u\rangle \geq \frac{\varrho}{\tau^2}\big\|u\big\|^2_{L^2(\Omega;\mathbb{R}^d)} + \frac{\min_{|e|=1}\mathbb{D}(e)^2}{\tau}\big\|e(u)\big\|^2_{L^2(\Omega;\mathbb{R}^{d\times d})}$$
$$+ \tau\big\|e(u)\big\|^\eta_{W^{1,\eta}(\Omega;\mathbb{R}^{d\times d})} + \int_\Omega \mathcal{A}(e(u),\vartheta):e(u)\,\mathrm{d}x - C_\varphi|\Omega|, \tag{4.6}$$

$$\langle A_2(u,\vartheta), \vartheta\rangle \geq \frac{1}{\tau}\big\|\vartheta\big\|^2_{L^2(\Omega)} + \kappa_0\big\|\nabla\vartheta\big\|^2_{L^2(\Omega;\mathbb{R}^d)}$$
$$+ \int_\Omega\Big(\frac{2-\sqrt{\tau}}{2\tau^2}\Big(\mathbb{D}\big(e(u)\big)^2 + 2\mathbb{D}e(u):e(u_\tau^{k-1})\Big) + \mathbb{G}\big(\nabla e(u)\big)^2$$
$$+ 2\mathbb{G}\nabla e(u)\vdots\nabla e(u_\tau^{k-1}) + \mathcal{A}\big(e(u),\vartheta\big):e\Big(\frac{u-u_\tau^{k-1}}{\tau}\Big)\Big)\vartheta$$
$$+ \big(\mathcal{K}_1(e(u),\vartheta)\nabla e(u)\big)\cdot\nabla\vartheta\,\mathrm{d}x. \tag{4.7}$$

Note that $\varphi'(e):e \geq -C_\varphi$ from (3.17b) has already been used in (4.6). Let us briefly discuss the estimation of the possibly non-negative terms under this test.

The last term $\int_\Omega \mathcal{A}(e(u),\vartheta):e(u)\,\mathrm{d}x$ in (4.6) can be estimated as (4.10) below but even easier because the factor ϑ is not here. As for (4.7), one can estimate the $\mathcal{K}_1$-term by using Hölder's inequality for 3 factors with exponents 4, η, and 2 and by using (3.17e)

$$\int_\Omega \mathcal{K}_1(e(u),\vartheta)\nabla e(u)\cdot\nabla\vartheta\mathrm{d}x \ge -\int_\Omega C_{\mathcal{K}_1}\sqrt{1+|\vartheta|}\big|\nabla e(u)\big|\,\big|\nabla\vartheta\big|\,\mathrm{d}x$$
$$\ge -\int_\Omega C_{\mathcal{K}_1}\Big(1+\sqrt{|\vartheta|}\Big)\big|\nabla e(u)\big|\,\big|\nabla\vartheta\big|\,\mathrm{d}x$$
$$\ge C_{\delta,\eta} - \delta\big\|\vartheta\big\|_{L^2(\Omega)}^2 - \delta\big\|\nabla e(u)\big\|_{L^\eta(\Omega;\mathbb{R}^{d\times d\times d})}^\eta - \delta\big\|\nabla\vartheta\big\|_{L^2(\Omega;\mathbb{R}^d)}^2 \tag{4.8}$$

with $\delta > 0$ arbitrarily small and again some $C_{\delta,\eta}$ depending on η and δ; here we needed that $\eta > 4$. The heat sources with quadratic growth can be estimated as

$$\int_\Omega \big(\mathbb{D}(e(u))^2+\mathbb{G}|\nabla e(u)|^2\big)\vartheta\,\mathrm{d}x \ge C_{\delta,\eta} - \delta\big\|e(u)\big\|_{L^\eta(\Omega;\mathbb{R}^{d\times d})}^\eta$$
$$-\delta\big\|\nabla e(u)\big\|_{L^\eta(\Omega;\mathbb{R}^{d\times d\times d})}^\eta - \delta\big\|\vartheta\big\|_{L^2(\Omega)}^2; \tag{4.9}$$

here we used that $\eta > 4$. By (3.17f), the $\mathcal{A}$-term in (4.1b) can be estimated as

$$\int_\Omega \mathcal{A}(e(u),\vartheta){:}e(u)\vartheta\,\mathrm{d}x \ge -C_{\mathcal{A}}\int_\Omega \Big(1+|e(u)|^{2^*/q}+|\vartheta|^{s/q}\Big)\big|e(u)\big|\,\big|\vartheta\big|\,\mathrm{d}x$$
$$\ge C_{\delta,q,s} - \delta\big\|e(u)\big\|_{L^\eta(\Omega;\mathbb{R}^{d\times d})}^\eta - \delta\|\vartheta\|_{L^2(\Omega)}^2 \tag{4.10}$$

with $C_{\delta,q,s}$ dependent on q and s from (3.17f); here we used that $\eta > 2\max(2^*{+}q/q, q/(q{-}s))$ and needed $s < q$ as used in (3.17f). The term $\int_\Omega \mathcal{A}(e(u),\vartheta){:}e(u_\tau^{k-1})\vartheta\,\mathrm{d}x$ is even easier. The remaining terms involving u_τ^{k-1} have at most linear decay and can be handled easily by the Young inequality. Altogether, when choosing $\delta > 0$ sufficiently small so that the corresponding terms are dominated by the coercive terms and when using still Korn's inequality, we proved the coercivity of the whole operator A in the sense $\langle A(u,\vartheta),(u,\vartheta)\rangle \ge \epsilon\|u\|_{W^{2,\eta}(\Omega;\mathbb{R}^d)}^2 + \epsilon\|\vartheta\|_{W^{1,2}(\Omega)}^2 - C$ with some small $\epsilon > 0$ depending also on τ.

The pseudomonotonicity of the operator A is obvious because all nonlinear terms are either of the highest-order but monotone, which concerns especially the term $\mathrm{div}^2(\tau|\nabla e(u_\tau^k)|^{\eta-2}\nabla e(u_\tau^k))$, or of a lower order with the exception of the $\mathcal{K}_1$-term in (4.1b) but this term is linear in $\nabla e(u_\tau^k)$, hence weakly continuous. Then the claimed existence of a solution to $A(u,\vartheta) = F$ follows by the mentioned Brézis theorem and, of course, we put $u_\tau^k = u$ and $\vartheta_\tau^k = \vartheta$.

Eventually, one can test (4.1b) by $(\vartheta_\tau^k)^-$, which reveals $\vartheta_\tau^k \ge 0$. For this, we can realize that, for a moment in the proof before, we can define $\mathcal{A}(e,\vartheta) = 0$ for $\vartheta \le 0$, and then find that this possible re-definition is, in fact, irrelevant for this particular solution. □

Let us define the piecewise affine interpolants u_τ and ϑ_τ by

$$\big[u_\tau,\vartheta_\tau\big](t) := \tfrac{t-(k-1)\tau}{\tau}\big(u_\tau^k,\vartheta_\tau^k\big) + \tfrac{k\tau-t}{\tau}\big(u_\tau^{k-1},\vartheta_\tau^{k-1}\big) \quad \text{for } t\in[(k-1)\tau,k\tau] \tag{4.11}$$

with $k=0,\ldots,K_\tau:=T/\tau$. Besides, we define also the backward piecewise constant interpolant $\bar u_\tau$ and $\bar\vartheta_\tau$ by

$$\big[\bar u_\tau,\bar\vartheta_\tau\big](t) := \big(u_\tau^k,\vartheta_\tau^k\big) \qquad \text{for } t\in((k-1)\tau,k\tau], \quad k=1,\ldots,K_\tau. \tag{4.12}$$

We will also need the piecewise affine interpolant of the (piecewise constant) velocity $\frac{\partial u_\tau}{\partial t}$, which we denote by $\big[\frac{\partial u_\tau}{\partial t}\big]^{\mathrm i}$, i.e.

$$\Big[\frac{\partial u_\tau}{\partial t}\Big]^{\mathrm i}(t) := \tfrac{t-(k-1)\tau}{\tau}\frac{\partial u_\tau}{\partial t}(k\tau) + \tfrac{k\tau-t}{\tau}\frac{\partial u_\tau}{\partial t}(k\tau-\tau) \quad \text{for } t\in((k-1)\tau,k\tau]. \tag{4.13}$$

Note that $\frac{\partial}{\partial t}\big[\frac{\partial u_\tau}{\partial t}\big]^{\mathrm i}$ is piecewise constant with the values $\frac{u_\tau^k-2u_\tau^{k-1}+u_\tau^{k-2}}{\tau^2}$ on the particular intervals $((k-1)\tau,k\tau)$.

Lemma 4.2 (Energy estimates). *Let the assumptions of Lemma 4.1 hold, and let* $\|u_{0\tau}\|_{W^{2,\eta}(\Omega;\mathbb R^d)} = \mathcal O(\tau^{-1/\eta})$. *Let also be sufficiently small* τ, *namely* $\tau\le\min(4,(\min_{|e|=1}\mathbb D(e)^2/(4C_\varphi))^2)$ *with* C_φ *from* (3.17b) *if* $C_\varphi>0$, *otherwise* $0<\tau\le\min(4,T)$ *can be arbitrary if* φ *is convex. Then, with some* C_0 *independent on* τ, *it holds:*

$$\big\|u_\tau\big\|_{W^{1,\infty}(I;L^2(\Omega;\mathbb R^d))\cap L^\infty(I;W^{2,2}(\Omega;\mathbb R^d))} \le C_0, \tag{4.14a}$$

$$\big\|\bar\vartheta_\tau\big\|_{L^\infty(I;L^1(\Omega))} \le C_0, \tag{4.14b}$$

$$\big\|e(\bar u_\tau)\big\|_{L^\infty(I;W^{1,\eta}(\Omega;\mathbb R^{d\times d}))} \le C_0\tau^{-1/\eta}. \tag{4.14c}$$

Sketch of the proof. We test the mechanical part (4.1a) by $(u_\tau^k-u_\tau^{k-1})/\tau$. We use the semiconvexity of φ, and hence the convexity of the functional $e\mapsto\varphi(e)+\frac12\tau^{-1/2}\mathbb D(e)^2+\frac{\tau}{\eta}|e|^\eta$ for sufficiently small τ as assumed. Then we have a discrete chain rule at disposal:

$$\begin{aligned}
&\bigg(\varphi'(e(u_\tau^k)) + \mathbb De\bigg(\frac{u_\tau^k-u_\tau^{k-1}}{\tau}\bigg) + \tau|e(u_\tau^k)|^{\eta-2}e(u_\tau^k)\bigg){:}e\bigg(\frac{u_\tau^k-u_\tau^{k-1}}{\tau}\bigg)\\
&= \bigg(\varphi'(e(u_\tau^k)) + \frac{1}{\sqrt\tau}\mathbb De(u_\tau^k) + \tau|e(u_\tau^k)|^{\eta-2}e(u_\tau^k)\bigg){:}e\bigg(\frac{u_\tau^k-u_\tau^{k-1}}{\tau}\bigg)\\
&\quad - \frac{1}{\sqrt\tau}\mathbb De(u_\tau^{k-1}){:}e\bigg(\frac{u_\tau^k-u_\tau^{k-1}}{\tau}\bigg) + (1-\sqrt\tau)\mathbb D\bigg(e\bigg(\frac{u_\tau^k-u_\tau^{k-1}}{\tau}\bigg)\bigg)^2\\
&\ge \frac1\tau\bigg(\varphi(e(u_\tau^k)) + \frac{1}{2\sqrt\tau}\mathbb D\big(e(u_\tau^k)\big)^2 + \frac\tau\eta|e(u_\tau^k)|^\eta\\
&\quad - \varphi(e(u_\tau^{k-1})) - \frac{1}{2\sqrt\tau}\mathbb D\big(e(u_\tau^{k-1})\big)^2 - \frac\tau\eta|e(u_\tau^{k-1})|^\eta\bigg)\\
&\quad - \frac{1}{\sqrt\tau}\mathbb De(u_\tau^{k-1}){:}e\bigg(\frac{u_\tau^k-u_\tau^{k-1}}{\tau}\bigg) + (1-\sqrt\tau)\mathbb D\bigg(e\bigg(\frac{u_\tau^k-u_\tau^{k-1}}{\tau}\bigg)\bigg)^2
\end{aligned}$$

$$= \frac{\varphi(e(u_\tau^k)) - \varphi(e(u_\tau^{k-1}))}{\tau} + \frac{|e(u_\tau^k)|^\eta - |e(u_\tau^{k-1}|^\eta}{\eta} + \Big(1 - \frac{\sqrt{\tau}}{2}\Big)\mathbb{D}\Big(e\Big(\frac{u_\tau^k - u_\tau^{k-1}}{\tau}\Big)\Big)^2 \tag{4.15}$$

provided τ is sufficiently small as already specified; here we also used that $1-\frac{\sqrt{\tau}}{2} \geq 0$. Similarly, and even more simply, we use also the convexity of the kinetic energy $\frac{\varrho}{2}|\cdot|^2$ and also of $E \mapsto \frac{1}{2}\mathbb{H}(E)^2 + \frac{\tau}{\eta}|E|^\eta$. Thus, summing up over the time levels $l = 1, \ldots, k$, we arrive at the following estimate for the mechanical energy

$$\frac{\varrho}{2}\Big\|\frac{u_\tau^k - u_\tau^{k-1}}{\tau}\Big\|^2_{L^2(\Omega;\mathbb{R}^d)} + \sum_{l=1}^k \int_\Omega \Big(1 - \frac{\sqrt{\tau}}{2}\Big)\,\mathbb{D}\bigg(e\Big(\frac{u_\tau^l - u_\tau^{l-1}}{\tau}\Big)\bigg)^2 + \mathbb{G}\bigg(\nabla e\Big(\frac{u_\tau^l - u_\tau^{l-1}}{\tau}\Big)\bigg)^2 \,\mathrm{d}x + \int_\Omega \varphi(u_\tau^k) + \frac{1}{2}\mathbb{H}\big(\nabla e(u_\tau^k)\big)^2 + \frac{\tau}{\eta}|e(u_\tau^k)|^\eta + \frac{\tau}{\eta}|\nabla e(u_\tau^k)|^\eta \mathrm{d}x$$
$$\leq \tau \sum_{l=1}^k \bigg(\int_\Omega g_\tau^l \cdot \frac{u_\tau^l - u_\tau^{l-1}}{\tau} - \mathcal{A}(e(u_\tau^l), \vartheta_\tau^l){:}e\Big(\frac{u_\tau^l - u_\tau^{l-1}}{\tau}\Big)\,\mathrm{d}x + \int_\Gamma h_\tau^l \cdot \frac{u_\tau^l - u_\tau^{l-1}}{\tau}\,\mathrm{d}S\bigg) + \int_\Omega \frac{\varrho}{2}|v_0|^2 + \varphi(u_{0\tau}) + \frac{1}{2}\mathbb{H}\big(\nabla e(u_{0\tau})\big)^2 + \frac{\tau}{\eta}|e(u_{0\tau})|^\eta + \frac{\tau}{\eta}|\nabla e(u_{0\tau})|^\eta \,\mathrm{d}x. \tag{4.16}$$

Adding also the heat part (4.1b) tested by 1 and summed up over the time levels $l = 1, \ldots, k$, we arrive at the the following estimate for the total energy

$$\frac{\varrho}{2}\Big\|\frac{u_\tau^k - u_\tau^{k-1}}{\tau}\Big\|^2_{L^2(\Omega;\mathbb{R}^d)} + \big\|\vartheta_\tau^k\big\|_{L^1(\Omega)} + \int_\Omega \varphi(u_\tau^k) + \frac{1}{2}\mathbb{H}\big(\nabla e(u_\tau^k)\big)^2 + \frac{\tau}{\eta}|e(u_\tau^k)|^\eta + \frac{\tau}{\eta}|\nabla e(u_\tau^k)|^\eta \mathrm{d}x$$
$$\leq \tau \sum_{l=1}^k \bigg(\int_\Omega g_\tau^l \cdot \frac{u_\tau^l - u_\tau^{l-1}}{\tau}\,\mathrm{d}x + \int_\Gamma h_\tau^l \cdot \frac{u_\tau^l - u_\tau^{l-1}}{\tau} + f_\tau^l \,\mathrm{d}S\bigg) + \big\|\mathfrak{C}(e(u_{0\tau}), \theta_0)\big\|_{L^1(\Omega)} + \int_\Omega \frac{\varrho}{2}|v_0|^2 + \varphi(u_{0\tau}) + \frac{1}{2}\mathbb{H}\big(\nabla e(u_{0\tau})\big)^2 + \frac{\tau}{\eta}|e(u_{0\tau})|^\eta + \frac{\tau}{\eta}|\nabla e(u_{0\tau})|^\eta \,\mathrm{d}x. \tag{4.17}$$

This imitates the energy balance (3.12) with the regularizing terms added and again we can see cancellation of the dissipative and the $\mathcal{A}$-adiabatic terms. Note that here we benefit from having carefully designed (4.1b) with suitable coefficients $1-\sqrt{\tau}/2$ in the $\mathbb{D}$-term. Eventually, using the by-part summation of the boundary term

$$\tau \sum_{l=1}^k h_\tau^l \cdot \frac{u_\tau^l - u_\tau^{l-1}}{\tau} = h_\tau^k \cdot u_\tau^k - h_\tau^1 \cdot u_{0\tau} - \tau \sum_{l=1}^k \frac{h_\tau^{l-1} - h_\tau^{l-2}}{\tau} \cdot u_\tau^l \tag{4.18}$$

and the qualification (3.17h), from (4.17), the a-priori estimates (4.14) follow from the discrete Gronwall inequality. □

Proposition 4.3 (Further estimates). *Under the assumption of Lemma 4.2, for some constants C, C_r, C_s independent of $\tau \le 1$, it also holds:*

$$\big\|u_\tau\big\|_{W^{1,2}(I;W^{2,2}(\Omega;\mathbb{R}^d))} \le C, \tag{4.19a}$$

$$\big\|\nabla\bar\vartheta_\tau\big\|_{L^r(Q;\mathbb{R}^d)} \le C_r \qquad \textit{with} \;\; r < \frac{d+2}{d+1}, \tag{4.19b}$$

$$\big\|\bar\vartheta_\tau\big\|_{L^s(Q)} \le C_s \qquad \textit{with} \;\; s < \frac{d+2}{d}. \tag{4.19c}$$

Proof. We test (4.1b) by a special nonlinearity $\chi(\vartheta_\tau^k) := 1-(1{+}\vartheta_\tau^k)^{-\varepsilon}$ as suggested in [19], which has simplified the original technique from [5,6]. After summation for $k=1,\ldots,T/\tau$, we obtain

$$\begin{aligned}
\kappa_0\varepsilon\int_Q \frac{|\nabla\bar\vartheta_\tau|^2}{(1{+}\bar\vartheta_\tau)^{1+\varepsilon}}\,\mathrm{d}x\mathrm{d}t &= \kappa_0\int_Q \chi'(\bar\vartheta_\tau)|\nabla\bar\vartheta_\tau|^2\,\mathrm{d}x\mathrm{d}t\\
&\le \int_Q \chi'(\bar\vartheta_\tau)\mathcal{K}_0(e(\bar u_\tau),\bar\vartheta_\tau)\nabla\bar\vartheta_\tau{\cdot}\nabla\bar\vartheta_\tau\,\mathrm{d}x\mathrm{d}t = \int_Q \mathcal{K}_0(e(\bar u_\tau),\bar\vartheta_\tau)\nabla\bar\vartheta_\tau{\cdot}\nabla\chi(\bar\vartheta_\tau)\,\mathrm{d}x\mathrm{d}t\\
&\le \int_Q \mathcal{K}_0(e(\bar u_\tau),\bar\vartheta_\tau)\nabla\bar\vartheta_\tau{\cdot}\nabla\chi(\bar\vartheta_\tau)\,\mathrm{d}x\mathrm{d}t + \int_\Omega \widehat\chi(\bar\vartheta_\tau(T,\cdot))\,\mathrm{d}x\\
&\le \int_\Omega \widehat\chi(\vartheta_0)\,\mathrm{d}x + \int_\Sigma \bar f_\tau\chi(\bar\vartheta_\tau)\,\mathrm{d}S\mathrm{d}t + \int_Q \bar r_\tau\chi(\bar\vartheta_\tau)\\
&\qquad -\mathcal{K}_1(e(\bar u_\tau),\bar\vartheta_\tau)\nabla e(\bar u_\tau)\cdot\nabla\chi(\bar\vartheta_\tau)\,\mathrm{d}x\mathrm{d}t\\
&\le \big\|\mathfrak{C}(e(u_{0\tau}),\theta_0)\big\|_{L^1(\Omega)} + \big\|\bar f_\tau\big\|_{L^1(\Sigma)} + \big\|\bar r_\tau\big\|_{L^1(Q)}\\
&\qquad + \frac{\varepsilon T C_0^2}{4\delta} + \varepsilon\delta C_{\mathcal{K}_1}\int_Q \frac{|\nabla\bar\vartheta_\tau|^2}{(1{+}\bar\vartheta_\tau)^{1+\varepsilon}}\,\mathrm{d}x\mathrm{d}t
\end{aligned} \tag{4.20}$$

where κ_0 is from (3.17c), and where $\widehat\chi$ is a primitive of χ such that $\widehat\chi(0)=0$, and where we abbreviated the heat sources $\bar r_\tau := (1{-}\sqrt{\tau}/2)\mathbb{D}(e(\frac{\partial u_\tau}{\partial t}))^2 + \mathbb{G}(\nabla e(\frac{\partial u_\tau}{\partial t}))^2 + \mathcal{A}(e(\bar u_\tau),\bar\vartheta_\tau){:}e(\frac{\partial u_\tau}{\partial t})$. The inequality in the 4th line of (4.20) arises just from the mentioned test of (4.1b) by $\chi(\vartheta_\tau^k)$ when one uses monotonicity of χ, hence convexity of $\widehat\chi$, so that the "discrete chain rule" holds:

$$\frac{\widehat\chi(\vartheta_\tau^k)-\widehat\chi(\vartheta_\tau^{k-1})}{\tau} \le \frac{\vartheta_\tau^k-\vartheta_\tau^{k-1}}{\tau}\chi(\vartheta_\tau^k).$$

In the last line of (4.20), C_0 is from the already proved estimate (4.14a) and we have used that $\widehat\chi(\vartheta_0)\le\vartheta_0 =: \mathfrak{C}(e(u_{0\tau}),\theta_0)$, and that always $0\le\chi\le 1$, as well as

$$\begin{aligned}
\int_Q \mathcal{K}_1(e(\bar u_\tau),\bar\vartheta_\tau)\nabla e(\bar u_\tau)\cdot\nabla\chi(\bar\vartheta_\tau)\,\mathrm{d}x\mathrm{d}t &= \varepsilon\int_Q \nabla e(\bar u_\tau){\vdots}\frac{\mathcal{K}_1(e(\bar u_\tau),\bar\vartheta_\tau)\otimes\nabla\bar\vartheta_\tau}{(1{+}\bar\vartheta_\tau)^{1+\varepsilon}}\,\mathrm{d}x\mathrm{d}t\\
&\le \varepsilon\int_Q \frac{1}{4\delta}|\nabla e(\bar u_\tau)|^2 + \delta\frac{|\mathcal{K}_1(e(\bar u_\tau),\bar\vartheta_\tau)|^2}{1{+}\bar\vartheta_\tau}\frac{|\nabla\bar\vartheta_\tau|^2}{(1{+}\bar\vartheta_\tau)^{1+\varepsilon}}\,\mathrm{d}x\mathrm{d}t
\end{aligned}$$

$$\le \varepsilon\int_Q \frac{1}{4\delta}|\nabla e(\bar u_\tau)|^2 + \delta C_{\mathcal{K}_1}\frac{|\nabla\bar\vartheta_\tau|^2}{(1{+}\bar\vartheta_\tau)^{1+\varepsilon}}\,\mathrm{d}x\mathrm{d}t$$
$$\le \frac{\varepsilon T}{4\delta}\|\bar u_\tau\|^2_{L^\infty(I;W^{2,2}(\Omega;\mathbb{R}^d))} + \varepsilon\delta C_{\mathcal{K}_1}\int_Q \frac{|\nabla\bar\vartheta_\tau|^2}{(1{+}\bar\vartheta_\tau)^{1+\varepsilon}}\,\mathrm{d}x\mathrm{d}t \tag{4.21}$$

with $C_{\mathcal{K}_1}$ from (3.17e). Choosing δ small enough, namely $\delta = \kappa_0/(2C_{\mathcal{K}_1})$, from (4.20) we obtain after a little algebra an estimate

$$L := \int_Q \frac{|\nabla\bar\vartheta_\tau|^2}{(1{+}\bar\vartheta_\tau)^{1+\varepsilon}}\,\mathrm{d}x\mathrm{d}t \le \frac{2E_\tau}{\kappa_0\varepsilon} + \frac{\varepsilon T C_0^2 C_{\mathcal{K}_1}}{\kappa_0^2} + \frac{2}{\kappa_0\varepsilon}\int_Q |\bar r_\tau|\,\mathrm{d}x\mathrm{d}t$$
$$\le C_1 + C_1\int_Q\Big|\mathbb{D}\Big(e\Big(\frac{\partial u_\tau}{\partial t}\Big)\Big)^2 + \mathbb{G}\Big(\nabla e\Big(\frac{\partial u_\tau}{\partial t}\Big)\Big)^2 + \mathcal{A}\Big(e(\bar u_\tau),\bar\vartheta_\tau\Big){:}e\Big(\frac{\partial u_\tau}{\partial t}\Big)\Big|\,\mathrm{d}x\mathrm{d}t \tag{4.22}$$

with $E_\tau := \big\|\mathfrak{C}(e(u_{0\tau}),\theta_0)\big\|_{L^1(\Omega)} + \big\|\bar f_\tau\big\|_{L^1(\Sigma)}$ and with a suitable constant C_1 depending on $\max_{\tau>0} E_\tau$, i.e. on $\|\mathfrak{C}(e(u_0),\theta_0)\|_{L^1(\Omega)}$ and $\|f\|_{L^1(\Sigma)}$, as well as on ε, κ_0, $C_{\mathcal{K}_1}$, T, and C_0 from (4.14a), namely $C_1 = \max(2E/\varepsilon{+}\varepsilon T C_0^2 C_{\mathcal{K}_1}/\kappa_0, 2/\varepsilon)/\kappa_0$.

Further, we test (4.1a) by $u_\tau^k{-}u_\tau^{k-1}$ and handle it as in (4.16), sum it for $k = 1,\dots,T/\tau$, and add this to (4.22) with a sufficiently big weight, namely $2C_1$, to see the dissipation $\xi(e(\frac{\partial u_\tau}{\partial t}),\nabla e(\frac{\partial u_\tau}{\partial t}))$ on the left-hand side. Thus, after making still the discrete by-part integration $\int_0^T \bar h_\tau{\cdot}\frac{\partial u_\tau}{\partial t}\,\mathrm{d}t = h_\tau(T){\cdot}u_\tau(T) - h_\tau(\tau){\cdot}u_{0\tau} - \int_\tau^T \frac{\partial h_\tau}{\partial t}{\cdot}\bar u_\tau\,\mathrm{d}t$, cf. (4.18), we obtain

$$\int_Q \mathbb{D}\Big(e\Big(\frac{\partial u_\tau}{\partial t}\Big)\Big)^2 + \mathbb{G}\Big(\nabla e\Big(\frac{\partial u_\tau}{\partial t}\Big)\Big)^2\,\mathrm{d}x\mathrm{d}t + L$$
$$\le \frac{4C_1}{\min(C_1,1)(2{-}\sqrt{\tau})}\Big(E_{\eta,\tau} + \Big|\int_Q \bar g_\tau{\cdot}\frac{\partial u_\tau}{\partial t}\,\mathrm{d}x\mathrm{d}t + \int_\Sigma \bar h_\tau{\cdot}\frac{\partial u_\tau}{\partial t}\,\mathrm{d}S\mathrm{d}t\Big|\Big)$$
$$+\frac{2{+}4C_1}{\min(C_1,1)(2{-}\sqrt{\tau})}\int_Q \Big|\mathcal{A}(e(\bar u_\tau),\bar\vartheta_\tau){:}e\Big(\frac{\partial u_\tau}{\partial t}\Big)\Big|\,\mathrm{d}x\mathrm{d}t$$
$$\le \frac{4C_1}{\min(C_1,1)(2{-}\sqrt{\tau})}\Big(E_{\eta,\tau} + \big\|\bar g_\tau\big\|_{L^1(I;L^2(\Omega;\mathbb{R}^d))}\Big\|\frac{\partial u_\tau}{\partial t}\Big\|_{L^\infty(I;L^2(\Omega;\mathbb{R}^d))}$$
$$+\Big\|\frac{\partial h_\tau}{\partial t}\Big\|_{L^1(I;L^{p_h}(\Gamma;\mathbb{R}^d))}\big\|\bar u_\tau\big\|_{L^\infty(I;L^{p_h'}(\Gamma;\mathbb{R}^d))}$$
$$+\big\|h_\tau(\tau)\big\|_{L^\infty(I;L^{p_h}(\Gamma;\mathbb{R}^d))}\Big(\big\|u_\tau(T)\big\|_{L^{p_h'}(\Gamma;\mathbb{R}^d)} + \big\|u_{0\tau}\big\|_{L^{p_h'}(\Gamma;\mathbb{R}^d)}\Big)$$
$$+\frac{2{+}4C_1}{\min(C_1,1)(2{-}\sqrt{\tau})}\int_Q \Big|\mathcal{A}(e(\bar u_\tau),\bar\vartheta_\tau){:}e\Big(\frac{\partial u_\tau}{\partial t}\Big)\Big|\,\mathrm{d}x\mathrm{d}t$$
$$\le C_2 + C_2\Big\|\mathcal{A}(e(\bar u_\tau),\bar\vartheta_\tau){:}e\Big(\frac{\partial u_\tau}{\partial t}\Big)\Big\|_{L^1(Q)} \tag{4.23}$$

with $E_{\eta,\tau} := \int_\Omega \frac{\varrho}{2}|v_0|^2 + \varphi(u_{0\tau}) + \frac{1}{2}\mathbb{H}\big(\nabla e(u_{0\tau})\big)^2 + \frac{\tau}{\eta}|e(u_{0\tau})|^\eta + \frac{\tau}{\eta}|\nabla e(u_{0\tau})|^\eta\,\mathrm{d}x$ and, counting the already obtained estimate (4.14a) and the assumption $\|u_{0\tau}\|_{W^{2,\eta}(\Omega;\mathbb{R}^d)} = \mathcal{O}(\tau^{-1/\eta})$, with some C_2 depending on C_1, $\max_{\tau>0} E_{\eta,\tau}$, $\|g\|_{L^1(I;L^2(\Omega;\mathbb{R}^d))}$, $\|h\|_{W^{1,1}(I;L^{p_h}(\Gamma;\mathbb{R}^d))}$, and on C_0 from (4.14a). Here we used that $2-\sqrt{\tau} \geq 1 > 0$. Furthermore, for $1 \leq r < 2$, by Hölder's inequality, we have

$$
\begin{aligned}
\int_Q \big|\nabla\bar\vartheta_\tau\big|^r \mathrm{d}x\mathrm{d}t &= \int_Q \big|\nabla\bar\vartheta_\tau\big|^r \big|1+\bar\vartheta_\tau\big|^{-(1+\varepsilon)r/2}\big|1+\bar\vartheta_\tau\big|^{(1+\varepsilon)r/2}\mathrm{d}x\mathrm{d}t \\
&\leq \bigg(\int_Q \big|\nabla\bar\vartheta_\tau\big|^2\big|1+\bar\vartheta_\tau\big|^{-1-\varepsilon}\mathrm{d}x\mathrm{d}t\bigg)^{r/2}\bigg(\int_Q \big|1+\bar\vartheta_\tau\big|^{(1+\varepsilon)r/(2-r)}\mathrm{d}x\mathrm{d}t\bigg)^{(2-r)/2} \\
&= L^{r/2}\bigg(\int_0^T \big\|1+\bar\vartheta_\tau(t,\cdot)\big\|^{(1+\varepsilon)r/(2-r)}_{L^{(1+\varepsilon)r/(2-r)}(\Omega)}\mathrm{d}t\bigg)^{(2-r)/2}.
\end{aligned}
\tag{4.24}
$$

Then we use the Gagliardo–Nirenberg inequality to interpolate $L^{(1+\varepsilon)r/(2-r)}(\Omega)$ between $L^1(\Omega)$ and $W^{1,r}(\Omega)$ equipped with the equivalent norm $\|\cdot\|_{W^{1,r}(\Omega)} \cong \|\cdot\|_{L^1(\Omega)} + \|\nabla\cdot\|_{L^r(\Omega;\mathbb{R}^d)}$, i.e.

$$
\big\|\bar\vartheta_\tau\big\|_{L^{\frac{(1+\varepsilon)r}{2-r}}(\Omega)} \leq G_{r,\lambda,\varepsilon}\big\|\bar\vartheta_\tau\big\|^{1-\lambda}_{L^1(\Omega)}\Big(\big\|\bar\vartheta_\tau\big\|_{L^1(\Omega)} + \big\|\nabla\bar\vartheta_\tau\big\|_{L^r(\Omega;\mathbb{R}^d)}\Big)^\lambda. \tag{4.25}
$$

We obtain:

$$
\begin{aligned}
&\big\|1+\bar\vartheta_\tau(t,\cdot)\big\|_{L^{\frac{(1+\varepsilon)r}{2-r}}(\Omega)} \\
&\quad\leq G_{r,\lambda,\varepsilon}\Big(\big\|1+\bar\vartheta_\tau(t,\cdot)\big\|_{L^1(\Omega)} + \big\|\nabla\bar\vartheta_\tau(t,\cdot)\big\|_{L^r(\Omega;\mathbb{R}^d)}\Big)^\lambda\big\|1+\bar\vartheta_\tau(t,\cdot)\big\|^{1-\lambda}_{L^1(\Omega)} \\
&\quad\leq G_{r,\lambda,\varepsilon}\Big(|\Omega|+C+\big\|\nabla\bar\vartheta_\tau(t,\cdot)\big\|_{L^r(\Omega;\mathbb{R}^d)}\Big)^\lambda\big(|\Omega|+C_0\big)^{1-\lambda}
\end{aligned}
\tag{4.26}
$$

with C_0 from (4.14b). The Gagliardo–Nirenberg inequality (4.25) holds provided

$$
\frac{2-r}{(1+\varepsilon)r} \geq \lambda\Big(\frac{1}{r} - \frac{1}{d}\Big) + 1 - \lambda \qquad \text{with} \qquad 0 < \lambda \leq 1. \tag{4.27}
$$

We raise (4.26) to the power $(1+\varepsilon)r/(2-r)$, integrate it over $[0,T]$, use it in (4.24), and choose $\lambda := (2-r)/(1+\varepsilon)$. Thus we obtain

$$
\begin{aligned}
&\bigg(\int_0^T \big\|1+\bar\vartheta_\tau(t,\cdot)\big\|^{(1+\varepsilon)r/(2-r)}_{L^{(1+\varepsilon)r/(2-r)}(\Omega)}\mathrm{d}t\bigg)^{\frac{2-r}{2}} \\
&\leq \bigg(\int_0^T G_{r,\lambda,\varepsilon}^{\frac{(1+\varepsilon)r}{2-r}}\big(|\Omega|+C_0\big)^{\frac{(1-\lambda)(1+\varepsilon)r}{2-r}}\Big(|\Omega|+C_0+\|\nabla\bar\vartheta_\tau(t,\cdot)\|_{L^r(\Omega;\mathbb{R}^d)}\Big)^{\frac{\lambda(1+\varepsilon)r}{2-r}}\mathrm{d}t\bigg)^{\frac{2-r}{2}} \\
&\leq \bigg(\int_0^T G_{r,\lambda,\varepsilon}^{\frac{(1+\varepsilon)r}{2-r}}\big(|\Omega|+C_0\big)^{\frac{(1-\lambda)(1+\varepsilon)r}{2-r}}\Big(|\Omega|+C_0+\|\nabla\bar\vartheta_\tau(t,\cdot)\|_{L^r(\Omega;\mathbb{R}^d)}\Big)^{r}\mathrm{d}t\bigg)^{\frac{2-r}{2}} \\
&\leq C_4 + C_4\bigg(\int_Q |\nabla\bar\vartheta_\tau|^r\mathrm{d}x\mathrm{d}t\bigg)^{\frac{2-r}{2}}
\end{aligned}
\tag{4.28}
$$

with C_4 depending on C_0, $|\Omega|$, and $G_{r,\lambda,\varepsilon}$ from (4.25). Merging (4.24) with (4.28) gives $\|\nabla\bar\vartheta_\tau\|^r_{L^r(Q;\mathbb{R}^d)} \le L^{r/2}\big(C_4 + C_4\|\nabla\bar\vartheta_\tau\|^{r(1-r/2)}_{L^r(Q;\mathbb{R}^d)}\big)$. Further, using (4.22) gives $\|\nabla\bar\vartheta_\tau\|^r_{L^r(Q;\mathbb{R}^d)} \le (C_1 + C_1\|\bar r_\tau\|_{L^1(\Omega)})^{r/2}\big(C_4 + C_4\|\nabla\bar\vartheta_\tau\|^{r(1-r/2)}_{L^r(Q;\mathbb{R}^d)}\big)$, i.e. the estimate $\|\nabla\bar\vartheta_\tau\|^r_{L^r(Q;\mathbb{R}^d)}/\big(1 + \|\nabla\bar\vartheta_\tau\|^{r(1-r/2)}_{L^r(Q;\mathbb{R}^d)}\big) \le C_1C_4\big(1+\|\bar r_\tau\|_{L^1(Q)}\big)^{r/2}$. Altogether,

$$C_1^{2/r}C_4^{2/r}\big(1+\|\bar r_\tau\|_{L^1(Q)}\big) \ge \left(\frac{\|\nabla\bar\vartheta_\tau\|^r_{L^r(Q;\mathbb{R}^d)}}{1+\|\nabla\bar\vartheta_\tau\|^{r(1-r/2)}_{L^r(Q;\mathbb{R}^d)}}\right)^{2/r} \ge \|\nabla\bar\vartheta_\tau\|^r_{L^r(Q;\mathbb{R}^d)} - C_5. \tag{4.29}$$

The latter inequality in (4.29) holds for C_5 large enough because, obviously, the function $\xi\mapsto\big(\frac{\xi^r}{1+\xi^{r(1-r/2)}}\big)^{2/r}$ has the growth $(r-r(1{-}r/2))(2/r)=r$ for $\xi\to\infty$. Substituting our choice of $\lambda := (2{-}r)/(1{+}\varepsilon)$ into (4.27), one gets after some algebra the conditions

$$r \le \frac{2+d-\varepsilon d}{1+d}. \tag{4.30}$$

Note that $0{<}\lambda{<}1$ needed in (4.27) is automatically ensured by $1{\le}r{<}2$ and $\varepsilon{>}0$.

Let us perform another interpolation:

$$\begin{aligned}\Big\|e\Big(\frac{\partial u_\tau}{\partial t}\Big)\Big\|_{L^{q'}(Q;\mathbb{R}^{d\times d})} &\le \Big\|\nabla\frac{\partial u_\tau}{\partial t}\Big\|_{L^{q'}(Q;\mathbb{R}^{d\times d})}\\ &\le G_{q,\mu}\Big\|\frac{\partial u_\tau}{\partial t}\Big\|^{1-\mu}_{L^2(Q;\mathbb{R}^d)}\bigg(\Big\|\frac{\partial u_\tau}{\partial t}\Big\|_{L^2(Q;\mathbb{R}^d)} + \Big\|\nabla e\Big(\frac{\partial u_\tau}{\partial t}\Big)\Big\|_{L^2(Q;\mathbb{R}^{d\times d\times d})}\bigg)^{\mu},\end{aligned} \tag{4.31}$$

where the last inequality is based on the Korn and the Gagliardo–Nirenberg inequalities interpolating $W^{1,q'}(\Omega)$ in between $L^2(\Omega)$ and $W^{2,2}(\Omega)$, which needs

$$\frac{1}{q'} = \frac{1}{d} + \mu\Big(\frac12 - \frac2d\Big) + (1{-}\mu)\frac12 \qquad\text{and}\qquad \frac12 \le \mu \le 1. \tag{4.32}$$

Relying on the uniform-in-time bound of $\|\frac{\partial u_\tau}{\partial t}\|_{L^2(\Omega;\mathbb{R}^d)}$, cf. (4.14a), we can further estimate

$$\begin{aligned}\int_Q &\Big|\mathcal{A}(e(\bar u_\tau),\bar\vartheta_\tau){:}e\Big(\frac{\partial u_\tau}{\partial t}\Big)\Big|\,\mathrm{d}x\mathrm{d}t\\ &\le \big\|\mathcal{A}(e(\bar u_\tau),\bar\vartheta_\tau)\big\|_{L^q(Q;\mathbb{R}^{d\times d})}\Big\|e\Big(\frac{\partial u_\tau}{\partial t}\Big)\Big\|_{L^{q'}(Q;\mathbb{R}^{d\times d})}\\ &\le G_{q,\mu}T^{\frac{1-\mu}{2}}C_0^{1-\mu}\big\|\mathcal{A}(e(\bar u_\tau),\bar\vartheta_\tau)\big\|_{L^q(Q;\mathbb{R}^{d\times d})}\\ &\quad\times\bigg(\sqrt{T}C_0 + \Big\|\nabla e\Big(\frac{\partial u_\tau}{\partial t}\Big)\Big\|_{L^{q'}(Q;\mathbb{R}^{d\times d\times d})}\bigg)^{\mu}\\ &\le C_6 + C_6\big\|\mathcal{A}(e(\bar u_\tau),\bar\vartheta_\tau)\big\|^q_{L^q(Q;\mathbb{R}^{d\times d})} + \delta\Big\|\nabla e\Big(\frac{\partial u_\tau}{\partial t}\Big)\Big\|^{\mu q'}_{L^2(Q;\mathbb{R}^{d\times d\times d})}\end{aligned} \tag{4.33}$$

with C_6 depending on μ, q, T, $G_{q,\mu}$ from (4.31), C_0 from (4.14a), and on $\delta > 0$ to be choosen. It leads to the choice $\mu q' = 2$, which allows for absorbing the last term of (4.33) on the left-hand side of (4.23) when choosing δ sufficiently small, e.g. $\delta = 1/(2C_2 c_{\mathbb{G}})$ with $c_{\mathbb{G}} > 0$ the positive-definitness constant of the tensor $\mathbb{G}$. Substituting $\mu = 2/q'$ into the former condition in (4.32) gives, after some algebra $q' = (2d{+}8)/(d{+}2)$ so that $q = q'/(q'{-}1)$ results as

$$q = \frac{2d+8}{d+6}. \tag{4.34}$$

By Hölder inequality using (3.17f), we can further estimate

$$\begin{aligned}\big\|\mathcal{A}(e(\bar{u}_\tau), \bar{\vartheta}_\tau)\big\|^q_{L^q(Q;\mathbb{R}^{d\times d})} &\le 3^{q-1} C^q_{\mathcal{A}} \int_Q 1 + |e|^{2^*} + |\vartheta|^s \,\mathrm{d}x\mathrm{d}t \\ &\le C_7 + C_7 \big\|e(\bar{u}_\tau)\big\|^{2^*}_{L^{2^*}(Q;\mathbb{R}^{d\times d})} + C_7 \big\|\bar{\vartheta}_\tau\big\|^s_{L^s(Q)};\end{aligned} \tag{4.35}$$

with C_7 depending on q, $|\Omega|$, and $C_{\mathcal{A}}$ from (3.17f). Further, we estimate the last term (4.35) by using once more the Gagliardo–Nirenberg inequality to interpolate $L^s(\Omega)$ in between $L^1(\Omega)$ and $W^{1,r}(\Omega)$:

$$\big\|\bar{\vartheta}_\tau\big\|_{L^s(\Omega)} \le N_{s,r,\zeta} \big\|\bar{\vartheta}_\tau\big\|^{1-\zeta}_{L^1(\Omega)} \Big(\big\|\bar{\vartheta}_\tau\big\|_{L^1(\Omega)} + \big\|\nabla\bar{\vartheta}_\tau\big\|_{L^r(\Omega;\mathbb{R}^d)}\Big)^\zeta \tag{4.36}$$

for

$$\frac{1}{s} \ge \zeta\Big(\frac{1}{r} - \frac{1}{d}\Big) + 1 - \zeta, \qquad 0 \le \zeta < 1, \tag{4.37}$$

and by (4.24), (4.28), and the last estimate in (4.29), we have $\|\nabla\bar{\vartheta}_\tau\|_{L^r(\Omega;\mathbb{R}^d)} \le C_4^{2/r} L + C_5$ so that we can estimate

$$\begin{aligned}C_7 \big\|\bar{\vartheta}_\tau\big\|^s_{L^s(\Omega)} &\le C_7 N^s_{s,r,\zeta} C_3^{(1-\zeta)s} \Big(C_3 + \big\|\nabla\bar{\vartheta}_\tau\big\|_{L^r(\Omega;\mathbb{R}^d)}\Big)^{\zeta s} \\ &\le C_8 + \frac{1}{2C_4^{2/r}} \big\|\nabla\bar{\vartheta}_\tau\big\|^r_{L^r(\Omega;\mathbb{R}^d)} \le C_5 + C_8 + \frac{L}{2}\end{aligned} \tag{4.38}$$

with L from (4.22) and with C_8 depending on $N_{s,r,\zeta}$, C_3, C_4, C_7, s, and ζ. Here we needed $\zeta s < r$, i.e.

$$\frac{1}{s} > \frac{\zeta}{r}. \tag{4.39}$$

The last term in (4.38) is to be absorbed in the left-hand side of (4.23). Optimal ζ makes the right-hand sides of (4.37) and (4.39) mutually equal, which gives $\zeta = d/(d{+}1)$. Note that always $0 < \zeta < 1$, as required for (4.36). Taking into account $r < (d{+}2)/(d{+}1)$, from (4.37) (or, equally, from (4.39)) we obtain

$$s < \frac{d+2}{d}. \tag{4.40}$$

This is indeed satisfied if $s < q$ as used in (3.17f) provided $d \le 3$, which was also needed in the proof of Lemma 4.1. It eventually gives the estimate (4.19b) and, from the first and the second terms in (4.23), also the estimate (4.19a). By interpolation with (4.14b), we then obtain also (4.41a). □

Proposition 4.4 (Further estimates II). *Under the assumption of Lemma 4.2, for some constant C independent of τ, it also holds:*

$$\Big\|\varrho\frac{\partial}{\partial t}\Big[\frac{\partial u_\tau}{\partial t}\Big]^{\rm i}\Big\|_{L^2(I;W^{2,2}(\Omega;\mathbb{R}^d)^*)+L^{\eta'}(I;W^{2,\eta}(\Omega;\mathbb{R}^d)^*)}\le C, \tag{4.41a}$$

$$\Big\|\varrho\frac{\partial}{\partial t}\Big[\frac{\partial u_\tau}{\partial t}\Big]^{\rm i}-\tau{\rm div}\big(|e(\bar u_\tau)|^{\eta-2}e(\bar u_\tau)\big) +\tau{\rm div}^2\big(|\nabla e(\bar u_\tau)|^{\eta-2}\nabla e(\bar u_\tau)\big)\Big\|_{L^2(I;W^{2,2}(\Omega;\mathbb{R}^d)^*)}\le C, \tag{4.41b}$$

$$\Big\|\frac{\partial\vartheta_\tau}{\partial t}\Big\|_{L^1(I;W^{3,2}(\Omega)^*)}\le C. \tag{4.41c}$$

It is to explain that the "dual" norm in (4.41a) and (4.41b) of expressions like $a-{\rm div}\,\sigma+{\rm div}^2\mathfrak{h}$ refers to the linear functional

$$z\mapsto\int_Q a{\cdot}z+\sigma{:}e(z)+\mathfrak{h}\vdots\nabla e(z)\,{\rm d}x{\rm d}t +\int_\Sigma\Big(\big(\sigma{+}{\rm div}\,\mathfrak{h}\big){\cdot}\nu+{\rm div}_{\rm S}(\mathfrak{h}\cdot\nu)-({\rm div}_{\rm S}\nu)\big(\mathfrak{h}{:}(\nu{\otimes}\nu)\big)\Big)\cdot z -\big(\mathfrak{h}{:}(\nu{\otimes}\nu)\big)\cdot\frac{\partial z}{\partial\nu}\,{\rm d}S{\rm d}t, \tag{4.42}$$

cf. also (3.15).

Proof of Proposition 4.4. These "dual" estimates can be obtained routinely as a consequence of the previously derived ones; note that (4.41b) relies on the growth restriction used in (3.17f) so that $\mathcal{A}(e(\bar u_\tau),\bar\vartheta_\tau)$ is bounded in $L^q(Q;\mathbb{R}^{d\times d})$. □

Proposition 4.5 (Convergence for $\tau\to0$). *Let (3.17) hold and η used in (4.1) be sufficiently large as specified in Lemma 4.1, and let also $u_{0\tau}\to u_0$ in $W^{2,2}(\Omega;\mathbb{R}^d)$ and even $\|u_{0\tau}\|_{W^{2,\eta}(\Omega;\mathbb{R}^d)}=o(\tau^{-1/\eta})$. Then there is a subsequence such that*

$$u_\tau\to u\quad\text{strongly in } W^{1,2}(I;W^{2,2}(\Omega;\mathbb{R}^d)), \tag{4.43a}$$

$$\bar\vartheta_\tau\to\vartheta\quad\text{strongly in }\ L^s(Q)\ \text{ with any }\ s<(d{+}2)/d, \tag{4.43b}$$

and any (u,ϑ) obtained in this way is a weak solution to the initial-boundary-value problem (3.9)–(3.10).

Proof. Choose a weakly* converging subsequence $\{(u_\tau,\bar\vartheta_\tau)\}_{\tau>0}$ in the topology of the estimates (4.14a)–(4.19a-c). We now use (and modify) some tricks from [41, Formulas (3.61)–(3.67)]: By (4.41b), we can choose this subsequence so that also

$$\varrho\frac{\partial}{\partial t}\Big[\frac{\partial u_\tau}{\partial t}\Big]^{\rm i}-\tau{\rm div}\big(|e(\bar u_\tau)|^{\eta-2}e(\bar u_\tau)\big)+\tau{\rm div}^2\big(|\nabla e(\bar u_\tau)|^{\eta-2}\nabla e(\bar u_\tau)\big)\to\zeta \tag{4.44}$$

weakly in $L^2(I;W^{2,2}(\Omega;\mathbb{R}^d)^*)$ for some $\zeta\in L^2(I;W^{2,2}(\Omega;\mathbb{R}^d)^*)$. For any w smooth in space with a compact support in Q and piecewise affine in time on the partition of $[0,T]$ with some time-step $\tau_0>0$, it holds

$$\begin{aligned}\langle\zeta,w\rangle &= \lim_{\tau\to0}\Big\langle \varrho\frac{\partial}{\partial t}\Big[\frac{\partial u_\tau}{\partial t}\Big]^{\rm i} - \tau{\rm div}\big(|e(\bar u_\tau)|^{\eta-2}e(\bar u_\tau)\big)\\ &\qquad +\tau{\rm div}^2\big(|\nabla e(\bar u_\tau)|^{\eta-2}\nabla e(\bar u_\tau)\big),w\Big\rangle\\ &= -\lim_{\tau\to0}\int_Q \varrho\Big[\frac{\partial u_\tau}{\partial t}\Big]^{\rm i}\cdot\frac{\partial w}{\partial t} - \tau\big|e(\bar u_\tau)\big|^{\eta-2}e(\bar u_\tau){:}e(w)\\ &\qquad +\tau|\nabla e(\bar u_\tau)|^{\eta-2}\nabla e(\bar u_\tau)\,\vdots\,\nabla e(w)\,{\rm d}x{\rm d}t = -\int_Q\varrho\frac{\partial u}{\partial t}\cdot\frac{\partial w}{\partial t}\,{\rm d}x{\rm d}t; \end{aligned}\tag{4.45}$$

note that w simply vanishes on $[0,\tau]$ for τ small enough. The second equality in (4.45) is due to the Green formula and the by-part-summation formula (see (4.51) below for a full version with nonvanishing end points) while the last equality in (4.45) is based on the convergence $[\frac{\partial u_\tau}{\partial t}]^{\rm i}\to\frac{\partial u}{\partial t}$ weakly in $L^2(I;W^{2,2}(\Omega;\mathbb{R}^d))$ and on the estimates

$$\begin{aligned}&\Big|\int_Q\tau|e(\bar u_\tau)|^{\eta-2}e(\bar u_\tau){:}e(w){\rm d}x{\rm d}t\Big|\\ &\qquad\le\tau\big\|e(\bar u_\tau)\big\|^{\eta-1}_{L^\eta(Q;\mathbb{R}^{d\times d})}\big\|e(w)\big\|_{L^\eta(Q;\mathbb{R}^{d\times d})}=\mathcal{O}(\tau^{1/\eta})\to0,\end{aligned}\tag{4.46}$$

$$\begin{aligned}&\Big|\int_Q\tau|\nabla e(\bar u_\tau)|^{\eta-2}\nabla e(\bar u_\tau)\,\vdots\,\nabla e(v){\rm d}x{\rm d}t\Big|\\ &\qquad\le\tau\big\|\nabla e(\bar u_\tau)\big\|^{\eta-1}_{L^\eta(Q;\mathbb{R}^{d\times d\times d})}\big\|\nabla e(v)\big\|_{L^\eta(Q;\mathbb{R}^{d\times d\times d})}=\mathcal{O}(\tau^{1/\eta})\to0\end{aligned}\tag{4.47}$$

based on (4.14c). Extending $\langle\zeta,w\rangle=-\int_Q\varrho\frac{\partial u}{\partial t}\cdot\frac{\partial w}{\partial t}\,{\rm d}x{\rm d}t$ obtained in (4.45) for all w smooth with a compact support, this shows that ζ is, in the sense of distributions, equal to $\varrho\frac{\partial^2u}{\partial t^2}$. In particular, we have shown that

$$\varrho\frac{\partial^2 u}{\partial t^2}=\ddot u\in L^2(I;W^{2,2}(\Omega;\mathbb{R}^d)^*).\tag{4.48}$$

Combining (4.44) with (4.48) allows also for

$$\begin{aligned}\varrho\frac{\partial u_\tau}{\partial t}(T)&=\varrho v_0+\int_0^T\varrho\frac{\partial}{\partial t}\Big[\frac{\partial u_\tau}{\partial t}\Big]^{\rm i}{\rm d}t\\ &=\varrho v_0+\int_0^T\varrho\frac{\partial}{\partial t}\Big[\frac{\partial u_\tau}{\partial t}\Big]^{\rm i}-\tau{\rm div}\big(|e(\bar u_\tau)|^{\eta-2}e(\bar u_\tau)\big)\\ &\quad+\tau{\rm div}^2\big(|\nabla e(\bar u_\tau)|^{\eta-2}\nabla e(\bar u_\tau)\big)\,{\rm d}t\\ &\quad+\int_0^T\tau{\rm div}\big(|e(\bar u_\tau)|^{\eta-2}e(\bar u_\tau)\big)-\tau{\rm div}^2\big(|\nabla e(\bar u_\tau)|^{\eta-2}\nabla e(\bar u_\tau)\big)\,{\rm d}t\\ &\to\varrho v_0+\int_0^T\zeta\,{\rm d}t=\varrho v_0+\int_0^T\varrho\frac{\partial^2u}{\partial t^2}\,{\rm d}t=\varrho\frac{\partial u}{\partial t}(T)\quad\text{weakly in }W^{2,2}(\Omega;\mathbb{R}^d)^*.\end{aligned}\tag{4.49}$$

Due to the estimate (4.14a), $\frac{\partial u_\tau}{\partial t}(T)$ is bounded also in $L^2(\Omega;\mathbb{R}^d)$ and thus we obtain also

$$\frac{\partial u_\tau}{\partial t}(T) \to \frac{\partial u}{\partial t}(T) \qquad \text{weakly in } \ L^2(\Omega;\mathbb{R}^d). \tag{4.50}$$

Furthermore, by the (suitably interpolated) Aubin-Lions' lemma, combining (4.14b), (4.19b) and (4.41c), one obtains (4.43b).

Now we test the boundary-value problem (4.1a)–(4.2a,b) by a smooth v_τ^k, sum it for $k = 1, \dots T/\tau$, and use the by-part summation

$$\sum_{k=1}^{T/\tau} \big(u^k - 2u^{k-1} + u^{k-2}\big) \cdot z^k = \big(u^{T/\tau} - u^{T/\tau-1}\big) \cdot z^{T/\tau} \\ -\big(u^0 - u^{-1}\big) \cdot v^1 - \sum_{k=2}^{T/\tau} \big(u^{k-1} - u^{k-2}\big) \cdot \big(z^k - z^{k-1}\big) \tag{4.51}$$

to obtain the identity

$$\int_\Omega \varrho \frac{\partial u_\tau}{\partial t}(T) \cdot v_\tau(T)\,\mathrm{d}x - \int_\tau^T \!\! \int_\Omega \varrho \frac{\partial u_\tau}{\partial t}(\cdot - \tau) \cdot \frac{\partial v_\tau}{\partial t}\,\mathrm{d}x\mathrm{d}t \\ + \int_Q \Big(\mathbb{D}e\Big(\frac{\partial u_\tau}{\partial t}\Big) + \varphi'(e(\bar u_\tau)) \\ + \mathcal{A}(e(\bar u_\tau), \bar\vartheta_\tau) + \tau\big|e(\bar u_\tau)\big|^{\eta-2} e(\bar u_\tau)\Big) : e(\bar v_\tau) + \bar{\mathfrak{h}}_\tau \vdots \nabla e(\bar v_\tau)\,\mathrm{d}x\mathrm{d}t \\ = \int_\Omega v_0 \cdot v_\tau(\tau)\,\mathrm{d}x + \int_Q \bar g_\tau \cdot \bar v_\tau\,\mathrm{d}x\mathrm{d}t + \int_\Sigma \bar h_\tau \cdot \bar v_\tau\,\mathrm{d}S\mathrm{d}t, \tag{4.52}$$

where $\bar{\mathfrak{h}}_\tau = \mathbb{G}\nabla e(\frac{\partial u_\tau}{\partial t}) + \mathbb{H}\nabla e(\bar u_\tau) + \tau|\nabla e(\bar u_\tau)|^{\eta-2}\nabla e(\bar u_\tau)$ and $\bar v_\tau$ and v_τ is respectively the piecewise constant and the piecewise affine interpolant in time of $(v_\tau^k)_{k=0}^{T/\tau}$.

We know that

$$\frac{\partial u_\tau}{\partial t}(\cdot - \tau) \to \frac{\partial u}{\partial t} \qquad \text{strongly in } L^2(Q;\mathbb{R}^d), \tag{4.53a}$$

$$\varphi'(e(\bar u_\tau)) \to \varphi'(e(u)) \qquad \text{strongly in } L^{2^{*\prime}}(Q;\mathbb{R}^{d\times d}), \tag{4.53b}$$

$$\mathcal{A}(e(\bar u_\tau), \bar\vartheta_\tau) \to \mathcal{A}(e(u), \vartheta) \qquad \text{strongly in } L^q(Q;\mathbb{R}^{d\times d}), \tag{4.53c}$$

$$\mathbb{D}e\Big(\frac{\partial u_\tau}{\partial t}\Big) \to \mathbb{D}e\Big(\frac{\partial u}{\partial t}\Big) \qquad \text{weakly in } L^2(Q;\mathbb{R}^{d\times d}), \tag{4.53d}$$

$$\mathbb{G}\nabla e\Big(\frac{\partial u_\tau}{\partial t}\Big) \to \mathbb{G}\nabla e\Big(\frac{\partial u}{\partial t}\Big) \qquad \text{weakly in } L^2(Q;\mathbb{R}^{d\times d\times d}); \tag{4.53e}$$

for the above strong convergences, we used Aubin-Lions' theorem (possibly with an interpolation). In particular, we employed the growth reserve due to strict inequalities in (3.17b) and (3.17f). Eventually, using again (4.46)–(4.47), we can see that the nonlinear regularizing η-terms disappear in the limit for $\tau \to 0$.

It is now essential that we can legally substitute $v = \frac{\partial u}{\partial t}$ into (3.16a). For this, it is important that $\varrho\frac{\partial^2 u}{\partial t^2}$ in duality with $\frac{\partial u}{\partial t}$, and also that $\mathcal{A}(e(u), \vartheta) \in L^q(Q;\mathbb{R}^{d\times d})$ as well as $\varphi'(e(u)) \in L^\infty(I; L^{2^{*\prime}}(\Omega;\mathbb{R}^{d\times d}))$ are in duality with

$e(\frac{\partial u}{\partial t}) \in L^{q'}(Q;\mathbb{R}^{d\times d})\cap L^2(I;L^{2^*}(\Omega;\mathbb{R}^{d\times d}))$. By this substitution, after by-part integration in time, we obtain the *mechanical-energy equality*

$$\begin{aligned}\int_\Omega \frac{\varrho}{2}\Big|\frac{\partial u}{\partial t}(T)\Big|^2+\varphi\big(e(u(T))\big)+\frac12\mathbb{H}\big(\nabla e(u(T))\big)^2\mathrm{d}x\\ +\int_Q \mathbb{D}\Big(e\Big(\frac{\partial u}{\partial t}\Big)\Big)^2+\mathbb{G}\Big(\nabla e\Big(\frac{\partial u}{\partial t}\Big)\Big)^2\mathrm{d}x\mathrm{d}t\\ =\int_\Omega \frac{\varrho}{2}|v_0|^2+\varphi(e(u_0))+\frac12\mathbb{H}\big(\nabla e(u_0)\big)^2\mathrm{d}x\\ +\int_Q g\cdot\frac{\partial u}{\partial t}-\mathcal{A}(e(u),\vartheta){:}e\Big(\frac{\partial u}{\partial t}\Big)\,\mathrm{d}x\mathrm{d}t+\int_\Sigma h\cdot\frac{\partial u}{\partial t}\,\mathrm{d}x\mathrm{d}S\end{aligned}\tag{4.54}$$

with $\varphi(e)=\psi(e,0)$ from (3.7).

For the limit passage in the heat equation, we need to prove the strong L^2-convergence of $e(\frac{\partial u_\tau}{\partial t})$ and of $\nabla e(\frac{\partial u_\tau}{\partial t})$. We use

$$\begin{aligned}&\int_Q \mathbb{D}\Big(\frac{\partial e(u)}{\partial t}\Big)^2+\mathbb{G}\Big(\nabla e\Big(\frac{\partial u}{\partial t}\Big)\Big)^2\,\mathrm{d}x\mathrm{d}t\\ &\le \liminf_{\tau\to0}\int_Q \mathbb{D}\Big(\frac{\partial e(u_\tau)}{\partial t}\Big)^2+\mathbb{G}\Big(\nabla e\Big(\frac{\partial u_\tau}{\partial t}\Big)\Big)^2\,\mathrm{d}x\mathrm{d}t\\ &\le \limsup_{\tau\to0}\int_Q \Big(1-\frac{\sqrt{\tau}}{2}\Big)\mathbb{D}\Big(\frac{\partial e(u_\tau)}{\partial t}\Big)^2+\mathbb{G}\Big(\nabla e\Big(\frac{\partial u_\tau}{\partial t}\Big)\Big)^2\,\mathrm{d}x\mathrm{d}t\\ &\le \limsup_{\tau\to0}\int_\Omega \frac{\varrho}{2}|v_0|^2+\varphi(e(u_{0\tau}))+\frac12\mathbb{H}\big(\nabla e(u_{0\tau})\big)^2-\frac{\varrho}{2}\Big|\frac{\partial u_\tau}{\partial t}(T)\Big|^2\\ &\qquad-\varphi\big(e(u_\tau(T))\big)-\frac12\mathbb{H}\big(\nabla e(u_\tau(T))\big)^2+\frac{\tau}{\eta}\big|e(u_{0\tau})\big|^\eta+\frac{\tau}{\eta}\big|\nabla e(u_{0\tau})\big|^\eta\mathrm{d}x\\ &\qquad+\int_Q \bar g_\tau\cdot\frac{\partial u_\tau}{\partial t}-\mathcal{A}(e(\bar u_\tau),\bar\vartheta_\tau){:}e\Big(\frac{\partial u_\tau}{\partial t}\Big)\,\mathrm{d}x\mathrm{d}t+\int_\Sigma \bar h_\tau\cdot\frac{\partial u_\tau}{\partial t}\,\mathrm{d}x\mathrm{d}S\\ &\le \int_\Omega \frac{\varrho}{2}|v_0|^2+\varphi(e(u_0))+\frac12\mathbb{H}\big(\nabla e(u_0)\big)^2\\ &\qquad-\frac{\varrho}{2}\Big|\frac{\partial u}{\partial t}(T)\Big|^2-\varphi\big(e(u(T))\big)-\frac12\mathbb{H}\big(\nabla e(u_\tau(T))\big)^2\,\mathrm{d}x\\ &\qquad+\int_Q g\cdot\frac{\partial u}{\partial t}-\mathcal{A}(e(u),\vartheta){:}e\Big(\frac{\partial u}{\partial t}\Big)\,\mathrm{d}x\mathrm{d}t+\int_\Sigma h\cdot\frac{\partial u}{\partial t}\,\mathrm{d}x\mathrm{d}S\\ &=\int_Q \mathbb{D}\Big(\frac{\partial e(u)}{\partial t}\Big)^2+\mathbb{G}\Big(\nabla e\Big(\frac{\partial u_\tau}{\partial t}\Big)\Big)^2\,\mathrm{d}x\mathrm{d}t.\end{aligned}\tag{4.55}$$

Note that the third inequality is due to (4.16) while the last equality is exactly (4.54). Thus $\lim_{\tau\to0}\int_Q \mathbb{D}(e(\frac{\partial u_\tau}{\partial t}))^2\,\mathrm{d}x\mathrm{d}t=\int_Q \mathbb{D}(e(\frac{\partial u}{\partial t}))^2\,\mathrm{d}x\mathrm{d}t$ and also $\lim_{\tau\to0}\int_Q \mathbb{G}(\nabla e(\frac{\partial u_\tau}{\partial t}))^2\,\mathrm{d}x\mathrm{d}t = \int_Q \mathbb{G}(\nabla e(\frac{\partial u}{\partial t}))^2\,\mathrm{d}x\mathrm{d}t$. By (3.17g), the quadratic forms $\mathbb{D}(\cdot)^2$ and $\mathbb{G}(\cdot)^2$ are coercive and thus we obtain both $e(\frac{\partial u_\tau}{\partial t})\to e(\frac{\partial u}{\partial t})$ strongly in $L^2(Q;\mathbb{R}^{d\times d})$ and also $\nabla e(\frac{\partial u_\tau}{\partial t})\to\nabla e(\frac{\partial u}{\partial t})$ strongly in

$L^2(Q;\mathbb{R}^{d\times d\times d})$. Then the limit passage in the semi-linear heat equation is immediate. □

Note that the information (3.18a)–(3.18b) does not exactly follow from the estimates (4.41), partly also because $W^{1,1}$ is not weakly* compact. Anyhow (3.18a) has been obtained in (4.48), while the fact $W^{1,1}(I;W^{3,2}(\Omega)^*)$ in (3.18b) can be recovered from the limit heat equation itself.

5. Final discussion

Finally we outline some generalizations or modifications and some little more specific examples.

Remark 5.1 (Higher-dimensional case). In fact, Theorem 3.2 would hold for $d \geq 4$ provided the growth restriction (3.17f) would be strengthened by considering (4.40).

Remark 5.2 (Weakening of uniform positive-definiteness (3.17c) of $\mathcal{K}_0$). The condition (3.17c) might sometimes be found restrictive. Hencefore, it is worth realizing that

$$\mathcal{K}_0(e,\vartheta)\xi\cdot\xi \geq \kappa_0\frac{|\xi|^2}{(1{+}\vartheta)^a} \tag{5.1}$$

with some $a > 0$ is possible: the exponent ε in the formulas (4.20)–(4.28) replaces by $\varepsilon{+}a$ and the integrability of $\nabla\vartheta \in L^r(Q;\mathbb{R}^d)$ obtained in (4.30) is then worse, namely $r \leq \frac{2+d-\varepsilon d-ad}{1+d}$. It gives a restriction $a < 1/d$ not to be in conflict with $r \geq 1$. Of course, then also s calculated from (4.39) is modified.

Remark 5.3 (Weakening the growth restriction (3.17f) of $\mathcal{A}$). The restriction $s < q$ in (3.17f), i.e. the sub-linear growth of $\mathcal{A}(e,\cdot)$, was needed only in the estimate (4.10) to guarantee coercivity of the nonlinear operator arising in the time-discretisation. A finer treatment relying only on a weaker condition (4.40) might be executed: first approximate $\mathcal{A}$ so that the growth (3.17f) is satisfied, prove existence of a solution as above, then pass to the limit with this approximation using all the above a-priori estimates, in particular also the interpolation relying only on the weaker restriction (4.40).

Example 5.4 (Conventional model). One can re-consider the partly linear ansatz (2.12), yielding the conventional model from Remark 2.1, and check the assumptions for

$$\mathbb{K}(e,\theta) = \sum_{i=1}^{I}\kappa_i(e)\theta^{a_i} \qquad \text{and} \qquad \phi_2(\theta) = \begin{cases} c_0\theta^c/(c^2{-}c) & \text{if } c \neq 1,\\ c_0\theta\ln(\theta{-}1) & \text{if } c = 1,\end{cases} \tag{5.2}$$

considered for $c_0 > 0$ and some exponent $c > 0$. Now $\mathfrak{c}(e,\theta) = \mathfrak{c}(\theta) = -\theta[\phi_2]''_{\theta\theta}(\theta) = c_0\theta^{c-1}$ so that $\mathfrak{C}(\theta) = c_0\theta^c/c$, and thus we would have simply $\Theta(e,\vartheta) = k_0\vartheta^{1/c}$ with $k_0 = (c/c_0)^{1/c}$. In particular, we can observe that

$\Theta'_e = 0$, $\mathcal{K}_1 = 0$, and, by using $\Theta'_\vartheta = k_0\vartheta^{(1-c)/c}/c$,

$$\mathcal{K}_0(e,\vartheta) = \frac{k_0}{c}\mathbb{K}(e, k_0\vartheta^{1/c})\vartheta^{(1-c)/c} = \sum_{i=1}^{I} \frac{k_0^{a_i+1}}{c}\kappa_i(e)\vartheta^{(a_i+1-c)/c}, \quad (5.3)$$

$$\mathcal{A}(e,\vartheta) = \phi'_1(e)\Theta(\vartheta) = k_0\phi'_1(e)\vartheta^{1/c} \quad (5.4)$$

with ϕ_1 from (2.12). Let us verify the assumptions (3.17) for the physically relevant dimension $d = 3$: The growth condition (3.17d) needs $0 \le (a_i{+}1{-}c)/c < 1/3$, i.e. $c \le a_i{+}1 < \frac{4}{3}c$ while the uniform positive-definiteness (3.17c) of $\mathcal{K}_0$ needs $a_i{+}1{-}c = 0$ for at least some i, say, for $i = 1$. In fact, this restriction might be avoided when applying Remark 5.2. Of course, (3.17e) is now trivial since $\mathcal{K}_1 = 0$. Counting $q > 14/9$, $s < 5/3$, and $2^* = 6$, the growth condition (3.17f), i.e. $|\mathcal{A}(e,\vartheta)| = |k_0\phi'_1(e)\vartheta^{1/c}| \le (1 + |e|^{27/7-\epsilon} + \vartheta^{1-\epsilon})$, which will work for $c > 1$. If Remark 5.3 would be adopted, then even $c > 14/15$ could be allowed. Altogether, we can see quite a big freedom in this case. In fact, in [42, Remark 6] this bound was improved for $c > 15/18$ by allowing higher-order terms similar to ours. In general, a comparison with Remark 2.1 reveals that the data qualification here was weakened but this is obviously a not surprising effect of considering the higher-order terms.

Example 5.5 (Nonlinear coupling). In a general coupling, it is not entirely easy to reveal what the restrictions (3.17) really imposes on the original data ψ and $\mathbb{K}$. Anyhow, one can relatively easily check a perturbation of the previous Example 5.4 by augmenting (2.12) by adding a general smooth compactly supported coupling term. In particular, this term vanishes for $\theta \ge \theta_{\max}$ with some $\theta_{\max} > 0$, which causes the heat capacity $\mathfrak{c}(e,\theta)$ to be independent of e at least for $\theta \ge \theta_{\max}$. Due to the formula (3.2), this perturbation influences $\mathfrak{C}$ also for $\theta \ge \theta_{\max}$ but we can at least rely on the splitting $\mathfrak{C}(e,\theta) = \mathfrak{C}_0(e) + \mathfrak{C}_1(\theta)$ with $\mathfrak{C}_0(e) = \mathfrak{C}(e,\theta_{\max})$ for $\theta \ge \theta_{\max}$. Then $\Theta(e,\vartheta) = \mathfrak{C}_1^{-1}(\vartheta{-}\mathfrak{C}_0(e))$. Now $\mathcal{K}_1$ does not vanish. Checking at least the asymptotics for large θ, one can see that nothing is changed for $\mathcal{K}_0$ and also the additional term which occurs in $\mathcal{A}$ defined in (3.4d), i.e. $\mathfrak{C}'_e(e,\Theta(e,\vartheta))$, does not change the asymptotics because $\mathfrak{C}'_e$ is bounded for high ϑ. Since $\Theta(e,\vartheta) = \mathfrak{C}_1^{-1}(\vartheta{-}\mathfrak{C}_0(e))$, the asymptotics of $\Theta'_e(e,\cdot)$ is the same as $\Theta'_\vartheta(e,\cdot)$, and thus the asymptotics of $\mathcal{K}_1$ defined by (3.4c), is again as (5.3) for large ϑ, and (3.17e) gives the restriction $(a_i{+}1{-}c)/c < 1/2$ which, however, has anyhow to be satisfied due to (3.17d) if $d \ge 2$.

Remark 5.6 (Macroscopical response). In macroscopical bodies, the highest-gradient $\mathbb{G}$- and $\mathbb{H}$-terms are very small and represent rather a singular perturbation only, without influencing essentially the macroscopical response. This phenomenon is well known from variational calculus. In evolution problems, it is more difficult to prove it rigorously. In the isothermal case, the vanishing influence of the viscous hyperstress if $\mathbb{G} \to 0$ was shown in [39] if $d = 1$ and, under higher-order hyperstresses, if $d \ge 2$ too. The anisothermal case does not seem study from this viewpoint yet, although recent regularity results from [38] indicates the same effect at least for the heat capacity $\mathfrak{c}$ growing linearly. Anyhow, here it is reflected by the fact that, under the external loading qualified as (3.17)g,i, some a-priori estimates survive even if $\mathbb{G}\searrow 0$ and $\mathbb{H}\searrow 0$. After inspecting the above proofs, one can see that it concerns only the first part of the estimate (4.14a) and the estimate (4.14b) and, if restricting the growth of $\mathcal{A}$ so that the $\mathcal{A}$-term in (4.23) can be interpolated with the help of the $\mathbb{D}$-term instead of the $\mathbb{G}$-term in the left-hand side of (4.23), also the weakened estimate (4.19a), namely

$$\big\|u\big\|_{W^{1,\infty}(I;L^2(\Omega;\mathbb{R}^d))\,\cap\, W^{1,2}(I;W^{1,2}(\Omega;\mathbb{R}^d))}\le C, \tag{5.5a}$$

$$\big\|\vartheta\big\|_{L^\infty(I;L^1(\Omega))}\le C. \tag{5.5b}$$

All other estimates rely essentially on $\mathbb{G}>0$ and $\mathbb{H}>0$. Yet, if $\psi'''_{e\theta\theta}=0$ as discussed in Sect. 1 and in Example 5.4, then $\mathcal{K}_1=0$ and C_1 in (4.22) would be independent of $\mathbb{G}$ and $\mathbb{H}$, and then also (4.19b,c) would hold uniformly.

Remark 5.7 (No inertia allows for $\mathbb{G}=0$). If $\varrho=0$ (=the so-called quasistatic case), we can execute the modified interpolation as outlined in Remark 5.6 and consider $\mathbb{G}=0$. This scenario is a physically natural case especially if, instead of the strain e, we would consider about other physical variable as, e.g., magnetization, dielectric polarization, or various phase fields.

Remark 5.8 (Finite strains). A general nonlinear ansatz (2.1) allows for considering the full gradient ∇u in place of $e(u)$ in $\psi(\cdot,\theta)$ and thus a modification to frame-indifferent free energies at finite strains. This requires to admit that $\psi(\cdot,\theta)$ and thus also $\varphi=\psi(\cdot,0)$ are nonconvex. Also the corresponding capillarity-like term $\mathbb{H}\nabla^2u$ can be made frame indifferent. The dissipation rate (2.2) can be interpreted as locally-in-time frame-indifferent. The mathematical analysis is then basically unchanged. Yet, it should be emphasized that realistic large-strain models would require still a non-polynomial growth of $\psi(\cdot,\theta)$ to capture the phenomenon of local nonpenetration, i.e. the blow-up of the free energy when $\det(\mathbb{I}{+}\nabla u)\searrow 0$. Moreover, it is well recognized that a fully frame-indifferent dissipation rate would have to bear a very nonlinear form [1,15] not allowing for a (simple) mathematical analysis.

Remark 5.9 (Nonlocal hyperstress). To avoid still higher gradients and more complicated boundary conditions, one may consider rather nonlocal hyperstress. This leads to a modification of the free energy (2.1) to

$$\hat\psi(e,\theta,\nabla e):=\psi(e,\theta)+\mathfrak{H}(\nabla e)\qquad\text{where}$$
$$\big[\mathfrak{H}(\nabla e)\big](x)=\frac14\int_\Omega H(x,\tilde x)\big(\nabla e(x){-}\nabla e(\tilde x)\big)^2\,\mathrm{d}\tilde x\quad\text{with}\quad H(x,\tilde x)=\frac{\mathbb{H}}{|x{-}\tilde x|^{d+2\rho}} \tag{5.6}$$

with some $0<\rho<1$. Note that this special choice of the kernel H yields $E\mapsto(\int_\Omega[\mathfrak{H}(E)](x)\,\mathrm{d}x)^{1/2}=(\int_{\Omega\times\Omega}\frac{\mathbb{H}(E(x)-E(\tilde x))^2}{4|x-\tilde x|^{d+2\rho}}\,\mathrm{d}\tilde x\mathrm{d}x)^{1/2}$ the standard seminorm in the Sobolev–Slobodetskiĭ space $W^{\rho,2}(\Omega;\mathbb{R}^{d\times d\times d})$. The hyperstress $\mathfrak{h}$ would now take the form

$$\mathfrak{h}=\mathfrak{h}(e,\dot e,\nabla e,\nabla\dot e,\theta)\,:=\hat\psi'_{\nabla e}(e,\theta,\nabla e)+\frac12\xi'_{\nabla\dot e}(\dot e,\nabla\dot e)=\mathfrak{H}'\nabla e+\mathbb{G}\nabla\dot e, \tag{5.7}$$

where the ansatz (2.1)–(2.2) has been taken into account. Assuming $\mathbb{H}$ symmetric, the term $\mathfrak{H}'\nabla e$ in (5.7) means

$$\big[\mathfrak{H}'\nabla e\big](x) = \mathbb{H}\int_\Omega \frac{\nabla e(x)-\nabla e(\tilde x)}{|x-\tilde x|^{d+2\rho}}\,\mathrm{d}\tilde x. \tag{5.8}$$

The somewhat restrictive assumptions (3.17c-e) on $\mathcal{K}_0$, $\mathcal{K}_1$, and $\mathcal{A}$ can then be weakened as follows:

$$\exists\, C_{\mathcal{K}_0}, C_{\mathcal{K}_1}\in\mathbb{R},\ \epsilon>0\ \ \forall\,(e,\vartheta)\in\mathbb{R}^{d\times d}_{\rm sym}\times\mathbb{R}:$$
$$\big|\mathcal{K}_0(e,\vartheta)\big| \le C_{\mathcal{K}_0}\big(1+|e|^{p^*/(d+2)-\epsilon}+|\vartheta|^{1/d-\epsilon}\big), \tag{5.9a}$$
$$\big|\mathcal{K}_1(e,\vartheta)\big| \le C_{\mathcal{K}_1}\sqrt{1+|\vartheta|}\big(1+|e|^{p^*}\big)^{(p-2)/(2p)}, \qquad p=\frac{2d}{d-2\rho}, \tag{5.9b}$$
$$\exists C_{\mathcal{A}}<\infty,\ \ \text{for } q>(2d+8)/(d+6),\ s<q:$$
$$\forall\,(e,\vartheta)\in\mathbb{R}^{d\times d}_{\rm sym}\times\mathbb{R}:\quad \big|\mathcal{A}(e,\vartheta)\big|\le C_{\mathcal{A}}\big(1+|e|^{p^*/q}+|\vartheta|^{s/q}\big), \tag{5.9c}$$

or possibly only $s<(d+2)/d$ in (5.9c) if the two-stage limit passage devised in Remark 5.3 would be implemented. The analysis can be performed in the same lines provided we invent also the corresponding nonlocal term in the viscous hyperstress. The estimates (4.14a), (4.19a) and (4.41), and also (4.43a) and (4.48) use the Sobolev–Slobodetskiĭ space $W^{2+\rho,2}(\Omega;\mathbb{R}^d)$ instead of $W^{2,2}(\Omega;\mathbb{R}^d)$.

One can modify the estimate (4.8) by using Hölder's inequality for 4 factors with exponents 4, $\frac{2p\eta}{p^*(p-2)}$, η, and 2 and by using (5.9b)

$$\begin{aligned}
&\bigg|\int_\Omega \mathcal{K}_1(e(u^k_\tau),\vartheta^k_\tau)\nabla e(u^k_\tau)\cdot\nabla\vartheta^k_\tau\mathrm{d}x\bigg|\\
&\quad\le \int_\Omega C_{\mathcal{K}_1}\sqrt{1+|\vartheta^k_\tau|}\big(1+|e^k_\tau|^{p^*}\big)^{\frac{p-2}{2p}}\big|\nabla e(u^k_\tau)\big|\,\big|\nabla\vartheta^k_\tau\big|\,\mathrm{d}x\\
&\quad\le \int_\Omega C_{\mathcal{K}_1}\Big(1+\sqrt{|\vartheta^k_\tau|}\Big)\big(1+|e^k_\tau|^{p^*}\big)^{\frac{p-2}{2p}}\,\big|\nabla e(u^k_\tau)\big|\,\big|\nabla\vartheta^k_\tau\big|\,\mathrm{d}x\\
&\quad\le C_{\delta,\eta}+\delta\big\|\vartheta^k_\tau\big\|^2_{L^2(\Omega)}+\delta\big\|e(u^k_\tau)\big\|^\eta_{L^\eta(\Omega;\mathbb{R}^{d\times d})}\\
&\qquad\quad+\delta\big\|\nabla e(u^k_\tau)\big\|^\eta_{L^\eta(\Omega;\mathbb{R}^{d\times d\times d})}+\delta\big\|\nabla\vartheta^k_\tau\big\|^2_{L^2(\Omega;\mathbb{R}^d)}
\end{aligned} \tag{5.10}$$

with $\delta>0$ arbitrarily small and some $C_{\delta,\eta}$ depending on η and δ. Again η should be big enough, i.e. here $\eta>4+2p^*(p-2)/p$ which guarantees that the sum of the inverse of the mentioned 4 exponents in (5.10) is less than 1.

Further, one can modify (4.21) as

$$\begin{aligned}
&\int_Q \mathcal{K}_1(e(\bar u_\tau),\bar\vartheta_\tau)\nabla e(\bar u_\tau)\cdot\nabla\chi(\bar\vartheta_\tau)\,\mathrm{d}x\mathrm{d}t=\varepsilon\int_Q\nabla e(\bar u_\tau)\vdots\frac{\mathcal{K}_1(e(\bar u_\tau),\bar\vartheta_\tau)\otimes\nabla\bar\vartheta_\tau}{(1+\bar\vartheta_\tau)^{1+\varepsilon}}\,\mathrm{d}x\mathrm{d}t\\
&\quad\le\varepsilon\int_Q C_{\delta,p}|\nabla e(\bar u_\tau)|^p+C_{\delta,p}\frac{|\mathcal{K}_1(e(\bar u_\tau),\bar\vartheta_\tau)|^{2p/(p-2)}}{(1+\bar\vartheta_\tau)^{p/(p-2)}}+\delta\frac{|\nabla\bar\vartheta_\tau|^2}{(1+\bar\vartheta_\tau)^{1+\varepsilon}}\,\mathrm{d}x\mathrm{d}t\\
&\quad\le\frac{\varepsilon TC^p}{C_{\delta,p}}+\int_Q C_{\delta,p}C_{\mathcal{K}_1}\big(1+|e(\bar u_\tau)|^{p^*}\big)\,\mathrm{d}x\mathrm{d}t+\varepsilon\delta\int_Q\frac{|\nabla\bar\vartheta_\tau|^2}{(1+\bar\vartheta_\tau)^{1+\varepsilon}}\,\mathrm{d}x\mathrm{d}t
\end{aligned} \tag{5.11}$$

with $C_{\mathcal{K}_1}$ from (5.9b), using also the embedding

$$W^{2+\rho,2}(\Omega) \subset W^{2,p}(\Omega) \subset W^{1,p^*}(\Omega) \qquad \text{provided} \;\; p \geq \frac{2d}{d{-}2\rho}, \tag{5.12}$$

which is just where the choice of p in (5.9b) came from.

Acknowledgments

The author is thankful to Alexander Mielke, Paolo Podio-Guidugli, Jürgen Sprekels, Ulisse Stefanelli, and Giuseppe Tomassetti for fruitful and inspiring discussion in particular about higher-order boundary conditions, about energy conservation, and about nonlinear thermodynamical coupling. Special thanks are also to Marita Thomas for very careful reading of early versions of the manuscript and a lot of valuable suggestions.

References

[1] Antman, S.S.: Physically unacceptable viscous stresses. Z. Angew. Math. Phys. **49**(6), 980–988 (1998)

[2] Ball, J.M., Currie, J.C., Olver, P.J.: Null Lagrangeans, weak continuity, and variational problems of arbitrary order. J. Funct. Anal. **41**, 135–174 (1981)

[3] Blanchard, D., Guibé, O.: Existence of a solution for a nonlinear system in thermo viscoelasticity. Adv. Differ. Equ. **5**, 1221–1252 (2000)

[4] Brézis, H.: Équations et inéquations non-linéaires dans les espaces vectoriel en dualité. Ann. Inst. Fourier **18**, 115–176 (1968)

[5] Boccardo, L., Dall'aglio, A., Gallouët, T., Orsina, L.: Nonlinear parabolic equations with measure data. J. Funct. Anal. **147**, 237–258 (1997)

[6] Boccardo, L., Gallouët, T.: Non-linear elliptic and parabolic equations involving measure data. J. Funct. Anal. **87**, 149–169 (1989)

[7] Brokate, M., Sprekels, J.: Hystreresis and Phase Transitions. Springer, New York (1996)

[8] Bulíček, M., Feireisl, E., Málek, J.: A Navier–Stokes–Fourier system for incompressible fluids with temperature dependent material coefficients. Nonlinear. Anal. Real World Appl. **10**, 992–1015 (2009)

[9] Bulíček, M., Kaplický, P., Málek, J.: An L^2-maximal regularity result for the evolutionary Stokes–Fourier system. Appl. Anal. **90**, 31–45 (2011)

[10] Colli, P., Krejčí, P., Rocca, E., Sprekels, J.: A nonlocal quasilinear multi-phase system with nonconstant specific heat and heat conductivity. J. Differ. Equ. **251**, 1354–1387 (2011)

[11] Colli, P., Sprekels, J.: Global existence for a three-dimensional model for the thermo-mechanical evolution of shape memory alloys. Nonlinear Anal. **18**, 873–888 (1992)

[12] Colli, P., Sprekels, J.: Global solution to the full one-dimensional Frémond model for shape memory alloys. Math. Methods Appl. Sci. **18**, 371–385 (1995)

[13] Dafermos, C.M.: Global smooth solutions to the initial boundary value problem for the equations of one-dimensional thermoviscoelasticity. SIAM J. Math. Anal. **13**, 397–408 (1982)

[14] Dafermos, C.M., Hsiao, L.: Global smooth thermomechanical processes in one-dimensional nonlinear thermoviscoelasticity. Nonlinear Anal. **6**, 435–454 (1982)

[15] Demoulini, S.: Weak solutions for a class of nonlinear systems of viscoelasticity. Arch. Ration. Mech. Anal. **155**, 299–334 (2000)

[16] Eck, C.: Existence of solutions to a thermo-viscoelastic contact problem with Coulomb friction. Math. Models Methods Appl. Sci. **12**, 1491–1511 (2002)

[17] Eck, C., Jarušek, J., Krbec, M.: Unilateral Contact Problems. Chapman & Hall/CRC, Boca Raton (2005)

[18] Fabrizio, M., Lazzari, B., Nibbi, R.: Thermodynamics of non-local materials: extra fluxes and internal powers. Contin. Mech. Thermodyn. **23**, 509–525 (2011)

[19] Feireisl, E., Petzeltová, H., Rocca, E.: Existence of solutions to a phase transition model with microscopic movements. Math. Methods Appl. Sci. **32**, 1345–1369 (2009)

[20] Fried, E., Gurtin, M.E.: Tractions, balances, and boundary conditions for non-simple materials with application to liquid flow at small-length scales. Arch. Ration. Mech. Anal. **182**, 513–554 (2006)

[21] Hamiel, Y., Lyakhovsky, V., Ben-Zion, Y.: The elastic strain energy of damaged solids with applications to non-linear deformation of crystalline rocks. Pure Appl. Geophys. (2011). doi:10.1007/s00024-011-0265-7

[22] Krejčí, P., Rocca, E., Sprekels, J.: A nonlocal phase-field model with nonconstant heat conductivity. Interf. Free Bound. **9**, 285–306 (2007)

[23] Krejčí, P., Sprekels, J.: Temperature-dependent hysteresis in one-dimensional thermovisco-elastoplasticity. Appl. Math. **43**, 173–205 (1998)

[24] Krejčí, P., Sprekels, J.: On a system of nonlinear PDEs with temperature-dependent hysteresis in one-dimensional thermoplasticity. J. Math. Anal. Appl. **209**, 25–46 (1997)

[25] Krejčí, P., Stefanelli, U.: Existence and nonexistence for the full thermomechanical Souza-Auricchio model of shape memory wires. Math. Mech. Solids **16**, 349–365 (2011)

[26] Kohn, R.V., Müller, S.: Surface energy and microstructure in coherent phase transitions. Commun. Pure Appl. Math. **47**, 405–435 (1994)

[27] Lyakhovsky, V., Ben-Zion, Y., Agnon, A.: Distributed damage, faulting, and friction. J. Geophys. Res. **102**, 27635–27649 (1997)

[28] Lyakhovsky, V., Reches, Z., Weiberger, R., Scott, T.E.: Nonlinear elastic behaviour of damaged rocks. Geophys. J. Int. **130**, 157–166 (1997)

[29] Müller, S.: Variational models for microstructure and phase transitions. In: Hildebrandt, S., Struwe, M. (eds.) Calculus of Variations and Geometric Evolution Problems. Lecturer Notes in Mathematics, vol. 1713, pp.85–210. Springer, Berlin (1999)

[30] Nečas, J.: Dynamic in the nonlinear thermo-visco-elasticity. In: Schulze, B.-W., Triebel, H. (eds.) Partial Differential Equations, pp. 197–203. Teubner, Leipzig (1989)

[31] Nečas, J., Novotný, A., Šverák, V.: On the uniqueness of solution to the nonlinear thermo-visco-elasticity. Math. Nachr. **149**, 319–324 (1990)

[32] Niezgódka, M., Pawłow, I.: A generalized Stefan problem in several space variables. Appl. Math. Optim. **9**, 193–224 (1983)

[33] Podio-Guidugli, P.: Contact interactions, stress, and material symmetry, for nonsimple elastic materials. Theor. Appl. Mech. **28–29**, 261–276 (2002)

[34] Podio-Guidugli, P., Vergara Caffarelli, G.: Surface interaction potentials in elasticity. Arch. Ration. Mech. Anal. **109**, 343–381 (1990)

[35] Podio-Guidugli, P., Vianello, M.: Hypertractions and hyperstresses convey the same mechanical information. Contin. Mech. Thermodyn. **22**, 163–176 (2010)

[36] Pawłow, I., Zajączkowski, W.M.: Global existence to a three-dimensional nonlinear thermoelasticity system arising in shape memory materials. Math. Methods Appl. Sci. **28**, 407–442 (2005)

[37] Pawłow, I., Zajączkowski, W.M.: Unique global solvability in two-dimensional non-linear thermoelasticity. Math. Methods Appl. Sci. **28**, 551–592 (2005)

[38] Pawłow, I., Zajączkowski, W.M.: Global regular solutions to a Kelvin-Voigt type thermoviscoelastic system. (to appear)

[39] Rajagopal, K.R., Roubíček, T.: On the effect of dissipation in shape-memory alloys. Nonlinear Anal. Real World Appl. **4**, 581–597 (2003)

[40] Roubíček, T.: Models of microstructure evolution in shape memory alloys. In: Ponte Castaneda, P., Telega, J.J., Gambin, B. (eds.) Nonlinear Homogenization and its Application to Composites, Polycrystals and Smart Materials. NATO Sci. Series **II/170**, pp. 269–304. Kluwer, Dordrecht (2004)

[41] Roubíček, T.: Incompressible ionized non-Newtonean fluid mixtures. SIAM J. Math. Anal. **39**, 863–890 (2007)

[42] Roubíček, T.: Thermo-visco-elasticity at small strains with L^1-data. Q. Appl. Math. **67**, 47–71 (2009)

[43] Roubíček, T.: Thermodynamics of rate independent processes in viscous solids at small strains. SIAM J. Math. Anal. **40**, 256–297 (2010)

[44] Roubíček, T., Stefanelli, U.: Phenomenological model of Souza-Auricchio's type for magnetic shape-memory alloys. In preparation

[45] Roubíček, T., Tomassetti, G. Phase transformations in electrically conductive ferromagnetic shape-memory alloys, their thermodynamics and analysis. Arch. Ration. Mech. Anal. (in print)

[46] Šilhavý, M.: Phase transitions in non-simple bodies. Arch. Ration. Mech. Anal. **88**, 135–161 (1985)

[47] Toupin, R.A.: Elastic materials with couple stresses. Arch. Ration. Mech. Anal. **11**, 385–414 (1962)

Tomáš Roubíček
Mathematical Institute
Charles University
Sokolovská 83
186 75 Praha 8
Czech Republic

Tomáš Roubíček
Institute of Thermomechanics of the ASCR
Dolejškova 5
182 00 Praha 8
Czech Republic
e-mail: tomas.roubicek@mff.cuni.cz

Received: 20 February 2012.
Accepted: 24 October 2012.