QUARTERLY OF APPLIED MATHEMATICS
VOLUME LII, NUMBER 4
DECEMBER 1994, PAGES 721-737

THERMOMECHANICAL EVOLUTION OF A MICROSTRUCTURE

By

KARL-HEINZ HOFFMANN (*Technische Universität München, Germany*)

AND

TOMÁŠ ROUBÍČEK (*Academy of Sciences, Prague, Czech Republic*)

Abstract. A nonisothermal microstructure evolution model, governed by a Helmholtz free energy which need not be convex as a function of deformations, is formulated by using a convexified geometry proposed already in [13]. A multidimensional but scalar case is treated. It is shown that, as a special case, this model includes the usual nonlinear thermo-visco-elasticity. In the case of an actual appearance of a microstructure, the existence of a weak solution to a partial linearized model is shown by a semi-implicit time discretization.

0. Introduction and notation. The aim of this paper is to extend the isothermal evolution model of a microstructure, proposed in [13], for the case of nonisothermal processes. There are several alternative models already proposed in the literature, namely, the phase-field type model by Frémond [6] (cf. also Colli, Frémond, and Visintin [2] or Hoffmann, Niezgódka, and Zheng Songmu [7]) and the Landau-Ginzburg-Devonshire type model by Falk [5] (cf. also Alt, Hoffmann, Niezgódka, and Sprekels [1], Niezgódka and Sprekels [12], or Sprekels [7]). The model proposed here might be considered "philosophically" as a certain combination of the models by Frémond and by Falk because it involves, on the one hand, a single free energy for all phases like Falk's model and, on the other hand, a microstructure which can "macroscopically" describe local portions of particular phases in a mixture very much like the parameters β in Frémond's model. The microstructure will be modelled here by the (suitably generalized) Young measures.

First, let us introduce briefly some less standard notation needed for the model from [13]. Unfortunately, we will have to confine ourselves to the scalar case because the existence of a nontrivial convex compactification (on which our theory is essentially based) is an open problem in vectorial multidimensional problems. The model from [13] involves as a state at a current time a couple $q = (u, m) \in Q$ where $u \in H_0^1(\Omega)$ is a displacement and $m \in \mathrm{Car}(\Omega; \mathbb{R}^n)^*$ is a "microstructure" which can be understood as a certain generalization of a Young measure representing, roughly

Received June 15, 1992.
1991 *Mathematics Subject Classification*. Primary 80A15, 49J40, 35R45.

speaking, oscillations (and possibly also concentrations) in the spatial gradient ∇u. Here $\Omega \subset \mathbb{R}^n$ is a bounded Lipschitz domain and $\mathrm{Car}(\Omega; \mathbb{R}^n)$ denotes the linear space of all Carathéodory functions $h\colon \Omega \times \mathbb{R}^n \to \mathbb{R}$ (that means $h(\cdot, v)$ are measurable and $h(x, \cdot)$ are continuous) with at most quadratic growth, i.e., $|h(x, v)| \le a_h(x) + b_h|v|^2$ for some $a_h \in L^1(\Omega)$ and $b_h < +\infty$. We will consider $\mathrm{Car}(\Omega; \mathbb{R}^n)$ as a locally convex space endowed by the collection of seminorms $\{|\cdot|_r\}_{r\in\mathbb{R}}$ with $|h|_r = \sup_{\|v\|_{L^2(\Omega)}\le r} |\int_\Omega h(x, v(x))\,dx|$. Let us abbreviate $F = L^2(\Omega) \times \mathrm{Car}(\Omega; \mathbb{R}^n)$ and $P = F^* = L^2(\Omega) \times \mathrm{Car}(\Omega; \mathbb{R}^n)^*$, where P is considered in the weak $\times$ weak* topology (the star will denote the dual space); $L^2(\Omega)$ is identified with its own dual. We imbed $H_0^1(\Omega)$ into P by means of a mapping $i\colon H_0^1(\Omega) \to P$ defined by $u \mapsto (u, e(\nabla u))$ with the imbedding $e\colon L^2(\Omega; \mathbb{R}^n) \to \mathrm{Car}(\Omega; \mathbb{R}^n)^*$ defined by $e(v)(h) = \int_\Omega h(x, v(x))\,dx$. Note that i is continuous when $H_0^1(\Omega)$ is endowed with its norm topology. Then we define the state space Q as the closure of $i(H_0^1(\Omega))$ in P. It was proved in [13] that Q is a convex, σ-compact, locally compact, closed subset of P. For $p = (u, m) \in P$, we put $|||p||| = |\langle(u, m), (0, h_N)\rangle|^{1/2} = |\langle m, h_N\rangle|^{1/2}$, where $h_N = \mathrm{Car}(\Omega; \mathbb{R}^n)$ is defined by $h_N(x, v) = |v|^2$ and $\langle\cdot, \cdot\rangle$ denotes (and will always denote) the canonical duality pairing between the respective spaces. Obviously, $|||i(u)||| = \|u\|_{H_0^1(\Omega)} \equiv (\int_\Omega |\nabla u|^2\,dx)^{1/2}$ for all $u \in H_0^1(\Omega)$; hence, $|||\cdot|||$ is a continuous extension on Q of the usual norm in the Sobolev space $H_0^1(\Omega)$.

Let $S_1\colon L^2(\Omega; \mathbb{R}^n) \to \mathrm{Car}(\Omega; \mathbb{R}^n)$ be defined by $\xi = (\xi_1, \dots, \xi_n) \mapsto h_\xi$, where $h_\xi(x, v) = \xi(x)\cdot v \equiv \sum_{i=1}^n \xi_i(x) v_i$. The adjoint mapping $S_1^*\colon \mathrm{Car}(\Omega; \mathbb{R}^n)^* \to L^2(\Omega; \mathbb{R}^n)$ assigns each microstructure $m \in \mathrm{Car}(\Omega; \mathbb{R}^n)^*$ its "mean value" (also called a first momentum). For $q = (u, m) \in Q$ we have always $S_1^* m = \nabla u$.

Let us still recall the definition from [14] of the substitution (denoted by "$\bullet$") of a microstructure $m \in \mathrm{Car}(\Omega; \mathbb{R}^n)^*$ into a Carathéodory function $h \in \mathrm{Car}(\Omega; \mathbb{R}^n)$, which represents a generalization of the classical Nemytskiĭ operator. Denoting by $\mathscr{M}(\Omega) \equiv L^1(\Omega)^{**}$ the space of finitely additive, absolutely continuous measures with a bounded variation on Ω, we define $h \bullet m \in \mathscr{M}(\Omega)$ by the relation

$$\langle h \bullet m, f\rangle = \langle m, hf\rangle \quad \forall f \in L^\infty(\Omega). \tag{0.1}$$

The following regularity holds: if $hf \in \mathrm{Car}(\Omega; \mathbb{R}^n)$ for all $f \in L^r(\Omega)$ with some $r < +\infty$, then $h \bullet m \in L^{r/(r-1)}(\Omega)$. Then obviously $(gh) \bullet m = g(h \bullet m)$ for all $g \in L^r(\Omega)$. For any $v \in L^2(\Omega; \mathbb{R}^n)$, we have obviously $[h \bullet e(v)](x) = h(x, v(x))$ for a.a. $x \in \Omega$; therefore, the mapping $m \mapsto h \bullet m$ is actually the extension of the Nemytskiĭ operator generated by h. Let us emphasize that the geometry of $\mathrm{Car}(\Omega; \mathbb{R}^n)^*$ makes this extended operator linear, although the original Nemystiĭ operator $L^2(\Omega; \mathbb{R}^n) \to L^1(\Omega)$ is generally nonlinear (of course, with respect to the usual geometry of $L^2(\Omega; \mathbb{R}^n)$).

1. A steady-state situation. We want to derive by a standard thermomechanical way a model of a nonisothermal evolution of a microstructure and, at the same time, to use the convex geometry imposed on Q from P similarly to what was done in [13].

Let us begin in this section with a steady-state situation. We admit a heterogenous material.

The state of our system will now include a temperature θ, being thus a triple $(\theta, q) = (\theta, u, m) \in L^2(\Omega) \times Q$. As usual, the starting point will be a Helmholtz free energy density $\psi_x : \mathbb{R}^+ \times \mathbb{R} \times \mathbb{R}^n \to \mathbb{R}$ at a current point $x \in \Omega$ which will be chosen in the form

$$\psi_x(\theta, u, v) = \varphi_0(x)u + \varphi_1(x, v) + \alpha(x, \theta)\varphi_2(x, v) - c_0(x)\theta \ln \theta, \qquad (1.1)$$

where $c_0 \in L^\infty(\Omega)$ is a heat capacity (greater than some positive constant), $\varphi_0 \in L^2(\Omega)$ is the external loading, φ_1, $\varphi_2 \in \mathrm{Car}(\Omega; \mathbb{R}^n)$ determine a temperature-dependent elastic potential, and $\alpha(x, \cdot): \mathbb{R} \to \mathbb{R}$ is a continuous function corresponding to thermal expansion. We suppose a coercivity of φ_1

$$\varphi_1(x, v) \geq a|v|^2 - b(x) \qquad (1.2)$$

with some a positive and $b \in L^1(\Omega)$ and at most linear growth of both φ_2 and α:

$$|\varphi_2(x, v)| \leq C|v| + b(x), \qquad (1.3)$$

$$|\alpha(x, \theta)| \leq C|\theta| + b(x), \qquad (1.4)$$

with some $C < +\infty$ and $b \in L^2(\Omega)$. Obviously, (1.3) with (1.4) guarantee $\alpha(\theta)\varphi_2 \in \mathrm{Car}(\Omega; \mathbb{R}^n)$ whenever $\theta \in L^2(\Omega)$, where $\alpha(\theta)\varphi_2$ abbreviates naturally the function $(x, v) \mapsto \alpha(x, \theta(x))\varphi_2(x, v)$.

EXAMPLE 1.1. The standard one-dimensional nonlinear thermo-elasticity model (see, e.g., [4]) can be obtained by the choice $n = 1$ and

$$\begin{aligned} \varphi_1(x, v) &= \tfrac{1}{2}v^2, \\ \varphi_2(x, v) &= v. \end{aligned} \qquad (1.5)$$

EXAMPLE 1.2. The one-dimensional shape memory alloy model of a Devonshire type uses $\varphi_1(x, v) = v^6 - v^4$ and $\varphi_2(x, v) = v^2$; cf. [1, 5, 7, 17]. To satisfy our growth requirements $\varphi_1 \in \mathrm{Car}(\Omega; \mathbb{R}^n)$ and (1.3), we need to modify both φ_1 and φ_2, for example, in the following manner:

$$\begin{aligned} \varphi_1(x, v) &= (v^6 - v^4)/(1 + av^4), \\ \varphi_2(x, v) &= v^2/(1 + a|v|) \end{aligned} \qquad (1.6)$$

for some small $a > 0$. Note that the potential $\varphi_1 + a(\theta)\varphi_2$ is a nonconvex function of v wherever $\alpha(\theta) < 0$, which necessarily creates a nontrivial microstructure in the steady state.

Continuing the standard thermomechanical approach, we define the entropy density as

$$s_x(\theta, u, v) = -\frac{\partial \psi_x}{\partial \theta}(\theta, u, v) = \alpha_0'(x, \theta)\varphi_2(x, v) + c_0(x) + c_0(x)\ln\theta \qquad (1.7)$$

and the internal energy density as

$$\begin{aligned} e_x(\theta, u, v) &= \psi_x(\theta, u, v) + \theta s_x(\theta, u, v) \\ &= \varphi_0(x)u + \varphi_1(x, v) + (\alpha(x, \theta) - \theta\alpha_0'(x, \theta))\varphi_2(x, v) + c_0(x)\theta. \end{aligned} \qquad (1.8)$$

For $q = (u, m) \in Q$ and appropriate $\theta \in L^2(\Omega)$ we still define the corresponding distribution of free energy, entropy, and internal energy as (in general) measures on Ω:

$$\psi(\theta, q) \equiv \psi(\theta, u, m) = \varphi_0 u + (\varphi_1 + \alpha(\theta)\varphi_2) \bullet m - c_0\theta \ln \theta,$$
$$s(\theta, q) \equiv s(\theta, u, m) = -(\alpha'(\theta)\varphi_2) \bullet m + c_0 + c_0 \ln \theta$$
$$e(\theta, q) \equiv e(\theta, u, m) = \varphi_0 u + (\varphi_1 + \alpha(\theta)\varphi_2 - \theta\alpha'(\theta)\varphi_2) \bullet m + c_0\theta.$$

The total Helmholtz free energy $\Psi\colon L^2(\Omega) \times H_0^1(\Omega) \to \mathbb{R}$ is then classically defined by

$$\Psi(\theta, u) = \int_\Omega \psi_x(\theta(x), u(x), \nabla u(x))\, dx.$$

Here we define also the extended total free energy for $q \in P$ by

$$\Psi_{\mathrm{e}}(\theta, q) \equiv \Psi_{\mathrm{e}}(\theta, u, m) = \int_\Omega \psi(\theta, u, m)(x)\, dx$$
$$= \langle u, \varphi_0\rangle + \langle m, \varphi_1\rangle + \langle m, \alpha(\theta)\varphi_2\rangle - \int_\Omega c_0(x)\theta(x) \ln \theta(x)\, dx$$

(for the last equality use Eq. (0.1) with $f \equiv 1$) and the augmented free energy by

$$\Psi_{\mathrm{a}}(\theta, q) \equiv \Psi_{\mathrm{a}}(\theta, u, m) = \begin{cases} \Psi_{\mathrm{e}}(\theta, u, m) & \text{if } (u, m) \in Q, \\ +\infty & \text{otherwise}. \end{cases}$$

Note that Ψ_{e} represents on Q a continuous extension of Ψ because $\Psi_{\mathrm{e}}(\theta, i(u)) \equiv \Psi(\theta, u)$ for every $u \in H_0^1(\Omega)$. In view of this fact, let us omit the subscript "e" for "extended" without any misunderstanding. Analogously, we also define the extended total entropy $\mathscr{S}$ and total internal energy $\mathscr{E}$ by

$$\mathscr{S}(\theta, q) \equiv \mathscr{S}(\theta, u, m) = \int_\Omega s(\theta, u, m)(x)\, dx,$$
$$\mathscr{E}(\theta, q) \equiv \mathscr{E}(\theta, u, m) = \int_\Omega e(\theta, u, m)(x)\, dx.$$

2. A nonlinear thermomechanical evolution model. Now we want to formulate a time evolution of the state $(\theta, q) = (\theta(t), q(t))$ for $t \in [0, 1]$. As usual, a dot will abbreviate the time derivative, e.g. $\dot\theta \equiv \frac{d}{dt}\theta$.

For the momentum equation we need to prescribe the kinetic and the Rayleigh dissipative energies, denoted respectively by T and R_θ with $\theta \in L^2(\Omega)$; here we admit the dissipation to be temperature dependent. We take them as in [13, Examples 2.1 and 2.2]. As for $T\colon P \to \mathbb{R}$, we consider a positive mass density $\varrho \in L^\infty(\Omega)$; take the linear operator $S_0\colon L^2(\Omega) \to F$ defined by $S_0\xi = (\sqrt{\varrho}\xi, 0)$; and, for $\dot q = (\dot u, \dot m) \in P$, put

$$T(\dot q) = \frac{1}{2}\|S_0^*\dot q\|^2_{L^2(\Omega)} = \frac{1}{2}\int_\Omega \varrho(x)\dot u(x)^2\, dx. \tag{2.1}$$

As for $R_\theta\colon P \to \mathbb{R}$, we consider a measure space I of abstract indices i and a measurable function $d_\theta\colon I \times \Omega \times \mathbb{R}^n \to \mathbb{R}$ describing the dissipative mechanism such that

$$\begin{aligned}&\text{meas}(I) < +\infty,\\&\forall x \in \Omega:\ \text{the collection } \{d_\theta(i, x, \cdot)\}_{i\in I} \text{ is equi-continuous},\\&\exists d_0 \in L^2(I \times \Omega),\ d_1 \in \mathbb{R} : |d_\theta(i, x, v)| \le d_0(i, x) + d_1|v|,\end{aligned} \tag{2.2}$$

and then take the operator $S_\theta\colon L^2(I \times \Omega) \to F$ defined by $S_\theta\xi = (0, h_{\theta,\xi})$ with $h_{\theta,\xi} \in \text{Car}(\Omega; \mathbb{R}^n)$ prescribed by $h_{\theta,\xi}(x, v) = \int_I \xi(i, x)d_\theta(i, x, v)\,di$, and put

$$R_\theta(\dot q) = \tfrac{1}{2}\|S_\theta^*\dot q\|^2_{L^2(I\times\Omega)}. \tag{2.3}$$

It is not difficult to verify that (2.2) ensures actually $h_{\theta,\xi} \in \text{Car}(\Omega; \mathbb{R}^n)$ and the linear operator $\xi \mapsto h_{\theta,\xi}\colon L^2(I \times \Omega) \to \text{Car}(\Omega; \mathbb{R}^n)$ to be bounded. Besides, for $p = (u, m) \in P$, we can see that $[S_\theta^* p](i, x) = [d(i, \cdot, \cdot) \bullet m](x)$ for a.a. $(i, x) \in I \times \Omega$ because $\langle S_\theta^* p, \xi\rangle = \int_\Omega [h_{\theta,\xi} \bullet m](x)\,dx = \int_{I\times\Omega} \xi(i, x)[d(i, \cdot, \cdot) \bullet m](x)\,dx\,di$ for any $\xi \in L^2(I \times \Omega)$. We will still define the distribution of the dissipative energy $r_\theta(\dot q) \in L^1(\Omega)$ by

$$r_\theta(\dot q)(x) = \int_I [d_\theta(i, \cdot, \cdot) \bullet \dot m]^2(x)\,di, \qquad \text{with } \dot q = (\dot u, \dot m). \tag{2.4}$$

It is obvious that $R_\theta(\dot q) = \int_\Omega r_\theta(\dot q)(x)\,dx$. The bilinear forms corresponding to T and R_θ will be denoted by $\widehat{T}\colon P \times P \to \mathbb{R}$ and $\widehat{R}_\theta\colon P \times P \to \mathbb{R}$, respectively. That means, e.g., $\widehat{T}(p_1, p_2) = \langle S_0^* p_1, S_0^* p_2\rangle$. Since only elements from $S_0(L^2(\Omega)) \cup S_\theta(L^2(I \times \Omega)) \subset F$ come into account to evaluate $T(p)$ and $R_\theta(p)$, we may in fact consider $p \in S_0(L^2(\Omega))^* \cap S_\theta(L^2(I \times \Omega))^*$. We will adopt this convention when writing expressions of the type $T(\dot q(t))$ and $R_\theta(\dot q(t))$, and we will then understand $\dot q(t)$ as the weak* limit of $\varepsilon^{-1}(q(t+\varepsilon) - q(t))$ in $S_0(L^2(\Omega))^* \cap S_\theta(L^2(I \times \Omega))^*$ only.

Following [13], we consider the generalized momentum equation for evolution of $q = (u, m)$ in the differential inclusion form

$$DT\ddot q + +DR_\theta \dot q + \partial_q \Psi_{\text{a}}(\theta, q) \ni 0, \tag{2.5}$$

where $\partial_q\Psi_{\text{a}}(\theta, q)$ denotes the subgradient of the convex function $\Psi_{\text{a}}(\theta, \cdot)\colon P \to \mathbb{R} \cap \{+\infty\}$ at a point $q \in Q$ and DT, $DR_\theta\colon P \to P^*$ are the Gâteaux derivatives of T and R_θ, respectively; for example, $DR_\theta = S_\theta^{**}S_\theta^*$. We need still an equation for the internal energy balance, which is naturally considered in the form

$$\dot e(\theta, q) + \nabla \cdot j = A_1 + A_2, \tag{2.6}$$

where j is a heat flux and A_1 and A_2 are heat sources balancing respectively the dissipation of the mechanical energy and the temperature dependence of the elastic potential $\varphi_1 + \alpha(\theta)\varphi_2$. We will determine A_1, A_2 from the standard energy-preservation requirement

$$\mathscr{E}(\theta(t), q(t)) + T(\dot q(t)) = \text{const.} \quad \text{for } t \in [0, 1] \tag{2.7}$$

by the following, a bit formal calculations (cf. Remark 2.2). Supposing, for a moment, that $\Psi(\theta, \cdot)$ is smooth, we get by multiplying (2.5) by $\dot q$ the energetic identity $\frac{d}{dt}(T(\dot q) + \Psi(\theta, q)) + \mathscr{S}(\theta, q)\dot\theta + R(\dot q) = 0$. Integrating (2.6) over Ω gives

$\frac{d}{dt}\mathscr{E}(\theta, q) = \int_\Omega (A_1(x) + A_2(x))\, dx$ provided $j = 0$ on $\partial\Omega$. Altogether,

$$\frac{d}{dt}(\mathscr{E}(\theta, q) + T(\dot{q})) = -\frac{d}{dt}\Psi(\theta, q) - \mathscr{S}(\theta, q)\dot{\theta} - R(\dot{q}) + \int_\Omega (A_1(x) + A_2(x))\, dx$$
$$= \int_\Omega (A_1(x) - r_\theta(\dot{q})(x))\, dx$$
$$+ \int_\Omega (A_2(x) - \dot{\psi}(\theta, q)(x) - s(\theta, q)(x)\dot{\theta}(x))\, dx.$$

A comparison with (2.7) suggests naturally taking $A_1 = r_\theta(\dot{q})$ and $A_2 = \dot{\psi}(\theta, q) + s(\theta, q)\dot{\theta}$. Putting these into (2.6) and realizing that $\dot{e}(\theta, q) = \dot{\psi}(\theta, q) + s(\theta, q)\dot{\theta} + \dot{s}(\theta, q)\theta$, we obtain the equation

$$\theta \dot{s}(\theta, q) + \nabla \cdot j = r_\theta(\dot{q}). \tag{2.8}$$

In view of Eq. (1.7) we have $\dot{s}(\theta, q) = [\alpha''(\theta)\dot{\theta}\varphi_2] \bullet m + [\alpha'(\theta)\varphi_2] \bullet \dot{m} + c_0\dot{\theta}/\theta$. Putting this into Eq. (2.6) gives

$$c_0\dot{\theta} + \nabla \cdot j = r_\theta(\dot{q}) + [\theta\alpha''(\theta)\dot{\theta}\varphi_2] \bullet m + [\theta\alpha'(\theta)\varphi_2] \bullet \dot{m}. \tag{2.9}$$

Now we want to verify the entropy condition

$$\frac{d}{dt}\mathscr{S}(\theta, q) \geq 0. \tag{2.10}$$

From Eq. (2.8) we get immediately

$$\frac{d}{dt}\mathscr{S}(\theta, q) = \int_\Omega \frac{r_\theta(\dot{q})}{\theta}\, dx + \int_{\partial\Omega} \frac{j}{\theta} \cdot dS - \int_\Omega \frac{j \cdot \nabla\theta}{\theta^2}\, dx, \tag{2.11}$$

and we can see that (2.10) will actually be fulfilled if we put $j = -\lambda\nabla\theta$ and isolate our system by imposing the boundary condition $\partial\theta/\partial\nu = 0$ on $\partial\Omega$, where ν is a normal to $\partial\Omega$. Here $\lambda \in L^\infty(\Omega)$ is a positive heat conductivity coefficient. Then we can eventually rewrite Eq. (2.9) to get the desired heat transfer equation:

$$c_0\dot{\theta} - \nabla \cdot (\lambda\nabla\theta) = r_\theta(\dot{q}) + [\theta\alpha''(\theta)\dot{\theta}\varphi_2] \bullet m + [\theta\alpha'(\theta)\varphi_2] \bullet \dot{m}. \tag{2.12}$$

Of course, we have supposed $\theta > 0$, but from the maximum principle for Eq. (2.12) we can see that (at least) every sufficiently regular solution will satisfy this hypothesis provided the initial state satisfies it. As a result, our model seems to be thermodynamically correct.

To simplify our problem a bit, we will assume $\alpha(x, \cdot)$ to be affine, which is, in view of Eq. (1.1), of the same generality as if we assume $\alpha(x, \cdot)$ to be linear, i.e.,

$$\alpha(x, \theta) = \alpha_0(x)\theta \quad \text{with } \alpha_0 \in L^\infty(\Omega). \tag{2.13}$$

This considerably simplifies the internal energy (1.8): $e_x(\theta, u, v) = \varphi_0(x)u + \varphi_1(x, v) + c_0(x)\theta$, and also the right-hand side of Eq. (2.12): $r_\theta(\dot{q}) + [\alpha_0\theta\varphi_2] \bullet \dot{m}$.

We complete our problem by imposing also some initial conditions on the state:

$$\theta(0, \cdot) = \theta_0 \in L^2(\Omega), \tag{2.14}$$
$$q(0) = q_0 \in Q, \tag{2.15}$$

and initial conditions on the impulse:

$$\dot{q}(0) = p_0 \in P; \tag{2.16}$$

and eventually for the boundary condition for the heat flux, for simplicity, let us take

$$\frac{\partial \theta}{\partial \nu} = 0 \quad \text{on } \partial\Omega. \tag{2.17}$$

Now we are to formulate our problem weakly. Let us agree that $S_0^* p$ will mean the function $(t, x) \mapsto [S_0^* p(t)](x)$, and similarly $S_\theta^* p$ means the function $(t, i, x) \mapsto [S_{\theta(t)}^* p(t)](i, x)$. For $\theta \in L^2((0, 1) \times \Omega)$ let us define the set $\mathscr{A}_\theta$ of admissible trajectories by

$$\mathscr{A}_\theta = \{q\colon [0, 1] \to Q;\ S_0^* \dot{q} \in L^2((0, 1) \times \Omega),\ S_\theta^* \dot{q} \in L^2((0, 1) \times I \times \Omega),$$
$$\phi_{01} \bullet q \in L^1((0, 1) \times \Omega),\ \phi_2 \bullet \dot{q} \in L^2((0, 1) \times \Omega)\},$$

where $\phi_{01} = (\varphi_0, \varphi_1) \in F$ and $\phi_2 = (0, \varphi_2) \in F$, and, for example, $\phi_{01} \bullet q$ with $q = (u, m)$ means $\varphi_0 u + \varphi_1 \bullet m$ so that $\int_\Omega [\phi_{01} \bullet q(t)](x)\, dx = \langle q(t), \phi_{01}\rangle$. Note that the Gâteaux derivative of $\Psi(\theta, \cdot)\colon P \to \mathbb{R}$ is constant over P and equal to $\phi_{01} + \alpha(\theta)\phi_2 \in F \subset F^{**}$, and then $\partial_q \Psi_{\mathrm{a}}(\theta, q) = \phi_{01} + \alpha(\theta)\phi_2 + N_Q(q)$ where $N_Q(q) = \{\phi \in P^* = F^{**};\ \forall \tilde{q} \in Q\colon \langle \phi, \tilde{q} - q\rangle \geq 0\}$ is the normal cone to Q at the point $q \in Q$. We shall obtain our weak formulation by multiplying (2.5) and (2.12) by suitable test functions and making the per-parts integration in time, and eventually using Green's formula for Eq. (2.12).

DEFINITION 2.1. The trajectory $(\theta, q) = (\theta, u, m) \in [L^2(0, 1; H_0^1(\Omega)) \cap C^0(0, 1; L^2(\Omega))] \times \mathscr{A}_\theta$ is the weak solution of the Cauchy problem for the system (2.5) and (2.12) with Eqs. (2.13)–(2.17) if q fulfills Eq. (2.15) and, for all $\tilde{q} \in \mathscr{A}_\theta$ with $\tilde{q}(1) = q(1)$,

$$\int_0^1 (-\widehat{T}(\dot{q}, \dot{\tilde{q}} - \dot{q}) + \widehat{R}_\theta(\dot{q}, \tilde{q} - q) + \langle \tilde{q} - q, \phi_{01} + \alpha_0 \theta \phi_2\rangle)\, dt \geq \widehat{T}(p_0, \tilde{q}(0) - q_0), \tag{2.18}$$

and, for every $z \in C^\infty([0, 1] \times \overline{\Omega})$ with $z(1, \cdot) = 0$,

$$\int_0^1 \int_\Omega (-c_0 \theta \dot{z} + \lambda \nabla\theta \cdot \nabla z - r_\theta(\dot{q}) z - [\alpha_0 \theta \varphi_2 z] \bullet \dot{m})\, dx\, dt = \int_\Omega \theta_0 z(0, \cdot)\, dx. \tag{2.19}$$

Let us note that all the integrals in (2.18) and (2.19) have actually a sense; realize that, for $\theta \in L^2((0, 1) \times \Omega)$ and $q \in \mathscr{A}_\theta$, we have always $\langle q(t), \alpha_0 \theta(t, \cdot)\phi_2\rangle = \int_\Omega \alpha_0(x)\theta(t, x)[\phi_2 \bullet q](x)\, dx \in L^2(0, 1)$, and $r_\theta(\dot{q}) = \int_I S_\theta^*(\dot{q})^2\, di \in L^1((0, 1) \times \Omega)$ (cf. (2.4)), and moreover $[\alpha_0 \theta \varphi_2 z] \bullet \dot{m} = \alpha_0 \theta z [\phi_2 \bullet \dot{q}] \in L^1((0, 1) \times \Omega)$. Also the initial condition (2.14) has a good sense because $\theta \in C^0(0, 1; L^2(\Omega))$, while Eq. (2.15) has a rather "energetic" sense only, i.e., only the components $S_0^* q$, $S_\theta^* q$, and $\phi_2 \bullet q$ are factually essential (cf. also [13]).

We will show that our model is factually an immediate generalization of a classical nonlinear thermo-visco-elasticity model, which will justify it a bit. Let us emphasize that the proper applications of our model are in cases with nonconvex free energy like the shape memory alloys from Example 1.2 where a nontrivial microstructure is essentially inevitable, but in spite of this, for a very special data, our model basically coincides with (a scalar version of) the nonlinear thermo-visco-elasticity system (see [4, 9, 10], for example):

$$\begin{aligned} \varrho \ddot{u} - \nabla \cdot (\mu \nabla \dot{u}) - \Delta u - \mathbf{1} \cdot \nabla(\alpha_0 \theta) &= -\varphi_0 , \\ c_0 \dot{\theta} - \nabla \cdot (\lambda \nabla \theta) - \mu |\nabla \dot{u}|^2 - \mathbf{1} \cdot \alpha_0 \theta \nabla \dot{u} &= 0 . \end{aligned} \tag{2.20}$$

Here $\mu \in L^\infty(\Omega)$ is a viscosity coefficient, the other coefficients ϱ, λ, c_0, α_0, and φ_0 have the same meaning as previously established, and $\mathbf{1} = (1, \dots, 1) \in \mathbb{R}^n$ is used to scalarize the respective terms; this is certainly rather artificial since we here consider u to be scalar-valued while the proper multidimensional model should involve u as $\mathbb{R}^n$-valued (cf., e.g., [9, 10]). We complete Eqs. (2.20) with the initial conditions

$$\begin{aligned} u(0, \cdot) &= u_0 , \\ \dot{u}(0, \cdot) &= u_1 , \\ \theta(0, \cdot) &= \theta_0 . \end{aligned} \tag{2.21}$$

A couple $(\theta, u) \in [L^2(0, 1; H^1(\Omega)) \cap C^0(0, 1; L^2(\Omega))] \times H^1(0, 1; H^1_0(\Omega))$ will be called a weak solution of Eqs. (2.20), (2.21) with the boundary conditions (2.17) and $u = 0$ on $\partial\Omega$ if the following integral identities are satisfied for every $z \in C^\infty([0, 1] \times \overline{\Omega})$ with $z(1, \cdot) = 0$:

$$\int_0^1 \int_\Omega (-\varrho \dot{u} \dot{z} + \mu \nabla \dot{u} \cdot \nabla z + \mathbf{1} \cdot \alpha_0 \theta \nabla z + \nabla u \cdot \nabla z + \varphi_0 z) \, dx \, dt = \int_\Omega u_1 z(0, \cdot) \, dx , \tag{2.22}$$

$$\int_0^1 \int_\Omega (-c_0 \theta \dot{z} + \lambda \nabla \theta \cdot \nabla z - \mu |\nabla \dot{u}|^2 z - \mathbf{1} \cdot \alpha_0 \theta \nabla \dot{u} z) \, dx \, dt = \int_\Omega \theta_0 z(0, \cdot) \, dx . \tag{2.23}$$

To show that our model can cover Eqs. (2.20), (2.21), we take the following special data:

$$\begin{aligned} \varphi_1(x, v) &= \tfrac{1}{2} |v|^2 , \\ \varphi_2(x, v) &= \sum_{i=1}^n v_i ; \end{aligned} \tag{2.24}$$

$$\begin{aligned} I &= \{1, \dots, n\} , \\ d_\theta(i, x, v) &= \sqrt{\mu(x)} v_i . \end{aligned} \tag{2.25}$$

Note that Eqs. (2.25) obviously satisfy (2.2) and determine, independently of θ, $S_\theta \colon L^2(\Omega; \mathbb{R}^n) \to F \colon \xi \mapsto h_\xi$ with $h_\xi(x, v) = \sqrt{\mu(x)} \sum_{i=1}^n \xi_i(x) v_i$; therefore, $S_\theta \equiv \sqrt{\mu} S_1$. Besides, for $n = 1$, Eqs. (2.24) coincide apparently with Eqs. (1.5).

PROPOSITION 2.1. If (θ, u, m) is a solution due to Definition 2.1 with the data (2.24) and (2.25), then

$$m(t) = e(\nabla u(t)) \qquad \text{for a.a. } t \in [0, 1] \tag{2.26}$$

and (θ, u) is a weak solution of Eqs. (2.20), (2.21) provided the initial conditions (2.14)–(2.16) and (2.21) are related by $q_0 = (u_0, m_0)$ and $p_0 = (u_1, m_1)$ with m_0 and m_1 arbitrary.

Proof. Let us first prove that

$$\langle m, \varphi_1\rangle \geq \langle e(\nabla u), \varphi_1\rangle \quad \forall (u, m) \in Q. \tag{2.27}$$

By the very definition of Q, there is a net $\{u_\iota\} \subset H_0^1(\Omega)$ such that $e(\nabla u_\iota) \to m$ in $\mathrm{Car}(\Omega; \mathbb{R}^n)^*$. In particular, $\langle e(\nabla u_\iota), \varphi_1\rangle \to \langle m, \varphi_1\rangle$ and also $\langle \nabla u_\iota, \xi\rangle = \langle e(\nabla u_\iota), S_1\xi\rangle \to \langle m, S_1\xi\rangle = \langle S_1^* m, \xi\rangle = \langle \nabla u, \xi\rangle$ for any $\xi \in L^2(\Omega; \mathbb{R}^n)$, which shows that $u_\iota \to u$ weakly in $H_0^1(\Omega)$. However, by the convexity of $\varphi_1(x, \cdot)$ we have always $\liminf\langle e(\nabla u_\iota), \varphi_1\rangle = \liminf \int_\Omega \varphi_1(x, \nabla u_\iota(x))\, dx \geq \int_\Omega \varphi_1(x, \nabla u(x))\, dx = \langle e(\nabla u), \varphi_1\rangle$, which proves (2.27).

Moreover, the equality in (2.27) appears if and only if $m = e(\nabla u)$. Indeed, the equality in (2.27) implies $\|u_\iota\|^2_{H_0^1(\Omega)} = 2\langle e(\nabla u_\iota), \varphi_1\rangle \to 2\langle e(\nabla u), \varphi_1\rangle = \|u\|^2_{H_0^1(\Omega)}$. By the uniform convexity of the norm of $H_0^1(\Omega)$, we get $u_\iota \to u$ strongly in $H_0^1(\Omega)$. For any $h \in \mathrm{Car}(\Omega; \mathbb{R}^n)$ we then have $\langle e(\nabla u_\iota), h\rangle = \int_\Omega h(x, \nabla u_\iota(x))\, dx \to \int_\Omega h(x, \nabla u(x))\, dx = \langle e(\nabla u), h\rangle$ because of the continuity of the Nemytskiĭ operator $L^2(\Omega; \mathbb{R}^n) \to L^1(\Omega)$ generated by the Carathéodory integrand h. Since h is arbitrary, $m = e(\nabla u)$.

Let us still notice that $\widehat{T}(q_1, q_2) = \int_\Omega \varrho u_1 u_2\, dx$ and $\widehat{R}(q_1, q_1) = \int_\Omega \mu \nabla u_1 \nabla u_2\, dx$ for any $q_i = (u_i, m_i) \in Q$, $i = 1, 2$.

To prove Eq. (2.26) let us put $\tilde{q} = i(u)$ into (2.18) with a possible change only for $t = 1$. This makes zero all terms but the one with φ_1, which gives $\int_0^1 \langle e(\nabla u) - m, \varphi_1\rangle\, dt \geq 0$. In view of (2.27) it implies $\langle e(\nabla u(t)), \varphi_1\rangle = \langle m(t), \varphi_1\rangle$ for a.a. $t \in [0, 1]$. As shown above, this means just Eq. (2.26).

Now, take $\varepsilon > 0$ and plug $\tilde{q} = i(u + \varepsilon z)$ into (2.18). We also use (2.27) and the identity $\langle e(\nabla u + \varepsilon \nabla z) - e(\nabla u), \varphi_1\rangle = \int_\Omega (\varepsilon \nabla u \cdot \nabla z + \frac{1}{2}\varepsilon^2 |\nabla z|^2)\, dx$. This gives eventually the inequality

$$\int_0^1 \int_\Omega \left(-\varepsilon \varrho \dot{u} \dot{z} + \varepsilon \mu \nabla \dot{u} \cdot \nabla z + \mathbf{1} \cdot \varepsilon \alpha_0 \theta \nabla z + \varepsilon \nabla u \cdot \nabla z + \varepsilon \varphi_0 z + \frac{\varepsilon^2}{2} |\nabla z|^2\right) dx\, dt \\ \geq \int_\Omega \varepsilon \varrho u_1 z(0, \cdot)\, dx. \tag{2.28}$$

Now we derive (2.28) by $\varepsilon > 0$ and then pass to the limit as ε tends to zero, which causes the term $\frac{1}{2} \int_\Omega \varepsilon |\nabla z|^2\, dt$ to vanish. We get eventually

$$\int_0^1 \int_\Omega (-\varrho \dot{u} \dot{z} + \mu \nabla \dot{u} \cdot \nabla z + \mathbf{1} \cdot \alpha_0 \theta \nabla z + \nabla u \cdot \nabla z + \varphi_0 z)\, dt \geq \int_\Omega \varrho u_1 z(0, \cdot)\, dx. \tag{2.29}$$

This must hold also for $-z$ instead of z, which yields even equality in (2.29). This is just Eq. (2.22), however.

The identity (2.23) follows immediately from Eq. (2.19) when one realizes that $r_\theta(\dot{q}) = \sum_{i=1}^n [S_\theta^*(\dot{q})^2](i, \cdot) = \mu |\nabla \dot{u}|^2$ and $[\alpha_0 \theta \varphi_2] \bullet \dot{m} = \mathbf{1} \cdot \alpha_0 \theta \nabla \dot{u}$. □

Certainly, an important theoretical justification for our model is a uniqueness at least in special cases (let us emphasize that in the general case uniqueness naturally cannot be expected; cf. [13, Sec. 5]):

COROLLARY 2.1. Let the data (2.24)–(2.25) be taken. If Eqs. (2.20), (2.21) possess a unique weak solution (θ, u), then the solution (θ, u, m) due to Definition 2.1 is also unique.

REMARK 2.1. The relation (2.26) means that the microstructure m has a trivial character. If m could be represented as a Young measure, Eq. (2.26) would mean that this measure is a Dirac one almost everywhere. Of course, such trivial microstructure is related only with the special data (2.24), (2.25) and certainly cannot be expected in general cases; cf. also the numerical experiments performed in [15] for a certain isothermal one-dimensional case with a nonconvex free energy.

REMARK 2.2. We derive Eq. (2.12) from the energy balance a bit formally as if Ψ_a would be smooth. However, it is known that for systems like (2.5) with Ψ_a nonsmooth and taking values including ∞, such an energy balance can be violated by "inelastic interactions" of the mass in the second-order term $\varrho\ddot{u}$ with the distributed obstacle on u (cf. [16]). On the other hand, it seems (and numerical experiments in [15] confirm this) that this does not appear in our system where u is, in fact, not constrained and $q = (u, m) \in Q$ represents a constraint on m only, but m does not appear explicitly in the kinetic energy form T.

REMARK 2.3. In case the initial data are regular enough, the uniqueness of Eqs. (2.22), (2.23) assumed in Corollary 2.1 has been proved in [4] for $n = 1$, while for $n = 3$ it follows when adapting the results of [10] to our scalar case.

3. A partially linearized problem. We have seen in Proposition 2.1 that our problem includes basically also multidimensional nonlinear thermo-visco-elasticity which is itself a difficult problem solved only recently by a somewhat nonconstructive manner in [9]. A one-dimensional nonlinear thermo-visco-elasticity allows a more constructive approach (see [4]), but in both cases a spatial regularity of $\nabla\dot{u}$ is employed. In the general case, this would perhaps corresponds here to some spatial regularity of $\dot{m}$, which is presently a rather unclear matter. Besides, the fixed-point technique used in [9] requires the solution of the momentum equation for a given θ to be unique (and dependent on θ continuously in an appropriate sense). This approach would have to be modified for our problem because the uniqueness of the solution of (2.5) for a fixed θ generally does not hold (cf. the example in [13, Sec. 5]). Yet, the modified, multivalued versions of the Schauder fixed-point theorem require the set of solutions of (2.5) to be convex, which is unfortunately not evident here. On the contrary, situations like that investigated in [16] show that the second-order inequality of the type (2.5) can have a nonconvex set of solutions provided an "obstacle" acts on u and provided no other conditions of a local energy preservation type are imposed. (This, however, still gives our problem, which has no explicit obstacle on u, a chance; cf. also Remark 2.2.)

All of this forces us to accept the standard simplification which, in the case of Eqs. (2.20), neglects the term $\mu|\nabla\dot{u}|^2$ while replacing $\alpha_0\theta\nabla\dot{u}$ by $\alpha_0\theta_0\nabla\dot{u}$ (cf., e.g., [3, 11]). This relies on the assumptions that the process is sufficiently slow so that it produces a negligible amount of heat by dissipation of the mechanical energy and that the temperature θ does not differ too much from the initial temperature θ_0. The system thus resulting from Eqs. (2.20) is obviously linear.

In our case we modify analogously the heat equation (2.12); that means we replace it by

$$c_0\dot{\theta} - \nabla\cdot(\lambda\nabla\theta) = [\alpha_0\theta_0\varphi_2]\bullet\dot{m}. \tag{3.1}$$

A similar simplification has been used also in [2, Eq. (2.7)]. Besides. to simplify considerably the derivation of the a priori estimates and to enable a direct usage of [13] for the convergence proof, we shall confine ourselves to a temperature-independent dissipation. Then we dare omit the subscript θ in the relevant objects; i.e., instead of S_θ, R_θ, $\widehat{R}_\theta$, and $\mathscr{A}_\theta$, we will write respectively S, R, $\widehat{R}$, and $\mathscr{A}$.

DEFINITION 3.1. The trajectory $(\theta, q) = (\theta, u, m) \in (L^2(0,1;H^1(\Omega)) \cap C^0(0,1;L^2(\Omega)) \times \mathscr{A}$ is the weak solution of the Cauchy problem for the partially linearized system (2.5) and (3.1) with Eqs. (2.13)–(2.17) and with $R_\theta \equiv R$ if q fulfills Eq. (2.15) and, for all $\tilde{q} \in \mathscr{A}$ with $\tilde{q}(1) = q(1)$,

$$\int_0^1 (-\widehat{T}(\dot{q}, \dot{\tilde{q}} - \dot{q}) + \widehat{R}(\dot{q}, \tilde{q} - q) + \langle \tilde{q} - q, \phi_{01} + \alpha_0\theta\phi_2\rangle)\,dt \geq \widehat{T}(p_0, \tilde{q}(0) - q_0), \tag{3.2}$$

and, for every $z \in C^\infty([0,1]\times\overline{\Omega})$ with $z(1,\cdot) = 0$,

$$\int_0^1\int_\Omega (-c_0\theta\dot{z} + \lambda\nabla\theta\cdot\nabla z - [\alpha_0\theta_0\varphi_2 z]\bullet\dot{m})\,dx\,dt = \int_\Omega \theta_0 z(0,\cdot)\,dx. \tag{3.3}$$

In view of (3.2), the resulting system remains nonlinear; that is why we called it an only partially linearized problem.

This simplification enables us to prove the existence of the weak solution even by a constructive method, which is a good basis for a numerical realization. We use a semi-implicit discretization in time with an equidistant partition of the time-interval $[0, 1]$; $\tau > 0$ will be a time step, τ^{-1} an integer. For simplicity we keep the spatial variables continuous, but a further discretization in space can easily be made by a finite element method; we refer to [15] for discretization of (3.2) while the spatial discretization of Eq. (3.3) would then be rather standard.

DEFINITION 3.2. For $\tau > 0$, the approximate solution $(\theta_\tau, q_\tau) = (\theta_\tau, u_\tau, q_\tau)$ of the partially linearized Cauchy problem will be $(\theta_\tau, q_\tau) \in C^0(0,1;H^1(\Omega)\times Q)$ which is a piecewise linear trajectory on particular time intervals $[(k-1)\tau, k\tau]$, $k = 1, \ldots, \tau^{-1}$, and such that $q_\tau^k = q_\tau(k\tau) \in Q$ satisfies the following recurrent variational inequalities:

$$\widehat{T}(q_\tau^k - 2q_\tau^{k-1} + q_\tau^{k-2}, \tilde{q} - q_\tau^k) + \tau\widehat{R}(q_\tau^k - q_\tau^{k-1}, \tilde{q} - q_\tau^k) + \tau^2\langle \tilde{q} - q_\tau^k, \phi_{01} + \alpha_0\theta_\tau^{k-1}\phi_2\rangle \geq 0 \tag{3.4}$$

for all $\tilde{q} \in Q$ and $k = 1, \ldots, \tau^{-1}$ and $\theta_\tau^k = \theta_\tau(k\tau) \in H^1(\Omega)$ satisfies the following recurrent integral identity:

$$\int_\Omega (c_0(\theta_\tau^k - \theta_\tau^{k-1})z + \tau\lambda\nabla\theta_\tau^k\cdot\nabla z - (\alpha_0\theta_0\varphi_2 z)\bullet(m_\tau^k - m_\tau^{k-1}))\,dx = 0 \tag{3.5}$$

for all $z \in H^1(\Omega)$, with $\theta_\tau^0 = \theta_0$, $q_\tau^0 = q_0$, and $q_\tau^{-1} = q_0 - \tau p_0$.

REMARK 3.1. Conditions (3.4), (3.5) represent merely a semi-implicit discretization of the system (2.5) and (3.1):

$$DT\left(\frac{q_\tau^k - 2q_\tau^{k-1} + q_\tau^{k-2}}{\tau^2}\right) + DR\left(\frac{q_\tau^k - q_\tau^{k-1}}{\tau}\right) + \partial_q \Psi_{\mathrm{a}}(\theta_\tau^{k-1}, q_\tau^k) \ni 0,$$

$$c_0 \frac{\theta_\tau^k - \theta_\tau^{k-1}}{\tau} - \nabla\cdot(\lambda\nabla\theta_\tau^k) = [\alpha_0\theta_0\varphi_2] \bullet \frac{m_\tau^k - m_\tau^{k-1}}{\tau}.$$

Let us note the temperature in the discretized momentum inclusion is retarded; hence the scheme is not fully implicit.

PROPOSITION 3.1. Let (1.2)–(1.4) be valid. Then for each $\tau > 0$ there is at least one approximate solution (θ_τ, q_τ) due to Definition 3.2.

Proof. For $k = 1$ the variational inequality (3.4) is equivalent with the following minimization problem:

$$\begin{cases} \text{minimize } J(q) \\ \text{subject to } q \in Q, \end{cases} \tag{3.6}$$

where $J : P \to \mathbb{R}$ is defined by

$$J(q) = T(q) + \tau R(q) + \tau^2 \Psi(\theta_\tau^0, q) - \widehat{T}(q, 2q_\tau^0 - q_\tau^{-1}) - \tau\widehat{R}(q, q_\tau^0).$$

Clearly, (3.6) is a convex minimization problem (with a convex, continuous criterion and the convex, closed domain Q). By (1.2)–(1.4) and by the continuity of the imbedding $H_0^1(\Omega) \subset L^2(\Omega)$, we have $\Psi(\theta_\tau^0, u) \geq \frac{1}{2}a\|u\|^2_{H_0^1(\Omega)} - c$ with a from (1.2) and some c depending on $\|\theta_\tau^0\|_{L^2(\Omega)}$. Thus the extended total free energy is coercive on Q: $\Psi(\theta_\tau^0, q) \geq \frac{1}{2}a|||q|||^2 - c$. Taking into account also that, for all $q \in Q$, $T(q) - \widehat{T}(q, 2q_\tau^0 - q_\tau^{-1}) \geq -T(2q_\tau^0 - q_\tau^{-1}) > -\infty$ and $R(q) - \widehat{R}(q, q_\tau^0) \geq -R(q_\tau^0) > -\infty$ and that Q is locally compact, the existence of a solution $q = q_\tau^1$ of the above minimization problem directly follows.

Having θ_τ^0 and q_τ^1, we can obtain the solution θ_τ^1 of Eq. (3.5) by standard arguments.

Then we can proceed recursively for $k = 2, \dots, \tau^{-1}$ to construct successively q_τ^2, θ_τ^2, q_τ^3, etc. $\square$

We will assume the Rayleigh dissipative energy R coercive in the sense that, for some a_1, $a_2 > 0$ and all $q_i = (u_i, m_i) \in Q$, $i = 1, 2$,

$$R(q_1 - q_2) \geq a_1\|u_1 - u_2\|^2_{H_0^1(\Omega)} + a_2\|\varphi_2 \bullet (m_1 - m_2)\|^2_{L^2(\Omega)}. \tag{3.7}$$

In other words, (3.7) makes the system dumped enough (by the way, more than it was necessary in [13]). Note that always $\varphi_2 \bullet m \in L^2(\Omega)$ because of (1.3); hence (3.7) is actually realizable. Note that (2.2) implies R is bounded on Q in the sense that, for some $C < +\infty$ and all $q \in Q$,

$$R(q) \leq C(1 + |||q|||^2) \tag{3.8}$$

holds because

$$\|S^* i(u)\|_{L^2(I\times\Omega)} = \sup_{\|\xi\|_{L^2(I\times\Omega)}\le 1} \langle i(u), S\xi\rangle$$
$$\le \|(i,x)\mapsto d(i,x,\nabla u(x))\|_{L^2(I\times\Omega)} \le C(1+\|u\|_{H_0^1(\Omega)})$$

for any $u \in H_0^1(\Omega)$. For $a \in L^\infty(\Omega)$, let us still abbreviate $\underline{a} = \operatorname{ess\,inf}_{x\in\Omega} a(x)$ and $\overline{a} = \operatorname{ess\,sup}_{x\in\Omega} a(x)$. Moreover, the superscript "I" used, e.g., in (3.13) denotes a piecewise linear continuous interpolation in time.

PROPOSITION 3.2. Let (1.2)–(1.4), (2.2), and (3.7) be valid, $\underline{c}_0 > 0$, $\underline{\lambda} > 0$, $\theta_0 \in L^\infty(\Omega)$, and $(\theta_\tau, q_\tau) = (\theta_\tau, u_\tau, m_\tau)$ be an approximate solution due to Definition 3.2. Then the following a priori estimates hold with C_i independent of τ:

$$\max_{t\in[0,1]} |||q_\tau(t)||| \le C_0, \tag{3.9}$$

$$\|\dot{u}_\tau\|_{L^\infty(0,1;L^2(\Omega))\cap L^2(0,1;H_0^1(\Omega))} \le C_1, \tag{3.10}$$

$$\|S^*\dot{q}_\tau\|_{L^2((0,1)\times I\times\Omega)} \le C_2, \tag{3.11}$$

$$\|\theta_\tau\|_{L^\infty(0,1;H^1(\Omega))\cap H^1(0,1;L^2(\Omega))} \le C_3, \tag{3.12}$$

$$\|\varrho\dot{u}_\tau^{\mathrm{I}}\|_{H^1(0,1;H^{-1}(\Omega))} \le C_4. \tag{3.13}$$

Proof. We can put $\tilde{q} = q_\tau^{k-1} \in Q$ into (3.4) and $z = \theta_\tau^k$ into (3.5), which gives

$$\begin{aligned}&\widehat{T}(q_\tau^k - 2q_\tau^{k-1} + q_\tau^{k-2}, q_\tau^k - q_\tau^{k-1}) + 2\tau R(q_\tau^k - q_\tau^{k-1}) + \tau^2\langle q_\tau^k - q_\tau^{k-1}, \phi_{01}\rangle \\ &\quad\le \tau^2\langle q_\tau^{k-1} - q_\tau^k, \alpha_0\theta_\tau^{k-1}\phi_2\rangle\end{aligned} \tag{3.14}$$

and

$$\langle c_0\theta_\tau^k - c_0\theta_\tau^{k-1}, \theta_\tau^k\rangle + \tau\langle\lambda\nabla\theta_\tau^k, \nabla\theta_\tau^k\rangle = \langle\alpha_0\theta_0\theta_\tau^k, \varphi_2\bullet(m_\tau^k - m_\tau^{k-1})\rangle. \tag{3.15}$$

We divide (3.14) by τ^2 and then sum it with Eq. (3.15). To estimate the left-hand side, we use the obvious estimates $\widehat{T}(q_\tau^k - 2q_\tau^{k-1} + q_\tau^{k-2}, q_\tau^k - q_\tau^{k-1}) \ge T(q_\tau^k - q_\tau^{k-1}) - T(q_\tau^{k-1} - q_\tau^{k-2})$ and $\langle c_0\theta_\tau^k - c_0\theta_\tau^{k-1}, \theta_\tau^k\rangle \ge \frac{1}{2}\langle c_0\theta_\tau^k, \theta_\tau^k\rangle - \frac{1}{2}\langle c_0\theta_\tau^{k-1}, \theta_\tau^{k-1}\rangle$, while for the right-hand side we use the estimates

$$\begin{aligned}\langle q_\tau^{k-1} - q_\tau^k, \alpha_0\theta_\tau^{k-1}\phi_2\rangle &= \langle\alpha_0\theta_\tau^{k-1}, \varphi_2\bullet(m_\tau^{k-1} - m_\tau^k)\rangle \\ &\le \frac{\tau\overline{\alpha}_0^2}{2\varepsilon}\|\theta_\tau^{k-1}\|_{L^2(\Omega)}^2 + \frac{\tau\varepsilon}{a_2}R\left(\frac{q_\tau^k - q_\tau^{k-1}}{\tau}\right)\end{aligned}$$

and similarly

$$\langle\alpha_0\theta_0\theta_\tau^k, \varphi_2\bullet(m_\tau^k - m_\tau^{k-1})\rangle \le \frac{\tau\overline{\alpha}_0^2}{2\varepsilon}\|\theta_0\|_{L^\infty(\Omega)}^2\|\theta_\tau^k\|_{L^2(\Omega)}^2 + \frac{\tau\varepsilon}{a_2}R\left(\frac{q_\tau^k - q_\tau^{k-1}}{\tau}\right).$$

Then we take $\varepsilon < a_2$ to absorb both the terms $\tau\varepsilon a_2^{-1}R((q_\tau^k - q_\tau^{k-1})/\tau)$ in the left-hand side and then use the discrete Gronwall inequality to treat the remaining terms. This gives eventually

$$\|\theta_\tau^k\|_{L^2(\Omega)}^2 + T\left(\frac{q_\tau^k - q_\tau^{k-1}}{\tau}\right) + \langle q_\tau^k, \phi_{01}\rangle + \sum_{i=1}^k \tau\left(\|\nabla\theta_\tau^i\|_{L^2(\Omega)}^2 + R\left(\frac{q_\tau^i - q_\tau^{i-1}}{\tau}\right)\right) \le C$$

for all $k = 1, \dots, \tau^{-1}$ with some C depending on $\underline{c}_0$, $\underline{\lambda}$, $\|\theta_0\|^2_{L^\infty(\Omega)}$, $\overline{\alpha}_0$, $T(p_0)$, and $\langle q_0, \phi_{01}\rangle$. This gives immediately the estimate (3.11) and the first part of (3.10), and also (3.9) when taking (1.2) into consideration.

The rest of (3.10) follows from (3.7) with (3.11): $\|\dot{u}_\tau\|^2_{L^2(0,1;H^1_0(\Omega))} \le a_1^{-1}\int_0^1 R(\dot{q}_\tau)\,dt \le (2a_1)^{-1}C_2^2 \le C_1$.

The estimate (3.12) will be obtained by putting $z = \theta^k_\tau - \theta^{k-1}_\tau$ into Eq. (3.5), which yields

$$\left\|\sqrt{c_0}\frac{\theta^k_\tau - \theta^{k-1}_\tau}{\tau}\right\|^2_{L^2(\Omega)} + \frac{1}{2\tau}\|\sqrt{\lambda}\nabla\theta^k_\tau\|^2_{L^2(\Omega)} - \frac{1}{2\tau}\|\sqrt{\lambda}\nabla\theta^{k-1}_\tau\|^2_{L^2(\Omega)}$$
$$\le \left\langle \alpha_0\theta_0\left(\varphi_2 \bullet \frac{m^k_\tau - m^{k-1}_\tau}{\tau}\right), \frac{\theta^k_\tau - \theta^{k-1}_\tau}{\tau}\right\rangle$$
$$\le \frac{\varepsilon\overline{\alpha}_0^2}{2}\|\theta_0\|^2_{L^\infty(\Omega)}\left\|\frac{\theta^k_\tau - \theta^{k-1}_\tau}{\tau}\right\|^2_{L^2(\Omega)} + \frac{1}{\varepsilon a_2}R\left(\frac{q^k_\tau - q^{k-1}_\tau}{\tau}\right).$$

For $0 < \varepsilon < 2\underline{c}_0(\overline{\alpha}_0\|\theta_0\|_{L^\infty(\Omega)})^{-2}$ we can absorb the first right-hand-side term, and by summing successively for k and using (3.11), which implies $\int_0^t R(\dot{q}_\tau)\,dt \le \frac{1}{2}C_2^2$, we get the estimate of $\dot{\theta}_\tau$ in $L^2((0,1)\times\Omega)$ and of $\nabla\theta_\tau$ in $L^\infty(0,1;L^2(\Omega))$.

We go on to (3.13). Let us take some $z \in H^1_0(\Omega)$ and put $\tilde{q} = i(u^k_\tau - z)$ into (3.4); for the imbedding $i: H^1_0(\Omega) \to Q$, see Sec. 0. Realizing that $\frac{\partial}{\partial t}\dot{u}^{\mathrm{I}}_\tau$ is piecewise constant in time with $\sqrt{\varrho}\frac{\partial}{\partial t}\dot{u}^{\mathrm{I}}_\tau = \tau^{-2}S_0^*(q^k_\tau - 2q^{k-1}_\tau + q^{k-2}_\tau)$ for $t \in ((k-1)\tau, k\tau)$ (see Eq. (2.1)) and using the identity $\sqrt{\varrho}z = S_0^*(i(u^k_\tau - z) - q^k_\tau)$, we get the estimate

$$\left\langle \varrho\frac{\partial}{\partial t}\dot{u}^{\mathrm{I}}_\tau, z\right\rangle_{L^2(\Omega)} \le \widehat{R}(\dot{q}_\tau, q^S_\tau - \tilde{q}) + \langle q^S_\tau - \tilde{q}, \phi_{01} + \alpha_0\theta^S_{-\tau}\phi_2\rangle$$

where q^S_τ is the step function corresponding to q_τ; this means $q^S_\tau(t) = q_\tau(k\tau)$ for $t \in ((k-1)\tau, k\tau)$, and $\theta^S_{-\tau}$ denotes the retarded step function corresponding to θ_τ, that means $\theta^S_{-\tau}(t) = \theta^S_\tau(t-\tau)$ with $\theta^S_{-\tau}(t) = \theta_0$ for $0 < t < \tau$. The rest is nearly the same as in [13, Sec. 3] provided we use additionally an estimate of $\langle q^S_\tau - \tilde{q}, \alpha_0\theta^S_{-\tau}\phi_2\rangle$: in view of (1.3) we can see that

$$|\langle i(u), \alpha_0\theta\phi_2\rangle| = \left|\int_\Omega \alpha_0\theta\varphi_2(\nabla u)\,dx\right| \le c\left(1 + \|\theta\|^2_{L^2(\Omega)} + \|u\|^2_{H^1_0(\Omega)}\right)$$

with some c depending on $\overline{a}_0$ and C and b from (1.3), and therefore by continuity we have $|\langle q, \alpha_0\theta\phi_2\rangle| \le c(1 + \|\theta\|^2_{L^2(\Omega)} + |||q|||^2)$ for any $q \in Q$, which enables the estimate

$$\langle q^S_\tau - \tilde{q}, \alpha_0\theta^S_{-\tau}\phi_2\rangle \le c(2 + 2\|\theta^S_{-\tau}\|^2_{L^2(\Omega)} + |||q^S_\tau|||^2 + \|u^S_\tau - z\|^2_{H^1_0(\Omega)}$$
$$\le c(2 + 2C_3^2 + 3C_0^2 + 2\|z\|^2_{H^1_0(\Omega)}). \quad \square$$

Now we will show the convergence of the approximate solutions (θ_τ, q_τ). We denote

$$L^\infty_{\mathrm{w}}(0, 1; P) = \{q: [0, 1] \to P; \forall f \in F: \langle q(\cdot), f\rangle \in L^\infty(0, 1)\},$$

endowed with the topology induced projectively via all the mappings $q \mapsto \langle q(\cdot), f\rangle$ from the weak* topology of $L^\infty(0, 1)$.

PROPOSITION 3.3. Let the assumptions (1.2)–(1.4), (2.2), and (3.7) be fulfilled, $\theta_0 \in L^\infty(\Omega)$, $\underline{c}_0 > 0$, $\underline{\lambda} > 0$, and $\underline{\varrho} > 0$, and the measure on I have a countable base in the sense of [8, Sec. 52]. Then there exists a subsequence $\{(\theta_\tau, q_\tau)\}_{\tau>0} \equiv \{(\theta_\tau, u_\tau, m_\tau)\}_{\tau>0}$ (denoted by the same indices, for simplicity) and a cluster point $q = (u, m)$ in $L^\infty_{\mathrm{w}}(0, 1; P)$ of the sequence $\{q_\tau\}_{\tau>0}$ such that Eq. (2.15) is fulfilled and

$$\langle q_\tau(\cdot), \phi_{01}\rangle \to \langle q(\cdot), \phi_{01}\rangle \qquad \text{weakly* in } L^\infty(0, 1), \tag{3.16}$$

$$\dot u_\tau \to \dot u \qquad \text{weakly* in } L^\infty(0, 1; L^2(\Omega)), \tag{3.17}$$

$$S^*\dot q_\tau \to S^*\dot q \qquad \text{weakly in } L^2((0, 1)\times I\times\Omega), \tag{3.18}$$

$$\varphi_2 \bullet m_\tau \to \varphi_2 \bullet m \qquad \text{weakly in } H^1(0, 1; L^2(\Omega)), \tag{3.19}$$

$$\theta_\tau \to \theta \qquad \text{weakly* in } L^\infty(0, 1; H^1(\Omega)) \cap H^1(0, 1; L^2(\Omega)), \tag{3.20}$$

and every (θ, q) thus obtained is the weak solution due to Definition 3.1.

Proof. It suffices simply to modify the proof from [13, Theorem 4.1]. Since (3.20) follows immediately from (3.12) and the convergence in the terms $c_0\dot\theta_\tau$ and $\nabla\cdot(\lambda\nabla\theta_\tau)$ in the heat transfer equation is standard, we must only prove (3.19) and the convergence of the terms $\int_0^1 \langle \tilde q - q^S_\tau, \alpha_0\theta^S_{-\tau}\varphi_2\rangle\,dt$ and $\int_0^1\int_\Omega \alpha_0\theta_0 z[\varphi_2 \bullet \dot m_\tau]\,dx\,dt$ resulting from our semi-implicit discretization. As for (3.18), let us only remark that the countable base of the measure on I ensures $L^2(I)$ (and thus also $L^2((0, 1)\times I\times\Omega)$) to be separable (cf. [8, Sec. 52]), which ensures metrizability of the weak topology restricted on bounded subsets.

As for (3.19), let us note that (3.7) and (3.11) imply $\|\varphi_2 \bullet \dot m_\tau\|_{L^2((0,1)\times\Omega)} \le (2a_2)^{-1/2}C_2$; therefore, we may suppose that the subsequence has been chosen so that

$$\varphi_2 \bullet m_\tau \to y \quad \text{weakly in } H^1(0, 1; L^2(\Omega)). \tag{3.21}$$

We want to show $y = \varphi_2\bullet m$. Let us take some $\xi \in L^2(\Omega)$, hence $\xi\varphi_2 \in \mathrm{Car}(\Omega; \mathbb{R}^n)$. By the definition of the topology of $L^\infty_{\mathrm{w}}(0, 1; P)$, $\int_0^1 \eta(t)\langle m(t), \xi\varphi_2\rangle\,dt$ is a cluster point of the sequence of $\int_0^1 \eta(t)\langle m_\tau(t), \xi\varphi_2\rangle\,dt$ for every $\eta \in L^1(0, 1)$. Obviously, $\int_0^1 \eta(t)\langle m_\tau(t), \xi\varphi_2\rangle\,dt = \int_0^1\int_\Omega \eta(t)\xi(x)[\varphi_2 \bullet m_\tau(t)](x)\,dx\,dt$. Since the linear hull of the collection $\{\eta\xi\}$ is dense in $L^1(0, 1; L^2(\Omega))$, we have $\varphi_2 \bullet m$ as a weak* cluster point of $\varphi_2 \bullet m_\tau$ in $L^\infty(0, 1; L^2(\Omega))$. Comparing it with (3.21) yields $y = \varphi_2 \bullet m$ and hence (3.19).

By the estimate $\|\varphi_2 \bullet m_\tau^S - \varphi_2 \bullet m_\tau\|_{L^2((0,1)\times\Omega)} = \sqrt{\tau/3}\|\varphi_2 \bullet \dot{m}_\tau\|_{L^2((0,1)\times\Omega)} \le \sqrt{\tau/(6a_2)}C_2$, we can see that $\varphi_2 \bullet m_\tau^S$ converges weakly in $L^2((0,1)\times\Omega)$ to the same limit as $\varphi_2 \bullet m_\tau$, that means to $\varphi_2 \bullet m$. In view of (3.19), $\theta_\tau \to \theta$ strongly in $L^2((0,1)\times\Omega)$, and evidently also $\theta_{-\tau}^S$ converges to the same limit because $\|\theta_{-\tau}^S - \theta_\tau\|_{L^2((0,1)\times\Omega)} = \sqrt{\tau/3}\|\dot{\theta}_\tau\|_{L^2((0,1)\times\Omega)} \le \sqrt{\tau/3}C_3$. All this allows the limit passage $\int_0^1 \langle \tilde{q} - q_\tau^S, \alpha_0\theta_{-\tau}^S\phi_2\rangle\, dt = \langle \alpha_0\theta_{-\tau}^S, \varphi_2 \bullet (\tilde{m} - m_\tau^S)\rangle \to \langle \alpha_0\theta, \varphi_2 \bullet (\tilde{m} - m)\rangle = \int_0^1 \langle \tilde{q} - q, \alpha_0\theta\phi_2\rangle\, dt$.

The remaining limit passage $\int_0^1 \int_\Omega \alpha_0\theta_0 z[\varphi_2 \bullet \dot{m}_\tau]\, dx\, dt \to \int_0^1 \int_\Omega \alpha_0\theta_0 z[\varphi_2 \bullet \dot{m}]\, dx\, dt$ is a direct consequence of (3.19). □

Let us still note that joining Propositions 3.1 and 3.3 gives the existence of a weak solution of the partially linearized model due to Definition 3.1.

Acknowledgment. The second author gratefully acknowledges the support of this research by the Alexander von Humboldt Foundation during his stay at the Institute for Mathematics at the University of Augsburg.

References

[1] H. W. Alt, K.-H. Hoffman, M. Niezgódka, and J. Sprekels, *A numerical study of structural phase transitions in shape memory alloys*, Preprint No. 90, Institut für Mathematik, Universität Augsburg, 1985

[2] P. Colli, M. Frémond, and A. Visintin, *Thermo-mechanical evolution of shape memory alloys*, Quart. Appl. Math. **48**, 31–47 (1990)

[3] C. M. Dafermos, *On the existence and the asymptotic stability of solutions to the equations of linear thermoelasticity*, Arch. Rational Mech. Anal. **29**, 241–271 (1968)

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

[5] F. Falk, *Landau theory and martensitic phase transitions* (L. Delaey and M. Chandrasekaran, eds.), Proc. Internat. Conf. on Martensitic Transformations, Les Editions de Physique, Les Ulis, 1982

[6] M. Frémond, *Matériaux à mémoire de forme*, C. R. Acad. Sci. Paris Sér. II Méc. Phys. Chim. Sci. Univers. Sci. Terre **304**, 239–244 (1987)

[7] K.-H. Hoffmann, M. Niezgódka, and Zheng Songmu, *Existence and uniqueness of global solutions to an extended model of the dynamical development in shape memory alloys*, Nonlinear Anal. **15**, 977–990 (1990)

[8] A. N. Kolmogorov and S. V. Fomin, *Elements of the Theory of Functions and Functional Analysis*, Graylock Press, Baltimore, 1961

[9] J. Nečas, *Dynamic in the nonlinear thermo-visco-elasticity*, Symposium "Partial Differential Equations" Holzhau 1988 (B.-W. Schulze and H. Triebel, eds.), Teubner-Texte Math. **112**, Teubner, Leipzig, 1989, pp. 197–203

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

[11] R. E. Nickell and J. L. Sackman, *Variational principles for linear coupled thermoelasticity*, Quart. Appl. Math. **26**, 11–26 (1968)

[12] M. Niezgódka and J. Sprekels, *Existence of solutions for a mathematical model of structural phase transitions in shape memory alloys*, Math. Methods Appl. Sci. **10**, 197–223 (1988)

[13] T. Roubíček, *Evolution of a microstructure: a convexified model*, Math. Methods Appl. Sci. **16**, 625–642 (1993)

[14] T. Roubíček, *Optimality conditions for nonconvex variational problems relaxed in terms of Young measures* (submitted)

[15] T. Roubíček, *Finite element approximation of a microstructure evolution*, Math. Methods Appl. Sci. **17**, 377–393 (1994)

[16] M. Schatzman, *A hyperbolic problem of second order with unilateral constraints: the vibrating string with a concave obstacle*, J. Math. Anal. Appl. **73**, 138–191 (1980)

[17] J. Sprekels, *Global existence for thermomechanical processes with nonconvex free energies of Ginzburg-Landau form*, J. Math. Anal. Appl. **141**, 333–348 (1989)