# Finite thermoelastoplasticity and creep under small elastic strains


Tomáš Roubíček
*Mathematical Institute, Charles University,*
*Sokolovská 83, CZ–186 75 Praha 8, Czech Republic*
and
*Institute of Thermomechanics, Czech Academy of Sciences,*
*Dolejškova 5, CZ–182 08 Praha 8, Czech Republic*

Ulisse Stefanelli
*Faculty of Mathematics, University of Vienna,*
*Oskar-Morgenstern-Platz 1, A-1090 Vienna, Austria*
and
*Istituto di Matematica Applicata e Tecnologie Informatiche E. Magenes - CNR,*
*v. Ferrata 1, I-27100 Pavia, Italy*





*Abstract.* A mathematical model for an elastoplastic continuum subject to large strains is presented. The inelastic response is modeled within the frame of rate-dependent gradient plasticity for nonsimple materials. Heat diffuses through the continuum by the Fourier law in the actual deformed configuration. Inertia makes the nonlinear problem hyperbolic. The modelling assumption of small elastic Green-Lagrange strains is combined in a thermodynamically consistent way with the possibly large displacements and large plastic strain. The model is amenable to a rigorous mathematical analysis. The existence of suitably defined weak solutions and a convergence result for Galerkin approximations is proved.



*Key Words.* Thermoplastic materials, finite strains, creep, Maxwell viscoelastic rheology, heat transport, Lagrangian description, energy conservation, frame indifference, Galerkin approximation, convergence, weak solutions.

*AMS Subject Classification:* 35Q74, 35Q79, 65M60 74A15, 74A30, 74C15, 74J30, 80A20.


## 1 Introduction

*Thermoelastodynamic problems* in combination with inelastic effects are ubiquitous in applications and have triggered an intense research activity cutting across material science, engineering, and mathematics [18, 25, 44]. A number of rigorous mathematical results are available in case of small strains, whereas the literature for finite-strain problems is relatively less developed. Still, large strains arise naturally in a variety of different thermomechanical contexts including, for instance, plastic deformations, rolling, and impacts, and in a number of different materials, from metals to polymers.

The focus of this note is to contribute a finite-strain thermomechanical model for an inelastic continuum. The evolution of the medium will be described by its deformation $y$ from the reference configuration, its plastic strain $P$, and its absolute temperature $\theta$. We consider a fully dynamic problem and address viscoelastic rheologies of *Maxwell* type, including *creep* [37,42] or *(visco)plasticity* [30]. These are both permanent deformation dynamics. Plasticity is an activated effect, for its onset corresponds to the reaching of a given yield stress. On the contrary, one usually speaks of creep in case material relazation is always active upon loading, without any yield threshold [2]. In this paper, creep is modeled by the Maxwell viscoelastic rheology; note however that different rheologies are sometimes associated in the literature with creep as well. Within the classical frame of Generalized Standard Materials [12,16], the model results from the combination of energy and momentum conservation with the plastic flow rule and will be checked to be thermodynamically consistent.

Our crucial modeling tenet is that the large deformations are stored by the plastic strain $P$, whereas the *elastic part* of the total deformation strain remains close to the identity. This assumptions is often not restrictive, as elastoplastic materials often sustain relatively small elastic deformations before plasticizing. This is for instance the case of ordinary metals, which usually plasticize around a few strain percents, as well as of geophysical applications, where soils and rocks sustain small elastic deformations before sliding and cracking [22]. On the other hand, such smallness assumption allows for a satisfactory mathematical treatment in terms existence and approximability of solutions.

Note that we explicitly include in the model nonlocal effects in terms of higher order gradients of $\nabla y$ and $P$. This sets our model within the frame of *2nd-grade nonsimple* [46] and *gradient plastic* materials [36]. We allow for no hardening in the shear deformation, which is a typical attribute in particular of creep, i.e. the viscoelastic rheology of Maxwell type.

The elastoplastic evolution of the medium is combined with heat production and thus heat transfer. A weak thermal coupling through the temperature-dependence of the dissipation potential is considered. This models the possible temperature dependence of the plastic yield stress or the viscous moduli responsible for creep. An important feature of the model is that heat convection governed by the *Fourier law* occurs *in the actual deformed configuration.* This is in most applications much more physical than considering heat conduction in the material reference configuration, especially at large strains. Here again the smallness assumption on the elastic strain plays a crucial role in avoiding analytical technicalities related with the control of the inverse of deformation gradient, cf. (2.8) vs. (2.10) below.

Existence results in finite-strain elastoplasticity are at the moment restricted to the isothermal case. In the incremental setting, the early result in [27] has been extended in [29] by including a term in $\operatorname{curl} P$ in the energy, see also [45] for the case of compatible plastic deformations. In the quasistatic case, the existence of *energetic solutions* [32] has been proved in [23] in the frame of gradient plasticity, see also [15, 26, 28]. A discussion on a possible finite-element approximation is in [33] and rigorous linearization results are in [13, 34]. Viscoplasticity is addressed in [30] instead.

The novelty of this paper is that of dealing with the *nonisothermal* and *dynamic* case. To the best of our knowledge, the analysis of *both* these features is unprecedented in the frame of finite elastoplasticity. Our main result, Theorem 3.2, states the existence of suitably weak solutions. This relies on the smallness assumption on the elastic strain, which in turn allows us to tame the nonlinear nature of the coupling of thermal and mechanical effects. The existence proof is based on a regularization and Galerkin approximation procedure, which could serve as basis for numerical investigation.

The plan of the paper is as follows. In Section 2 we formulate the model. In particular, we specify the form of the total energy and of the dissipation. This brings to the formulation of an evolution system of partial differential equations and inclusions. The thermodynamical consistency of the model and various possible modifications are also discussed. Section 3 presents a variational notion of solution as well as the main analytical statement. The existence proof is then detailed in Section 4.

## 2 The model and its thermodynamics

We devote this section to presenting our model for elastoplastic continua with heat transfer. This is formulated in Lagrangian coordinates with $\Omega \subset \mathbb{R}^d$ ($d = 2$ or 3) being a bounded smooth reference (fixed) configuration. The variables of the model are

| | |
|---|---|
| $y : \Omega \to \mathbb{R}^d$ | deformation, |
| $P : \Omega \to \mathrm{GL}^+(d) := \{P \in \mathbb{R}^{d\times d};\ \det P > 0\}$ | plastic part of the inelastic strain, |
| $\theta : \Omega \to (0, \infty)$ | absolute temperature. |

The model will result by combining momentum and energy conservation with the dynamics of internal variables. In order to specify the latter and provide constitutive relations, we introduce a free energy and a dissipation (pseudo)potential in the following subsections.

We consider the standard multiplicative decomposition [19, 21]

$$F = F_{\rm el}P \quad \text{with } F = \nabla y. \tag{2.1}$$

Our basic modelling assumption is that the *elastic* part of the *strain is small* in the sense that the elastic Green-Lagrange strain $E_{\rm el} := \frac{1}{2}(F_{\rm el}^\top F_{\rm el} - \mathbb{I})$ is small is small as well. Here, the superscript $\top$ stands for transposition and $\mathbb{I}$ is the identity matrix. Note that $E_{\rm el}$ is sometimes called *Green-St. Venant* strain as well. In fact, the smallness of $E_{\rm el}$ is equivalent to $F^\top F \sim P^\top P$ and is weaker than $F \sim P$, since it allows large elastic rotations, a circumstance which may be relevant in many applications. Such a smallness assumption is well-fitted to the case of metals or rocks, where comparably small elastic strains already activate the plastification, or to polymers or soils undergoing considerable deformation through creep. Moreover, the smallness assumption allows us to simplify the mathematical treatment of the model.

Most notably, the *elastic energy density* $\psi_{\rm E}(F_{\rm el})$ of the medium can be assumed to have a controlled polynomial growth, see (3.1a) below. In addition, we assume $\psi_{\rm E}$ to be *frame-indifferent*, namely

$$\forall Q \in \mathrm{SO}(d) : \qquad \psi_{\rm E}(QF_{\rm el}) = \psi_{\rm E}(F_{\rm el}). \tag{2.2}$$

Here, we used the notation $\mathrm{SO}(d)$ for the matrix special-orthogonal group $\mathrm{SO}(d) := \{Q \in \mathrm{GL}^+(d);\ QQ^\top = Q^\top Q = \mathbb{I}\}$. Equivalently, $\psi_{\rm E}(F_{\rm el})$ can be expressed as a function of the elastic Cauchy-Green tensor $F_{\rm el}^\top F_{\rm el}$ only.

The multiplicative decomposition (2.1) allows us to express the free energy in terms of the total strain tensor $\nabla y$ and the plastic strain $P$ via the substitution $F_{\rm el} = FP^{-1}$. In most materials, processes such as plastification or creep lead dominantly to shear deformation but not to volumetric variation. As such, the isochoric constraint $\det P \sim 1$ is often considered.

This is taken into account by this model, where values of $\det P$ to be close to 1 are energetically favored by a specific hardening-like term, cf. Remark 2.6 below. Such hardening term (denoted by $\psi_{\rm H}$ in (2.3) below) may control the full plastic strain $P$, which would correspond to the case of coventional kinematic hardening in the shear plastic strain (which is a relevant model typically in metals). In any case, the stored energy will control $\det P$ to be positive, which is important to guarantee the (uniform) invertibility of $P$ needed in the mentioned expression $F_{\rm el} = FP^{-1}$.

The mechanical stored energy $\Psi_{\rm M}$ will have elastic and hardening parts $\psi_{\rm E}$ and $\psi_{\rm H}$ and will be augmented by gradient terms and a thermal contribution $\psi_{\rm T}$ (considered for simplicity to depend solely on temperature, i.e., in particular thermal expansion is here neglected), cf. Remark 2.5 for some extension. By integrating on the reference configuration $\Omega$, the *free energy* of the body is expressed by

$$\Psi(\nabla y, P, \theta) = \Psi_{\rm M}(\nabla y, P) + \Psi_{\rm T}(\theta) \quad \text{ with } \ \Psi_{\rm T}(\theta) = \int_\Omega \psi_{\rm T}(\theta)\,{\rm d}x$$
$$\text{and} \quad \Psi_{\rm M}(\nabla y, P) = \int_\Omega \psi_{\rm E}(\nabla y\, P^{-1}) + \psi_{\rm H}(P) + \frac{1}{2}\kappa_0\big|\nabla^2 y\big|^2 + \frac{1}{q}\kappa_1|\nabla P|^q\,{\rm d}x. \tag{2.3}$$

In particular, the $\kappa_0$-term qualifies the material as *2nd-grade nonsimple*, also called *multipolar* or *complex*, see the seminal [46] and [11, 35, 38, 39, 43, 47]. On the other hand, the $\kappa_1$-term describes nonlocal plastic effects and is inspired to the by-now classical *gradient plasticity* theory [9, 10, 36]. In particular, its occurrence turns out to be crucial in order to prevent the formation of plastic microstructures and ultimately ensures the necessary compactness for the analysis. Note that in the finite-plasticity context, the introduction of suitable regularizing terms on the plastic variables seems at the moment unavoidable [15, 23, 29]. The exponent $q$ in the $\kappa_1$-term is given and fixed to be larger than $d$, which eases some points of the analysis. Note however that the choice $\kappa_1(\nabla P)|\nabla P|^2$ for some $\kappa_1(\nabla P) \sim 1 + |\nabla P|^{q-2}$ could be considered as well.

The *frame-indifference* of the mechanical stored energy (2.2) translates in terms of $\Psi$ as

$$\forall Q \in {\rm SO}(d): \qquad \Psi(Q\nabla y, P, \theta) = \Psi(\nabla y, P, \theta).$$

In particular let us note that the gradient terms are frame-indifferent as well.

The partial functional derivatives of $\Psi$ give origin to corresponding driving forces. We use the symbol "$\partial_w$" to indicate both partial differentiation with respect to the variable $w$ of a smooth functional or subdifferentiation of a convex functional. In case of a single-argument smooth functional, we write shortly $(\cdot)'$.

The *second Piola-Kirchhoff stress* $\varSigma_{\rm el}$, here augmented by a contribution arising from the gradient $\kappa_0$-term, is defined as

$$\varSigma_{\rm el} = \partial_{\nabla y}\Psi = \psi_{\rm E}'(\nabla y P^{-1})P^{-\top} - \kappa_0{\rm div}\nabla^2 y. \tag{2.4a}$$

Furthermore, the *driving stress* for the plastification, again involving a contribution arising from the gradient $\kappa_1$-term, reads

$$\varSigma_{\rm in} = \partial_P\Psi = \nabla y^\top\psi_{\rm E}'(\nabla y P^{-1}){:}(P^{-1})' + \psi_{\rm H}'(P) - {\rm div}(\kappa_1|\nabla P|^{q-2}\nabla P). \tag{2.4b}$$

Here and in what follows, we use the (standard) notation "$\cdot$" and "$:$" and "$\vdots$" for the contraction product of vectors, 2nd-order, and 3rd tensors, respectively. The term $(P^{-1})'$ is a 4th-order tensor, namely

$$(P^{-1})' = \Big(\frac{\mathrm{Cof}' P^\top}{\det P} - \frac{\mathrm{Cof} P^\top \otimes \mathrm{Cof} P}{(\det P)^2}\Big)$$

where $\mathrm{Cof}P = (\det P)P^{-\top}$ is the classical *cofactor matrix*. The product $\nabla y^\top \psi'_{\rm E}(\nabla y P^{-1}){:}(P^{-1})'$ turns out to be a 2nd-order tensor, as expected. Both variations of $\Psi$ above are taken with respect to the corresponding $L^2$ topologies.

Eventually, the *entropy* $\eta$, the *heat capacity* $c_{\rm v}$, and the *thermal part* $\vartheta$ of the internal energy are classically recovered as

$$\eta = -\psi'_\theta = -\psi'_{\rm T}(\theta), \qquad c_{\rm v} = -\theta\psi''_{\theta\theta} = -\theta\psi''_{\rm T}(\theta), \quad \text{and} \quad \vartheta = \psi_{\rm T}(\theta) - \theta\psi'_{\rm T}(\theta). \tag{2.5}$$

Note in particular that $\dot\vartheta = \psi'_{\rm T}(\theta)\dot\theta - \dot\theta\psi'_{\rm T}(\theta) - \theta\psi''_{\rm T}(\theta)\dot\theta = c_{\rm v}(\theta)\dot\theta$. The *entropy equation* reads

$$\theta\dot\eta + \mathrm{div}\, j = \text{dissipation rate}. \tag{2.6}$$

We assume the heat flux $j$ to be governed by the *Fourier* law $j = -\mathscr{K}\nabla\theta$ where is $\mathscr{K}$ the *effective heat-transport tensor*, see (2.10) below. Substituting $\eta$ from (2.5) into (2.6), we obtain the *heat-transfer equation*

$$c_{\rm v}(\theta)\dot\theta - \mathrm{div}(\mathscr{K}\nabla\theta) = \text{dissipation rate}.$$

Note that, $c_{\rm v}$ depends on the temperature only through $\psi_{\rm T}$.

The model consists of the following system of semilinear equations

$$\varrho\ddot y = \mathrm{div}\, \varSigma_{\rm el} + g(y), \qquad \text{(momentum equilibrium)} \tag{2.7a}$$

$$\partial_R\mathfrak{R}\big(\theta; \dot P P^{-1}\big) + \varSigma_{\rm in} P^\top \ni 0, \qquad \text{(flow rule for the inelastic strain)} \tag{2.7b}$$

$$c_{\rm v}(\theta)\dot\theta = \mathrm{div}\big(\mathscr{K}(P,\theta)\nabla\theta\big) + \partial_R\mathfrak{R}\big(\theta; \dot P P^{-1}\big){:}(\dot P P^{-1}) \qquad \text{(heat-transfer equation)} \tag{2.7c}$$

where $\mathfrak{R} = \mathfrak{R}(\theta; R)$ is the (possibly nonsmooth) *(pseudo)potential* related to dissipative forces of viscoplastic origin, $R$ is the placeholder for the plastification rate $\dot P P^{-1}$, and $\partial_R\mathfrak{R}$ stands for the subdifferential of the function $\mathfrak{R}(\theta; \cdot)$, which is assumed to be convex. The right-hand side of (2.7a) features the pull-back $g{\circ}y$ of the *actual* gravity force $g : \mathbb{R}^d \to \mathbb{R}^d$. This is the most simple example of a *nondead* load, i.e., a load given in terms of the actual deformed configuration of the body.

The *effective heat-transfer tensor* $\mathscr{K}$ is to be related with the symmetric *heat-conductivity tensor* $\mathbb{K} = \mathbb{K}(\theta)$ which is a given material property, see (3.1c) below. The need for such effective quantities stems from the fact that driving forces are to be considered Eulerian in nature, so that a pull-back to the reference configuration is imperative. We use the adjective *effective* to indicate quantities that are defined on the reference configuration but act on the actual one. A first choice in this direction is

$$\mathscr{K}(F,\theta) = (\det F)F^{-1}\mathbb{K}(\theta)F^{-\top} = (\mathrm{Cof}F^\top)\mathbb{K}(\theta)F^{-\top} = \frac{(\mathrm{Cof}F^\top)\mathbb{K}(\theta)\mathrm{Cof}F}{\det F}. \tag{2.8}$$

This is just the usual pull-back transformation of 2nd-order covariant tensors. In the isotropic case $\mathbb{K}(\theta)=k(\theta)\mathbb{I}$, relation (2.8) can also be written by using the right Cauchy-Green tensor $C$ as

$$\mathscr{K}(C,\theta)=\det C^{1/2}k(\theta)C^{-1}\quad\text{with}\quad C=F^\top F,\tag{2.9}$$

cf. [7, Formula (67)] or [14, Formula (3.19)] for mass transport instead of heat. In fact, the effective transport-coefficient tensor is a function of $C$ in general anisotropic cases as well, cf. [20, Sect. 9.1]. In view of this, we now use our smallness assumption $E_{\rm el}\sim 0$, which yields only $F^\top F\sim P^\top P$, in order to infer that we can, in fact, substitute $F$ with $P$ into (2.8) as a good modelling ansatz, even though there need not be $P\sim F$. Thus, the relation (2.8) turns into

$$\mathscr{K}(P,\theta)=\frac{({\rm Cof}P^\top)\mathbb{K}(\theta){\rm Cof}P}{\det P}.\tag{2.10}$$

This expression bears the advantage of being independent of $(\nabla y)^{-1}$, which turns out useful in relation with estimation and passage to the limit arguments, cf. [20, 41].

The plastic flow rule (2.7b) complies with the so-called *plastic-indifference* requirement. Indeed, the evolution is insensible to prior plastic deformations, for the stored energy

$$\widehat\psi_{\rm M}=\widehat\psi_{\rm M}(F,P):=\psi_{\rm E}(F_{\rm el})\qquad\text{with}\quad F_{\rm el}=FP^{-1}$$

and the dissipation potential

$$\widehat{\mathfrak{R}}=\widehat{\mathfrak{R}}(P,\theta;\dot P)=\mathfrak{R}(\theta;\dot PP^{-1})$$

complies with the following invariant properties

$$\widehat\psi_{\rm M}(FQ,PQ)=\widehat\psi_{\rm M}(F,P),\quad \psi_{\rm H}(PQ)=\psi_{\rm H}(P),\ \text{ and }\ \widehat{\mathfrak{R}}(PQ,\theta;\dot PQ)=\widehat{\mathfrak{R}}(P,\theta;\dot P)\tag{2.11}$$

for any $Q\in{\rm SO}(d)$, as discussed by Mandel [24] and later in [26, 30]. In particular, we can equivalently test the flow rule (2.7b) by $\dot PP^{-1}$ or rewrite it as

$$\partial_R\mathfrak{R}\big(\theta;\dot PP^{-1}\big)P^{-\top}+\varSigma_{\rm in}=\partial_{\dot P}\widehat{\mathfrak{R}}\big(P,\theta;\dot P\big)+\varSigma_{\rm in}\ni 0\tag{2.12}$$

and test it on by $\dot P$ obtaining (at least formally) that

$$\partial_R\mathfrak{R}\big(\theta;\dot PP^{-1}\big){:}\dot PP^{-1}=-\varSigma_{\rm in}P^\top{:}\dot PP^{-1}=-\varSigma_{\rm in}P^\top P^{-\top}{:}\dot P=-\varSigma_{\rm in}{:}\dot P,\tag{2.13}$$

where we used also the algebra $AB{:}C=A{:}CB^\top$.

System (2.7) has to be complemented by suitable boundary and initial conditions. As for the former we prescribe

$$\varSigma_{\rm el}\nu-{\rm div}_{\rm s}\big(\kappa_0\nabla^2y\big)+Ny=Ny_\flat(t),\qquad \kappa_0\nabla^2y{:}(\nu\otimes\nu)=0,\tag{2.14a}$$
$$P=\mathbb{I},\qquad\text{and}\qquad \mathscr{K}(P,\theta)\nabla\theta{\cdot}\nu+K\theta=K\theta_\flat(t)\qquad\text{on }\ \partial\Omega.\tag{2.14b}$$

Relations (2.14a) correspond to a Robin-type mechanical condition. In particular, $\nu$ is the external normal at $\partial\Omega$, ${\rm div}_{\rm s}$ denotes the surface divergence defined as a trace of the surface gradient (which is a projection of the gradient on the tangent space through the projector

$\mathbb{I} - \nu \otimes \nu$), and $N$ is the elastic modulus of idealized *boundary* springs. In particular, $y_\flat$ is the (possibly time-dependent) position of the elastic support. Similarly we prescribe in (2.14b) Robin-type boundary condition for temperature, where $K$ is the boundary heat-transfer coefficient and $\theta_\flat$ is the external temperature. We assume $P$ to be the identity at $\partial\Omega$, meaning that plasticization occurs in the bulk only. This is chosen here for the sake of simplicity and could be weakened by imposing $P = \mathbb{I}$ on a portion of $\partial\Omega$ or even by a Neumann condition, this last requiring however a more delicate estimation argument, see Remark 4.5. Eventually, initial conditions read

$$y(0) = y_0, \qquad \dot{y}(0) = v_0, \qquad P(0) = P_0, \qquad \theta(0) = \theta_0. \tag{2.15}$$

The full model (2.7) with (2.14)-(2.15) is thermodynamically consistent. This can be checked by testing relations (2.7a) and (2.7b) by $\dot{y}$ and $\dot{P}P^{-1}$, respectively. By adding up these contributions and using (2.13) we obtain the *mechanical energy balance*

$$\frac{\mathrm{d}}{\mathrm{d}t}\bigg(\int_\Omega \frac{\varrho}{2}|\dot{y}|^2\,\mathrm{d}x + \Psi_{\rm M}(\nabla y, P) + \int_\Gamma \frac{1}{2}N|y|^2\,\mathrm{d}S\bigg) + \int_\Omega \partial_R\mathfrak{R}\big(\theta; \dot{P}P^{-1}\big):(\dot{P}P^{-1})\,\mathrm{d}x$$
$$= \int_\Omega g(y)\cdot\dot{y}\,\mathrm{d}x + \int_\Gamma N y_\flat\cdot\dot{y}\,\mathrm{d}S. \tag{2.16}$$

Note that this computation is formal and that such mechanical energy balance can be rigorously justified in case of smooth solutions only, which may not exists as $y$ lacks time regularity due to the possible occurrence of shock-waves in the nonlinear hyperbolic system (2.7a). Also the power of the external mechanical load in (2.16), i.e. $y_\flat\cdot\dot{y}$, is not well defined if $\nabla\dot{y}$ is not controlled. This term will be hence treated by by-part integration in time later on, see (4.11).

By adding to (2.16) the space integral of the heat equation (2.7c) we obtain the *total energy balance*

$$\frac{\mathrm{d}}{\mathrm{d}t}\bigg(\underbrace{\int_\Omega \frac{\varrho}{2}|\dot{y}|^2 + \vartheta\,\mathrm{d}x}_{\substack{\text{kinetic and heat}\\ \text{energies in the bulk}}} + \underbrace{\Psi_{\rm M}(\nabla y, P)}_{\substack{\text{mechanical energy}\\ \text{in the bulk}}} + \underbrace{\int_\Gamma \frac{1}{2}N|y|^2\mathrm{d}S}_{\substack{\text{mechanical energy}\\ \text{on the boundary}}}\bigg)$$
$$= \underbrace{\int_\Omega g(y)\cdot\dot{y}\,\mathrm{d}x}_{\substack{\text{power of}\\ \text{bulk load}}} + \underbrace{\int_\Gamma N y_\flat\cdot\dot{y}\,\mathrm{d}S}_{\substack{\text{power of surface}\\ \text{load on }\Gamma}} + \underbrace{\int_\Gamma K(\theta - \theta_\flat)\,\mathrm{d}S}_{\substack{\text{heat flux}\\ \text{thru }\Gamma}}. \tag{2.17}$$

From (2.6) with the heat flux $j = -\mathscr{K}\nabla\theta$ and with the dissipation rate ( which here equals the heat production rate) $r := \partial_R\mathfrak{R}\big(\theta; \dot{P}P^{-1}\big):(\dot{P}P^{-1})$, one can read the *entropy imbalance*

$$\frac{\mathrm{d}}{\mathrm{d}t}\int_\Omega \eta\,\mathrm{d}x = \int_\Omega \frac{r + \mathrm{div}(\mathscr{K}\nabla\theta)}{\theta}\,\mathrm{d}x = \int_\Omega \frac{r}{\theta} - \mathscr{K}\nabla\theta\cdot\nabla\frac{1}{\theta}\,\mathrm{d}x + \int_\Gamma \frac{\mathscr{K}\nabla\theta}{\theta}\cdot\nu\,\mathrm{d}S \tag{2.18}$$
$$= \int_\Omega \underbrace{\frac{r}{\theta} + \frac{\mathscr{K}\nabla\theta\cdot\nabla\theta}{\theta^2}}_{\substack{\text{entropy production}\\ \text{rate in the bulk }\Omega}}\mathrm{d}x + \int_\Gamma \underbrace{\frac{K(\theta_\flat - \theta)}{\theta}}_{\substack{\text{entropy flux through}\\ \text{the boundary }\Gamma}}\,\mathrm{d}S \geq \int_\Gamma K\Big(\frac{\theta_\flat}{\theta} - 1\Big)\theta\,\mathrm{d}S,$$

provided $\theta > 0$ and $\mathscr{K}$ is positive semidefinite. In particular, if the system is thermally isolated, i.e. $K = 0$, (2.18) states that the overall entropy is nondecreasing in time. This shows consistency with the 2nd Law of Thermodynamics.

Eventually, the 3rd Law of Thermodynamics (i.e. non-negativity of temperature), holds as soon as the initial/boundary conditions are suitably qualified so that $r \geq 0$. In fact, we do not consider any adiabatic-type effects, which might cause cooling.

We conclude the presentation of the model with a number of remarks and comments on our modeling choices and possible extensions.

**Remark 2.1** (*Kelvin-Voigt viscosity*). Viscous mechanical dynamics, i.e., Kelvin-Voigt-type viscosity, could be considered as well. This gives rise to a viscous contribution $\sigma_{\rm vi}(F, \dot{F})$ to the second Piola-Kirchhoff stress of the form $FS(U, \dot{U})$ where $S$ a symmetric tensor and $U^2 = F^\top F$, i.e. $F = QU$ with $Q \in {\rm SO}(d)$, cf. [1]. Equivalently, one can express such viscous contribution as $\sigma_{\rm vi}(F, \dot{F}) = F\hat{S}(C, \dot{C})$ for some given function $\hat{S}$. For $\sigma_{\rm vi}$ to have a potential, namely $\sigma_{\rm vi}(F, \dot{F}) = \partial_{\dot{F}}\zeta(F, \dot{F})$, frame indifference imposes that $\zeta(F, \dot{F}) = \zeta(QF, Q(F{+}A\dot{F}))$ for all $Q \in {\rm SO}(d)$ and all $A \in \mathbb{R}^{d\times d}$ antisymmetric. This in turn forces $\sigma_{\rm vi}$ to be strongly nonlinear. In case of nonsimple materials, such nonlinear viscosity is mathematically tractable, although its analysis is very delicate [31]. In combination with the Maxwellian rheology (=creep), this combination is sometimes referred as the *Jeffreys rheology*.

**Remark 2.2** (*Scaling of the plastic gradient*). Let us now inspect the relation between the parameter $\kappa_1$ and the length scale of plasticity in the material. To this aim, we consider $d = 2$ and resort to a stratified situation where $F$ and $P$ are constant the $x_1$ direction. We consider a shear band of width $2\ell$ along the plane $x_1 = 0$. By letting $\Omega$ be a rectangle of unit size, $q = 2$, and $\kappa_0 = 0$ for simplicity, and letting $F_{\rm el} = \mathbb{I}$ and thus $\nabla y = F = P$, the simplest profile of $P$ compatible with (2.3) is continuous and piecewise affine in the $x_2$-coordinate. Assume a time-dependent shear (caused by boundary conditions) with a constant velocity in the $x_2$-direction, namely

$$
y(x) = \begin{pmatrix} y_1(x_1,x_2) \\ y_2(x_1,x_2) \end{pmatrix} \quad \text{with} \quad y_1(x_1,x_2) = \begin{cases} x_1 + t & \text{if } x_2 > \ell, \\ x_1 + 2t - 2t\dfrac{x_2}{\ell} + t\dfrac{x_2^2}{\ell^2} & \text{if } 0 \leq x_2 \leq \ell, \\ x_1 - 2t - 2t\dfrac{x_2}{\ell} - t\dfrac{x_2^2}{\ell^2} & \text{if } 0 \geq x_2 \geq -\ell, \\ x_1 - t & \text{if } x_2 < -\ell, \end{cases}
$$

$$
\text{and} \quad y_2(x_1,x_2) = x_2.
$$

The corresponding plastic strain reads

$$
P = F = \nabla y = \begin{pmatrix} 1 & \partial y_1/\partial x_2 \\ 0 & 1 \end{pmatrix} \qquad \text{with} \quad \frac{\partial y_1}{\partial x_2} = \begin{cases} 2t\dfrac{x_2{-}\ell}{\ell^2} & \text{if } 0 \leq x_2 \leq \ell, \\ 0 & \text{if } |x_2| > \ell, \\ -2t\dfrac{x_2{+}\ell}{\ell^2} & \text{if } 0 \geq x_2 \geq -\ell. \end{cases} \tag{2.19}
$$

Therefore, in view of (2.19) the plastic-strain gradient $\nabla P$ has only one nonvanishing entry, namely

$$\left|\frac{\partial P_{12}}{\partial x_2}\right| \sim \begin{cases} 0 & \text{if } |x_2| > \ell, \\ 2t/\ell^2 & \text{if } |x_2| \le \ell. \end{cases}$$

Correspondingly, a bounded energy contribution from the term $\kappa_1|\nabla P|^2$ reveals the scaling

$$\ell \sim \kappa_1^{1/3} t^{2/3}$$

In particular, the occurrence of the plastic-strain gradient in the model has a hardening effect, for the slip zone widens as $t^{2/3}$ by accommodating large plastic slips. One could control such core size by letting $\kappa_1$ decay in time (which would be rather disputable), or resorting to computing the plastic-strain gradient in the actual configuration. Alternatively, one may consider including the plastic-strain gradient term into the dissipation potential rather than into the free energy. All these options seem to give rise to new nonlinearities in (2.7) which seriously complicate the analysis. We hence prefer to stay with the case of a constant and very small $\kappa_1$ (and thus a very narrow slip zone), so that the validity of this simplified model can be guaranteed on the relevant time scales.

**Remark 2.3** (*Alternatives in gradient terms*)**.** The gradient terms in the free energy (2.3) are all Lagrangian in nature. This choice is here dictated by the sake of mathematical simplicity. On the other hand, for $\nabla y$ and $P$ one could resort to computing gradients with respect to the actual configurations. These choices however give rise to additional nonlinearities which seem to prevent the possibility of developing a complete existence theory. It would be also physically relevant to consider the elastic strain gradient $\nabla F_{\rm el}$ instead of the total strain gradient $\nabla F = \nabla^2 y$ in the free energy (2.3). This would lead to replacing $-\kappa_0\,{\rm div}\nabla^2 y$ by $-{\rm div}\nabla(\nabla y\,{\rm Cof}P^\top)$ in (2.4a), but it would also give rise to extra contributions to (2.4b,d). Ultimately, these terms bring to additional mathematical difficulties which seem presently out of reach.

**Remark 2.4** (*Self-interpenetration*)**.** In large-strain theories, self-interpenetration of matter is usually excluded by constraining possible deformations. A possibility in this direction is to resort to the classical Ciarlet-Nečas condition [5] which reads

$$\int_\Omega \det(\nabla y(x))\,{\rm d}x \le |y(\Omega)|.$$

In general terms, encompassing such condition seems however to be out of reach, as it is usually the case for dynamic inelasticity with constraints.

**Remark 2.5** (*More general thermal coupling*)**.** A general coupling between $F$ and $\theta$ seems to call for viscosity in $F$, which is here not considered. On the other hand, one can consider making $\psi_{\rm T}$ in (2.3) dependent on $P$ as well, i.e. $\psi_{\rm T} = \psi_{\rm T}(P,\theta)$. The heat capacity would then depend also on $P$, i.e. $c_{\rm v}(P,\theta) = -\theta\partial_{\theta\theta}\psi_{\rm T}(P,\theta)$ and the heat-transfer equation (2.7c) would be augmented by an *adiabatic term*, leading to

$$c_{\rm v}(P,\theta)\dot\theta = {\rm div}\big(\mathscr{K}(P,\theta)\nabla\theta\big) + \partial_R\mathfrak{R}\big(\theta;\dot P P^{-1}\big){:}(\dot P P^{-1}) + \big(\partial_P\psi_{\rm T}(P,\theta) - \partial_P\psi_{\rm T}(P,0)\big)\dot P.$$

Due to the last term, the analysis is then substantially more complicated and rests on a a suitable modification of Lemma 4.3 below, cf. also [20, 32, 40] for this technique.

**Remark 2.6** (*Isochoric plasticity/creep*)**.** Isochoric plasticity or creep would correspond to the nonaffine holonomic constraint $\det P = 1$. The flow rule (2.12) would involve a *reaction force* to this constraint, resulting into

$$\partial_R \mathfrak{R}\big(\theta; \dot{P}P^{-1}\big)P^{-\top} + \Sigma_{\rm in} \ni \Lambda {\rm Cof} P \qquad \text{and} \qquad \det P = 1,$$

where $\Lambda$ is a $\mathbb{R}^{d\times d}$-valued Lagrange multiplier, note that we used also the formula ${\rm Cof} P = (\det P)'$. The mathematical analysis of such system seems open. A relevant model, which would be *approximately isochoric* for small $\delta > 0$, can be tackled by the analysis presented in the following Sections 3–4 when considering the hardening term of the type

$$\psi_{\rm H}(P) := \begin{cases} \dfrac{\delta}{\max(1, \det P)^r} + \dfrac{(\det P - 1)^2}{2\delta} & \text{if } \ \det P > 0, \\ +\infty & \text{if } \ \det P \geq 0\,; \end{cases}$$

note that the minimum of this potential is attained just at the set ${\rm SL}(d)$ of the isochoric plastic strains, and that it complies with condition (3.1b) ahead for $r \geq qd/(q-d)$ and also with the plastic-indifference condition (2.11).

# 3 Existence of weak solutions

This section introduces the definition of weak solution to the problem and brings to the statement of the our main existence result, namely Theorem 3.2. Let us start by fixing some notation.

We will use the standard notation $C(\cdot)$ for the space of continuous bounded functions, $L^p$ for Lebesgue spaces, and $W^{k,p}$ for Sobolev spaces whose $k$-th distributional derivatives are in $L^p$. Moreover, we will use the abbreviation $H^k = W^{k,2}$ and, for all $p \geq 1$, we let the conjugate exponent $p' = p/(p-1)$ (with $p' = \infty$ if $p = 1$), and we use the notation $p^*$ for the Sobolev exponent $p^* = pd/(d-p)$ for $p < d$, $p^* < \infty$ for $p = d$, and $p^* = \infty$ for $p > d$. Thus, $W^{1,p}(\Omega) \subset L^{p^*}(\Omega)$ or $L^{p^{*\prime}}(\Omega) \subset W^{1,p}(\Omega)^*$= the dual to $W^{1,p}(\Omega)$. In the vectorial case, we will write $L^p(\Omega;\mathbb{R}^d) \cong L^p(\Omega)^d$ and $W^{1,p}(\Omega;\mathbb{R}^d) \cong W^{1,p}(\Omega)^d$.

Given the fixed time interval $I = [0, T]$, we denote by $L^p(I; X)$ the standard Bochner space of Bochner-measurable mappings $I \to X$, where $X$ is a Banach space. Moreover, $W^{k,p}(I; X)$ denotes the Banach space of mappings in $L^p(I; X)$ whose $k$-th distributional derivative in time is also in $L^p(I; X)$.

Let us list here the assumptions on the data which are used in the following:

$$\psi_{\rm E}:\mathbb{R}^{d\times d}\to\mathbb{R}^+ \ \text{ continuously differentiable, } \ \sup_{F\in\mathbb{R}^{d\times d}}\frac{|\psi_{\rm E}'(F)|}{1{+}|F|^{2^*/2-1}}<\infty, \tag{3.1a}$$

$$\psi_{\rm H}:\mathbb{R}^{d\times d}\to[0,+\infty] \ \text{ continuously differentiable on } \ {\rm GL}^+(d),$$
$$\psi_{\rm H}(P)\ge\begin{cases}\epsilon/(\det P)^r & \text{if } \ \det P>0,\\ \ \ +\infty & \text{if } \ \det P\le 0,\end{cases}\qquad r\ge\frac{qd}{r{-}q},\qquad q>d, \tag{3.1b}$$

$$\varrho>0,\ \ \mathbb{K}:\mathbb{R}\to\mathbb{R}^{d\times d} \ \text{ continuous, bounded, and uniformly positive-definite,} \tag{3.1c}$$

$$\mathfrak{R}(\theta;R)=\mathfrak{R}_1(\theta;R)+\mathfrak{R}_2(\theta;R), \tag{3.1d}$$

$$\mathfrak{R}_1(\theta;R)=\sigma_{\rm Y}(\theta)|R| \ \text{ where } \ \sigma_{\rm Y}:\mathbb{R}^+\to\mathbb{R}^+ \ \text{ is continuous and bounded,} \tag{3.1e}$$

$$\mathfrak{R}_2:\mathbb{R}_+\times\mathbb{R}^{d\times d}\to\mathbb{R}^+ \ \text{ continuously differentiable,} \tag{3.1f}$$

$$\exists a_{\mathfrak{R}}>0\ \forall\theta\in\mathbb{R},\ \ R,R_1,R_2\in\mathbb{R}^{d\times d}:$$
$$(\partial_R\mathfrak{R}(\theta;R_1)-\partial_R\mathfrak{R}(\theta;R_2)){:}(R_1{-}R_2)\ge a_{\mathfrak{R}}|R_1{-}R_2|^2, \tag{3.1g}$$

$$a_{\mathfrak{R}}|R|^2\le\mathfrak{R}(\theta;R)\le(1+|R|^2)/a_{\mathfrak{R}},\qquad \mathfrak{R}(\theta;-R)=\mathfrak{R}(\theta;R), \tag{3.1h}$$

$$c_{\rm v}:\mathbb{R}_+\to\mathbb{R}_+ \ \text{ continuous, bounded, with positive infimum,} \tag{3.1i}$$

$$y_0{\in}H^2(\Omega)^d,\ \ v_0{\in}L^2(\Omega)^d,\ \ P_0{\in}W^{1,q}(\Omega)^{d\times d},\ \ \psi_{\rm H}(P_0){\in}L^1(\Omega),\ \ q>d, \tag{3.1j}$$

$$\theta_0{\in}L^1(\Omega),\ \ \theta_0\ge0, \tag{3.1k}$$

$$g\in C(\mathbb{R}^d)^d,\ \ \theta_\flat\in L^1(\Sigma),\ \ \theta_\flat\ge0, \tag{3.1l}$$

where $q$ in (3.1b,j) refers to the exponent used in (2.3). The nonnegative function $\sigma_{\rm Y}=\sigma_{\rm Y}(\theta)$ is in the position of a temperature-dependent *yield stress*, i.e., a threshold triggering plastification.

Note that the polynomial growth assumption on $\psi_{\rm E}'$ from (3.1a) is not particularly restrictive for $d=2$ but requires $\psi_{\rm E}'(F)\lesssim|F|^2$ for $d=3$. Such a restricted growth is however compatible with the assumption that $F_{\rm el}$ is close to the identity. Indeed, if $\psi_{\rm E}$ were a linearization of a nonlinear elastic energy density at $F_{\rm el}=\mathbb{I}$ one would have

$$\psi_{\rm E}(F_{\rm el})=\frac12(F_{\rm el}{-}\mathbb{I}){:}{\rm D}^2\psi_{\rm E}(\mathbb{I}){:}(F_{\rm el}{-}\mathbb{I})$$

so that the bound in (3.1a) trivially holds. Note that the fourth-order tensor ${\rm D}^2\psi_{\rm E}(\mathbb{I})$ plays here the role of (a multiple of) the elasticity-moduli tensor in linearized elasticity.

In the case when $\mathfrak{R}(\theta;\cdot)$ is nonsmooth at 0, its subdifferential are indeed set-valued and thus (3.1h) is to be satisfied for any selection from the involved subdifferentials. On the other hand, the heat-production rate $\partial_R\mathfrak{R}\big(\theta;\dot PP^{-1}\big){:}(\dot PP^{-1})$ in (2.7c) remains single-valued as $\partial_R\mathfrak{R}\big(\theta;\cdot\big)$ is multivalued just in $\dot PP^{-1}=0$.

Testing (2.7) by smooth functions and using Green formula in space (even twice for (2.7a) together with a surface Green formula over $\Gamma$), the boundary conditions (2.14), by-part integration in time for (2.7a,b), and the definition of the convex subdifferential $\partial_R\mathfrak{R}\big(\theta;\cdot\big)$ for (2.7b), we arrive at the following.

**Definition 3.1** (Weak formulation of (2.7) with (2.14)-(2.15))**.** *We call the triple* $(y,P,\theta)$

*with*

$$y \in L^\infty(I;H^2(\Omega)^d)\cap H^1(I;L^2(\Omega)^d), \tag{3.2a}$$
$$P \in L^\infty(I;W^{1,q}(\Omega)^{d\times d})\cap H^1(I;L^2(\Omega)^{d\times d}) \quad \textit{with} \quad \frac{1}{\det P}\in L^\infty(Q), \tag{3.2b}$$
$$\theta \in L^1(I;W^{1,1}(\Omega)), \qquad \theta\ge 0 \ \ \textit{a.e. on} \ \ Q, \tag{3.2c}$$

*a weak solution to the initial-boundary-value problem* (2.7) *with* (2.14)-(2.15) *if*

$$\mathrm{div}\big(\kappa_1|\nabla P|^{q-2}\nabla P\big)\in L^2(Q)^{d\times d} \tag{3.2d}$$

*and if the following hold:*

(i) *The weak formulation of the momentum balance* (2.7a) *with* (2.4a)

$$\int_Q \Big(\psi_{\rm E}'(\nabla yP^{-1}){:}(\nabla\tilde y\,P^{-1}) - \varrho\dot y{\cdot}\dot{\tilde y} + \kappa_0\nabla^2 y\vdots\nabla^2\tilde y\,\mathrm{d}x\,\mathrm{d}t + \int_\Sigma Ny{\cdot}\tilde y\,\mathrm{d}S\,\mathrm{d}t = \int_Q g(y){\cdot}\tilde y\,\mathrm{d}x\,\mathrm{d}t + \int_\Omega v_0{\cdot}\tilde y(0)\,\mathrm{d}x + \int_\Sigma Ny_\flat{\cdot}\tilde y\,\mathrm{d}S\,\mathrm{d}t \tag{3.3a}$$

*holds for any $\tilde y$ smooth with $\tilde y(T)=0$.*

(ii) *The weak formulation of the plastic flow rule* (2.7b) *in the form* (2.12) *with* (2.4b)

$$\int_Q \Big(\mathfrak{R}(\theta;R) + \psi_{\rm E}'(\nabla yP^{-1})P^\top{:}(\nabla y((P^{-1})')(R{-}\dot PP^{-1}) - \mathrm{div}\big(\kappa_1|\nabla P|^{q-2}\nabla P\big)P^\top{:}(R{-}\dot PP^{-1})\Big)\,\mathrm{d}x\,\mathrm{d}t \ge \int_Q \mathfrak{R}\big(\theta;\dot PP^{-1}\big)\,\mathrm{d}x\,\mathrm{d}t \tag{3.3b}$$

*holds for any $R$ smooth.*

(iii) *The weak formulation of the heat equation* (2.7c)

$$\int_Q \mathscr{K}(P,\theta)\nabla\theta{\cdot}\nabla\tilde\theta - C_{\rm v}(\theta)\dot{\tilde\theta} - \partial_R\mathfrak{R}\big(\theta;\dot PP^{-1}\big){:}(\dot PP^{-1})\tilde\theta\,\mathrm{d}x\,\mathrm{d}t + \int_\Sigma K\theta\tilde\theta\,\mathrm{d}S\,\mathrm{d}t = \int_\Sigma K\theta_\flat\tilde\theta\,\mathrm{d}S\,\mathrm{d}t + \int_\Omega C_{\rm v}(\theta_0)\tilde\theta(0)\,\mathrm{d}x \tag{3.3c}$$

*holds for any $\tilde\theta$ smooth with $\tilde\theta(T)=0$ and with $C_{\rm v}(\cdot)$ a primitive function to $c_{\rm v}(\cdot)$.*

(iv) *The remaining initial conditions $y(0)=y_0$ and $P(0)=P_0$ are satisfied.*

Let us note that, due to (3.2b), we have also $P^{-1} = \mathrm{Cof}P^\top/\det P \in L^\infty(Q)^{d\times d}$ so that in particular $\mathrm{div}\big(\kappa_1|\nabla P|^{q-2}\nabla P\big)P^\top{:}\dot PP^{-1}\in L^1(Q)$ due to (3.2d) and thus (3.3b) has a good sense.

Our main analytical result is an existence theorem for weak solutions. This is to be seen as a mathematical consistency property of the proposed model. It reads as follows.

**Theorem 3.2** (Existence of weak solutions)**.** *Let the assumptions* (3.1) *hold. Then, there exists a weak solution* $(y, P, \theta)$ *in the sense of Definition* 3.1 *with* $P$ *valued in* $\mathrm{GL}^+(d)$*. Moreover, the energy conservation* (2.17) *holds on the time intervals* $[0, t]$ *for all* $t \in I$*, i.e.*

$$\int_\Omega \frac{\varrho}{2}|\dot{y}(t)|^2 + C_{\rm v}(\theta(t))\,\mathrm{d}x + \Psi_{\rm M}(\nabla y(t), P(t)) + \int_\Gamma \frac{1}{2}N|y(t)|^2 \mathrm{d}S$$
$$= \int_0^t\!\!\int_\Omega g(y)\cdot\dot{y}\,\mathrm{d}x\,\mathrm{d}t + \int_0^t\!\!\int_\Gamma Ny_\flat\cdot\dot{y} + K(\theta{-}\theta_\flat)\,\mathrm{d}S\,\mathrm{d}t$$
$$+ \int_\Omega \frac{\varrho}{2}|v_0|^2 + C_{\rm v}(\theta_0)\,\mathrm{d}x + \Psi_{\rm M}(\nabla y_0, P_0) + \int_\Gamma \frac{1}{2}N|y_0|^2 \mathrm{d}S\,. \tag{3.4}$$

We will prove this result in Propositions 4.1-4.4 by a regularization, transformation, and approximation procedure. This also provides a (conceptual) algorithm that is numerically stable and converges as the discretization and the regularization parameters tend to 0.

# 4 Galerkin approximation, stability, convergence

We devote this section to the proof of the existence result, namely Theorem 3.2. As already mentioned, we apply a constructive method providing an approximation of the problem. This results from combining a regularization in terms of the small parameter $\varepsilon > 0$ and a Galerkin approximation, described by the small parameter $h > 0$ instead. The regularization is aimed on the one hand at smoothing the potential $\mathfrak{R}(\theta, R)$ in a neighborhood of $R = 0$ and on the other hand at making the heat-production rate and boundary and initial temperatures bounded, see below. We obtain the existence of approximated solutions, their stability (a-priori estimates), and their convergence to weak solutions, at least in terms of subsequences. The general philosophy of a-priori estimation relies on the fact that temperature plays a role in connection with dissipative mechanisms only: adiabatic effects are omitted and most estimates on the mechanical part of the system are independent of temperature and its discretization. The estimates and the convergence rely on the independence of the heat capacity from mechanical variables, cf. also Remark 2.5 above.

Let us begin by detailing the regularization. First, we smoothen the convex (but generally nonsmooth) potential $\mathfrak{R}(\theta;\cdot) : \mathbb{R}^{d\times d} \to \mathbb{R}^+$ in order to be able to use the conventional theory of ordinary-differential equations for the Galerkin approximation. To this goal, we use the splitting $\mathfrak{R}(\theta; R) = \mathfrak{R}_1(\theta; R) + \mathfrak{R}_2(\theta; R)$ from (3.1d). Then, we exploit the general Yosida-regularization construction to the nonsmooth part, i.e. $\mathfrak{R}_\varepsilon(\theta; R) := \mathfrak{R}_{1,\varepsilon}(\theta; R) + \mathfrak{R}_2(\theta; R)$ with

$$\mathfrak{R}_{1,\varepsilon}(\theta; R) := \min_{\widetilde{R}\in\mathbb{R}^{d\times d}} \left(\mathfrak{R}_1(\theta; \widetilde{R}) + \frac{1}{2\varepsilon}|\widetilde{R}{-}R|^2\right) = \begin{cases} \dfrac{\sigma_{\rm Y}(\theta)}{2\varepsilon}|R|^2 & \text{if } |R| \le \varepsilon \\[2mm] \sigma_{\rm Y}(\theta)\left(|R|{-}\dfrac{\varepsilon}{2}\right) & \text{if } |R| > \varepsilon, \end{cases} \tag{4.1}$$

where $\sigma_{\rm Y}$ is from (3.1e). Properties (3.1d-f) of the original potential $\mathfrak{R}$ entail analogous properties for the regularization, in particular

$$\exists \varepsilon_0 > 0\ \forall\, 0 < \varepsilon \le \varepsilon_0 : \forall \theta{\in}\mathbb{R},\ R_1, R_2{\in}\mathbb{R}^{d\times d} : \tag{4.2a}$$

$$(\partial_R\mathfrak{R}_\varepsilon(\theta; R_1) - \partial_R\mathfrak{R}_\varepsilon(\theta; R_2)){:}(R_1{-}R_2) \ge \frac{1}{2}a_{\mathfrak{R}}|R_1{-}R_2|^2, \tag{4.2b}$$

$$\frac{1}{2}a_{\mathfrak{R}}|R|^2 \le \mathfrak{R}_\varepsilon(\theta; R) \le (1 + |R|^2)/a_{\mathfrak{R}} \tag{4.2c}$$

where $a_{\mathfrak{R}} > 0$ is from (3.1h). In addition, we can prove that

$$\forall\theta\in\mathbb{R},\ \ R\in\mathbb{R}^{d\times d},\ \ R\neq 0: \qquad \lim_{\varepsilon\to 0,\ \widetilde{\theta}\to\theta} \partial_R\mathfrak{R}_\varepsilon(\widetilde{\theta}; R) = \partial_R\mathfrak{R}(\theta; R). \tag{4.2d}$$

Note that $\mathfrak{R}(\theta;-R) = \mathfrak{R}(\theta;R)$ assumed in (3.1h) implies $\partial_R\mathfrak{R}_\varepsilon(\theta;0) = 0$ so that the limit in (4.2d) exists even if $R = 0$ and we have

$$\forall\theta\in\mathbb{R}: \qquad \lim_{\varepsilon\to 0,\ \widetilde{\theta}\to\theta} \partial_R\mathfrak{R}_\varepsilon(\widetilde{\theta}; 0) = 0 \in \partial_R\mathfrak{R}(\theta; 0). \tag{4.2e}$$

In order to simplify the convergence proof, we apply the so-called enthalpy transformation to the heat equation. This consists in rescaling temperature by introducing a new variable

$$\vartheta = C_{\rm v}(\theta) \tag{4.3}$$

where $C_{\rm v}$ is the primitive of $c_{\rm v}$ vanishing in 0. Note that $\dot{\vartheta} = c_{\rm v}(\theta)\dot{\theta}$ and that $C_{\rm v}$ is increasing so that its inverse $C_{\rm v}^{-1}$ exists and $\nabla\theta = \nabla C_{\rm v}^{-1}(\vartheta) = \nabla\vartheta/c_{\rm v}(\theta) = \nabla\vartheta/c_{\rm v}(C_{\rm v}^{-1}(\vartheta))$. Upon letting

$$\mathfrak{K}(P,\vartheta) := \frac{\mathscr{K}(P, C_{\rm v}^{-1}(\vartheta))}{c_{\rm v}(C_{\rm v}^{-1}(\vartheta))},$$

we rewrite and regularize the system (2.7) by

$$\varrho\ddot{y} = {\rm div}\,\Sigma_{\rm el} + g(y), \tag{4.4a}$$

$$\partial_R\mathfrak{R}_\varepsilon\big(C_{\rm v}^{-1}(\vartheta); \dot{P}P^{-1}\big)P^{-\top} + \Sigma_{\rm in} = 0, \tag{4.4b}$$

$$\dot{\vartheta} = {\rm div}\big(\mathfrak{K}(P,\vartheta)\nabla\vartheta\big) + \frac{\partial_R\mathfrak{R}_\varepsilon\big(\theta; \dot{P}P^{-1}\big){:}(\dot{P}P^{-1})}{1+\varepsilon|\dot{P}P^{-1}|^2}, \tag{4.4c}$$

where $\Sigma_{\rm el}$ and $\Sigma_{\rm in}$ are again from (2.4). Note that we used (4.4b) in the form (2.12), which allows the test by $\dot{P}$, in contrast to (2.7b) which is to be tested by the product $\dot{P}P^{-1}$ which is not legitimate at the level of the Galerkin discretisation. Let us note that due to the boundedness/growth assumptions (3.1h), the dissipation rate has a quadratic growth and the regularization of the heat-production rate in (4.4c) is bounded. We are hence in the position of resorting to a $L^2$-theory instead of the $L^1$-theory for the regularized heat problem.

The boundary conditions are correspondingly modified, i.e. $\mathscr{K}(P,\theta)\nabla\theta\cdot\nu + K\theta = K\theta_\flat(t)$ in (2.14b) and $\theta(0) = \theta_0$ in (2.15) modify respectively as

$$\mathfrak{K}(P,\vartheta)\nabla C_{\rm v}^{-1}(\vartheta)\cdot\nu + KC_{\rm v}^{-1}(\vartheta) = K\theta_{\flat\varepsilon} \qquad \text{with} \quad \theta_{\flat\varepsilon} := \frac{\theta_\flat}{1+\varepsilon\theta_\flat}, \tag{4.5a}$$

$$\vartheta(0) = \vartheta_{0\varepsilon} := C_{\rm v}(\theta_{0\varepsilon}) \qquad \text{with} \quad \theta_{0\varepsilon} = \frac{\theta_0}{1+\varepsilon\theta_0}. \tag{4.5b}$$

In particular, $\theta_{\flat,\varepsilon}$ and $\theta_{0\varepsilon}$ in (4.5) are bounded.

As announced, we use a Galerkin approximation in space for (4.4). (which, in its evolution variant, is sometimes referred to as *Faedo-Galerkin method*). For possible numerical implementation, one can imagine a conformal finite element formulation, with $h > 0$ denoting the *mesh size*. Assume for simplicity that the sequence of nested finite-dimensional subspaces $V_h \subset H^2(\Omega)$ invading $H^1(\Omega)$ is given. This will make all Laplacians defined in the usual strong sense even on the discrete level, allowing for some simplification in the estimates. For simplicity, we assume that all initial conditions $(y_0, P_0, \vartheta_{0\varepsilon})$ belong to all finite-dimensional subspaces so that no additional approximation of such conditions is needed.

The outcome of the Galerkin approximation is an an initial-value problem for a system of ordinary differential equations (ODEs). In (4.8d) below, we denote $|\cdot|_h$ the seminorm on $L^2(I; H^1(\Omega)^*)$ defined by

$$|\xi|_h^* := \sup\left\{ \int_Q \xi v \,\mathrm{d}x\,\mathrm{d}t \;:\; \|v\|_{L^2(I;H^1(\Omega))} \le 1,\; v(t) \in V_h \;\text{ for a.e. } t \in I \right\}. \tag{4.6}$$

Similar seminorms (with the same notation) are defined on spaces tensor-valued functions. On $L^2$-spaces we let

$$|\xi|_h := \sup\left\{ \int_Q \xi{:}v \,\mathrm{d}x\,\mathrm{d}t \;:\; \|v\|_{L^2(Q)^{d\times d}} \le 1,\; v(t) \in V_h^{d\times d} \;\text{ for a.e. } t \in I \right\}, \tag{4.7}$$

to be used for (4.8e) below. This family of seminorms makes the linear spaces $L^2(I; H^1(\Omega)^*)$ and $L^2(Q)^{d\times d}$ and $L^2(Q)$, metrizable locally convex spaces (Fréchet spaces). Henceforth, we use the symbol $C$ to indicate a positive constant, possibly depending on data but independent of regularization and discretization parameters. Dependences on such parameters will be indicated in indices. Our stability result reads as follows.

**Proposition 4.1** (Discrete solution and a-priori estimates)**.** *Let assumptions* (3.1) *hold and $\varepsilon, h > 0$ be fixed. Then, the Galerkin approximation of* (4.4) *with the initial/boundary conditions* (2.14)-(2.15) *modified by* (4.5) *admits a solution on the whole time interval $I = [0, T]$. By denoting such solution as $(y_{\varepsilon h}, P_{\varepsilon h}, \vartheta_{\varepsilon h})$, the following estimates hold*

$$\big\|y_{\varepsilon h}\big\|_{L^\infty(I;H^2(\Omega)^d)\cap W^{1,\infty}(I;L^2(\Omega)^d)} \le C, \tag{4.8a}$$

$$\big\|P_{\varepsilon h}\big\|_{L^\infty(I;W^{1,q}(\Omega)^{d\times d})\cap H^1(I;L^2(\Omega)^{d\times d})} \le C \quad \text{and} \quad \Big\|\frac{1}{\det P_{\varepsilon h}}\Big\|_{L^\infty(Q)} \le C, \tag{4.8b}$$

$$\big\|\vartheta_{\varepsilon h}\big\|_{L^2(I;H^1(\Omega))} \le C_\varepsilon, \tag{4.8c}$$

$$\big|\dot{\vartheta}_{\varepsilon h}\big|_{h_0}^* \le C_\varepsilon \quad \text{for } \; h_0 \ge h > 0, \tag{4.8d}$$

$$\big|\mathrm{div}(|\nabla P_{\varepsilon h}|^{q-2}\nabla P_{\varepsilon h})\big|_{h_0} \le C \quad \text{for } \; h_0 \ge h > 0, \tag{4.8e}$$

*where $C$ and $C_\varepsilon$ are some constant independent of $h$ and $h_0$, $C$ being independent also of $\varepsilon$.*

*Sketch of the proof.* The existence of a global solution to the Galerkin approximation follows directly by the usual successive-continuation argument applied to the underlying system of ODEs.

Let us now move to a priori estimation. We start by recovering the mechanical energy balance, see (2.16). In particular, we use $\dot{y}_{\varepsilon h}$, $\dot{P}_{\varepsilon h}$, and $\dot{\zeta}_{\varepsilon h}$ as test functions into each corresponding equation discretized by the Galerkin method. More specifically, using $\dot{y}_{\varepsilon h}$ as test in the Galerkin approximation of (4.4a) with its boundary condition (2.14a), we obtain

$$\int_\Omega \frac{\varrho}{2}|\dot{y}_{\varepsilon h}(t)|^2 + \frac{\kappa_0}{2}|\nabla^2 y_{\varepsilon h}(t)|^2 \,\mathrm{d}x + \int_\Gamma \frac{1}{2}N|y_{\varepsilon h}(t)|^2 \,\mathrm{d}S + \int_0^t\!\!\int_\Omega \partial_{\nabla y}\widehat{\psi}_{\rm M}(\nabla y_{\varepsilon h}, P_{\varepsilon h}){:}\nabla\dot{y}_{\varepsilon h}\,\mathrm{d}x\,\mathrm{d}t$$
$$= \int_0^t\!\!\int_\Omega g(y){\cdot}\dot{y}_{\varepsilon h}\,\mathrm{d}x\,\mathrm{d}t + \int_0^t\!\!\int_\Gamma N y_\flat{\cdot}\dot{y}_{\varepsilon h}\,\mathrm{d}S\,\mathrm{d}t + \int_\Omega \frac{\varrho}{2}|v_0|^2 + \frac{\kappa_0}{2}|\nabla^2 y_0|^2\,\mathrm{d}x + \int_\Gamma \frac{1}{2}N|y_0|^2\,\mathrm{d}S. \tag{4.9}$$

By testing the Galerkin approximation of (4.4b) by $\dot{P}_{\varepsilon h}$ one gets

$$\int_\Omega \psi_{\rm H}(P_{\varepsilon h}(t)) + \frac{\kappa_1}{q}|\nabla P_{\varepsilon h}(t)|^q \mathrm{d}x + \int_0^t\!\!\int_\Omega \partial_R\mathfrak{R}_\varepsilon\big(\theta_{\varepsilon h};\dot{P}_{\varepsilon h}P_{\varepsilon h}^{-1}\big){:}\dot{P}_{\varepsilon h}P_{\varepsilon h}^{-1}$$
$$+ \partial_P\widehat{\psi}_{\rm M}(\nabla y_{\varepsilon h}, P_{\varepsilon h}){:}\dot{P}_{\varepsilon h}\mathrm{d}x\,\mathrm{d}t = \int_\Omega \psi_{\rm H}(P_0) + \frac{\kappa_1}{q}|\nabla P_0|^q\mathrm{d}x. \tag{4.10}$$

Taking the sum of (4.9)-(4.10) and using the calculus

$$\partial_{\nabla y}\widehat{\psi}_{\rm M}{:}\nabla\dot{y}_{\varepsilon h} + \partial_P\widehat{\psi}_{\rm M}{:}\dot{P}_{\varepsilon h} = \frac{\partial}{\partial t}\widehat{\psi}_{\rm M}(\nabla y_{\varepsilon h}, P_{\varepsilon h}),$$

we obtain the discrete analogue of (2.16).

The boundary term in (4.9) contains $\dot{y}$, which is not well defined on $\Gamma$. We overcome this obstruction by by-part integration

$$\int_0^t\!\!\int_\Gamma N y_\flat{\cdot}\dot{y}_{\varepsilon h}\,\mathrm{d}S\mathrm{d}t = \int_\Gamma N y_\flat(t){\cdot}y_{\varepsilon h}(t)\,\mathrm{d}S - \int_0^t\!\!\int_\Gamma N\dot{y}_\flat{\cdot}y_{\varepsilon h}\,\mathrm{d}S\,\mathrm{d}t - \int_\Gamma N y_\flat(0){\cdot}y_0\,\mathrm{d}S \tag{4.11}$$

so that this boundary term can be estimated by using the assumption (3.1l) on $y_\flat$.

These estimates allow us to obtain the bounds (4.8a,b). More in detail, the first estimate in (4.8b) follows from the coercivity (3.1h) of $\mathfrak{R}_\varepsilon$ so that we have also that $\dot{P}_{\varepsilon h}P_{\varepsilon h}^{-1}$ is bounded in $L^2(Q)^{d\times d}$. In particular, we have here used the boundary condition on the plastic strain (2.14b), see also Remark 4.5.

Exploiting (3.1b) we can use the Healey-Krömer Theorem [17, Thm. 3.1] for the plastic strain instead of the deformation gradient. In particular, [17, Thm. 3.1] states that any function $u \in W^{2,p}(\Omega;\mathbb{R}^d)$ with $\det\nabla u > 0$ such that $\int_\Omega |\det\nabla u|^{-q}\,\mathrm{d}x \le K$ is such that $\min_{x\in\bar\Omega}\det\nabla u =: \epsilon > 0$, provided $p > d$ and $q > pd/(p{-}d)$. In fact, this estimate holds uniformly with respect to $u$, as $\epsilon$ depends on $K$ and data only. In fact, by inspecting its proof, see also in [20,41], one easily realizes that this result holds for any matrix field, even if it does not come from a gradient of a vector field. In particular, one has that any $P \in W^{1,p}(\Omega;\mathbb{R}^{d\times d})$ with $\det P > 0$ such that $\int_\Omega |\det P|^{-q}\,\mathrm{d}x \le K$ fulfills $\min_{x\in\bar\Omega}\det P =: \epsilon > 0$. This gives the second estimate in (4.8b). It is important that it is available even on the Galerkin level, so that in fact the singularity of $\psi_{\rm H}$ is not seen during the evolution and the Lavrentiev phenomenon is excluded. Let us point out that, in the frame of our weak thermal coupling the assumption (3.1h), these estimates hold independently of temperature, and thus the constants in (4.8a,b) are independent of $\varepsilon$.

Let us now test the Galerkin approximation of the heat equation (4.4c) by $\vartheta_{\varepsilon h}$. This test is allowed at the level of Galerkin approximation, although it does not lead to the total energy balance. We obtain

$$\frac{\mathrm{d}}{\mathrm{d}t}\frac{1}{2}\int_\Omega \vartheta_{\varepsilon h}^2\,\mathrm{d}x+\int_\Omega \mathfrak{K}(P_{\varepsilon h},\vartheta_{\varepsilon h})\nabla\vartheta_{\varepsilon h}{\cdot}\nabla\vartheta_{\varepsilon h}\,\mathrm{d}x+\int_\Gamma K\vartheta_{\varepsilon h}^2\,\mathrm{d}S=\int_\Omega r_\varepsilon\vartheta_{\varepsilon h}\,\mathrm{d}x+\int_\Gamma K\theta_{\flat\varepsilon}\vartheta_{\varepsilon h}\,\mathrm{d}S. \tag{4.12}$$

After integration over $[0,t]$, we use the Gronwall inequality and exploit the control of the initial condition $|\theta_{0\varepsilon}|\le 1/\varepsilon$ due to the regularization (4.5). The last boundary term in (4.12) can be controlled as $|\theta_{\flat,\varepsilon}|\le 1/\varepsilon$, again due to the regularization (4.5). Using the positive definiteness of $\mathbb{K}$ in (3.1c) and recalling (2.10), we get the bound $\|(\mathrm{Cof}P_{\varepsilon h})\nabla\theta_{\varepsilon h}/\sqrt{\det P_{\varepsilon h}}\|_{L^2(Q)^d}\le C_\varepsilon$. Then also (4.8c) by using

$$\begin{aligned}\|\nabla\theta_{\varepsilon h}\|_{L^2(Q)^d}&=\Big\|\frac{P_{\varepsilon h}^\top\mathrm{Cof}P_{\varepsilon h}}{\det P_{\varepsilon h}}\nabla\theta_{\varepsilon h}\Big\|_{L^2(Q)^d}\\&\le\Big\|\frac{P_{\varepsilon h}}{\sqrt{\det P_{\varepsilon h}}}\Big\|_{L^\infty(Q)^{d\times d}}\Big\|\frac{\mathrm{Cof}P_{\varepsilon h}}{\sqrt{\det P_{\varepsilon h}}}\nabla\theta_{\varepsilon h}\Big\|_{L^2(Q)^d}\le C_\varepsilon,\end{aligned}\tag{4.13}$$

where the latter bound follows from (4.8b).

By comparison, we obtain the estimate (4.8d) of $\dot\vartheta_{\varepsilon h}$ in the seminorm (4.6). Again by comparison we obtain (4.8e), using (4.4b) with (2.4b) and taking advantage of the boundedness of the term $\partial_R\mathfrak{R}_\varepsilon(\theta_{\varepsilon h};\dot P_{\varepsilon h}P_{\varepsilon h}^{-1})P_{\varepsilon h}^{-\top}$ in $L^2(Q)^{d\times d}$ and similarly also of the first term in (2.4b). More specifically, the term $\nabla y_{\varepsilon h}^\top\psi_{\rm E}'(\nabla y_{\varepsilon h}P_{\varepsilon h}^{-1}){:}(P_{\varepsilon h}^{-1})'$ turns out to be bounded in $L^2(Q)^{d\times d}$ because $\nabla y_{\varepsilon h}^\top$ is bounded in $L^\infty(I;L^{2^*}(\Omega)^{d\times d})$ and $\psi_{\rm E}'$ is bounded in $L^\infty(I;L^{2^*2/(2^*-2)}(\Omega)^{d\times d})$ due to the growth condition (3.1a), and $(P_{\varepsilon h}^{-1})'$ is controlled in $L^\infty(Q)^{d\times d\times d\times d}$. Here, we emphasize that one cannot perform on relation (2.7b) the nonlinear test by $\mathrm{div}(|\nabla P_{\varepsilon h}|^{q-2}\nabla P_{\varepsilon h})$ to obtain the estimate (4.8e) in the full $L^2(Q)^{d\times d}$-norm. □

**Proposition 4.2** (Convergence of the Galerkin approximation for $h\to0$)**.** *Let assumptions* (3.1) *hold and let* $\varepsilon>0$ *be fixed. Then, for* $h\to0$*, there exists a not relabeled subsequence of* $\{(y_{\varepsilon h},P_{\varepsilon h},\vartheta_{\varepsilon h})\}_{h>0}$ *converging weakly* in the topologies indicated in* (4.8)a-g *to some* $(y_\varepsilon,P_\varepsilon,\vartheta_\varepsilon)$*. Every such limit triple is a weak solution to the regularized problem* (4.4) *with the initial/boundary conditions* (2.14)-(2.15) *modified by* (4.5)*. Moreover, the following a-priori estimate holds*

$$\big\|\mathrm{div}(|\nabla P_\varepsilon|^{q-2}\nabla P_\varepsilon)\big\|_{L^2(Q)^{d\times d}}\le C. \tag{4.14}$$

*Furthermore, the following strong convergences hold for* $h\to0$

$$\dot P_{\varepsilon h}P_{\varepsilon h}^{-1}\to\dot P_\varepsilon P_\varepsilon^{-1}\qquad\textit{strongly in}\ \ L^2(Q)^{d\times d}, \tag{4.15a}$$

$$\nabla P_{\varepsilon h}\to\nabla P_\varepsilon\qquad\textit{strongly in}\ \ L^q(Q)^{d\times d\times d}. \tag{4.15b}$$

*Proof.* The existence of weakly* converging not relabeled subsequences follows by the classical Banach selection principle. Let us indicate one such weak* limit by $(y_\varepsilon,P_\varepsilon,\vartheta_\varepsilon)$ and prove that it solves the regularized problem (4.4). Note that, estimates (4.14) follow from (4.8e), which are independent of $h$ and $h_0$, cf. [40, Sect. 8.4] for this technique. More in detail, one can consider a Hahn-Banach extension of the linear bounded functional occuring in (4.8e) from the linear subspace of $L^2(Q)^{d\times d}$ as in the definition (4.7) of the seminorm $|\cdot|_h$ to the whole space $L^2(Q)^{d\times d}$. This extension is bounded, sharing the same bound $C$ as in (4.8e).

Selecting, for a moment, another subsequence of these extensions which converges weakly in $L^2(Q)^{d\times d}$, one can eventually identify the limit again as $\mathrm{div}(|\nabla P_\varepsilon|^{q-2}\nabla P_\varepsilon)\in L^2(Q)^{d\times d}$ and see that, in fact, the whole originally selected subsequence converges as well.

In order to check that weak* limits are solutions, we are called to prove convergence of the dissipation-rate term, i.e. the heat-production rate, in the heat-transfer equation. This in turn requires that we prove the strong convergence of $P_{\varepsilon h}$, i.e. (4.15). To this aim, let $\tilde P_h$ be elements of the finite-dimensional subspaces which are approximating $P_\varepsilon$ with respect to strong $L^2$ topologies along with the corresponding time derivatives. Such approximants can be constructed by projections at the level of time derivatives. In particular, one can ask that $\tilde P_h\to P_\varepsilon$ strongly in $H^1(0,T;L^2(\Omega)^{d\times d})\cap L^1(0,T;W^{1,q}(\Omega)^{d\times d})$.

As for the strong convergence of $\nabla P_{\varepsilon h}$, we exploit the uniform monotonicity of the $q$-Laplacian. The Galerkin identity related to (4.4b)

$$\int_Q \nabla y_{\varepsilon h}^\top \psi_{\rm E}'(F_{{\rm el},\varepsilon h}){:}(P_{\varepsilon h}^{-1})'{:}\tilde P+\partial_R\mathfrak{R}_\varepsilon(\theta_{\varepsilon h};\dot P_{\varepsilon h}P_{\varepsilon h}^{-1}){:}(\tilde P P_{\varepsilon h}^{-1}) \\ +\kappa_1|\nabla P_{\varepsilon h}|^{q-2}\nabla P_{\varepsilon h}\vdots\nabla\tilde P\,\mathrm{d}x\,\mathrm{d}t=0 \qquad (4.16)$$

will be used here for $\tilde P:=P_{\varepsilon h}{-}\tilde P_h$ where $\tilde P_h\to P_\varepsilon$ strongly in $H^1(I;L^2(\Omega)^{d\times d})$. For some constant $c_{d,q}>0$, cf. [6, Lemma I.4.4], this allows for estimating as follows

$$\begin{aligned}&\lim_{h\to0} c_{d,q}\|\nabla P_{\varepsilon h}{-}\nabla P_\varepsilon\|^q_{L^q(Q)^{d\times d\times d}}\\ &\le \lim_{h\to0}\int_Q\big(|\nabla P_{\varepsilon h}|^{q-2}\nabla P_{\varepsilon h}-|\nabla P_\varepsilon|^{q-2}\nabla P_\varepsilon\big)\vdots\nabla\big(P_{\varepsilon h}-P_\varepsilon\big)\,\mathrm{d}x\,\mathrm{d}t\\ &=\lim_{h\to0}\frac{1}{\kappa_1}\int_Q \nabla y_{\varepsilon h}^\top\psi_{\rm E}'(F_{{\rm el},\varepsilon h}){:}(P_{\varepsilon h}^{-1})'{:}(P_{\varepsilon h}{-}\tilde P_h)\\ &\qquad +\partial_R\mathfrak{R}_\varepsilon(\theta_{\varepsilon h};\dot P_{\varepsilon h}P_{\varepsilon h}^{-1}){:}((P_{\varepsilon h}{-}\tilde P_h)P_{\varepsilon h}^{-1})+\psi_{\rm H}'(P_{\varepsilon h}){:}(P_{\varepsilon h}{-}\tilde P_h)\\ &\qquad +|\nabla P_{\varepsilon h}|^{q-2}\nabla P_{\varepsilon h}\vdots\nabla(\tilde P_h{-}P_\varepsilon)-|\nabla P_\varepsilon|^{q-2}\nabla P_\varepsilon\vdots\nabla\big(P_{\varepsilon h}-P_\varepsilon\big)\,\mathrm{d}x\,\mathrm{d}t=0, \end{aligned}\qquad (4.17)$$

where we used $P_{\varepsilon h}{-}\tilde P_h\to0$ strongly in $L^2(Q)^{d\times d}$ due to our estimates (4.8b) and classical Aubin-Lions compact-embedding theorem. Moreover, $\nabla y_{\varepsilon h}^\top$ is bounded in $L^\infty(I;L^{2^*}(\Omega)^{d\times d})$ and $\psi_{\rm E}'(F_{{\rm el},\varepsilon h})$ is bounded in $L^\infty(I;L^{2^*2/(2^*-2)}(\Omega)^{d\times d})$ due to the growth restriction (3.1a), so that $\nabla y_{\varepsilon h}^\top\psi_{\rm E}'(F_{{\rm el},\varepsilon h})$ is bounded in $L^2(Q)^{d\times d}$. This allows to pass to the limit in the term which contains $\nabla y_{\varepsilon h}^\top\psi_{\rm E}'(F_{{\rm el},\varepsilon h})$ and similarly also in the $\partial_R\mathfrak{R}_\varepsilon$-term, by taking into account that $\partial_R\mathfrak{R}_\varepsilon(\theta_{\varepsilon h};\dot P_{\varepsilon h}P_{\varepsilon h}^{-1})$ is bounded in $L^2(Q)^{d\times d}$. As for the last term, note that $\nabla P_{\varepsilon h}\to\nabla P_\varepsilon$ weakly in $L^q(Q)^{d\times d}$ while $|\nabla P_\varepsilon|^{q-2}\nabla P_\varepsilon\in L^{q'}(Q)^{d\times d}$ is fixed. Thus (4.15b) is proved. From this, we can also obtain the strong convergence of the $q$-Laplacian of $P_{\varepsilon h}$ in $L^{q'}(I;(W^{1,q}(\Omega)^{d\times d})^*)$, and thus, due to the bound (4.8e), also

$$\mathrm{div}(|\nabla P_{\varepsilon h}|^{q-2}\nabla P_{\varepsilon h})\to\mathrm{div}(|\nabla P_\varepsilon|^{q-2}\nabla P_\varepsilon)\quad\text{weakly in }\ L^2(Q)^{d\times d}. \qquad (4.18)$$

As for the strong convergence of $\dot P_{\varepsilon h}$ (or rather of $\dot P_{\varepsilon h}P_{\varepsilon h}^{-1}$ which occurs in the dissipation rate in (4.4c)), we use the strong monotonicity (4.2c) of $\partial_R\mathfrak{R}_\varepsilon(\theta;\cdot)$ and again (4.16) but now with the test function $\tilde P=\dot P_{\varepsilon h}-\dot{\tilde P}_h$. Taking $a_{\mathfrak{R}}>0$ from the uniform monotonicity assumption (3.1g), in view of (4.2b), we can estimate

$$
\begin{aligned}
&\limsup_{h\to 0}\frac{1}{2}a_{\mathfrak{R}}\big\|\dot P_{\varepsilon h}P_{\varepsilon h}^{-1}-\dot P_{\varepsilon}P_{\varepsilon}^{-1}\big\|^2_{L^2(Q)^{d\times d}}\\
&\quad\le\limsup_{h\to 0}\int_Q\big(\partial_R\mathfrak{R}_\varepsilon(\theta_{\varepsilon h};\dot P_{\varepsilon h}P_{\varepsilon h}^{-1})-\partial_R\mathfrak{R}_\varepsilon(\theta_{\varepsilon h};\dot P_{\varepsilon}P_{\varepsilon}^{-1})\big){:}(\dot P_{\varepsilon h}P_{\varepsilon h}^{-1}-\dot P_{\varepsilon}P_{\varepsilon}^{-1})\,\mathrm{d}x\,\mathrm{d}t\\
&\quad=\limsup_{h\to 0}\int_Q\partial_R\mathfrak{R}_\varepsilon(\theta_{\varepsilon h};\dot P_{\varepsilon h}P_{\varepsilon h}^{-1}){:}(\dot P_{\varepsilon h}P_{\varepsilon h}^{-1}-\dot{\tilde P}_{h}P_{\varepsilon h}^{-1})\,\mathrm{d}x\,\mathrm{d}t\\
&\qquad\quad+\lim_{h\to 0}\int_Q\partial_R\mathfrak{R}_\varepsilon(\theta_{\varepsilon h};\dot P_{\varepsilon h}P_{\varepsilon h}^{-1}){:}(\dot{\tilde P}_{h}P_{\varepsilon h}^{-1}-\dot P_{\varepsilon}P_{\varepsilon}^{-1})\,\mathrm{d}x\,\mathrm{d}t\\
&\qquad\qquad-\lim_{h\to 0}\int_Q\partial_R\mathfrak{R}_\varepsilon(\theta_{\varepsilon h};\dot P_{\varepsilon}P_{\varepsilon}^{-1}){:}(\dot P_{\varepsilon h}P_{\varepsilon h}^{-1}-\dot P_{\varepsilon}P_{\varepsilon}^{-1})\,\mathrm{d}x\mathrm{d}t\\
&\quad=\lim_{h\to 0}\int_Q\nabla y_{\varepsilon h}^\top\psi_{\rm E}'(F_{{\rm el},\varepsilon h}){:}(P_{\varepsilon h}^{-1})'{:}(\dot{\tilde P}_h-\dot P_{\varepsilon h})+\psi_{\rm H}'(P_{\varepsilon h}){:}(\dot{\tilde P}_h-\dot P_{\varepsilon h})\,\mathrm{d}x\mathrm{d}t\\
&\qquad\quad+\limsup_{h\to 0}\int_Q\kappa_1|\nabla P_{\varepsilon h}|^{q-2}\nabla P_{\varepsilon h}\,\vdots\,\nabla(\dot{\tilde P}_h-\dot P_{\varepsilon h})\,\mathrm{d}x\mathrm{d}t\\
&\quad=\lim_{h\to 0}\int_Q\nabla y_{\varepsilon h}^\top\psi_{\rm E}'(F_{{\rm el},\varepsilon h}){:}(P_{\varepsilon h}^{-1})'{:}(\dot{\tilde P}_h-\dot P_{\varepsilon h})+\psi_{\rm H}'(P_{\varepsilon h}){:}(\dot{\tilde P}_h-\dot P_{\varepsilon h})\\
&\qquad\quad+\kappa_1{\rm div}(|\nabla P_{\varepsilon h}|^{q-2}\nabla P_{\varepsilon h}){:}\dot{\tilde P}_h\,\mathrm{d}x\mathrm{d}t+\limsup_{h\to 0}\int_\Omega\frac{\kappa_1}{q}|\nabla P_0|^q-\frac{\kappa_1}{q}|\nabla P_{\varepsilon h}(T)|^q\,\mathrm{d}x\\
&\quad\le-\int_Q\kappa_1{\rm div}(|\nabla P_\varepsilon|^{q-2}\nabla P_\varepsilon){:}\dot P_\varepsilon\,\mathrm{d}x\,\mathrm{d}t+\int_\Omega\frac{\kappa_1}{q}|\nabla P_0|^q-\frac{\kappa_1}{q}|\nabla P_\varepsilon(T)|^q\,\mathrm{d}x=0,
\end{aligned}
\tag{4.19}
$$

where we used that $\nabla\dot P_{\varepsilon h}$ is well defined at the level of Galerkin approximations (although not in the limit) and we also used the fact that

$$
\begin{aligned}
&\liminf_{h\to 0}\int_Q|\nabla P_{\varepsilon h}|^{q-2}\nabla P_{\varepsilon h}\,\vdots\,\nabla(\dot P_{\varepsilon h}-\dot{\tilde P}_h)\,\mathrm{d}x\,\mathrm{d}t\\
&\quad=\liminf_{h\to 0}\int_\Omega\frac{1}{q}|\nabla P_{\varepsilon h}(T)|^q-\frac{1}{q}|\nabla P_0|^q\,\mathrm{d}x+\lim_{h\to 0}\int_Q{\rm div}(|\nabla P_{\varepsilon h}|^{q-2}\nabla P_{\varepsilon h}){:}\dot{\tilde P}_h\,\mathrm{d}x\,\mathrm{d}t\\
&\quad\ge\int_\Omega\frac{1}{q}|\nabla P_\varepsilon(T)|^q-\frac{1}{q}|\nabla P_0|^q\,\mathrm{d}x+\int_Q{\rm div}(|\nabla P_\varepsilon|^{q-2}\nabla P_\varepsilon){:}\dot P_\varepsilon\,\mathrm{d}x\,\mathrm{d}t
\end{aligned}
\tag{4.20}
$$

as well as the Green formula combined with the by-part integration in time:

$$
\int_Q{\rm div}(|\nabla P_\varepsilon|^{q-2}\nabla P_\varepsilon){:}\dot P_\varepsilon\,\mathrm{d}x\,\mathrm{d}t=\int_\Omega\frac{1}{q}|\nabla P_0|^q-\frac{1}{q}|\nabla P_\varepsilon(T)|^q\,\mathrm{d}x.
\tag{4.21}
$$

Note that the last term above makes sense as $t\mapsto\nabla P_\varepsilon(t)$ is weakly continuous to $L^q(\Omega)^{d\times d}$ due to (4.8b). The other integrals in (4.20) are also well-defined due to estimate (4.8e). Note also that we used (4.18) here. Moreover, it is possible to show that $-{\rm div}(|\nabla P|^{q-2}\nabla P)$ is indeed the subdifferential of the potential $P\mapsto\int_\Omega\frac{1}{q}|\nabla P|^q\,\mathrm{d}x$ in $L^2(\Omega)^{d\times d}$. In particular, the chain rule in (4.21) holds true, see [4]. This proves convergence (4.15a).

The convergence of the mechanical part for $h\to 0$ is now straightforward. As the highest-order term in (4.4a) is linear, weak convergence together and Aubin-Lions compactness for lower-order terms suffices. The limit passage in the quasilinear $q$-Laplacian as well as in the $\mathfrak{R}_\varepsilon$-term in (4.4b) follow from the already proved strong convergences (4.15).

Eventually, the limit passage in the semilinear heat-transfer equation (4.4c) can be ascertained due to the already proved strong convergences (4.15a,b), allowing indeed the passage to the limit in the (regularized) right-hand side. □

In order to remove the regularization by passing to the limit for $\varepsilon \to 0$, we cannot directly rely on the estimates (4.8c)-(4.8d), which are dependent on $\varepsilon > 0$. On the other hand, having already passed to the limit in $h$ we are now in the position of performing a number of nonlinear tests for the heat equation, which are specifically tailored to the $L^1$-theory.

**Lemma 4.3** (Further a-priori estimates for temperature)**.** *Let $\vartheta_\varepsilon$ be the (rescaled) temperature component of the weak solution to the regularized problem* (4.4)*, whose existence is proved in Proposition* 4.2*. Then,*

$$\vartheta_\varepsilon \geq 0 \quad a.e. \ in \ Q. \tag{4.22}$$

*Moreover, one has that*

$$\exists C_1 > 0: \quad \|\vartheta_\varepsilon\|_{L^\infty(I;L^1(\Omega))} \leq C_1, \tag{4.23a}$$
$$\forall 1 \leq s < (d{+}2)/(d{+}1) \ \exists C_s > 0: \quad \|\nabla\vartheta_\varepsilon\|_{L^s(Q)^d} \leq C_s \tag{4.23b}$$

*where the constants $C_1$, $C_s$ are independent of $\varepsilon$.*

*Proof.* The nonnegativity (4.22) is readily obtained by testing the heat equation in the regularized enthalpy form (4.4c) by $\min(0,\vartheta_\varepsilon)$, and using the assumptions that $\theta_\flat \geq 0$ and $\vartheta_0 \geq 0$, cf. (3.1k,l). Note that the normalization $C_{\rm v}(0) = 0$ for the primitive function $C_{\rm v}$ of $c_{\rm v}$ in (4.3) is here used. The nonnegativity (4.22) allows us to read the bound (4.23a) from the test of the heat equation by 1.

We now perform a second *nonlinear* test in order to gain an estimate on $\nabla\vartheta$ independent of $\varepsilon$. We follow the classical [3] in the simplified variant developed in [8]. This calls for testing the heat equation (4.4c) on $\chi_\omega(\vartheta_\varepsilon)$ where the increasing concave function $\chi_\omega : [0,+\infty) \to [0,1]$ is defined as $\chi_\omega(w) := 1 - 1/(1{+}w)^\omega$ for some small $\omega > 0$ to be chosen later. By using the nonnegativity of $C_{\rm v}$ and the fact that $0 \leq \chi_\omega(\vartheta_\varepsilon) \leq 1$, $\chi_\omega'(w) = \omega/(1{+}w)^{1+\omega}$, $r_\varepsilon \leq r$, and $\theta_\flat/(1{+}\varepsilon\theta_\flat) \leq \theta_\flat$, we obtain the estimate

$$\begin{aligned}\omega\int_Q \frac{a_{\mathbb K}}{(1{+}\vartheta_\varepsilon)^{1+\omega}}\Big|\frac{{\rm Cof}\,P_\varepsilon}{\sqrt{\det P_\varepsilon}}\nabla\vartheta_\varepsilon\Big|^2\,{\rm d}x\,{\rm d}t &\leq \int_Q \mathscr{K}(P_\varepsilon,\vartheta_\varepsilon)\nabla\vartheta_\varepsilon\cdot\nabla\chi_\omega(\vartheta_\varepsilon)\,{\rm d}x\,{\rm d}t\\
&\leq \int_Q \mathscr{K}(P_\varepsilon,\vartheta_\varepsilon)\nabla\vartheta_\varepsilon\cdot\nabla\chi_\omega(\vartheta_\varepsilon)\,{\rm d}x\,{\rm d}t + \int_\Sigma K\vartheta_\varepsilon\chi_\omega(\vartheta_\varepsilon)\,{\rm d}S\,{\rm d}t\\
&\leq \int_Q r_\varepsilon\,{\rm d}x\,{\rm d}t + \int_\Sigma K\frac{\theta_\flat}{1{+}\varepsilon\theta_\flat}{\rm d}S\,{\rm d}t + \int_\Omega C_{\rm v}(\vartheta_{0,\varepsilon})\,{\rm d}x\\
&\leq \int_Q r\,{\rm d}x\,{\rm d}t + \int_\Sigma K\theta_\flat{\rm d}S\,{\rm d}t + \int_\Omega C_{\rm v}(\theta_0)\,{\rm d}x,\end{aligned} \tag{4.24}$$

where $a_{\mathbb K} > 0$ stands for the positive-definiteness constant of $\mathbb K$, cf. (3.1c). By the Hölder inequality, we have that

$$\begin{aligned}\int_Q \Big|\frac{{\rm Cof}\,P_\varepsilon}{\sqrt{\det P_\varepsilon}}\nabla\vartheta_\varepsilon\Big|^s\,{\rm d}x\,{\rm d}t &= \int_Q (1{+}\vartheta_\varepsilon)^{(1+\omega)s/2}\frac{|({\rm Cof}\,P_\varepsilon)\nabla\vartheta_\varepsilon|^s}{(\det P_\varepsilon)^{s/2}(1{+}\vartheta_\varepsilon)^{(1+\omega)s/2}}\,{\rm d}x\,{\rm d}t\\
&\leq \Big(\int_Q (1{+}\vartheta_\varepsilon)^{(1+\omega)s/(2-s)}\,{\rm d}x\,{\rm d}t\Big)^{1-s/2}\Big(\int_Q \frac{|({\rm Cof}\,P_\varepsilon)\nabla\vartheta_\varepsilon|^2}{\det P_\varepsilon(1{+}\vartheta_\varepsilon)^{1+\omega}}\,{\rm d}x\,{\rm d}t\Big)^{s/2}\end{aligned}$$

and the last factor on the right-hand side is bounded due to (4.24). We now use the Gagliardo-Nirenberg inequality

$$\|v\|_{L^{(1+\omega)s/(2-s)}(\Omega)} \leq C_{\rm GN} \|v\|^{1-\lambda}_{L^1(\Omega)} \|v\|^{\lambda}_{W^{1,s}(\Omega)} \tag{4.25}$$

with $\|v\|_{W^{1,s}(\Omega)} := \|v\|_{L^1(\Omega)} + \|\nabla v\|_{L^s(\Omega)^d}$. The latter inequality holds for all $\lambda \in (0,1)$ such that

$$\frac{2-s}{(1+\omega)s} \geq \lambda \left( \frac{1}{s} - \frac{1}{d} \right) + 1 - \lambda \tag{4.26}$$

and, correspondingly, for some $C_{\rm GN} > 0$ depending on $s$, $\lambda$, and $\Omega$. We shall apply inequality (4.25) along with the choices $v = 1 + \vartheta_\varepsilon(t,\cdot)$ and $\lambda = (2-s)/(1+\omega)$. Note that $\lambda \in (0,1)$ as $s \in [1,(d+2)/(d+1)) \subset [1,2)$. Moreover, by letting $\omega > 0$ small enough one has that condition (4.26) holds, again by virtue of $s < (d+2)/(d+1)$. Hence, inequality (4.25) gives

$$\begin{aligned} &\left( \int_0^T \big\|1 + \vartheta_\varepsilon(t,\cdot)\big\|^{(1+\omega)s/(2-s)}_{L^{(1+\omega)s/(2-s)}(\Omega)} \,\mathrm{d}t \right)^{1-s/2} \\ &\leq \left( \int_0^T C_{\rm GN}^{(1+\omega)s/(2-s)} C_0^{(1-\lambda)(1+\omega)s/(2-s)} \Big( C_0 + \big\|\nabla\vartheta_\varepsilon(t,\cdot)\big\|_{L^s(\Omega)^d} \Big)^{\lambda(1+\omega)s/(2-s)} \,\mathrm{d}t \right)^{1-s/2} \\ &\leq C_\delta + \delta \int_Q \big|\nabla\vartheta_\varepsilon\big|^s \,\mathrm{d}x\,\mathrm{d}t \end{aligned} \tag{4.27}$$

where $C_0 = |\Omega| + C_1$ and $C_1$ is from (4.23a). The constant $C_\delta$ depends on $C_0$, $C_{\rm GN}$, and the small $\delta$, cf. e.g. [40, Formula (12.20)], and we used the fact that

$$\lambda(1+\omega)s/(2-s)(1-s/2) = s(1-s/2) < s.$$

We combine this estimate with the bound

$$\begin{aligned} \big\|\nabla\vartheta_\varepsilon\big\|^s_{L^s(Q)^d} &= \Big\| \frac{P_\varepsilon^\top(\mathrm{Cof}\,P_\varepsilon)}{\det P_\varepsilon} \nabla\vartheta_\varepsilon \Big\|^s_{L^s(Q)^d} \\ &\leq \Big\| \frac{P_\varepsilon}{\sqrt{\det P_\varepsilon}} \Big\|^s_{L^\infty(Q)^{d\times d}} \Big\| \frac{\mathrm{Cof}\,P_\varepsilon}{\sqrt{\det P_\varepsilon}} \nabla\vartheta_\varepsilon \Big\|^s_{L^s(Q)^{d\times d}}. \end{aligned} \tag{4.28}$$

Note that the last term in the right-hand side is what occurs in the left-hand side of (4.24). By choosing $\delta$ small enough, we deduce from (4.24), (4.27), and (4.28) that

$$\Big\| \frac{\mathrm{Cof}\,P_\varepsilon}{\sqrt{\det P_\varepsilon}} \nabla\vartheta_\varepsilon \Big\|_{L^s(Q)^{d\times d}} \leq \tilde{C}_s \tag{4.29}$$

for any $1 \leq s < (d+2)/(d+1)$ and some positive $\tilde{C}_s$. From the latter bound we directly deduce estimate (4.23b) by using again (4.28). □

**Proposition 4.4** (Convergence of the regularization for $\varepsilon \to 0$)**.** *Under assumptions* (3.1)*, as $\varepsilon \to 0$ there exists a subsequence of $\{(y_\varepsilon, P_\varepsilon, \vartheta_\varepsilon)\}_{\varepsilon>0}$ (not relabelled) which converges weakly* in the topologies indicated in* (4.8a-f)*,* (4.14)*, and* (4.23) *to some $(y, P, \vartheta)$. Every such a limit triple is a weak solution to the original problem in the sense of Definition* 3.1*. Moreover, the following strong convergences hold*

$$\dot{P}_\varepsilon P_\varepsilon^{-1} \to \dot{P} P^{-1} \qquad \textit{strongly in} \ \ L^2(Q)^{d\times d}, \tag{4.30a}$$

$$\nabla P_\varepsilon \to \nabla P \qquad \textit{strongly in} \ \ L^q(Q)^{d\times d\times d}. \tag{4.30b}$$

*Eventually, the regularity* (3.2d) *and the energy conservation* (3.4) *hold.*

*Proof.* Again, by the Banach selection principle, we can extract a not relabelled subsequence converging with respect to the topologies from the estimates (4.8a,b) inherited for $(y_\varepsilon, P_\varepsilon)$, (4.14), and (4.23), and indicate its limit by $(y, P, \vartheta)$.

The improved, strong convergences (4.30) can be obtained by arguing as in the proof of (4.15) in Proposition 4.2.

One has just to modify the argument in (4.17) as:

$$\begin{aligned}\lim_{\varepsilon\to0} c_{d,q}\|\nabla P_\varepsilon-\nabla P\|^q_{L^q(Q)^{d\times d\times d}} &\le \lim_{\varepsilon\to0}\int_Q \big(|\nabla P_\varepsilon|^{q-2}\nabla P_\varepsilon - |\nabla P|^{q-2}\nabla P\big)\vdots\nabla\big(P_\varepsilon - P\big)\,\mathrm{d}x\,\mathrm{d}t\\ = \lim_{\varepsilon\to0}\frac{1}{\kappa_1}\int_Q \nabla y_\varepsilon^\top\psi'_{\mathrm{E}}(F_{\mathrm{el},\varepsilon}){:}(P_\varepsilon^{-1})'{:}(P_\varepsilon-P) &+ \partial_R\mathfrak{R}_\varepsilon(\theta_\varepsilon;\dot P_\varepsilon P_\varepsilon^{-1}){:}((P_\varepsilon-P)P_\varepsilon^{-1})\,\mathrm{d}x\,\mathrm{d}t\\ &- \int_Q |\nabla P|^{q-2}\nabla P\big)\vdots\nabla\big(P_\varepsilon - P\big)\,\mathrm{d}x\,\mathrm{d}t = 0. \end{aligned}\tag{4.31}$$

Here we used that the sequence $\partial_R\mathfrak{R}_\varepsilon(\theta_\varepsilon;\dot P_\varepsilon P_\varepsilon^{-1})$ is bounded in $L^2(Q)^{d\times d}$ (without caring about its limit) while $(P_\varepsilon-P)P_\varepsilon^{-1}\to0$ strongly in $L^2(Q)^{d\times d}$ (or even in $L^\infty(Q)^{d\times d}$, cf. the arguments used already for (4.17)). Thus (4.30b) is proved.

For (4.30a), one has just to modify the argument in (4.19), for the term $\nabla\dot P_\varepsilon$ is not well defined. Relying on the fact that $\mathrm{div}(\kappa_1|\nabla P_\varepsilon|^{q-2}\nabla P_\varepsilon)\in L^2(Q)^{d\times d}$, we have

$$\begin{aligned}&\limsup_{\varepsilon\to0}\frac12 a_{\mathfrak{R}}\big\|\dot P_\varepsilon P_\varepsilon^{-1} - \dot P P^{-1}\big\|^2_{L^2(Q)^{d\times d}}\\ &\le \limsup_{\varepsilon\to0}\int_Q\big(\partial_R\mathfrak{R}_\varepsilon(\theta_\varepsilon;\dot P_\varepsilon P_\varepsilon^{-1}) - \partial_R\mathfrak{R}_\varepsilon(\theta;\dot P P^{-1})\big){:}\big(\dot P_\varepsilon P_\varepsilon^{-1} - \dot P P^{-1}\big)\,\mathrm{d}x\mathrm{d}t\\ &= \limsup_{\varepsilon\to0}\int_Q \partial_R\mathfrak{R}_\varepsilon(\theta_\varepsilon;\dot P_\varepsilon P_\varepsilon^{-1}){:}(\dot P_\varepsilon-\dot P)P_\varepsilon^{-1}\,\mathrm{d}x\,\mathrm{d}t\\ &\qquad+\lim_{\varepsilon\to0}\int_Q \partial_R\mathfrak{R}_\varepsilon(\theta_\varepsilon;\dot P_\varepsilon P_\varepsilon^{-1}){:}\dot P(P_\varepsilon^{-1}-P^{-1})\,\mathrm{d}x\,\mathrm{d}t\\ &\qquad\quad-\lim_{\varepsilon\to0}\int_Q \partial_R\mathfrak{R}_\varepsilon(\theta_\varepsilon;\dot P P^{-1}){:}(\dot P_\varepsilon P_\varepsilon^{-1} - \dot P P^{-1}))\,\mathrm{d}x\mathrm{d}t\\ &\overset{\text{(a)}}{=}\lim_{\varepsilon\to0}\int_Q \nabla y_\varepsilon^\top\psi'_{\mathrm{E}}\big(\nabla y_\varepsilon P_\varepsilon^{-1}\big){:}(P_\varepsilon^{-1})'{:}(\dot P_\varepsilon-\dot P) + \psi'_{\mathrm{H}}(P_\varepsilon){:}(\dot P_\varepsilon-\dot P)\,\mathrm{d}x\mathrm{d}t\\ &\qquad-\liminf_{\varepsilon\to0}\int_Q \mathrm{div}\big(\kappa_1|\nabla P_\varepsilon|^{q-2}\nabla P_\varepsilon\big){:}(\dot P_\varepsilon-\dot P)\,\mathrm{d}x\,\mathrm{d}t\\ &\qquad\quad-\lim_{\varepsilon\to0}\int_Q \partial_R\mathfrak{R}_\varepsilon(\theta_\varepsilon;\dot P P^{-1}){:}(\dot P_\varepsilon P_\varepsilon^{-1} - \dot P P^{-1}))\,\mathrm{d}x\mathrm{d}t\\ &\overset{\text{(b)}}{=}\lim_{\varepsilon\to0}\int_Q \nabla y_\varepsilon^\top\psi'_{\mathrm{E}}\big(\nabla y_\varepsilon P_\varepsilon^{-1}\big){:}(P_\varepsilon^{-1})'{:}(\dot P_\varepsilon-\dot P) + \psi'_{\mathrm{H}}(P_\varepsilon){:}(\dot P_\varepsilon-\dot P)\\ &\qquad+\mathrm{div}(\kappa_1|\nabla P_\varepsilon|^{q-2}\nabla P_\varepsilon){:}\dot P\,\mathrm{d}x\mathrm{d}t + \limsup_{\varepsilon\to0}\int_\Omega\frac{\kappa_1}{q}|\nabla P_0|^q - \frac{\kappa_1}{q}|\nabla P_\varepsilon(T)|^q\,\mathrm{d}x\\ &\le \int_\Omega\frac{\kappa_1}{q}|\nabla P_0|^q - \frac{\kappa_1}{q}|\nabla P(T)|^q\,\mathrm{d}x - \int_Q \mathrm{div}(\kappa_1|\nabla P|^{q-2}\nabla P){:}\dot P\,\mathrm{d}x\mathrm{d}t = 0.\end{aligned}$$

In addition to the arguments analogous to (4.20)-(4.21), for equality (a) we have used the fact that

$$\Big|\int_Q \partial_R\mathfrak{R}_\varepsilon(\theta_\varepsilon;\dot P_\varepsilon P_\varepsilon^{-1}){:}\dot P(P_\varepsilon^{-1}{-}P^{-1})\,\mathrm{d}x\,\mathrm{d}t\Big|$$
$$\le \|\partial_R\mathfrak{R}_\varepsilon(\theta_\varepsilon;\dot P_\varepsilon P_\varepsilon^{-1})\|_{L^2(Q)^{d\times d}}\|\dot P\|_{L^2(Q)^{d\times d}}\|P_\varepsilon^{-1}{-}P^{-1}\|_{L^\infty(Q)^{d\times d}}\to 0. \tag{4.32}$$

This follows as $P_\varepsilon^{-1}\to P^{-1}$ strongly in $L^\infty(Q)^{d\times d}$ due to our estimates (4.8b) inherited for $P_\varepsilon$, as used already for (4.17). Moreover, for equality (b) we also used that

$$\lim_{\varepsilon\to0}\int_Q \partial_R\mathfrak{R}_\varepsilon\big(\theta_\varepsilon;\dot PP^{-1}\big){:}(\dot P_\varepsilon P_\varepsilon^{-1}-\dot PP^{-1})\,\mathrm{d}x\mathrm{d}t$$
$$=\lim_{\varepsilon\to0}\int_Q \partial_R\mathfrak{R}_{1,\varepsilon}\big(\theta_\varepsilon;\dot PP^{-1}\big){:}(\dot P_\varepsilon P_\varepsilon^{-1}-\dot PP^{-1})\,\mathrm{d}x\,\mathrm{d}t$$
$$+\lim_{\varepsilon\to0}\int_Q \partial_R\mathfrak{R}_2\big(\theta_\varepsilon;\dot PP^{-1}\big){:}(\dot P_\varepsilon P_\varepsilon^{-1}-\dot PP^{-1})\,\mathrm{d}x\,\mathrm{d}t=0. \tag{4.33}$$

The latter follows from (4.2d,e) and the fact that $\theta_\varepsilon\to\theta$ strongly in $L^1(Q)$ hence $\sigma_{\rm Y}(\theta_\varepsilon)\to\sigma_{\rm Y}(\theta)$ a.e. Indeed, we have that $\partial_R\mathfrak{R}_{1,\varepsilon}(\theta_\varepsilon;\dot PP^{-1})$ converges a.e. on $Q$ either to $\partial_R\mathfrak{R}(\theta;\dot PP^{-1})$ if $\dot PP^{-1}(t,x)\ne0$ or to 0 otherwise. As the sequence $\partial_R\mathfrak{R}_{1,\varepsilon}(\theta_\varepsilon;\dot PP^{-1})$ is bounded in $L^\infty(Q)^{d\times d}$, the Vitali theorem ensures that it converges strongly in $L^r(Q)^{d\times d}$ for all $r<\infty$. On the other hand, $\partial_R\mathfrak{R}_2(\theta_\varepsilon;\dot PP^{-1})\to\partial_R\mathfrak{R}_2(\theta;\dot PP^{-1})$ strongly in $L^2(Q)^{d\times d}$ just by the usual continuity of the underlying Nemytskiĭ mapping. Since we have the weak convergence $\dot P_\varepsilon P_\varepsilon^{-1}\to\dot PP^{-1}$ in $L^2(Q)^{d\times d}$, convergence (4.33) follows.

The passage to the limit then follows similarly as in the proof of Proposition 4.2. A little difference concerns the strong convergence of $\vartheta_\varepsilon$, which follows again by the Aubin-Lions Theorem but we use here a coarser topology than in Proposition 4.2. Namely, the convergence holds in $L^p(Q)$ with arbitrary $1\le p<1+2/d$, related to the estimates (4.23) when interpolated. This change is however immaterial with respect to the limit passage in the mechanical part (2.7a,b). Actually, some arguments are even simplified, for we do not need to approximate the limit into the finite-dimensional subspaces as we did in (4.19). The heat-production rate on the right-hand side of (4.4c) converges now strongly in $L^1(Q)$.

Eventually, regularity (3.2d) can be obtained from the estimates (4.14), which are uniform in $\varepsilon>0$. The energy conservation (3.4) follows directly from the energy conservation in the mechanical part, as essentially used above while checking the strong convergences (4.30). Indeed, one integrates (2.17) over $[0,t]$ and sums it to the heat equation tested on the constant 1. Note that this is amenable as the constant 1 can be put in duality with $\dot\vartheta$, so that the chain-rule applies. □

**Remark 4.5** (*Boundary conditions on $P$*)**.** We have assumed here the homogeneous Dirichlet condition $P=\mathbb{I}$ on $\Sigma$ for the sake of simplicity. One has however to mention that other boundary conditions could be considered. In particular, this could be done if the hardening $\psi_{\rm H}$ were coercive on the whole plastic tensor $P$, otherwise only at the expense of some additional intricacies. Indeed, a bound on $P_{\varepsilon h}$ could be obtained from that on $\dot P_{\varepsilon h}P_{\varepsilon h}^{-1}$ by suitably exploiting the coercivity of the elastic energy. This would however require to strengthen the corresponding growth assumptions.

## Acknowledgements

The authors acknowledge the hospitality and the support of the Erwin Schrödinger Institute of the University of Vienna, where most of this research has been performed, as well as the partial support from Austrian-Czech projects 16-34894L (FWF/CSF) and 7AMB16AT015 (FWF/MSMT CR). T.R acknowledges also the support of Czech Science Foundation (CSF) project 16-03823S and 17-04301S as well as through the institutional support RVO: 61388998 (ČR). U.S. is partially supported by the Austrian Science Fund (FWF) projects F 65, P 27052, and I 2375 and by the Vienna Science and Technology Fund (WWTF) project MA14-009. The authors are gratefully indebted to the anaonymous referee for the careful reading of the manuscript and many valuable suggestions.

## References

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

[2] Z. P. Bažant and M. Jirásek. *Creep and Hygrothermal Effects in Concrete Structures.* Solid Mechanics and Its Applications, vol 225. Springer, 2018. (2018)

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

[4] H. Brézis. *Operateur Maximaux Monotones et Semi-groupes de Contractions dans les Espaces de Hilbert.* North-Holland, Amsterdam, 1973.

[5] P. G. Ciarlet and J. Nečas. Injectivity and self-contact in nonlinear elasticity. *Arch. Ration. Mech. Anal.*, 97:171–188, 1987.

[6] E. DiBenedetto. *Degenerate parabolic equations*, Universitext, Springer-Verlag, New York, 1993.

[7] F. P. Duda, A. C. Souza, and E. Fried. A theory for species migration in a finitely strained solid with application to polymer network swelling. *J. Mech. Phys. Solids*, 58:515–529, 2010.

[8] E. Feireisl and J. Málek. On the Navier-Stokes equations with temperature-dependent transport coefficients. *Diff. Equations Nonlin. Mech.*, pages 14pp.(electronic), Art.ID 90616, 2006.

[9] N. A. Fleck and J. W. Hutchinson. Strain gradient plasticity. *Adv. Appl. Mech.*, 33:95–361, 1997.

[10] N. A. Fleck and J. W. Hutchinson. A reformulation of strain gradient plasticity. *J. Mech. Phys. Solids.* 49:2245–2271, 2001.

[11] E. Fried and M. E. Gurtin. Tractions, balances, and boundary conditions for nonsimple materials with application to liquid flow at small-lenght scales. *Arch. Ration. Mech. Anal.*, 182:513–554, 2006.

[12] P. Germain, Q. Nguyen, and P. Suquet. Continuum thermodynamics. *J. Applied Mechanics (Trans. ASME)*, 50:1010–102, 1983.

[13] A. Giacomini and A. Musesti. Quasi-static evolutions in linear perfect plasticity as a variational limit of finite plasticity: a one-dimensional case. *Math. Models Methods Appl. Sci.*, 23:1275–1308, 2013. .

[14] S. Govindjee and J. C. Simo. Coupled stress-diffusion: case II. *J. Mech. Phys. Solids*, 41:863–887, 1993.

[15] D. Grandi and U. Stefanelli. Finite plasticity in $P^{\top}P$. Part II: quasistatic evolution and linearization. *SIAM J. Math. Anal.*, 49:1356–1384, 2017.

[16] B. Halphen and Q. S. Nguyen. Sue les matériaux standards généralisés. *J. Méchanique*, 14:39–63, 1975.

[17] T. J. Healey and S. Krömer. Injective weak solutions in second-gradient nonlinear elasticity. *ESAIM: Control, Optim. & Cal. Var.*, 15:863–871, 2009.

[18] M. Jirásek and Z. Bažant. *Inelastic Analysis of Structures*. J. Wiley, Chichester, 2002.

[19] E. Kröner. Allgemeine Kontinuumstheorie der Versetzungen und Eigenspannungen. *Arch. Ration. Mech. Anal.*, 4:273–334, 1960.

[20] M. Kružík and T. Roubíček. *Mathematical Methods in Continuum Mechanics of Solids*. IMM Series. Springer, Cham/Heidelberg, to appear 2018.

[21] E. Lee and D. Liu. Finite-strain elastic-plastic theory with application to plain-wave analysis. *J. Applied Phys.*, 38:19–27, 1967.

[22] V. Lyakhovsky, Y. Hamiel, and Y. Ben-Zion. A non-local visco-elastic damage model and dynamic fracturing. *J. Mech. Phys. Solids*, 59:1752–1776, 2011.

[23] A. Mainik and A. Mielke. Global existence for rate-independent gradient plasticity at finite strain. *J. Nonlinear Sci.*, 19:221–248, 2009.

[24] J. Mandel. *Plasticité classique et viscoplasticité*, volume 97 of *CISM Courses and Lectures*. Springer, Berlin, 1972.

[25] G. A. Maugin. *The Thermomechanics of Plasticity and Fracture*. Cambridge University Press, Cambridge, 1992.

[26] A. Mielke. Finite elastoplasticity, Lie groups and geodesics on SL($d$). In P. Newton, A. Weinstein, and P. J. Holmes, editors, *Geometry, Mechanics, and Dynamics*, pages 61–90. Springer–Verlag, New York, 2002.

[27] A. Mielke. Energetic formulation of multiplicative elasto-plasticity using dissipation distances. *Contin. Mech. Thermodyn.*, 15:351–382, 2003.

[28] A. Mielke. A new approach to elasto-plasticity using energy and dissipation functionals. In J.M. Hill and R. Moore, editors, *Applied Mathematics Entering the 21st Century: Invited Talks from the ICIAM 2003 Congress*, pages 315–335. SIAM, 2004.

[29] A. Mielke and S. Müller. Lower semicontinuity and existence of minimizers for a functional in elastoplasticity. *Z. Angew. Math. Mech.*, 86:233–250, 2006.

[30] A. Mielke, R. Rossi, and G. Savaré. Global existence results for viscoplasticity at finite strain. *Arch. Ration. Mech. Anal.*, 227: 423–475, 2018.

[31] A. Mielke and T. Roubíček. Thermoviscoelasticity in Kelvin-Voigt rheology at large strains. In preparation.

[32] A. Mielke and T. Roubíček. *Rate-Independent Systems – Theory and Application.* Springer, New York, 2015.

[33] A. Mielke and T. Roubíček. Rate-independent elastoplasticity at finite strains and its numerical approximation. *Math. Models Meth. Appl. Sci.*, 26:2203–2236, 2016.

[34] A. Mielke and U. Stefanelli. Linearized plasticity is the evolutionary Γ-limit of finite plasticity. *J. Europ. Math. Soc. (JEMS)*, 15:923–948, 2013.

[35] R.D. Mindlin and N.N. Eshel. On first strain-gradient theories in linear elasticity. *Intl. J. Solid Struct.*, 4:109–124, 1968.

[36] H.-B. Mühlhaus and E. Aifantis. A variational principle for gradient plasticity. *Int. J. Solids Struct.*, 28:845–857, 1991.

[37] J. Plešek and J. Korouš. Explicit integration method with time step control for viscoplasticity and creep. *Advances in Engineering Software*, 33, 2002.

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

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

[40] T. Roubíček. *Nonlinear Partial Differential Equations with Applications.* Birkhäuser, Basel, 2nd edition, 2013.

[41] T. Roubíček and G. Tomassetti. Thermodynamics of magneto- and poro-elastic materials under diffusion at large strains. (Preprint arXiv no.1703.06267). *Zeit. Angew. Math. Phys. – in print, DOI 10.1007/s00033-018-0932-y.*

[42] S. Roy and J. N. Reddy. *Computational Modeling of Polymer Composites – A Study of Creep and Environmental Effects.* CRC Press, Boca Raton, 2014.

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

[44] J.C. Simo and T. J. R. Hughes. *Computational Inelasticity.* Springer, New York, 1998.

[45] U. Stefanelli. Existence for dislocation-free finite plasticity. *ESAIM: Control, Optim. & Cal. Var.*, to appear, 2018.

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

[47] N. Triantafyllidis and E.C. Aifantis. A gradient approach to localization of deformation. I. Hyperelastic materials. *J. Elast.*, 16:225–237, 1986.