COUPLING THE NAVIER–STOKES–FOURIER EQUATIONS WITH THE JOHNSON–SEGALMAN STRESS-DIFFUSIVE VISCOELASTIC MODEL: GLOBAL-IN-TIME AND LARGE-DATA ANALYSIS

MICHAL BATHORY, MIROSLAV BULÍČEK, AND JOSEF MÁLEK

Abstract. We prove that there exists a large-data and global-in-time weak solution to a system of partial differential equations describing an unsteady flow of an incompressible heat-conducting rate-type viscoelastic stress-diffusive fluid filling up a mechanically and thermally isolated container of any dimension. To overcome the principle difficulties connected with ill-posedness of the diffusive Oldroyd-B model in three dimensions, we assume that the fluid admits a strengthened dissipation mechanism, at least for excessive elastic deformations. All the relevant material coefficients are allowed to depend continuously on the temperature, whose evolution is captured by a thermodynamically consistent equation. In fact, the studied model is derived from scratch using only the balance equations for linear momentum and energy, the formulation of the second law of thermodynamics and the constitutive equation for the internal energy. The latter is assumed to be a linear function of temperature, which simplifies the model. The concept of our weak solution incorporates both the temperature and entropy inequalities, and also the local balance of total energy provided that the pressure function exists.

1. Introduction

Material properties of both synthetic and organic viscoelastic materials are very sensitive to temperature changes. Reliable predictions of corresponding processes by computational tools requires to incorporate complex thermal/mechanical effects into the description of the model. The understanding how thermal and mechanical processes are coupled and what is the structure of the complete temperature equation has been considered as an open issue till recently (see [54, 31]). A methodology that can be used to develop such a complete model (system of partial differential equations – PDEs) and that is followed in this study has its origin in [50, 51]. A complete (i.e. including elastic contribution to the internal energy) thermodynamicly consistent models for viscoelastic rate type fluids is developed in [31] where also further references to earlier studies including in particular [36, 56, 22, 32] are given. Incorporation of additional stress diffusive phenomena into this thermodynamic framework is then developed in [42].

2020 *Mathematics Subject Classification.* 35A23, 76A10, 76D03.

Key words and phrases. viscoelastic heat-conducting fluids, Johnson–Segalman, weak solution.

Michal Bathory is recipient of an APART-MINT Fellowship (No. 11976) of the Austrian Academy of Sciences (ÖAW) and also acknowledges support from the project No. 1652119 financed by the Charles University Grant Agency (GAUK). Michal Bathory, Miroslav Bulíček and Josef Málek thank to the project No. 20-11027X financed by the Czech Science foundation (GAČR). Miroslav Bulíček and Josef Málek are members of the Nečas Center for Mathematical Modelling.

The aim of this study is to establish mathematical foundation for a robust class of heat-conducting viscoelastic rate-type fluids with stress diffusion. In particular, we identify reasonable conditions on material functions/coefficients that are sufficient to prove long-time and large-data existence of weak solution. To develop analysis for complete thermal/mechanical systems of PDEs is considerably harder than to studying merely mechanical systems. To our best knowledge, there is only one existing analytical work dealing with such a problem, see [13], where however the elastic response is drastically reduced to a spherical stress governed by a scalar quantity. In our work, we do not make such an assumption and we work with the full d-dimensional elastic tensor. On the other hand, we assume that there is a linear relation between the internal energy and temperature. The main purpose for this assumption is to simplify the (already very technical) presentation of the existence analysis. Additionally, the linear relationship between the internal energy and temperature is used in applications involving viscoelastic fluids, such as polymer melts, see [52]. Taking this aside, our model contains no further simplification. A complete physical derivation of the model studied in this paper and a more detailed description of the participating physical quantities are given in Section 2. This opening section continues below with an informal formulation of the main result, a brief description of the PDE system and a basic overview of the relevant literature. In Section 3, we introduce necessary notation, derive informally a priori estimates that naturally leads to the definition of function spaces in which the existence theory is established. This section also contains precise definition of the solution to studied problem and the formulation of the main result. Its proof represents the content of remaining sections of the paper, see Sections 4–6. Their more detailed description is given at the end of Section 3.

Formulation of the problem. We consider an incompressible fluid with the constant density set to be one, for simplicity. The fluid is flowing inside an open bounded connected set $\Omega \subset \mathbb{R}^d$ with a Lipschitz boundary $\partial\Omega$. For an arbitrary (but fixed) time interval $(0,T)$, $T>0$, we set $Q := (0,T)\times\Omega$ and $\Sigma := (0,T)\times\partial\Omega$. Our main objective, in this study, is to develop a long-time and large-data existence theory for the following initial- and boundary-value problem.

For given

- right-hand side $\boldsymbol{g} : Q \to \mathbb{R}^d$,
- initial data $\boldsymbol{v}_0 : \Omega \to \mathbb{R}^d$, $\mathbb{B}_0 : \Omega \to \mathbb{R}^{d\times d}_{\mathrm{sym}}$ being positive definite and $\theta_0 : \Omega \to (0,\infty)$,
- constants $a \in \mathbb{R}$, $\alpha \geq 0$, $\mu > 0$ and $c_v > 0$,
- continuous functions $\nu, \lambda, \kappa : (0,\infty) \to (0,\infty)$ and $\mathbb{P} : (0,\infty)\times\mathbb{R}^{d\times d}_{\mathrm{sym}} \to \mathbb{R}^{d\times d}_{\mathrm{sym}}$,

we look for functions $\boldsymbol{v} : Q \to \mathbb{R}^d$, $\mathrm{p}, \theta, e, E, \eta, \xi : Q \to \mathbb{R}$ and $\mathbb{B}, \mathbb{S} : Q \to \mathbb{R}^{d\times d}_{\mathrm{sym}}$ fulfilling the (physical) restrictions

$$\theta > 0, \tag{1.1}$$

$$\mathbb{B}\boldsymbol{x}\cdot\boldsymbol{x} > 0 \quad \text{for all } \boldsymbol{x} \in \mathbb{R}^d \setminus \{0\}, \tag{1.2}$$

$$\mathbb{S} = 2\nu(\theta)\mathbb{D}\boldsymbol{v} + 2a\mu\theta\mathbb{B}, \tag{1.3}$$

$$e = c_v\theta, \tag{1.4}$$

$$E = \tfrac{1}{2}|\boldsymbol{v}|^2 + e, \tag{1.5}$$

$$\eta = c_v \ln\theta - f(\mathbb{B}), \quad \text{where} \quad f(\mathbb{B}) \coloneqq \mu(\operatorname{tr}\mathbb{B} - d - \ln\det\mathbb{B}), \tag{1.6}$$

$$\xi = \frac{2\nu(\theta)}{\theta}|\mathbb{D}\boldsymbol{v}|^2 + \kappa(\theta)|\nabla\ln\theta|^2 + \mathbb{P}(\theta,\mathbb{B})\cdot f'(\mathbb{B}) + \lambda(\theta)\nabla\mathbb{B}\cdot\nabla f'(\mathbb{B}), \tag{1.7}$$

and solving (in a suitable sense) the following system of PDEs in Q

$$\operatorname{div}\boldsymbol{v} = 0, \tag{1.8}$$

$$\partial_t\boldsymbol{v} + \boldsymbol{v}\cdot\nabla\boldsymbol{v} + \nabla\mathrm{p} - \operatorname{div}\mathbb{S} = \boldsymbol{g}, \tag{1.9}$$

$$\partial_t\mathbb{B} + \boldsymbol{v}\cdot\nabla\mathbb{B} + \mathbb{P}(\theta,\mathbb{B}) - \operatorname{div}(\lambda(\theta)\nabla\mathbb{B}) = \mathbb{W}\boldsymbol{v}\mathbb{B} - \mathbb{B}\mathbb{W}\boldsymbol{v} + a(\mathbb{D}\boldsymbol{v}\mathbb{B} + \mathbb{B}\mathbb{D}\boldsymbol{v}), \tag{1.10}$$

$$\partial_t e + \boldsymbol{v}\cdot\nabla e - \operatorname{div}(\kappa(\theta)\nabla\theta) = \mathbb{S}\cdot\mathbb{D}\boldsymbol{v}, \tag{1.11}$$

$$\partial_t E + \boldsymbol{v}\cdot\nabla E - \operatorname{div}(\kappa(\theta)\nabla\theta) = \operatorname{div}(-\mathrm{p}\boldsymbol{v} + \mathbb{S}\boldsymbol{v}) + \boldsymbol{g}\cdot\boldsymbol{v}, \tag{1.12}$$

$$\partial_t\eta + \boldsymbol{v}\cdot\nabla\eta - \operatorname{div}(\kappa(\theta)\nabla\ln\theta) + \operatorname{div}(\lambda(\theta)\nabla f(\mathbb{B})) = \xi \tag{1.13}$$

completed by the boundary conditions on Σ

$$\boldsymbol{v}\cdot\boldsymbol{n} = 0, \qquad (\mathbb{S}\boldsymbol{n} + \alpha\boldsymbol{v})_\tau = 0, \tag{1.14}$$

$$\boldsymbol{n}\cdot\nabla\mathbb{B} = 0, \tag{1.15}$$

$$\boldsymbol{n}\cdot\nabla\theta = 0, \tag{1.16}$$

and by the initial conditions fulfilled in Ω

$$\boldsymbol{v}(0,\cdot) = \boldsymbol{v}_0, \qquad \mathbb{B}(0,\cdot) = \mathbb{B}_0, \qquad \theta(0,\cdot) = \theta_0. \tag{1.17}$$

The physical meaning of the above unknowns is the following: $\boldsymbol{v}$ is the flow velocity, p is the pressure, $\mathbb{B}$ is the extra stress tensor (arising due to the elastic properties of the fluid), θ is the temperature, e is the internal energy, E is the total energy and η is the entropy. We shall now state informally our main result.

Main result. *If the material coefficients $\kappa(\theta)$ and $P(\cdot,\mathbb{B})$ grow sufficiently fast as $\theta\to\infty$ and $|\mathbb{B}|\to\infty$, respectively (with the other coefficients being merely bounded and positive), then there exists a generalized global-in-time solution of the system (1.1)–(1.17) for any initial data with finite total energy and entropy.*

In order to explain the equations above, let us first clarify some notation, see also the beginning of Section 3. The symbol $\boldsymbol{v}\cdot\nabla\boldsymbol{v}$ denotes a vector with the i-component $(\boldsymbol{v}\cdot\nabla\boldsymbol{v})_i = \sum_{k=1}^d v_k\partial_{x_k}v_i$. Similarly, $\boldsymbol{v}\cdot\nabla\mathbb{B}$ is a tensor with the ij-component $(\boldsymbol{v}\cdot\nabla\mathbb{B})_{ij} = \sum_{k=1}^d v_k\partial_{x_k}(\mathbb{B})_{ij}$. The first two terms of each equation (1.9)–(1.13) represent the material (or convective) derivative of the respective unknown and we shall sometimes use the abbreviation

$$\overset{\bullet}{u} \coloneqq \partial_t u + \boldsymbol{v}\cdot\nabla u.$$

Further, the symbol $\boldsymbol{n}$ denotes the outward unit normal vector at a given point of $\partial\Omega$ and $\boldsymbol{z}_\tau$ stands for the tangential part (with respect to $\partial\Omega$) of any vector $\boldsymbol{z}\in\mathbb{R}^d\cap\partial\Omega$, i.e. $\boldsymbol{z}_\tau \coloneqq \boldsymbol{z} - (\boldsymbol{z}\cdot\boldsymbol{n})\boldsymbol{n}$. Furthermore, for any vector $\boldsymbol{u}:\Omega\to\mathbb{R}^d$, the symbols $\mathbb{D}\boldsymbol{u}$ and $\mathbb{W}\boldsymbol{u}$ denote the symmetric and antisymmetric parts of a gradient $\nabla\boldsymbol{u} = (\partial_j\boldsymbol{u}_i)_{i,j=1}^d$ so that $\nabla\boldsymbol{u} = \mathbb{D}\boldsymbol{u} + \mathbb{W}\boldsymbol{u}$ with $(\mathbb{D}\boldsymbol{u})^T = \mathbb{D}\boldsymbol{u}$ and $(\mathbb{W}\boldsymbol{u})^T = -\mathbb{W}\boldsymbol{u}$.

The first two equations (1.8) and (1.9) resemble the incompressible Navier–Stokes system for the unknowns velocity field $\boldsymbol{v}$ and the pressure (constitutively undetermined part of the Cauchy stress) p, however, with an additional term $2a\mu\operatorname{div}(\theta\mathbb{B})$ coming from $\mathbb{S}$ and bringing to the problem two other quantities: the temperature θ and the tensor $\mathbb{B}$ representing the elastic response of the fluid. The presence of

this additional term prohibits one to use the usual methods known in the analysis of the Navier-Stokes-Fourier-like systems, as there is no longer an useful form of the balance of kinetic energy (the inner product $\mathbb{B}\cdot\mathbb{D}\boldsymbol{v}$ does not have a sign). Instead, the estimates on $\nabla\boldsymbol{v}$ are deduced only after taking into account the whole thermodynamical evolution of the system.

Since the dependence of the material parameters (namely the viscosity of the fluid) on the pressure p is neglected, we simplify the analysis by eliminating the pressure from the system completely, taking the Leray projection of (1.9) and searching for $\boldsymbol{v}$ in divergence-free function spaces. If needed (for example if we want to preserve the equation (1.12)), the pressure can be reconstructed at the last step. Then, it is known that Navier's slip boundary condition (1.14), or even more generally stick-slip boundary condition, allows one to prove that p is an integrable function (if the boundary of Ω is smooth enough so that $W^{2,r}$-regularity for the classical Neumann problem holds), see [15, 6, 11, 12, 5] for details. Recall that the integrability of the pressure is not known to be true in general for no-slip boundary condition. The integrability of p is not only important in itself, but is also useful for the validity of the weak formulation of (1.12).

To understand equation (1.10), it is better to define first the objective derivative of $\mathbb{B}$ as

$$\mathring{\mathbb{B}} \coloneqq \overset{\bullet}{\mathbb{B}} - (\mathbb{W}\boldsymbol{v}\mathbb{B} - \mathbb{B}\mathbb{W}\boldsymbol{v}) - a(\mathbb{D}\boldsymbol{v}\mathbb{B} + \mathbb{B}\mathbb{D}\boldsymbol{v}), \quad a\in\mathbb{R}. \tag{1.18}$$

This turns (1.10) into

$$\mathring{\mathbb{B}} + \mathbb{P}(\theta,\mathbb{B}) - \operatorname{div}(\lambda(\theta)\nabla\mathbb{B}) = 0, \tag{1.19}$$

which is a mathematical formulation of a generalized (due to an implicit form of P) Johnson–Segalman ([33]) viscoelastic model with stress diffusion (cf. [48] and references therein) and temperature dependent material parameters. The reason why $\mathring{\mathbb{B}}$ appears in (1.19) is that, unlike the material derivative, the objective derivative $\mathring{\mathbb{B}}$ (for any a) transforms correctly (as a tensor) under a time dependent rotation of the observer. When $a\in[-1,1]$, then $\mathring{\mathbb{B}}$ is precisely the Gordon-Schowalter derivative ([28]). It is known (see e.g. [49]) that by modifying the value of a, it is possible to capture a shear-thinning behaviour of the fluid. The case $a=0$ leads to the class of models with the corrotational objective derivative (cf. [59]), which has very special properties that simplify the analysis. The case $a=1$ in (1.18) coincides with the upper-convected objective derivative, which is probably the most popular choice in the literature. One of the main features of our analysis is that, we are able to treat (1.10) with any $a\in[-1,1]$ (or even $a\in\mathbb{R}$). As we shall see later, if $a\neq 0$, the summability of the nonlinear terms like $\mathbb{B}\mathbb{D}\boldsymbol{v}$ in (1.10) (and especially the related term in (1.11)) becomes the main difficulty. This is essentially the reason, why we formulate (1.10) with a general function $\mathbb{P}(\theta,\mathbb{B})$. The strategy is that if $\mathbb{P}(\theta,\mathbb{B})$ grows sufficiently fast as $|\mathbb{B}|\to\infty$, then $\mathbb{B}$ admits enough integrability to define a meaningful concept of solution of the system (1.1)–(1.17). Moreover, as the form of $\mathbb{P}$ can be attributed to the dissipation mechanism of the fluid, restricting its asymptotic growth should not be seen as a significant physical drawback of our model. Recall that, for the classical Oldroyd-B and Giesekus models, the function $\mathbb{P}$ takes the form

$$\mathbb{P}(\theta,\mathbb{B}) = \delta(\theta)(\mathbb{B}-\mathbb{I}) \quad \text{and} \quad \mathbb{P}(\theta,\mathbb{B}) = \delta(\theta)(\mathbb{B}^2-\mathbb{B}),$$

respectively. While these models are not covered by the analysis presented below, the existence result, in three dimensions, holds for

$$\mathbb{P}(\theta,\mathbb{B}) = \delta(\theta)(\mathbb{B}^{\alpha} - \mathbb{B}^{\alpha-1}), \quad \alpha > 2,$$

or

$$\mathbb{P}(\theta,\mathbb{B}) = \delta(\theta)\max\{1, \frac{|\mathbb{B}-\mathbb{I}|^2}{K^2}\}(\mathbb{B}-\mathbb{I}), \quad K > 0. \tag{1.20}$$

Note that the last model coincides with the Oldroyd-B model as long as $|\mathbb{B}-\mathbb{I}| \leq K$.

Due to (1.4), the balance of internal energy (1.11) is also the temperature equation. As we hinted above, the term $2a\mu\theta\mathbb{B}\cdot\mathbb{D}\boldsymbol{v}$ on the right-hand side of (1.11) is the most difficult term to control in the whole system (1.8)–(1.13) and it is also the term which is occasionally omitted in some "naive" approaches to thermoviscoelasticity, as pointed out in [31, Section 3]. Note also that this term does not have a clear sign and thus, one cannot conclude the positivity of temperature directly from (1.11) as in the Navier-Stokes-Fourier case. The equations (1.12) and (1.13) govern the evolution of two other unknowns E and η, respectively. Since these quantities together with θ are mutually connected by simple algebraic relations (1.4), (1.5) and (1.6), the equations (1.11)–(1.13) are interchangeable and each of them alone can be used as the equation for temperature evolution. To see this, note that (1.5) and (1.6) imply

$$\partial_t E = \boldsymbol{v}\cdot\partial_t\boldsymbol{v} + \partial_t e \tag{1.21}$$

$$\partial_t \eta = c_v\theta^{-1}\partial_t\theta - f'(\mathbb{B})\cdot\partial_t\mathbb{B}. \tag{1.22}$$

Within the considered system of equations (assuming that all involved operations are meaningful), one can verify that the equations (1.11), (1.12), (1.13) are mutually equivalent. We remark that this equivalence may no longer be in place when, on the level of generalized solutions, the integrability of the solution is not sufficient to define the critical nonlinear terms in (1.11) and (1.12), that is $\theta\mathbb{B}\cdot\mathbb{D}\boldsymbol{v}$ and $\theta\mathbb{B}\boldsymbol{v}$, respectively. For example, this would be the case where the initial datum $\mathbb{B}_0$ has low integrability, as then the available a priori estimates deteriorate (cf. (3.15) below). In such cases, one may be forced to discard (1.11), or even (1.12) from the notion of generalized solution and leave only (1.13), which is least restrictive but still sufficient (together with the global version of (1.12)) to keep track of the thermal evolution of the system. Generalized solutions relying on the weak formulation of balance of entropy were applied, e.g., in [25], [8], [23], [26] or in [24] for different fluid models. See also [9] for similar ideas in context of certain mixtures. For brevity, in this work, we shall avoid the low integrability case and work only in the setting, where both (1.11) and (1.13) (and (1.12) if the pressure can be defined) hold simultaneously, but only as the inequalities. Although these become automatically equalities if the solution is smooth enough (see (3.48) below), in general this is unknown.

State of the art. Regarding the existence analysis of a viscoelastic fluid model including the full temperature evolution, there is a recent study [14], where the authors develop a long-time and large-data existence theory for a rate-type incompressible viscoelastic fluid model with stress diffusion under the simplifying assumption that $\mathbb{B} = b\mathbb{I}$. This assumption leads to annihilation of irregular terms coming from the objective derivative and it also simplifies the momentum equation, where the coupling to the rest of the system is realized only via temperature and elastic

stress dependent viscosity. Other than that, to the authors' best knowledge, there is no existence theory in a setting that would be of similar generality as considered here. Thus, for the first time, we provide an existence analysis for a viscoelastic fluid model with a full thermal evolution and taking into account all components of the extra stress tensor. Moreover, the equation for the temperature we consider is derived from fundamental thermodynamical laws (similarly as in [14], [31], [42]) and consequently, the heating originates from both the viscous and elastic effects. Also, we would like to point out that the all material coefficients of the model depend on the temperature. Although we place some restrictions on the growth of these coefficients, these are only asymptotic and therefore unimportant from the point of view of physical applications. Furthermore, the model considered here has the property that the evolution of the temperature cannot be decoupled from the rest of the model even in the case of constant material coefficients.

Even if we confine to a much simpler class of isothermal processes, the existence theory there is far from being complete. Although there are several relevant global-in-time existence results for large data, in most cases, they are restricted in an essential way. For example, in [37] the authors provide an existence theory for a model with the corrotational Jaumann–Zaremba derivative (the case $a = 0$). This case is much easier than for the other choices of a since the corrotational part drops out upon multiplication by any matrix that commutes with $\mathbb{B}$. Moreover, it seems that the physically preferred case is $a = 1$, which corresponds to the upper convected (Oldroyd) derivative (see [41], [44], [45], [50] or [51]). Then, in [46], a proof of existence of a weak solution to FENE-P, Giesekus and PTT viscoelastic models is outlined. In fact, it is shown there that certain defect measures of the non-linear terms are compact. A complete proof in the case of two-dimensional flows of a Giesekus fluid is given in [10]. In the case of spherical elastic response when $\mathbb{B} = b\mathbb{I}$, we refer to [13] (and [6], [38] in the compressible case) for an analysis of such models. In the two-dimensional case, existence and regularity results can be found in [20]. An existence theory for related viscoelastic models (Peterlin class) was developed, e.g., in [39]. However, for these models, the energy storage mechanism depends only on the spherical part of the extra stress, which is a major simplification compared to our case. A notable exception is the thesis [34], where the author obtains a global weak solution to an Oldroyd-like diffusive model under certain growth assumptions on the material coefficients. However, the overall thermodynamical compatibility of the studied model is unclear. Furthermore, there are existence results for viscoelastic models involving various approximations that improve properties of the system, see e.g. [2] or [35].

The article [4] develops the existence theory for viscoelastic diffusive Oldroyd-B or Giesekus models. This result relies on a certain physical correction of the energy storage mechanism away from the stress-free state resulting at L^2 a priori estimates for $\nabla\mathbb{B}$. Interestingly, for such models, in two dimensions, uniqueness and full regularity of weak solution is available (at least in the spatially periodic case), see [16]. Various modifications of the classical Oldroyd-B model are also discussed in [18]. The article contains also existence results that are of local nature or for small (initial) data. Local-in-time existence of regular solutions to a viscoelastic Oldroyd-B model without diffusion was shown in [29]. It is also proved there that for small data there exists a global in time solution. For the steady case of a generalized Oldroyd-B model with small and regular data, see e.g. [1].

2. Thermodynamical compatibility of the model

In this section, we show the physical consistency of the system (1.1)–(1.17) as it follows naturally from the elementary balance equations for mass, momentum and energy and some reasonable constitutive assumptions. The latter can be efficiently encoded in just two scalar quantities describing how the energy is stored and dissipated in the material, see [50] and [51] for the origins of this method. Physical justification of viscoelastic fluid models similar to ours is carried out in many works, see [21], [31], [42], [44] or [45].

For the rest of this section, we make an implicit assumption that all functions depend smoothly on time and space position (if not specified otherwise), with the arguments (t, x) suppressed as usual.

Since the density of the fluid is assumed constant ($\varrho = 1$), the balance of mass

$$\dot{\varrho} + \varrho \operatorname{div} \boldsymbol{v} = 0$$

is reduced to (1.8). Next, the general form of the balance equations of momentum, total energy and specific entropy is

$$\dot{\boldsymbol{v}} = \operatorname{div} \mathbb{T}, \tag{2.1}$$

$$\dot{E} + \operatorname{div} \boldsymbol{j}_e = \operatorname{div}(\mathbb{T}\boldsymbol{v}), \tag{2.2}$$

$$\dot{\eta} + \operatorname{div} \boldsymbol{j}_\eta = \xi, \tag{2.3}$$

where $\mathbb{T}$ is the Cauchy stress tensor and $\boldsymbol{j}_e$ and $\boldsymbol{j}_\eta$ are energy and entropy fluxes, respectively. Tensor $\mathbb{T}$ is symmetric due to the conservation of angular momenta. Furthermore, the balance equation for the internal energy $e \coloneqq E - \frac{1}{2}|\boldsymbol{v}|^2$ is

$$\dot{e} + \operatorname{div} \boldsymbol{j}_e = \mathbb{T} \cdot \mathbb{D}\boldsymbol{v}, \tag{2.4}$$

as follows easily from (2.1) and (2.2).

Turning to thermodynamics, we assert the following fundamental relation (cf. [17, (1.8)]) between specific entropy, internal energy and positive definite tensor $\mathbb{B}$:

$$\eta = S(e, \mathbb{B}), \quad \text{where} \quad \partial_e S > 0. \tag{2.5}$$

In this case, the temperature θ is defined as usual by

$$\frac{1}{\theta} \coloneqq \partial_e S(e, \mathbb{B}). \tag{2.6}$$

Taking the material time derivative of both sides of (2.5) then leads to

$$\dot{\eta} = \frac{1}{\theta}\dot{e} + \partial_{\mathbb{B}} S(e, \mathbb{B}) \cdot \dot{\mathbb{B}}.$$

This in turn allows us to express the rate of entropy production in the general form via the balance equations (2.3) and (2.4) as follows:

$$\xi = \frac{1}{\theta}(\mathbb{T} \cdot \mathbb{D}\boldsymbol{v} - \operatorname{div} \boldsymbol{j}_e) + \operatorname{div} \boldsymbol{j}_\eta + \partial_{\mathbb{B}} S(e, \mathbb{B}) \cdot \dot{\mathbb{B}} \tag{2.7}$$

In the next step, we make special choices of $\mathbb{T}$, $\boldsymbol{j}_e$, $\boldsymbol{j}_\eta$ and S that lead to (1.9), (1.11)–(1.13) and verify, using the above formula and also (1.10), that $\xi \geq 0$.

The formula for specific entropy is chosen as

$$S(e, \mathbb{B}) \coloneqq c_v \ln e - f(\mathbb{B}), \tag{2.8}$$

where $c_v > 0$ is the specific heat constant and

$$f(\mathbb{B}) \coloneqq \mu(\operatorname{tr} \mathbb{B} - d - \ln \det \mathbb{B}), \quad \mu > 0, \tag{2.9}$$

is a function that characterizes the elastic properties of the fluid. If $\mu = 0$ or $\mathbb{B} = \mathbb{I}$, then (2.8) reduces to the classical Navier-Stokes-Fourier model, where one has

$$e = c_v \theta. \tag{2.10}$$

Note that as long as μ does not depend on temperature (which is the case in this work), this property actually remains valid even with our generalized assumption (2.8), as is immediately obvious from (2.8), (2.6) and (2.5).

Next, comparing (2.1), (2.4) and (2.3) with (1.9), (1.11) and (1.13), respectively, the constitutive choices for the fluxes are evidently as follows:

$$\mathbb{T} \coloneqq -\mathrm{p}\mathbb{I} + 2\nu(\theta)\mathbb{D}\boldsymbol{v} + 2a\mu\theta\mathbb{B}, \tag{2.11}$$

$$\boldsymbol{j}_e \coloneqq -\kappa(\theta)\nabla\theta, \tag{2.12}$$

$$\boldsymbol{j}_\eta \coloneqq -\kappa(\theta)\nabla \ln\theta + \lambda(\theta)\nabla f(\mathbb{B}), \tag{2.13}$$

where $\nu(\theta) > 0$, $\kappa(\theta) > 0$ and $\lambda(\theta) > 0$ are the kinematic viscosity, thermal conductivity and stress diffusion coefficients, respectively, and parameter a arises from the definition of the objective tensorial time derivative (1.18).

Finally, plugging the relations (2.11)–(2.13) and (1.19) into (2.7) and taking advantage of the identities

$$\begin{aligned}
&-\mathrm{p}\mathbb{I}\cdot\mathbb{D}\boldsymbol{v} = -\mathrm{p}\,\mathrm{div}\,\boldsymbol{v} = 0,\\
&\partial_{\mathbb{B}} S(e,\mathbb{B}) = -f'(\mathbb{B}) = -\mu(\mathbb{I}-\mathbb{B}^{-1}) \qquad \text{(see (A.21) below)},\\
&(\mathbb{I}-\mathbb{B}^{-1})\cdot(\mathbb{W}\boldsymbol{v}\mathbb{B} - \mathbb{B}\mathbb{W}\boldsymbol{v}) = (\mathbb{I}-\mathbb{B}^{-1})\mathbb{B}\cdot\mathbb{W}\boldsymbol{v} - \mathbb{B}(\mathbb{I}-\mathbb{B}^{-1})\cdot\mathbb{W}\boldsymbol{v} = 0,\\
&(\mathbb{I}-\mathbb{B}^{-1})\cdot(\mathbb{D}\boldsymbol{v}\mathbb{B} + \mathbb{B}\mathbb{D}\boldsymbol{v}) = 2(\mathbb{B}-\mathbb{I})\cdot\mathbb{D}\boldsymbol{v} = 2\mathbb{B}\cdot\mathbb{D}\boldsymbol{v},\\
&\nabla(\mathbb{I}-\mathbb{B}^{-1})\cdot\nabla\mathbb{B} = \mathbb{B}^{-1}\nabla\mathbb{B}\mathbb{B}^{-1}\cdot\nabla\mathbb{B} = |\mathbb{B}^{-\frac{1}{2}}\nabla\mathbb{B}\mathbb{B}^{-\frac{1}{2}}|^2
\end{aligned} \tag{2.14–2.15}$$

(here we used that $\mathbb{B}$ is a symmetric positive definite matrix, which follows from the same property of $\mathbb{B}_0$ as we shall see later) leads to

$$\begin{aligned}
\xi &= \frac{1}{\theta}\big(2\nu(\theta)|\mathbb{D}\boldsymbol{v}|^2 + 2a\mu\theta\mathbb{B}\cdot\mathbb{D}\boldsymbol{v} + \mathrm{div}(\kappa(\theta)\nabla\theta)\big) + \mathrm{div}(-\kappa(\theta)\nabla\ln\theta + \lambda(\theta)\nabla f(\mathbb{B}))\\
&\quad - \mu(\mathbb{I}-\mathbb{B}^{-1})\cdot(\mathbb{W}\boldsymbol{v}\mathbb{B} - \mathbb{B}\mathbb{W}\boldsymbol{v} + a(\mathbb{D}\boldsymbol{v}\mathbb{B} + \mathbb{B}\mathbb{D}\boldsymbol{v}) - P(\theta,\mathbb{B}) + \mathrm{div}(\lambda(\theta)\nabla\mathbb{B}))\\
&= \frac{2\nu(\theta)}{\theta}|\mathbb{D}\boldsymbol{v}|^2 + \kappa(\theta)|\nabla\ln\theta|^2 + \mu(\mathbb{I}-\mathbb{B}^{-1})\cdot P(\theta,\mathbb{B}) + \mu\lambda(\theta)|\mathbb{B}^{-\frac{1}{2}}\nabla\mathbb{B}\mathbb{B}^{-\frac{1}{2}}|^2,
\end{aligned}$$

which validates (1.7) and verifies the physical consistency of the model. Moreover, from the last expression, it is evident that $\xi \geq 0$ whenever $(\mathbb{I}-\mathbb{B}^{-1})\cdot P(\theta,\mathbb{B}) \geq 0$, in which case the second law of thermodynamics is always fulfilled.

3. Weak formulation & main result

In this section, we focus on mathematical properties of the problem (1.1)–(1.17). After we introduce the necessary notation, we formally derive a priori estimates that clarify the imposed restrictions on model parameters. They also indicate the functions spaces in which the long-time and large-data existence theory can be established. Then we provide the definition of weak solution to (1.1)–(1.17) and formulate the main result of the paper.

Notation and function spaces. The sets of symmetric, positive definite and positive semi-definite matrices are defined as follows:

$$\begin{aligned}
\mathbb{R}^{d\times d}_{\mathrm{sym}} &:= \{\mathbb{A}\in\mathbb{R}^{d\times d} : \mathbb{A}=\mathbb{A}^T\},\\
\mathbb{R}^{d\times d}_{>0} &:= \{\mathbb{A}\in\mathbb{R}^{d\times d}_{\mathrm{sym}} : \mathbb{A}\boldsymbol{x}\cdot\boldsymbol{x}>0 \text{ for all } 0\neq\boldsymbol{x}\in\mathbb{R}^d\},\\
\mathbb{R}^{d\times d}_{\geq 0} &:= \{\mathbb{A}\in\mathbb{R}^{d\times d}_{\mathrm{sym}} : \mathbb{A}\boldsymbol{x}\cdot\boldsymbol{x}\geq 0 \text{ for all } \boldsymbol{x}\in\mathbb{R}^d\}.
\end{aligned}$$

If $d=1$, we set $\mathbb{R}_{>0}:=\mathbb{R}^{1\times 1}_{>0}=(0,\infty)$ and $\mathbb{R}_{\geq 0}:=\mathbb{R}^{1\times 1}_{\geq 0}=[0,\infty)$. We use the symbol "$\cdot$" to denote the standard inner product in any multi-dimensional space, while the symbol "$\otimes$" denotes the outer product. Further, the symbol "$|\cdot|$" can be applied to either scalars, vectors or matrices, meaning always the Euclidean (or Frobenius) norm. The functions of matrices, such as matrix real powers, matrix logarithm and matrix exponential, are understood in the standard way, using the spectral decomposition for symmetric matrices, for instance. For various products of matrix-valued functions, we use an intuitive index-free notation. One can follow the rule that ∇ can only be contracted with another vector (or one-form), but never with columns or rows of some matrix, so for example: $\nabla\mathbb{A}\cdot\nabla\mathbb{B}=\sum_{i,j,k}\partial_i\mathbb{A}_{jk}\partial_i\mathbb{B}_{jk}$ or $(\boldsymbol{v}\otimes\mathbb{A})\cdot\nabla\mathbb{B}=\sum_{i,j,k}\boldsymbol{v}_i\mathbb{A}_{jk}\partial_i\mathbb{B}_{jk}$ or $|\mathbb{A}\nabla\mathbb{B}\mathbb{C}|^2=\sum_{i,j,k}\big(\sum_{l,m}\mathbb{A}_{il}\partial_k\mathbb{B}_{lm}\mathbb{C}_{mj}\big)^2$.

If not stated otherwise, the set $\Omega\subset\mathbb{R}^d$ is an open bounded set with a Lipschitz boundary (i.e. of the class $\mathcal{C}^{0,1}$) in the sense of [47, Sect. 2.1.1]. Let $\mathcal{O}\subset\mathbb{R}^m$ be an open bounded set (such as $(0,T)$, Ω or Q) and V be a subset of an Euclidean space. The symbol $(L^p(\mathcal{O};V),\|\cdot\|_{L^p(\mathcal{O};V)})$ denotes the Lebesgue space of functions $u:\mathcal{O}\to V$. The standard inner products in $L^2(\mathcal{O};V)$ and also in $L^2(\partial\Omega;V)$ are denoted as $(\cdot,\cdot)_{\mathcal{O}}$ and $(\cdot,\cdot)_{\partial\Omega}$, respectively. In the special case that $\mathcal{O}=\Omega$, we write just $\|\cdot\|_p$ instead of $\|\cdot\|_{L^p(\Omega;V)}$ and $(\cdot,\cdot)$ instead of $(\cdot,\cdot)_\Omega$.

The symbol $(W^{k,p}(\Omega;V),\|\cdot\|_{k,p})$, $1\leq p\leq\infty$, $k\in\mathbb{N}$, is used to denote the Sobolev spaces with their standard norm considered over the set Ω. If $p>1$, we set $W^{-k,p}(\Omega;V):=(W^{k,p'}(\Omega;V))^*$, where $p':=p/(p-1)$, $k\in\mathbb{N}$, and the star symbol "$*$" denotes the topological (continuous) dual space. For vector-valued functions, we introduce the following subspaces:

$$\begin{aligned}
W^{k,p}_{\boldsymbol{n}} &:= \{\boldsymbol{u}\in W^{k,p}(\Omega;\mathbb{R}^d) : \boldsymbol{u}\cdot\boldsymbol{n}=0\},\quad k\in\mathbb{N},\quad p<\infty,\\
W^{k,p}_{\boldsymbol{n},\mathrm{div}} &:= \{\boldsymbol{u}\in W^{k,p}_{\boldsymbol{n}} : \operatorname{div}\boldsymbol{u}=0\},\quad k\in\mathbb{N},\quad p<\infty,\\
W^{-k,2}_{\boldsymbol{n},\mathrm{div}} &:= (W^{k,2}_{\boldsymbol{n},\mathrm{div}})^*,\quad k\in\mathbb{N},\\
L^2_{\boldsymbol{n},\mathrm{div}} &:= \overline{W^{1,2}_{\boldsymbol{n},\mathrm{div}}}^{\|\cdot\|_2}.
\end{aligned}$$

The expression $\boldsymbol{u}\cdot\boldsymbol{n}$ is understood as a trace of a Sobolev function, for which we do not use any special notation. The meaning of the duality pairing $\langle\cdot,\cdot\rangle$ is always understandable in the given context.

Let X be a Banach space. The Bochner spaces $L^p(0,T;X)$ with $1\leq p\leq\infty$ consist of strongly measurable mappings $u:[0,T]\to X$ for which the norm

$$\|u\|_{L^p(0,T;X)} := \begin{cases} \left(\displaystyle\int_0^T \|u\|_X^p\right)^{\frac{1}{p}} & \text{if } 1\leq p<\infty,\\ \operatorname*{ess\,sup}_{(0,T)}\|u\|_X & \text{if } p=\infty,\end{cases}$$

is finite. If $X = L^q(\Omega;V)$ or $X = W^{k,q}(\Omega;V)$, with $1 \le q \le \infty$, $n \in \mathbb{N}$, we use the abbreviations $\|\cdot\|_{L^pL^q}$ or $\|\cdot\|_{L^pW^{k,q}}$, respectively, for the corresponding norms. Next, the space of weakly continuous functions is defined as

$$\mathcal{C}_w([0,T];X) \coloneqq \big\{u \in L^\infty(0,T;X) : \text{the function } \langle g,u\rangle \text{ is continuous in } [0,T] \text{ for every } g \in X^*\big\},$$

whereas the standard space of continuous X-valued functions on $[0,T]$ is denoted by $\mathcal{C}([0,T];X)$ and equipped with the norm

$$\|u\|_{\mathcal{C}([0,T];X)} \coloneqq \sup_{t\in[0,T]} \|u(t)\|_X.$$

In addition, if X is separable and reflexive, we define two more spaces. First, the space of X^*-valued Radon measures on $[0,T]$ is defined as

$$\mathcal{M}([0,T];X^*) \coloneqq (\mathcal{C}([0,T];X))^*.$$

Then, we set

$$BV([0,T];X^*) \coloneqq \big\{u \in L^\infty(0,T;X^*),\ \partial_t u \in \mathcal{M}([0,T];X^*)\big\}$$

to be the space of functions having X^*-valued bounded variation with respect to the time variable. Note that if $u \in BV([0,T];X^*)$ then it makes sense to define value from left and from right at any point t, i.e., there exist

$$u(t_+) \coloneqq \lim_{\tau\to t_+} u(\tau) \text{ for any } t \in [0,T) \quad \text{and} \quad u(t_-) \coloneqq \lim_{\tau\to t_-} u(\tau) \text{ for any } t \in (0,T],$$

where the limits are considered in the strong topology of X^*. For properties of BV mappings in Bochner spaces, we refer e.g. to [30].

Assumptions on material coefficients. The mathematical properties of the system (1.1)–(1.17) depend crucially on the behaviour of the material coefficients, which we now specify. We will require that

$$\nu, \kappa, \lambda, P \quad \text{are continuous functions in } \mathbb{R}, \mathbb{R}, \mathbb{R} \text{ and } \mathbb{R}\times\mathbb{R}^{d\times d}_{\mathrm{sym}}, \text{ respectively,} \tag{3.1}$$

and there are numbers $q, r > 0$, $C, C_\alpha > 0$ and $\omega_P > 0$, such that, for all $s \in \mathbb{R}$, the following conditions hold:

$$C^{-1} \le \nu(s) \le C, \tag{3.2}$$

$$C^{-1}(1+s^r) \le \kappa(s) \le C(1+s^r), \tag{3.3}$$

$$C^{-1} \le \lambda(s) \le C, \tag{3.4}$$

$$\mathbb{P}(s,\mathbb{A}) = \mathbb{P}(s,\mathbb{A})^T \qquad \text{for all } \mathbb{A} \in \mathbb{R}^{d\times d}_{\mathrm{sym}}, \tag{3.5}$$

$$|\mathbb{P}(s,\mathbb{A})| \le C(1+|\mathbb{A}|^{q+1}) \qquad \text{for all } \mathbb{A} \in \mathbb{R}^{d\times d}_{\mathrm{sym}}, \tag{3.6}$$

$$\mathbb{P}(s,\mathbb{A})\cdot\mathbb{A}^\alpha \ge C_\alpha|\mathbb{A}|^{q+1+\alpha} - C \qquad \text{for all } \alpha > 0 \text{ and } \mathbb{A} \in \mathbb{R}^{d\times d}_{>0}, \tag{3.7}$$

$$\mathbb{P}(s,\mathbb{A})\cdot\mathbb{I} \ge -C \qquad \text{for all } \mathbb{A} \in \mathbb{R}^{d\times d}_{>0}, \tag{3.8}$$

$$\mathbb{P}(s,\mathbb{A})\cdot(\mathbb{I}-\mathbb{A}^{-1}) \ge 0 \qquad \text{for all } \mathbb{A} \in \mathbb{R}^{d\times d}_{>0}, \tag{3.9}$$

$$\mathbb{P}(s,\mathbb{A}+\omega_P\mathbb{I})\boldsymbol{x}\cdot\boldsymbol{x} \le 0 \qquad \text{for all } \mathbb{A} \in \mathbb{R}^{d\times d}_{\mathrm{sym}} \text{ and } \boldsymbol{x} \in \mathbb{R}^d \text{ such that } \mathbb{A}\boldsymbol{x}\cdot\boldsymbol{x} \le 0. \tag{3.10}$$

Assumption (3.2) is quite standard for fluids. Restriction (3.3) means that κ is a bounded function near zero and has an r-growth near infinity. Assumption (3.4) is chosen just for simplicity. Condition (3.5) is necessary for validity of (1.10).

Assumptions (3.6) and (3.7) mean that $\mathbb{P}(\cdot,\mathbb{A})$ behaves asymptotically as $\mathbb{A}^{q+1}$, which is a crucial information to get sufficient a priori estimates. Condition (3.8) simplifies the analysis at one step and means basically that the leading order term of $\mathbb{P}(\cdot,\mathbb{A})$ appears with the positive sign, compare e.g. with the Oldroyd-B and Giesekus model, where $\mathbb{P}(\cdot,\mathbb{A}) = \mathbb{A} - \mathbb{I}$ and $\mathbb{P}(\cdot,\mathbb{A}) = \mathbb{A}^2 - \mathbb{A}$, respectively. Property (3.9) is important for the validity of the second law of thermodynamics in our model. Again, both Oldroyd-B and Giesekus models fulfill this requirement. Finally, the assumption (3.10) restricts the behaviour of $\mathbb{P}(\cdot,\mathbb{A})$ when $\mathbb{A}$ is not positive definite or if its eigenvalues are too small. We remark that this technical condition concerns the case $s \le 0$ or $\mathbb{A} \in \mathbb{R}^{d\times d}\setminus\mathbb{R}^{d\times d}_{>0}$ that actually never arises in the studied problem. An explicit example of function $\mathbb{P}$ satisfying (3.5)–(3.10) would be

$$\mathbb{P}(s,\mathbb{A}) = \delta(s)(1+|\mathbb{A}-\mathbb{I}|^{q-\beta})\mathbb{A}^{\beta}(\mathbb{A}-\mathbb{I}),$$

where δ is a continuous positive real function and $\beta \in [0,q]$. Indeed, note that, for any $\mathbb{A} \in \mathbb{R}^{d\times d}_{>0}$, we can write

$$\begin{aligned}\mathbb{A}^{\beta}(\mathbb{A}-\mathbb{I})\cdot(\mathbb{I}-\mathbb{A}^{-1}) &= \mathbb{A}^{\frac{\beta}{2}}\mathbb{A}^{\frac{\beta}{2}}(\mathbb{A}^{\frac12}-\mathbb{A}^{-\frac12})\mathbb{A}^{\frac12}\cdot(\mathbb{I}-\mathbb{A}^{-1})\\ &= \mathbb{A}^{\frac{\beta}{2}}(\mathbb{A}^{\frac12}-\mathbb{A}^{-\frac12})\cdot\mathbb{A}^{\frac{\beta}{2}}(\mathbb{I}-\mathbb{A}^{-1})\mathbb{A}^{\frac12} = |\mathbb{A}^{\frac{\beta}{2}}(\mathbb{A}^{\frac12}-\mathbb{A}^{-\frac12})|^2 \ge 0,\end{aligned}$$

implying (3.9). The properties (3.6), (3.7) and (3.8) follow easily from (A.20) in Appendix. Finally, we claim that (3.10) holds with $\omega_P = 1$. Indeed, let $0 \neq \boldsymbol{x} \in \mathbb{R}^d$ be an eigenvector of $\mathbb{A} \in \mathbb{R}^{d\times d}_{\mathrm{sym}}$, for which $\lambda := \mathbb{A}\boldsymbol{x}\cdot\boldsymbol{x}/|\boldsymbol{x}|^2 \le 0$. If $\mathbb{A}+\mathbb{I} \notin \mathbb{R}^{d\times d}_{>0}$ then we can redefine $\mathbb{P}(\cdot,\mathbb{A}+\mathbb{I})$ as needed. Otherwise, we have $\mathbb{A}+\mathbb{I} \in \mathbb{R}^{d\times d}_{>0}$, and thus $\lambda > -1$ and we can write

$$\begin{aligned}\mathbb{P}(s,\mathbb{A}+\mathbb{I})\boldsymbol{x}\cdot\boldsymbol{x} &= \delta(s)(1+|\mathbb{A}|^{q-\beta})(\mathbb{A}+\mathbb{I})^{\beta}\mathbb{A}\boldsymbol{x}\cdot\boldsymbol{x}\\ &= \delta(s)(1+|\mathbb{A}|^{q-\beta})(\lambda+1)^{\beta}\lambda|\boldsymbol{x}|^2 \le 0.\end{aligned}$$

Conditions on q and r. To make sure that the individual terms appearing in the weak formulation of the governing equations (defined below) are well defined, we need to restrict the parameters q and r by the conditions

$$r > 1 - \frac{2}{d} \quad \text{and} \quad q > 1 + \frac{2}{r-1+\frac{2}{d}}; \tag{3.11}$$

we recall that $d \ge 2$ is the dimension of the domain Ω.

Condition (3.11) is sufficient to define every term of the system (1.1)–(1.17) in a weak sense, with the exception of (1.12), which needs additional technical assumptions due to the presence of pressure (see the second part of Theorem 3.1 below). As such, condition (3.11) is actually sufficient for the existence of a weak solution, which is the content of our main result.

By imposing (3.11), we place some restrictions on the coefficients of the model which may not agree with experimental measurements. Note, however, that (3.3), (3.6) and (3.7) restrict only the asymptotic behaviour of the coefficients. For example, any continuous function κ defined on some interval (θ_0,θ_1), $0 < \theta_0 < \theta_1 < \infty$, can be modified in a neighbourhood of 0 and ∞ so that (3.3) holds. The interval (θ_0,θ_1) may represent the temperature range for which the model we are considering makes sense. When the fluid starts to freeze or boil, then we are clearly outside this range and it makes no sense to prescribe the coefficients ν, κ, δ and λ there. On the other hand, it is unclear whether one can deduce some absolute bounds for the temperature, besides $\theta > 0$, using only the information that is encoded in

the system. Thus, purely for mathematical reasons, we have to assume that these material coefficients are defined in some way also outside (θ_0, θ_1). A similar remark applies also for the other coefficients. For example, if $|\mathbb{A}|$ is too large, any realistic material eventually breaks down. Thus, we may set $\mathbb{P}(\cdot, \mathbb{A}) = \mathbb{A} - \mathbb{I}$, $|\mathbb{A}| \in [0, M)$, where M is large (to mimic the Oldroyd-B model, for example) and then extend this function continuously so that (3.7) holds with some large q, see (1.20).

A priori estimates. Let us now the motivate the definition of the weak solution to (1.1)–(1.13) by an informal derivation of the available a priori estimates. This clarifies the need for (3.11) and highlights the main idea of the existence proof. The starting point are the assumptions on the data:

$$E_0 \in L^1(\Omega; \mathbb{R}_{\geq 0}),\ \eta_0 \in L^1(\Omega; \mathbb{R}),\ \mathbb{B}_0 \in L^q(\Omega; \mathbb{R}^{d\times d}_{>0}),\ \boldsymbol{g} \in L^2(Q; \mathbb{R}^d). \tag{3.12}$$

In addition, we may suppose that

$$\theta \geq 0 \quad \text{and} \quad \mathbb{B}\boldsymbol{x} \cdot \boldsymbol{x} \geq 0 \quad \text{for all} \quad \boldsymbol{x} \in \mathbb{R}^d, \tag{3.13}$$

which is due to a suitable construction of the solution (cf. (5.62) below).

In what follows, the basic relations (1.4)–(1.7) and also (2.9) will be used without further reference. Moreover, the symbol C will be used to denote a positive constant that can change from line to line and can depend only on the data, domain Ω, time $T > 0$ and other constants appearing in (3.2)–(3.10).

Integrating (1.12) over Ω and applying the boundary conditions (1.14) and (1.16) drops the divergence terms, which, together with Young's inequality and $(3.13)_1$, leads to

$$\frac{\mathrm{d}}{\mathrm{d}t}\int_\Omega E = \int_\Omega \boldsymbol{g}\cdot\boldsymbol{v} \leq \frac{1}{2}\int_\Omega |\boldsymbol{g}|^2 + \int_\Omega E.$$

Hence, using $(3.12)_1$, we see that $E \in L^\infty(0,T; L^1(\Omega;\mathbb{R}))$, therefore also

$$\theta \in L^\infty(0,T; L^1(\Omega;\mathbb{R}_{\geq 0})) \quad \text{and} \quad \boldsymbol{v} \in L^\infty(0,T; L^2(\Omega;\mathbb{R}^d)). \tag{3.14}$$

Next, integrating the entropy inequality (1.13), and again applying the boundary conditions in the divergence terms, gives

$$\frac{\mathrm{d}}{\mathrm{d}t}\int_\Omega \eta(t) \geq 0.$$

Applying $(3.12)_2$ and $(3.13)_2$ ($\operatorname{tr}\mathbb{B} \geq 0$, to be precise), the last inequality yields

$$\int_\Omega (\ln\theta(t) + \ln\det\mathbb{B}(t)) > -C,$$

which is a very important inequality as it ensures that $\theta > 0$ and $\mathbb{B}$ is positive definite almost everywhere. Although one also gets $\xi \in L^1(0,T;L^1(\Omega))$ after integrating (1.13) and using $(3.14)_1$, this information turns out to be too weak. Instead, we can get better estimates directly from (1.10) and (1.11).

Due to the positive definiteness of $\mathbb{B}$, the equation (1.10) can be tested by the matrix power $\mathbb{B}^{q-1}$. (Though here one can also use $|\mathbb{B}|^{q-2}\mathbb{B}$ since $q \geq 1$ and the stress diffusion term is actually not important for the estimate itself.) Then, using (3.4), (3.7), Young's inequality and Lemma A.3 below, we eventually get

$$\frac{\mathrm{d}}{\mathrm{d}t}\int_\Omega \operatorname{tr}\mathbb{B}^q + \int_\Omega |\mathbb{B}|^{2q} + \int_\Omega |\nabla\mathbb{B}^{\frac{q}{2}}|^2 \leq C\int_\Omega |\mathbb{B}|^q|\mathbb{D}\boldsymbol{v}| + C \leq C\int_\Omega |\mathbb{D}\boldsymbol{v}|^2 + C. \tag{3.15}$$

Hence, integrating over $(0,T)$ and thanks to $(3.12)_3$, we have

$$\|\mathbb{B}\|_{L^{2q}L^{2q}} \leq C\|\mathbb{D}\boldsymbol{v}\|_{L^2L^2}^{\frac{1}{q}} + C. \tag{3.16}$$

Clearly, we need control over $\mathbb{D}\boldsymbol{v}$, but it has to be obtained differently than for the Navier–Stokes–Fourier systems, as we pointed out in the introduction.

Thanks to $\theta > 0$, we may test (1.11) by the function $-\theta^{-\beta}$ with $\beta \geq 0$. Eventually, applying $(3.14)_1$, (3.2) and (3.3), this leads to the estimate

$$\beta \int_Q \theta^{r-\beta-1}|\nabla\theta|^2 + \int_Q |\mathbb{D}\boldsymbol{v}|^2 \leq C\int_Q \theta|\mathbb{B}||\mathbb{D}\boldsymbol{v}| + C. \tag{3.17}$$

Using (3.14), (3.16) and the Hölder inequality, the above inequality gives

$$\begin{aligned}\beta\|\theta^{\frac{r-\beta+1}{2}}\|_{L^2W^{1,2}}^2 + \|\mathbb{D}\boldsymbol{v}\|_{L^2L^2}^2 &\leq C\|\theta\|_{L^{2q'}L^{2q'}}\|\mathbb{B}\|_{L^{2q}L^{2q}}\|\mathbb{D}\boldsymbol{v}\|_{L^2L^2} + C\\ &\leq C\|\theta\|_{L^{2q'}L^{2q'}}^{2q'} + \frac{1}{2}\|\mathbb{D}\boldsymbol{v}\|_{L^2L^2}^2 + C.\end{aligned} \tag{3.18}$$

The last term is absorbed by the left-hand side and for the first term we use the interpolation inequality

$$\|\theta\|_{2q'}^{2q'} \leq \|\theta\|_1^{2q' - \frac{d(r-\beta+1)(2q'-1)}{d(r-\beta)+2}} \|\theta\|^{\frac{d(r-\beta+1)(2q'-1)}{d(r-\beta)+2}}_{\frac{d(r-\beta+1)}{d-2}}$$

and (3.14) to deduce

$$\begin{aligned}&\beta\|\theta^{\frac{r-\beta+1}{2}}\|_{L^2W^{1,2}}^2 + \|\mathbb{D}\boldsymbol{v}\|_{L^2L^2}^2 \leq C\int_0^T \|\theta\|^{\frac{d(r-\beta+1)(2q'-1)}{d(r-\beta)+2}}_{\frac{d(r-\beta+1)}{d-2}} + C\\ &= C\int_0^T \|\theta^{\frac{r-\beta+1}{2}}\|^{\frac{2d(2q'-1)}{d(r-\beta)+2}}_{\frac{2d}{d-2}} + C \leq C\int_0^T \|\theta^{\frac{r-\beta+1}{2}}\|_{1,2}^{\frac{2d(2q'-1)}{d(r-\beta)+2}} + C.\end{aligned} \tag{3.19}$$

Hence, if

$$\frac{2d(2q'-1)}{d(r-\beta)+2} < 2, \tag{3.20}$$

the first term on the right-hand side can be absorbed by the left-hand side and thus, we get

$$\beta\|\theta^{\frac{r-\beta+1}{2}}\|_{L^2W^{1,2}}^2 + \|\mathbb{D}\boldsymbol{v}\|_{L^2L^2}^2 + \|\mathbb{B}\|_{L^{2q}L^{2q}}^{2q} \leq C. \tag{3.21}$$

Finally, the inequality (3.20) can be made true by choosing $\beta > 0$ sufficiently small if and only if q and r satisfy (3.11). Note that, in this case, we were able to estimate the right-hand side of (3.17), i.e., the "critical" term $\theta\mathbb{B}\cdot\mathbb{D}\boldsymbol{v}$ appearing in (1.11). It is easy to verify, using estimates (3.14) and (3.21) that all the other nonlinear terms appearing in the system (1.1)–(1.13) are integrable as well.

Definition of weak solution. Motivated by the above estimates, we now deliver the exact definition of a weak solution to (1.1)–(1.17).

Definition 3.1. *Let $T > 0$ and let $\Omega \subset \mathbb{R}^d$, $d \geq 2$, be a Lipschitz domain. Assume that the constants $a \in \mathbb{R}$, $\alpha \geq 0$, $c_v, \mu > 0$ and the functions $\nu, \kappa, \lambda, \mathbb{P}$ fulfil the assumptions* (3.1)–(3.9) *with the parameters q and r satisfying* (3.11) *and let $m \coloneqq \min\{2, \frac{4q}{q+2}\}$. Suppose that the initial data satisfy*

$$\boldsymbol{v}_0 \in L^2_{\boldsymbol{n},\mathrm{div}}(\Omega;\mathbb{R}^d), \quad \mathbb{B}_0 \in L^q(\Omega;\mathbb{R}^{d\times d}_{>0}), \quad \theta_0 \in L^1(\Omega;\mathbb{R}_{>0}), \tag{3.22}$$

$$\eta_0 \coloneqq c_v \ln\theta_0 - f(\mathbb{B}_0) \in L^1(\Omega;\mathbb{R}), \tag{3.23}$$

where f is given by (2.9)*, and that*

$$\boldsymbol{g} \in L^2(Q;\mathbb{R}^d). \tag{3.24}$$

Then, we say that the sextuplet $(\boldsymbol{v},\mathbb{B},\theta,e,E,\eta) : Q \to \mathbb{R}^d \times \mathbb{R}^{d\times d}_{>0} \times \mathbb{R}_{>0} \times \mathbb{R}_{>0} \times \mathbb{R}_{>0} \times \mathbb{R}$ is a weak solution of the initial-boundary value problem (1.1)–(1.17) *if all of the following conditions* (I)–(IV) *are satisfied:*

(I) *The functions $\boldsymbol{v}$, $\mathbb{B}$, θ and η fulfil the properties*

$$\boldsymbol{v} \in L^2(0,T;W^{1,2}_{\boldsymbol{n},\mathrm{div}}) \cap \mathcal{C}_w([0,T];L^2(\Omega;\mathbb{R}^d)), \tag{3.25}$$

$$\partial_t \boldsymbol{v} \in L^{\frac{d+2}{d}}(0,T;W^{-1,\frac{d+2}{d}}_{\boldsymbol{n},\mathrm{div}}), \tag{3.26}$$

$$\mathbb{B} \in L^m(0,T;W^{1,m}(\Omega;\mathbb{R}^{d\times d}_{>0})) \cap \mathcal{C}_w([0,T];L^q(\Omega;\mathbb{R}^{d\times d}_{>0})), \tag{3.27}$$

$$\mathbb{B} \in L^{2q}(Q;\mathbb{R}^{d\times d}_{>0}), \tag{3.28}$$

$$\mathbb{B}^{\frac{q}{2}} \in L^2(0,T;W^{1,2}(\Omega;\mathbb{R}^{d\times d}_{>0})), \tag{3.29}$$

$$\partial_t \mathbb{B} \in \big(L^{2q'}(0,T;W^{1,2q'}(\Omega;\mathbb{R}^{d\times d}))\big)^*, \tag{3.30}$$

$$\mathbb{B}^{-\frac{1}{2}}\nabla\mathbb{B}\mathbb{B}^{-\frac{1}{2}} \in L^2(Q;\mathbb{R}^d \times \mathbb{R}^{d\times d}_{\mathrm{sym}}), \tag{3.31}$$

$$\ln\det\mathbb{B} \in L^2(0,T;W^{1,2}(\Omega;\mathbb{R})) \cap L^\infty(0,T;L^1(\Omega;\mathbb{R})), \tag{3.32}$$

$$\theta \in L^\infty(0,T;L^1(\Omega;\mathbb{R}_{>0})) \cap L^{r+\frac{2}{d}+1-\varepsilon}(Q;\mathbb{R}_{>0}), \tag{3.33}$$

$$\theta^{\frac{r+1-\varepsilon}{2}} \in L^2(0,T;W^{1,2}(\Omega;\mathbb{R}_{>0})), \tag{3.34}$$

$$\ln\theta \in L^2(0,T;W^{1,2}(\Omega;\mathbb{R})) \cap L^\infty(0,T;L^1(\Omega;\mathbb{R})), \tag{3.35}$$

$$\eta \in L^m(0,T;W^{1,m}(\Omega;\mathbb{R})) \cap L^\infty(0,T;L^1(\Omega;\mathbb{R})) \tag{3.36}$$

for every $\varepsilon \in (0,1)$.

(II) *The relations* (1.1)–(1.7) *hold almost everywhere in Q.*

(III) *Equations* (1.9)–(1.13) *are satisfied in the following sense:*

$$\begin{aligned}
&\langle \partial_t \boldsymbol{v}, \boldsymbol{\varphi}\rangle - (\boldsymbol{v}\otimes\boldsymbol{v}, \nabla\boldsymbol{\varphi})_Q + (\mathbb{S}, \nabla\boldsymbol{\varphi})_Q + (\alpha \boldsymbol{v}_\tau, \boldsymbol{\varphi}_\tau)_\Sigma = (\boldsymbol{g},\boldsymbol{\varphi})_Q \\
&\text{for all } \boldsymbol{\varphi} \in L^{\frac{d}{2}+1}(0,T;W^{1,\frac{d}{2}+1}_{\boldsymbol{n},\mathrm{div}}),
\end{aligned} \tag{3.37}$$

$$\begin{aligned}
&\langle \partial_t \mathbb{B}, \mathbb{A}\rangle - (\mathbb{B}\otimes\boldsymbol{v}, \nabla\mathbb{A})_Q + (\mathbb{P}(\theta,\mathbb{B}), \mathbb{A})_Q + (\lambda(\theta)\nabla\mathbb{B}, \nabla\mathbb{A})_Q \\
&\quad = ((a\mathbb{D}\boldsymbol{v} + \mathbb{W}\boldsymbol{v})\mathbb{B}, \mathbb{A} + \mathbb{A}^T)_Q \\
&\text{for all } \mathbb{A} \in L^{2q'}(0,T;W^{1,2q'}(\Omega;\mathbb{R}^{d\times d})),
\end{aligned} \tag{3.38}$$

$$\begin{aligned}
&-(c_v\theta_0, \phi\varphi(0)) - (c_v\theta, \phi\partial_t\varphi)_Q - (c_v\theta\boldsymbol{v}, \nabla\phi\varphi)_Q + (\kappa(\theta)\nabla\theta, \nabla\phi\varphi)_Q \\
&\quad \geq (\mathbb{S}\cdot\mathbb{D}\boldsymbol{v}, \phi\varphi)_Q \\
&\text{for all } \varphi \in W^{1,\infty}((0,T);\mathbb{R}_{\geq 0}),\ \varphi(T) = 0, \ \text{ and all } \phi \in W^{1,\infty}(\Omega;\mathbb{R}_{\geq 0}),
\end{aligned} \tag{3.39}$$

$$\begin{aligned}
&-(\eta_0, \phi\varphi(0)) - (\eta, \phi\partial_t\varphi)_Q - (\eta\boldsymbol{v}, \nabla\phi\varphi)_Q \\
&\quad + \big(\kappa(\theta)\nabla\ln\theta - \lambda(\theta)\nabla f(\mathbb{B}), \nabla\phi\varphi\big)_Q \geq (\xi, \phi\varphi)_Q \\
&\text{for all } \varphi \in W^{1,\infty}((0,T);\mathbb{R}_{\geq 0}),\ \varphi(T) = 0, \ \text{ and all } \phi \in W^{1,\infty}(\Omega;\mathbb{R}_{\geq 0}),
\end{aligned} \tag{3.40}$$

$$\frac{\mathrm{d}}{\mathrm{d}t}\int_\Omega E + \alpha\int_{\partial\Omega}|\boldsymbol{v}|^2 = \int_\Omega \boldsymbol{g}\cdot\boldsymbol{v} \ \text{ a.e. in } [0,T]. \tag{3.41}$$

(IV) *The initial data are attained in the following way:*

$$\lim_{t\to 0+} \|\boldsymbol{v}(t)-\boldsymbol{v}_0\|_2 = 0, \tag{3.42}$$

$$\lim_{t\to 0+} \|\mathbb{B}(t)-\mathbb{B}_0\|_{q-\varepsilon} = 0 \quad \textit{for every} \quad \varepsilon\in(0,q-1], \tag{3.43}$$

$$\lim_{t\to 0+} \|\theta(t)-\theta_0\|_1 = 0, \tag{3.44}$$

$$\liminf_{t\to 0+} (\eta(t),\phi) \geq (\eta_0,\phi) \quad \textit{for all} \quad 0\leq\phi\in W^{1,\infty}(\Omega). \tag{3.45}$$

With this definition in hand, we now formulate our main result.

Theorem 3.1. *Suppose that all the assumptions of Definition 3.1 are fulfilled. Then, there exists a weak solution of the system* (1.1)–(1.17) *in the sense of Definition 3.1.*

In addition, if $d\leq 3$ *and* $\Omega\in\mathcal{C}^{1,1}$*, then there is a pressure* $\mathrm{p}\in L^{\frac{d+2}{d}}(Q;\mathbb{R})$ *such that the local balance of total energy* (1.12) *holds in the sense:*

$$\begin{aligned} -(\tfrac{1}{2}|\boldsymbol{v}_0|^2+c_v\theta_0,\phi\varphi(0)) - (E,\phi\partial_t\varphi)_Q + (\alpha|\boldsymbol{v}_\tau|^2,\phi\varphi)_\Sigma + (\kappa(\theta)\nabla\theta,\nabla\phi\varphi)_Q \\ = (E\boldsymbol{v}+\mathrm{p}\boldsymbol{v}-\mathbb{S}\boldsymbol{v},\nabla\phi\varphi)_Q \end{aligned} \tag{3.46}$$

for all $\varphi\in W^{1,\infty}((0,T);\mathbb{R})$, $\varphi(T)=0$, *and every* $\phi\in W^{1,\infty}(\Omega;\mathbb{R})$

and also (3.37) *can be generalized to*

$$\begin{aligned} \langle\partial_t\boldsymbol{v},\boldsymbol{\varphi}\rangle - (\boldsymbol{v}\otimes\boldsymbol{v},\nabla\boldsymbol{\varphi})_Q + (-p\mathbb{I}+\mathbb{S},\nabla\boldsymbol{\varphi})_Q + (\alpha\boldsymbol{v}_\tau,\boldsymbol{\varphi}_\tau)_\Sigma = (\boldsymbol{g},\boldsymbol{\varphi})_Q \\ \textit{for all} \quad \boldsymbol{\varphi}\in L^{\frac{d}{2}+1}(0,T;W^{1,\frac{d}{2}+1}_{\boldsymbol{n}}). \end{aligned} \tag{3.47}$$

We remark that if a weak solution admits enough regularity so that (3.37) can be tested by $\boldsymbol{v}$ and (3.39) can be localized in space, then (3.39) holds as an equality. Indeed, the localized version of (3.39) reads

$$c_v\partial_t\theta + c_v\boldsymbol{v}\cdot\nabla\theta - \operatorname{div}(\kappa(\theta)\nabla\theta) - \mathbb{S}\cdot\mathbb{D}\boldsymbol{v} \geq 0. \tag{3.48}$$

On the other hand, subtracting (3.37) tested by $\boldsymbol{v}$ from (3.41) yields

$$\int_\Omega (c_v\partial_t\theta - \mathbb{S}\cdot\mathbb{D}\boldsymbol{v}) = 0.$$

Since also

$$\int_\Omega (c_v\boldsymbol{v}\cdot\nabla\theta - \operatorname{div}(\kappa(\theta)\nabla\theta)) = 0$$

due to the boundary conditions $\boldsymbol{v}\cdot\boldsymbol{n}=0$ and $\nabla\theta\cdot\boldsymbol{n}=0$ on $\partial\Omega$, we conclude from the above that (3.48) must be an equality. Consequently, the entropy inequality (3.40) also becomes an equality, provided that one is able to justify $\mathbb{B}^{-1}$ and θ^{-1} as tests in (3.38) and (3.39). These considerations imply that a weak solution that admits sufficient regularity is also a solution of (1.1)–(1.17) in the classical sense.

The existence proof below is done only for $d\geq 3$ (the case $d=2$ is simpler). Also, it is clearly enough to focus on the case $\alpha>0$. In the simpler case $\alpha=0$ (corresponding to the free-slip boundary condition), one just has to use a different Korn–Poincaré inequality in case Ω is axially symmetric.

The general strategy of the proof is to approximate the system (1.9), (1.10), (1.11) using several parameters to obtain a proper Galerkin approximation generated by a smooth basis of eigenvectors and to show that the resulting (ODE) system has a solution. After that, our aim is to derive the entropy equation. At this point, possibly irregular terms containing θ and $\mathbb{B}$ are cut-off and $\boldsymbol{v}$ is smooth,

hence we easily obtain uniform estimates for the Galerkin approximations of $\mathbb{B}$ and θ, which might not be positive definite or positive, respectively. However, after taking the limit with these approximations and then proving certain maximum principles, we prove invertibility of θ and $\mathbb{B}$, which, in turn, enables us to derive the entropy equation. From this we read that the positivity of $\det \mathbb{B}$ and θ is preserved uniformly, which then enables us to remove the cut-off from the system. The proof of this is presented in Section 4. Note that at this point, the velocity is still kept in a finite ℓ-dimensional space. To the equation for the internal energy we add the regularization $-\omega\Delta_{r+2}\theta$ (the so-called $(r+2)$-Laplacian) in order to avoid weighted Sobolev spaces, where the density of smooth functions is not available in general.

Next, in Section 5, we first improve the uniform estimates by considering appropriate test functions in the equations for θ and $\mathbb{B}$. At this point such a procedure is rigorous. Finally, we let $\omega \to 0$ and $\ell \to \infty$ and we pass to the final limit, identify the non-linear terms and initial conditions, hereby obtaining a solution of the original problem. Finally, in Section 6, we prove the validity of the local energy equality provided $d \leq 3$.

4. Existence of a weak solution: the approximative problem

First we introduce a truncation, which is essential for the proof. We also prepare some simple estimates corresponding to this truncation that are used later in the proof. Recalling that ω_P is introduced in (3.10), we define, for any $\omega \in (0, \omega_P)$, the "cut-off" function g_ω in the following way:

$$g_\omega(\mathbb{A}, \tau) \coloneqq \frac{\max\{0, \Lambda(\mathbb{A}) - \omega\} \max\{0, \tau - \omega\}}{(|\Lambda(\mathbb{A})| + \omega)(1 + \omega|\mathbb{A}|^2)(|\tau| + \omega)(1 + \omega\tau^2)}, \quad \mathbb{A} \in \mathbb{R}^{d\times d}_{\mathrm{sym}}, \quad \tau \in \mathbb{R},$$

where $\Lambda(\mathbb{A})$ denotes the smallest eigenvalue of $\mathbb{A}$, i.e.,

$$\Lambda(\mathbb{A}) \coloneqq \min\{\lambda : \det(\mathbb{A} - \lambda\mathbb{I}) = 0\}.$$

Note that g_ω is a continuous function in $\mathbb{R}^{d\times d}_{\mathrm{sym}} \times \mathbb{R}$ and satisfies $0 \leq g_\omega(\mathbb{A}, \tau) < 1$ for every $(\mathbb{A}, \tau) \in \mathbb{R}^{d\times d}_{\mathrm{sym}} \times \mathbb{R}$. Moreover, if $\Lambda(\mathbb{A}) \leq \omega$ or $\tau \leq \omega$, then $g_\omega(\mathbb{A}, \tau) = 0$, whereas if $\Lambda(\mathbb{A}) > 0$ and $\tau > 0$, then $g_\omega(\mathbb{A}, \tau) \to 1$ as $\omega \to 0+$. Furthermore, we remark that

$$g_\omega(\mathbb{A}, \tau)(1 + |\mathbb{A}| + |\mathbb{A}|^2)(1 + \tau + \tau^2) \leq C(\omega). \tag{4.1}$$

The function g_ω is used below in the system (4.11)–(4.13) to control irregular terms of the original problem. We also truncate the initial functions $\mathbb{B}_0$ and θ_0 and set

$$\mathbb{B}_0^\omega(x) \coloneqq \begin{cases} \mathbb{B}_0(x) & \text{if } \Lambda(\mathbb{B}_0(x)) > \omega \text{ and } |\mathbb{B}_0(x)| < \sqrt{d}\,\omega^{-1}, \\ \mathbb{I} & \text{elsewhere;} \end{cases} \tag{4.2}$$

$$\theta_0^\omega(x) \coloneqq \begin{cases} \theta_0(x) & \text{if } \omega < \theta_0(x) < \omega^{-1}, \\ 1 & \text{elsewhere.} \end{cases} \tag{4.3}$$

With such definitions, these functions satisfy (a.e. in Ω)

$$\Lambda(\mathbb{B}_0^\omega) > \omega, \qquad \theta_0^\omega > \omega \tag{4.4}$$

$$|\mathbb{B}_0^\omega| < \sqrt{d}\,\omega^{-1}, \qquad |\theta_0^\omega| < \omega^{-1} \tag{4.5}$$

$$|\mathbb{B}_0^\omega| \leq \sqrt{d} + |\mathbb{B}_0|, \qquad \theta_0^\omega \leq 1 + \theta_0, \tag{4.6}$$

and, since $\ln 1 = 0$,

$$|\ln \det \mathbb{B}_0^\omega| \le |\ln \det \mathbb{B}_0|, \qquad |\ln \theta_0^\omega| \le |\ln \theta_0|. \tag{4.7}$$

Since $\mathbb{B}_0 \in L^q(\Omega;\mathbb{R}^{d\times d}_{>0})$, we also observe that the Lebesgue measure of the sets $\{\Lambda(\mathbb{B}_0) \le \omega\}$ and $\{|\mathbb{B}_0| \ge \omega^{-1}\}$ tends to zero as $\omega \to 0+$, and thus

$$\|\mathbb{B}_0^\omega - \mathbb{B}_0\|_q^q = \int_{\Lambda(\mathbb{B}_0)\le\omega} |\mathbb{I} - \mathbb{B}_0|^q + \int_{|\mathbb{B}_0|\ge\omega^{-1}} |\mathbb{I} - \mathbb{B}_0|^q \to 0. \tag{4.8}$$

Analogously, relying on $\theta_0 \in L^1(\Omega;\mathbb{R}_{>0})$, we also obtain

$$\|\theta_0^\omega - \theta_0\|_1 \to 0, \quad \omega \to 0+. \tag{4.9}$$

Next, we discretize the ω-truncated system in space by the Galerkin method.[1] Let $\{\boldsymbol{w}_i\}_{i=1}^\infty$, $\{\mathbb{W}_j\}_{j=1}^\infty$ and $\{w_k\}_{k=1}^\infty$ be bases of the spaces $W^{N,2}(\Omega;\mathbb{R}^d)\cap W^{1,2}_{\boldsymbol{n},\mathrm{div}}$, $W^{N,2}(\Omega;\mathbb{R}^{d\times d}_{\mathrm{sym}})$ and $W^{N,2}(\Omega;\mathbb{R})$, respectively, with the following properties:

- The bases are L^2-orthonormal and $W^{N,2}$-orthogonal.
- The number $N \in \mathbb{N}$ is chosen so large that the elements of the bases are Lipschitz (due to embeddings of Sobolev spaces).
- $w_1 = |\Omega|^{-\frac{1}{2}}$.
- For any $\ell, n \in \mathbb{N}$, there exist L^2-orthogonal projections
$$P_\ell : L^2(\Omega;\mathbb{R}^d) \to \mathrm{span}\{\boldsymbol{w}_i\}_{i=1}^\ell,$$
$$Q_n : L^2(\Omega;\mathbb{R}^{d\times d}) \to \mathrm{span}\{\mathbb{W}_j\}_{j=1}^n,$$
$$R_n : L^2(\Omega;\mathbb{R}) \to \mathrm{span}\{w_k\}_{k=1}^n$$
- P_ℓ, Q_n, R_n are L^2- and $W^{N,2}$-bounded, uniformly w.r.t. ℓ, n.

Existence of these bases and corresponding projections follows from standard results (see Appendix 4 in [40]) using the eigenvectors of the generalized Laplace or Stokes operators.

We fix $\ell, n \in \mathbb{N}$ and consider the problem of finding the functions $\alpha^i_{\ell n}$, $\beta^j_{\ell n}$, $\gamma^k_{\ell n}$ of time, where $i = 1,\dots,\ell$ and $j,k = 1,\dots,n$, such that the functions $\boldsymbol{v}_{\ell n}$, $\mathbb{B}_{\ell n}$, $\theta_{\ell n}$ and $\mathbb{S}^\omega_{\ell n}$ defined as

$$\boldsymbol{v}_{\ell n}(t,x) = \sum_{i=1}^\ell \alpha^i_{\ell n}(t)\boldsymbol{w}_i(x),\ \mathbb{B}_{\ell n}(t,x) = \sum_{j=1}^n \beta^j_{\ell n}(t)\mathbb{W}_j(x),\ \theta_{\ell n} = \sum_{k=1}^n \gamma^k_{\ell n}(t)w_k(x)$$

and

$$\mathbb{S}^\omega_{\ell n} \coloneqq 2\nu(\theta_{\ell n})\mathbb{D}\boldsymbol{v}_{\ell n} + 2a\mu g_\omega(\mathbb{B}_{\ell n},\theta_{\ell n})\theta_{\ell n}\mathbb{B}_{\ell n} \tag{4.10}$$

satisfy the following equations a.e. in $(0,T_0)$, $T_0 > 0$:

$$(\partial_t \boldsymbol{v}_{\ell n},\boldsymbol{w}_i) - (\boldsymbol{v}_{\ell n}\otimes\boldsymbol{v}_{\ell n},\nabla\boldsymbol{w}_i) + (\mathbb{S}^\omega_{\ell n},\nabla\boldsymbol{w}_i) + \alpha(\boldsymbol{v}_{\ell\boldsymbol{n}},\boldsymbol{\varphi})_{\partial\Omega} = (\boldsymbol{g},\boldsymbol{w}_i), \tag{4.11}$$

$$\begin{aligned}(\partial_t\mathbb{B}_{\ell n},\mathbb{W}_j) - (\mathbb{B}_{\ell n}\otimes\boldsymbol{v}_\ell,\nabla\mathbb{W}_j) + (\mathbb{P}(\theta_{\ell n},\mathbb{B}_{\ell n}),\mathbb{W}_j) + (\lambda(\theta_{\ell n})\nabla\mathbb{B}_{\ell n},\nabla\mathbb{W}_j)\\ = (2g_\omega(\mathbb{B}_{\ell n},\theta_{\ell n})(a\mathbb{D}\boldsymbol{v}_{\ell n} + \mathbb{W}\boldsymbol{v}_{\ell n})\mathbb{B}_{\ell n},\mathbb{W}_j),\end{aligned} \tag{4.12}$$

$$\begin{aligned}(c_v\partial_t\theta_{\ell n},w_k) - (c_v\theta_\ell\boldsymbol{v}_{\ell n},\nabla w_k) + ((\kappa(\theta_{\ell n}) + \omega|\nabla\theta_{\ell n}|^r)\nabla\theta_{\ell n},\nabla w_k)\\ = (\mathbb{S}^\omega_{\ell n}\cdot\mathbb{D}\boldsymbol{v}_{\ell n},w_k),\end{aligned} \tag{4.13}$$

for all $1 \le i \le \ell$, $1 \le j,k \le n$ and with the initial conditions

$$\boldsymbol{v}_{\ell n}(0) = P_\ell\boldsymbol{v}_0, \quad \mathbb{B}_{\ell n}(0) = Q_n\mathbb{B}_0^\omega, \quad \theta_{\ell n}(0) = R_n\theta_0^\omega \quad \text{in } \Omega. \tag{4.14}$$

[1] With this approach, we do not need the positive definiteness of the basis functions for $\mathbb{B}$.

By the L^2-orthonormality of the bases, we have

$$(\partial_t \boldsymbol{v}_{\ell n}, \boldsymbol{w}_i) = \sum_{m=1}^{\ell} \partial_t \alpha_{\ell n}^m (\boldsymbol{w}_m, \boldsymbol{w}_i) = (\alpha_{\ell n}^i)'$$

and similarly

$$(\partial_t \mathbb{B}_{\ell n}, \mathbb{W}_j) = (\beta_{\ell n}^j)' \quad \text{and} \quad (\partial_t \theta_{\ell n}, w_k) = (\gamma_{\ell n}^k)'.$$

Thus, (4.11)–(4.13) is a system of $\ell + 2n$ ordinary differential equations of the form

$$\left.\begin{aligned} (\alpha_{\ell n}^i)' &= F_1(t, \alpha_{\ell n}^1, \dots, \alpha_{\ell n}^\ell), && i = 1, \dots, \ell, \\ (\beta_{\ell n}^j)' &= F_2(\beta_{\ell n}^1, \dots, \beta_{\ell n}^n), && j = 1, \dots, n, \\ (\gamma_{\ell n}^k)' &= F_3(\gamma_{\ell n}^1, \dots, \gamma_{\ell n}^n), && k = 1, \dots, n. \end{aligned}\right\} \tag{4.15}$$

It is easy to see, using (3.1), that F_1, F_2 and F_3 are continuous with respect to the variables $\alpha_{\ell n}^i$, $\beta_{\ell n}^j$ and $\gamma_{\ell n}^k$ and measurable with respect to t, respectively. Moreover, the explicit dependence of F_1 on time is controlled by

$$|(\boldsymbol{g}, \boldsymbol{w}_i)| \le \|\boldsymbol{g}\|_2 \|\boldsymbol{w}_i\|_2 \in L^2(0, T; \mathbb{R}).$$

Thus, we can apply the Carathéodory existence theorem (see [19, Chapter 2, Theorem 1] or [60, Chapter 30]) and hereby obtain absolutely continuous functions $\alpha_{\ell n}^i$, $\beta_{\ell n}^j$, $\gamma_{\ell n}^k$, $1 \le i \le \ell$, $1 \le j, k \le n$, solving (4.15) on $(0, T_0)$, where $T_0 < T$ is the time of the first blow-up. In view of the a priori estimates derived below (see e.g. (4.18)), we are able to prove that

$$\sup_{t \in (0, T_0)} \Big(\sum_{i=1}^{\ell} (\alpha_{\ell n}^i(t))^2 + \sum_{j=1}^{n} (\beta_{\ell n}^j(t))^2 + \sum_{k=1}^{n} (\gamma_{\ell n}^k(t))^2 \Big) < \infty,$$

hence, there can be no blow-up and the functions $\boldsymbol{v}_{kl}, \mathbb{B}_{kl}, \theta_{kl}$ are defined on an arbitrary time interval, in particular on $[0, T]$.

Estimates uniform with respect to n. By multiplying the i-th equation in (4.11) by $\alpha_{\ell n}^i$, summing the result over all $i = 1, \dots, \ell$, integrating by parts and using the facts that the basis functions satisfy $\boldsymbol{v}_{\ell n} \cdot \boldsymbol{n} = 0$ on $\partial\Omega$ and $\operatorname{div} \boldsymbol{v}_{\ell n} = 0$ in Ω (hence the convective term vanishes), we obtain (a.e. in $(0, T)$)

$$\begin{aligned} \frac{1}{2} \frac{\mathrm{d}}{\mathrm{d}t} \|\boldsymbol{v}_{\ell n}\|_2^2 + \|\sqrt{2\nu(\theta_{\ell n})} \mathbb{D}\boldsymbol{v}_{\ell n}\|_2^2 + \alpha \|\boldsymbol{v}_{\ell n}\|_{L^2(\partial\Omega; \mathbb{R}^d)}^2 \\ = -(2a\mu g_\omega(\mathbb{B}_{\ell n}, \theta_{\ell n}) \theta_{\ell n} \mathbb{B}_{\ell n}, \mathbb{D}\boldsymbol{v}_{\ell n}) + (\boldsymbol{g}, \boldsymbol{v}_{\ell n}). \end{aligned} \tag{4.16}$$

Then we use (3.2), (4.1), Korn's and Young's inequality, and deduce

$$\begin{aligned} \frac{\mathrm{d}}{\mathrm{d}t} \|\boldsymbol{v}_{\ell n}\|_2^2 + \|\nabla \boldsymbol{v}_{\ell n}\|_2^2 + \alpha \|\boldsymbol{v}_{\ell n}\|_{L^2(\partial\Omega; \mathbb{R}^d)}^2 &\le C(\omega) \int_\Omega |\mathbb{D}\boldsymbol{v}_{\ell n}| + C \|\boldsymbol{g}\|_2 \|\nabla \boldsymbol{v}_{\ell n}\|_2 \\ &\le C(\omega) + C \|\boldsymbol{g}\|_2^2 + \frac{1}{2} \|\nabla \boldsymbol{v}_{\ell n}\|_2^2. \end{aligned}$$

Integration with respect to time and the use of (4.14) and (3.22) directly leads to

$$\sup_{t \in (0, T)} \|\boldsymbol{v}_{\ell n}(t)\|_2^2 + \int_0^T \|\nabla \boldsymbol{v}_{\ell n}\|_2^2 + \alpha \int_0^T \|\boldsymbol{v}_{\ell n}\|_{L^2(\partial\Omega; \mathbb{R}^d)}^2 \le C(\omega). \tag{4.17}$$

(the dependence of the constant C on the data is omitted as $\boldsymbol{g}$, $\boldsymbol{v}_0$, θ_0, or $\mathbb{B}_0$ are fixed functions in our setting). Utilizing the L^2-orthonormality of the basis vectors $\{\boldsymbol{w}_i\}_{i=1}^{\ell}$, estimate (4.17) yields

$$\sup_{t\in(0,T)}\sum_{i=1}^{\ell}(\alpha_{\ell n}^i(t))^2=\sup_{t\in(0,T)}\|\boldsymbol{v}_{\ell n}(t)\|_2^2\le C(\omega). \tag{4.18}$$

Hence, recalling $\boldsymbol{w}_i\in W^{1,\infty}(\Omega;\mathbb{R}^d)$, $i=1,\dots,\ell$ and then also the definition (4.10) and the estimate (4.1), we obtain

$$\|\boldsymbol{v}_{\ell n}\|_{L^\infty W^{1,\infty}}+\|\mathbb{S}_{\ell n}^{\omega}\|_{L^\infty L^\infty}\le C(\omega,\ell). \tag{4.19}$$

Using (4.19) in (4.11), we see that

$$\begin{aligned}\|(\alpha_{\ell n}^i)'\|_{L^2(0,T;\mathbb{R})}&=\|(\partial_t\boldsymbol{v}_{\ell n},\boldsymbol{w}_i)\|_{L^2(0,T;\mathbb{R})}\\&=\|(\boldsymbol{v}_{\ell n}\otimes\boldsymbol{v}_{\ell n}-\mathbb{S}_{\ell n}^{\omega},\nabla\boldsymbol{w}_i)-\alpha(\boldsymbol{v}_{\ell n},\boldsymbol{w}_i)_{\partial\Omega}+(\boldsymbol{g},\boldsymbol{w}_i)\|_{L^2(0,T;\mathbb{R})}\\&\le C(\omega,\ell)+C(\ell)\|\boldsymbol{g}\|_{L^2L^2}.\end{aligned} \tag{4.20}$$

Thus, we get

$$\|\partial_t\boldsymbol{v}_{\ell n}\|_{L^2W^{1,\infty}}=\|\textstyle\sum_{i=1}^{\ell}(\alpha_{\ell n}^i)'\boldsymbol{w}_i\|_{L^2W^{1,\infty}}\le C(\omega,\ell) \tag{4.21}$$

and, using the fundamental theorem of calculus and Hölder's inequality, also that

$$|\alpha_{\ell n}^i(t)-\alpha_{\ell n}^i(s)|\le\int_s^t|(\alpha_{\ell n}^i)'|\le C(\omega,\ell)|t-s|^{\frac12}\quad\text{for every } t,s\in[0,T] \tag{4.22}$$

and any $i=1,\dots,\ell$.

Next, we multiply the j-th equation in (4.12) by $\beta_{\ell n}^j$ and sum the result over $j=1,\dots,n$. Note that the convective term vanishes after integration by parts and use of $(1.14)_1$ and (1.8). Also the term including $\mathbb{W}\boldsymbol{v}_{\ell n}$ vanishes due to symmetry of $\mathbb{B}_{\ell n}^2$. Thus, we obtain

$$\begin{aligned}\frac12\frac{\mathrm{d}}{\mathrm{d}t}\|\mathbb{B}_{\ell n}\|_2^2+(\mathbb{P}(\theta_{\ell n},\mathbb{B}_{\ell n}),\mathbb{B}_{\ell n})+\|\sqrt{\lambda(\theta_{\ell n})}\nabla\mathbb{B}_{\ell n}\|_2^2&\\=(2ag_\omega(\mathbb{B}_{\ell n},\theta_{\ell n})\mathbb{D}\boldsymbol{v}_{\ell n}\mathbb{B}_{\ell n},\mathbb{B}_{\ell n})\quad\text{a.e. in }(0,T).&\end{aligned} \tag{4.23}$$

Then using (4.14), (3.7), (3.4) and (4.1) we obtain, after integration over $(0,t)$, $t\in(0,T)$, that

$$\|\mathbb{B}_{\ell n}(t)\|_2^2+\int_0^t\|\mathbb{B}_{\ell n}\|_{2+q}^{2+q}+\int_0^t\|\nabla\mathbb{B}_{\ell n}\|_2^2\le\|Q_n\mathbb{B}_0^{\omega}\|_2^2+C(\omega,\ell).$$

From this, using properties of Q_n and (4.5), we easily read that

$$\|\mathbb{B}_{\ell n}\|_{L^\infty L^2}+\|\mathbb{B}_{\ell n}\|_{L^{2+q}L^{2+q}}+\|\nabla\mathbb{B}_{\ell n}\|_{L^2L^2}\le C(\omega,\ell). \tag{4.24}$$

To estimate the time derivative of $\mathbb{B}_{\ell n}$, we take $\mathbb{A}\in L^{q+2}(0,T;W^{N,2}(\Omega))$ with $\|\mathbb{A}\|_{L^{q+2}W^{N,2}}\le1$ and use (4.12), Hölder's inequality, (4.24), (4.19), (3.4), (3.6),

(4.1), properties of Q_n and $(\min\{2, \frac{q+2}{q+1}\})' = q+2$ to get

$$\begin{aligned}
\langle \partial_t \mathbb{B}_{\ell n}, \mathbb{A}\rangle &= (\partial_t \mathbb{B}_{\ell n}, Q_n\mathbb{A})_Q \\
&= (\mathbb{B}_{\ell n}\otimes \boldsymbol{v}_{\ell n} - \lambda(\theta_{\ell n})\nabla \mathbb{B}_{\ell n}, \nabla Q_n\mathbb{A})_Q - (\mathbb{P}(\theta_{\ell n}, \mathbb{B}_{\ell n}), Q_n\mathbb{A})_Q \\
&\quad + (2g_\omega(\mathbb{B}_{\ell n}, \theta_{\ell n})(a\mathbb{D}\boldsymbol{v}_{\ell n} + \mathbb{W}\boldsymbol{v}_{\ell n})\mathbb{B}_{\ell n}, Q_n\mathbb{A})_Q \\
&\le C(\omega,\ell)\int_Q \big((|\mathbb{B}_{\ell n}| + |\nabla \mathbb{B}_{\ell n}|)|\nabla Q_n\mathbb{A}| + (|\mathbb{B}_{\ell n}|^{q+1}+1)|Q_n\mathbb{A}|\big) \\
&\le C(\omega,\ell)\int_0^T (\|\nabla \mathbb{B}_{\ell n}\|_1 + \|\mathbb{B}_{\ell n}\|_{q+1}^{q+1}+1)\|Q_n\mathbb{A}\|_{1,\infty} \\
&\le C(\omega,\ell)\int_0^T (\|\nabla \mathbb{B}_{\ell n}\|_2 + \|\mathbb{B}_{\ell n}\|_{q+2}^{q+1}+1)\|Q_n\mathbb{A}\|_{N,2} \\
&\le C(\omega,\ell)\|\mathbb{A}\|_{L^{q+2}W^{N,2}} \le C(\omega,\ell).
\end{aligned}$$

Hence, we can conclude

$$\|\partial_t \mathbb{B}_{\ell n}\|_{L^{\frac{q+2}{q+1}}W^{-N,2}} \le C(\omega,\ell). \tag{4.25}$$

Next, we multiply the k-th equation in (4.13) by $\gamma_{\ell n}^k$, sum the result over $k = 1,\dots,n$, use $(1.14)_1$, (1.8) and integration by parts in the convective term to get

$$\frac{c_v}{2}\frac{\mathrm{d}}{\mathrm{d}t}\|\theta_{\ell n}\|_2^2 + \|\sqrt{\kappa(\theta_{\ell n})}\nabla\theta_{\ell n}\|_2^2 + \omega\|\nabla\theta_{\ell n}\|_{r+2}^{r+2} = (\mathbb{S}_{\ell n}^\omega\cdot\mathbb{D}\boldsymbol{v}_{\ell n}, \theta_{\ell n}) \tag{4.26}$$

a.e. in $(0,T)$. Thus, integrating this inequality over time, using (4.19) and Young's, Gronwall's and Poincaré's inequalities, properties of R_n and (4.5), we deduce

$$\|\theta_{\ell n}(t)\|_{L^\infty L^2} + \|\sqrt{\kappa(\theta_{\ell n})}\nabla\theta_{\ell n}\|_{L^2L^2} + \|\theta_{\ell n}\|_{L^{r+2}W^{1,r+2}} \le C(\omega,\ell). \tag{4.27}$$

Furthermore, taking $\tau \in L^{r+2}(0,T;W^{N,2}(\Omega))$ with $\|\tau\|_{L^{r+2}W^{N,2}} \le 1$ and using (4.13), Young's inequality, Hölder's inequality, (3.3), (4.19), (4.27) and properties of R_n, we obtain

$$\begin{aligned}
\langle \partial_t\theta_{\ell n}, \tau\rangle &= (\partial_t\theta_{\ell n}, R_n\tau)_Q \\
&= (c_v\theta_{\ell n}\boldsymbol{v}_{\ell n} - \kappa(\theta_{\ell n})\nabla\theta_{\ell n} - \omega|\nabla\theta_{\ell n}|^r\nabla\theta_{\ell n}, \nabla R_n\tau)_Q + (\mathbb{S}_{\ell n}^\omega\cdot\mathbb{D}\boldsymbol{v}_{\ell n}, R_n\tau)_Q \\
&\le C(\omega,\ell)\int_Q \Big((|\theta_{\ell n}| + |\theta_{\ell n}|^{\frac{r}{2}}\big|\sqrt{\kappa(\theta_{\ell n})}\nabla\theta_{\ell n}\big| + |\nabla\theta_{\ell n}|^{r+1})|\nabla R_n\tau| + |R_n\tau|\Big) \\
&\le C(\omega,\ell)\int_0^T\int_\Omega \Big(|\theta_{\ell n}|^{r+1} + \Big|\sqrt{\kappa(\theta_{\ell n})}\nabla\theta_{\ell n}\Big|^{\frac{2r+2}{r+2}} + |\nabla\theta_{\ell n}|^{r+1} + 1\Big)\|R_n\tau\|_{1,\infty} \\
&\le C(\omega,\ell)\int_0^T \Big(\|\theta_{\ell n}\|_{r+2}^{r+1} + \|\sqrt{\kappa(\theta_{\ell n})}\nabla\theta_{\ell n}\|_2^{\frac{2r+2}{r+2}} + \|\nabla\theta_{\ell n}\|_{r+2}^{r+1} + 1\Big)\|R_n\tau\|_{N,2} \\
&\le C(\omega,\ell)\|\tau\|_{L^{r+2}W^{N,2}} \le C(\omega,\ell),
\end{aligned}$$

hence

$$\|\partial_t\theta_{\ell n}\|_{L^{\frac{r+2}{r+1}}W^{-N,2}} \le C(\omega,\ell). \tag{4.28}$$

The limit $n\to\infty$. For every $i = 1,\dots,\ell$, the sequence $\{\alpha_{\ell n}^i\}_{n=1}^\infty \subset \mathcal{C}([0,T];\mathbb{R})$ is bounded due to (4.18) and uniformly equicontinuous by (4.22). Hence, using the Arzelà–Ascoli theorem, for every $i = 1,\dots,\ell$, we obtain $\alpha_\ell^i \in \mathcal{C}([0,T];\mathbb{R})$ and a subsequence (not relabelled) such that

$$\alpha_{\ell n}^i \to \alpha_\ell^i \quad \text{strongly in } \mathcal{C}([0,T];\mathbb{R}) \tag{4.29}$$

as $n \to \infty$. Then, we define

$$\boldsymbol{v}_\ell \coloneqq \sum_{i=1}^{\ell} \alpha_\ell^i \boldsymbol{w}_i \in \mathcal{C}([0,T]; W^{1,\infty}(\Omega;\mathbb{R}^d) \cap W^{1,2}_{\boldsymbol{n},\mathrm{div}})$$

and note that

$$\boldsymbol{v}_{\ell n} \to \boldsymbol{v}_\ell \quad \text{strongly in } \mathcal{C}([0,T]; W^{1,\infty}(\Omega;\mathbb{R}^d)). \tag{4.30}$$

According to estimates (4.21), (4.24), (4.25), (4.27), (4.28) and using reflexivity of the underlying spaces and the Aubin–Lions lemma, there exist subsequences $\{\boldsymbol{v}_{\ell n}\}_{n=1}^\infty$, $\{\mathbb{B}_{\ell n}\}_{n=1}^\infty$, $\{\theta_{\ell n}\}_{n=1}^\infty$ and their limits $\boldsymbol{v}_\ell$, $\mathbb{B}_\ell$, θ_ℓ, such that

$$\partial_t \boldsymbol{v}_{\ell n} \overset{*}{\rightharpoonup} \partial_t \boldsymbol{v}_\ell \qquad \text{weakly* in } L^2(0,T; W^{1,\infty}(\Omega;\mathbb{R}^d)), \tag{4.31}$$

$$\boldsymbol{v}_{\ell n} \rightharpoonup \boldsymbol{v}_\ell \qquad \text{weakly in } L^2(0,T; L^2(\partial\Omega;\mathbb{R}^d)), \tag{4.32}$$

$$\mathbb{B}_{\ell n} \rightharpoonup \mathbb{B}_\ell \qquad \text{weakly in } L^2(0,T, W^{1,2}(\Omega;\mathbb{R}^{d\times d}_{\mathrm{sym}})), \tag{4.33}$$

$$\mathbb{B}_{\ell n} \to \mathbb{B}_\ell \qquad \text{strongly in } L^{2+q)}(Q;\mathbb{R}^{d\times d}_{\mathrm{sym}}) \text{ and a.e. in } Q, \tag{4.34}$$

$$\partial_t \mathbb{B}_{\ell n} \rightharpoonup \partial_t \mathbb{B}_\ell \qquad \text{weakly in } L^{\frac{q+2}{q+1}}(0,T; W^{-N,2}(\Omega;\mathbb{R}^{d\times d}_{\mathrm{sym}})), \tag{4.35}$$

$$\theta_{\ell n} \rightharpoonup \theta_\ell \qquad \text{weakly in } L^{r+2}(0,T, W^{1,r+2}(\Omega;\mathbb{R})), \tag{4.36}$$

$$\theta_{\ell n} \to \theta_\ell \qquad \text{strongly in } L^{r+2+\frac{4}{d})}(Q;\mathbb{R}^{d\times d}_{\mathrm{sym}}) \text{ and a.e. in } Q, \tag{4.37}$$

$$\partial_t \theta_{\ell n} \rightharpoonup \partial_t \theta_\ell \qquad \text{weakly in } L^{\frac{r+2}{r+1}}(0,T; W^{-N,2}(\Omega;\mathbb{R})). \tag{4.38}$$

Now, we explain how to take the limit in the non-linear terms appearing in (4.11), (4.12) and (4.13). To handle most of the terms, namely

$$\begin{gathered}\boldsymbol{v}_{\ell n}\otimes\boldsymbol{v}_{\ell n}, \quad \nu(\theta_{\ell n})\mathbb{D}\boldsymbol{v}_{\ell n}, \quad \mathbb{S}^\omega_{\ell n}, \quad \mathbb{P}(\theta_{\ell n},\mathbb{B}_{\ell n}), \quad \lambda(\theta_{\ell n})\nabla\mathbb{B}_{\ell n},\\ g_\omega(\mathbb{B}_{\ell n},\theta_{\ell n})(a\mathbb{D}\boldsymbol{v}_{\ell n}+\mathbb{W}\boldsymbol{v}_{\ell n})\mathbb{B}_{\ell n}, \quad \boldsymbol{v}_{\ell n}\cdot\nabla\theta_{\ell n}, \quad \mathbb{S}^\omega_{\ell n}\cdot\mathbb{D}\boldsymbol{v}_{\ell n},\end{gathered}$$

we use the following standard argument: all these terms can be seen as a product of a weakly converging sequence with a strongly converging sequence, obtained via Vitali's theorem, (3.1), continuity of g_ω and pointwise convergence of $\boldsymbol{v}_{\ell n}$, $\mathbb{B}_{\ell n}$ and $\theta_{\ell n}$. This argument is sufficient to take the limit $n \to \infty$ in the equations (4.11) and (4.12). In (4.12), we first multiply the equation by a function $\varphi \in \mathcal{C}^1([0,T];\mathbb{R})$, integrate over $(0,T)$, then take the limit and finally use the density of functions of the form $\varphi\mathbb{A}$, $\mathbb{A} \in \mathrm{span}\{\mathbb{W}_j\}_{j=1}^\infty$, in the space $L^{(q+2)'}(0,T; W^{N,2}(\Omega;\mathbb{R}^{d\times d}_{\mathrm{sym}}))$. This way, defining also

$$\mathbb{S}^\omega_\ell \coloneqq 2\nu(\theta_\ell)|\mathbb{D}\boldsymbol{v}_\ell|^2 + 2a\mu g_\omega(\mathbb{B}_\ell,\theta_\ell)\theta_\ell\mathbb{B}_\ell,$$

we obtain

$$\begin{gathered}(\partial_t\boldsymbol{v}_\ell,\boldsymbol{w}_i) - (\boldsymbol{v}_\ell\otimes\boldsymbol{v}_\ell,\nabla\boldsymbol{w}_i) + (\mathbb{S}^\omega_\ell,\nabla\boldsymbol{w}_i) + \alpha(\boldsymbol{v}_\ell,\boldsymbol{w}_i)_{\partial\Omega} = (\boldsymbol{g},\boldsymbol{w}_i)\\ \text{for every } i=1,\dots,\ell, \text{ and a.e. in } (0,T),\end{gathered} \tag{4.39}$$

and

$$\begin{aligned}&\langle\partial_t\mathbb{B}_\ell,\mathbb{A}\rangle - (\mathbb{B}_\ell\otimes\boldsymbol{v}_\ell,\nabla\mathbb{A})_Q + (\mathbb{P}(\theta_\ell,\mathbb{B}_\ell),\mathbb{A})_Q + (\lambda(\theta_\ell)\nabla\mathbb{B}_\ell,\nabla\mathbb{A})_Q\\ &\quad= (2g_\omega(\mathbb{B}_\ell,\theta_\ell)(a\mathbb{D}\boldsymbol{v}_\ell+\mathbb{W}\boldsymbol{v}_\ell)\mathbb{B}_\ell,\mathbb{A})_Q\\ &\qquad\qquad\text{for all } \mathbb{A}\in L^{q+2}(0,T; W^{N,2}(\Omega;\mathbb{R}^{d\times d}_{\mathrm{sym}})).\end{aligned} \tag{4.40}$$

However, the space of test functions in (4.40) can be enlarged using a standard density argument. Indeed, using Hölder's inequality, it is easy to see that every term of (4.40) (taking aside the time derivative) is well defined provided that

$$\mathbb{A} \in L^2(0,T;W^{1,2}(\Omega;\mathbb{R}^{d\times d}_{\mathrm{sym}})) \cap L^{q+2}(Q;\mathbb{R}^{d\times d}_{\mathrm{sym}})$$

and thus, we can read from (4.40) that

$$\partial_t \mathbb{B}_\ell \in \left(L^2(0,T;W^{1,2}(\Omega;\mathbb{R}^{d\times d}_{\mathrm{sym}})) \cap L^{q+2}(Q;\mathbb{R}^{d\times d}_{\mathrm{sym}})\right)^*.$$

Since we also have that

$$\mathbb{B}_\ell \in L^2(0,T;W^{1,2}(\Omega;\mathbb{R}^{d\times d}_{\mathrm{sym}})) \cap L^{q+2}(Q;\mathbb{R}^{d\times d}_{\mathrm{sym}}), \tag{4.41}$$

it follows from Lemma A.1 below that

$$\mathbb{B}_\ell \in \mathcal{C}([0,T];L^2(\Omega;\mathbb{R}^{d\times d}_{\mathrm{sym}})). \tag{4.42}$$

The value of $\mathbb{B}_\ell(0)$ can be identified by a standard argument, which we briefly outline here. Using $\mathbb{A}(t,x) = \psi(t)\mathbb{P}(x)$ in (4.40), where $\psi \in \mathcal{C}^1([0,T];\mathbb{R})$, $\psi(0)=1$, $\psi(T)=0$, and $\mathbb{P} \in W^{N,2}(\Omega;\mathbb{R}^{d\times d}_{\mathrm{sym}})$, one gets, after integration by parts, that

$$\begin{aligned}(\mathbb{B}_\ell(0),\mathbb{P}) = &-(\mathbb{B}_\ell,\mathbb{P}\partial_t\psi)_Q + (\boldsymbol{v}_\ell\cdot\nabla\mathbb{B}_\ell,\mathbb{P}\psi)_Q + (\mathbb{P}(\theta_\ell,\mathbb{B}_\ell),\mathbb{P}\psi)_Q\\ &- (\lambda(\theta_\ell)\nabla\mathbb{B}_\ell,\nabla\mathbb{P}\psi)_Q - (2g_\omega(\mathbb{B}_\ell,\theta_\ell)(a\mathbb{D}\boldsymbol{v}_\ell + \mathbb{W}\boldsymbol{v}_\ell)\mathbb{B}_\ell,\mathbb{P}\psi)_Q.\end{aligned} \tag{4.43}$$

On the other hand, exactly the same expression can be obtained also for $(\mathbb{B}_0^\omega,\mathbb{P})$ if one multiplies (4.12) by ψ, integrate over $(0,T)$ and by parts in the time derivative using (4.14) and uses completeness of $\{\mathbb{W}_j\}_{j=1}^\infty$ in $W^{N,2}(\Omega;\mathbb{R}^{d\times d}_{\mathrm{sym}})$ and the same arguments as before to take the limit $n\to\infty$. But since $\mathbb{P}$ was arbitrary and $W^{N,2}(\Omega;\mathbb{R}^{d\times d}_{\mathrm{sym}})$ is dense in $L^2(\Omega;\mathbb{R}^{d\times d}_{\mathrm{sym}})$, we conclude

$$\mathbb{B}_\ell(0) = \mathbb{B}_0^\omega. \tag{4.44}$$

We can use an analogous procedure to identify $\boldsymbol{v}_\ell(0)$, but here the situation is simpler since (4.30) directly implies $\boldsymbol{v}_\ell \in \mathcal{C}([0,T];W^{1,\infty}(\Omega;\mathbb{R}^d))$ and we obtain

$$\boldsymbol{v}_\ell(0) = P_\ell \boldsymbol{v}_0. \tag{4.45}$$

Our aim is now to take the limit in equation (4.13), where we need to justify the limit in the terms $\kappa(\theta_{\ell n})\nabla\theta_{\ell n}$ and $|\nabla\theta_{\ell n}|^r\nabla\theta_{\ell n}$ (the term $2\nu(\theta_{\ell n})|\mathbb{D}\boldsymbol{v}_{\ell n}|^2$ is easy due to (4.30)). For the first one, we use (3.3), (4.37) and Vitali's theorem to get

$$\sqrt{\kappa(\theta_{\ell n})} \to \sqrt{\kappa(\theta_\ell)} \quad \text{strongly in } L^{2+\frac{4}{r}}(Q;\mathbb{R}) \tag{4.46}$$

and then we combine this with (4.36), to obtain

$$\sqrt{\kappa(\theta_{\ell n})}\nabla\theta_{\ell n} \rightharpoonup \sqrt{\kappa(\theta_\ell)}\nabla\theta_\ell \quad \text{weakly in } L^1(Q;\mathbb{R}^d). \tag{4.47}$$

However, by (4.27) we know that (4.47) is valid also in $L^2(Q;\mathbb{R}^d)$ up to a subsequence, and hence, using again (4.46), we obtain

$$\kappa(\theta_{\ell n})\nabla\theta_{\ell n} = \sqrt{\kappa(\theta_{\ell n})}\sqrt{\kappa(\theta_{\ell n})}\nabla\theta_{\ell n} \rightharpoonup \sqrt{\kappa(\theta_\ell)}\sqrt{\kappa(\theta_\ell)}\nabla\theta_\ell = \kappa(\theta_\ell)\nabla\theta_\ell \tag{4.48}$$

weakly in $L^{\frac{r+2}{r+1}}(Q;\mathbb{R}^d)$.

Finally, due to (4.27), there exists $K \in L^{(r+2)'}(Q;\mathbb{R}^d)$ such that

$$|\nabla\theta_{\ell n}|^r\nabla\theta_{\ell n} \rightharpoonup K \quad \text{weakly in } L^{(r+2)'}(Q;\mathbb{R}^d). \tag{4.49}$$

Then, using also (4.48) and previous convergence results, we can take the limit in (4.13) and obtain, for all $\tau \in L^{r+2}(0,T;W^{N,2}(\Omega;\mathbb{R}))$, that

$$\langle c_v\partial_t\theta_\ell,\tau\rangle - (c_v\theta_\ell\boldsymbol{v}_\ell,\nabla\tau)_Q + (\kappa(\theta_\ell)\nabla\theta_\ell,\nabla\tau)_Q + \omega(K,\nabla\tau)_Q = (\mathbb{S}_\ell^\omega\cdot\mathbb{D}\boldsymbol{v}_\ell,\tau)_Q. \tag{4.50}$$

Recalling (4.38), (4.48) and (4.49), we easily conclude, using a density argument, that (4.50) is valid for all $\tau \in L^{r+2}(0,T;W^{1,r+2}(\Omega;\mathbb{R}))$ and that the time derivative extends to the functional $\partial_t\theta_\ell \in L^{(r+2)'}(0,T;W^{-1,(r+2)'}(\Omega;\mathbb{R}))$. Thus, using Lemma A.1 below, we also see that

$$\theta_\ell \in \mathcal{C}([0,T];L^2(\Omega;\mathbb{R})). \tag{4.51}$$

Furthermore, choosing $\tau = \theta_\ell$ in (4.50), rewriting the time derivative term and integrating by parts in the convective term leads to

$$\omega(K,\nabla\theta_\ell)_Q = \frac{c_v}{2}(\|\theta_\ell(0)\|_2^2 - \|\theta_\ell(T)\|_2^2) - \int_Q \kappa(\theta_\ell)|\nabla\theta_\ell|^2 + (\mathbb{S}_\ell^\omega \cdot \mathbb{D}\boldsymbol{v}_\ell, \theta_\ell)_Q. \tag{4.52}$$

We use this information to identify K as follows. We note that weak lower semicontinuity and (4.47) (which is valid in $L^2(Q;\mathbb{R}^d)$) imply

$$\int_Q \kappa(\theta_\ell)|\nabla\theta_\ell|^2 \le \liminf_{n\to\infty} \int_Q \kappa(\theta_{\ell n})|\nabla\theta_{\ell n}|^2. \tag{4.53}$$

Thus, if we integrate (4.26) over $(0,T)$ and use (4.53), (4.30), weak lower semicontinuity of $\|\cdot\|_2$ and the convergence results above to take the limes superior $n\to\infty$ and then apply (4.52), we get

$$\begin{aligned}
&\omega \limsup_{n\to\infty} \int_Q |\nabla\theta_{\ell n}|^{r+2} \\
&\quad = -\liminf_{n\to\infty} \frac{c_v}{2}\|\theta_{\ell n}(T)\|_2^2 + \frac{c_v}{2}\|\theta_0^\omega\|_2^2 - \liminf_{n\to\infty}\int_Q \kappa(\theta_{\ell n})|\nabla\theta_{\ell n}|^2 \\
&\qquad + \lim_{n\to\infty}(\mathbb{S}_{\ell n}^\omega \cdot \mathbb{D}\boldsymbol{v}_{\ell n}, \theta_{\ell n})_Q \\
&\quad \le -\frac{c_v}{2}\|\theta_\ell(T)\|_2^2 + \frac{c_v}{2}\|\theta_0^\omega\|_2^2 - \int_Q \kappa(\theta_\ell)|\nabla\theta_\ell|^2 + (\mathbb{S}_\ell^\omega \cdot \mathbb{D}\boldsymbol{v}_\ell, \theta_\ell)_Q \\
&\quad = \frac{c_v}{2}\|\theta_0^\omega\|_2^2 - \frac{c_v}{2}\|\theta_\ell(0)\|_2^2 + \omega(K,\nabla\theta_\ell)_Q.
\end{aligned} \tag{4.54}$$

To identify the initial condition for $\theta_\ell(0)$, it is enough to show that

$$\theta_\ell(t) \rightharpoonup \theta_0^\omega \quad \text{weakly in } L^2(\Omega;\mathbb{R}) \tag{4.55}$$

as $t \to 0+$ since then we can use (4.51) to conclude

$$\theta_\ell(0) = \theta_0^\omega \quad \text{a.e. in } \Omega \tag{4.56}$$

by the uniqueness of a (weak) limit. To prove (4.55), we return to (4.13), which we multiply by $\varphi \in W^{1,\infty}(0,T;\mathbb{R})$ fulfilling $\varphi(0) = 1$, $\varphi(T) = 0$ and integrate the result over $(0,T)$ to get

$$-(c_v\theta_0^\omega, w_k) - \int_0^T (c_v\theta_{\ell n}, w_k)\partial_t\varphi = \int_0^T h_n\varphi. \tag{4.57}$$

for all $k = 1,\dots,n$, where we integrated by parts and abbreviated

$$h_n = (c_v\theta_{\ell n}\boldsymbol{v}_{\ell n}, \nabla w_k) - (\kappa(\theta_{\ell n})\nabla\theta_{\ell n} + \omega|\nabla\theta_{\ell n}|^r\nabla\theta_{\ell n}, \nabla w_k) + (\mathbb{S}_{\ell n}^\omega \cdot \mathbb{D}\boldsymbol{v}_{\ell n}, w_k).$$

It follows from the results above (cf. the derivation of (4.50)) that

$$h_n \rightharpoonup h \quad \text{weakly in } L^{(r+2)'}(0,T;\mathbb{R}),$$

where

$$h = (c_v\theta_\ell\boldsymbol{v}_\ell, \nabla w_k) - (\kappa(\theta_\ell)\nabla\theta_\ell, \nabla w_k) - \omega(K,\nabla w_k) + (\mathbb{S}_\ell^\omega \cdot \mathbb{D}\boldsymbol{v}_\ell, w_k).$$

Thus, by taking the limit $n \to \infty$ in (4.57), we arrive at

$$-(c_v\theta_0^\omega, w_k) - \int_0^T (c_v\theta_\ell, w_k)\partial_t\varphi = \int_0^T h\varphi.$$

Making now a special choice

$$\varphi_\varepsilon(s) = \begin{cases} 1 & s \le t, \\ 1 - \frac{s-t}{\varepsilon} & s \in (t, t+\varepsilon), \\ 0 & s \ge t+\varepsilon, \end{cases}$$

where $t \in (0,T)$ and $0 < \varepsilon < T - t$, leads to

$$-(c_v\theta_0^\omega, w_k) + \frac{1}{\varepsilon}\int_t^{t+\varepsilon} (c_v\theta_\ell, w_k) = \int_0^{t+\varepsilon} h\varphi_\varepsilon.$$

Furthermore, we can take the limit $\varepsilon \to 0+$ in this equation using (4.51) on the left hand side and absolute continuity of integral on the right hand side to get

$$-(c_v\theta_0^\omega, w_k) + (c_v\theta_\ell(t), w_k) = \int_0^t f.$$

Finally, taking the limit $t \to 0+$ yields

$$\lim_{t\to 0+} (\theta_\ell(t), w_k) = (\theta_0^\omega, w_k),$$

for all $k = 1, \dots, n$, from which (4.55) follows by exploiting the density of the set $\operatorname{span}\{w_k\}_{k=1}^\infty$ in $L^2(\Omega;\mathbb{R})$. Hence, the identity (4.56) is proved and (4.54) hereby simplifies to

$$\limsup_{n\to\infty} \int_Q |\nabla\theta_{\ell n}|^{r+2} \le \int_Q K \cdot \nabla\theta_\ell. \tag{4.58}$$

Since the operator $\boldsymbol{u} \mapsto |\boldsymbol{u}|^r\boldsymbol{u}$ is monotone and continuous, it is standard to show, using (4.58) and the Minty method, that

$$K = |\nabla\theta_\ell|^r\nabla\theta_\ell \quad \text{a.e. in } Q.$$

Hence, we proved that

$$\langle c_v\partial_t\theta_\ell, \tau\rangle - (c_v\theta_\ell \boldsymbol{v}_\ell\nabla\tau)_Q + (\kappa(\theta_\ell)\nabla\theta_\ell + \omega|\nabla\theta_\ell|^r\nabla\theta_\ell, \nabla\tau)_Q = (\mathbb{S}_\ell^\omega \cdot \mathbb{D}\boldsymbol{v}_\ell, \tau)_Q \tag{4.59}$$

for all $\tau \in L^{r+2}(0,T;W^{1,r+2}(\Omega;\mathbb{R}))$.

Positive definiteness of $\mathbb{B}_\ell$ and positivity of θ_ℓ. Here we follow the method developed in [4]. We shall use the notation

$$h_+ = \max\{0,h\}, \qquad h_- = \min\{0,h\}.$$

We choose a fixed vector $\boldsymbol{x} \in \mathbb{R}^d$ with $|\boldsymbol{x}| = 1$, and $t \in (0,T)$. The idea is to use

$$\mathbb{A}_{\boldsymbol{x}} = \chi_{(0,t)}(b-\omega)_- \boldsymbol{x}\otimes\boldsymbol{x}, \quad \text{where} \quad b \coloneqq \mathbb{B}_\ell\boldsymbol{x}\cdot\boldsymbol{x}.$$

in (4.40). The function $\mathbb{A}_{\boldsymbol{x}}$ belongs to $L^2(0,T;W^{1,2}(\Omega;\mathbb{R}^{d\times d}_{\mathrm{sym}})) \cap L^{q+2}(Q;\mathbb{R}^{d\times d}_{\mathrm{sym}})$ and is thus a valid test function in (4.40). The key property of $\mathbb{A}_{\boldsymbol{x}}$ is that it vanishes whenever the smallest eigenvalue of $\mathbb{B}_\ell$ is greater than ω. Thus, we have

$$(\Lambda(\mathbb{B}_\ell) - \omega)_+(b-\omega)_- = 0,$$

which implies

$$g_\omega(\mathbb{B}_\ell, \theta_\ell)\mathbb{A}_{\boldsymbol{x}} = 0 \quad \text{a.e. in } Q. \tag{4.60}$$

Let us now evaluate separately the terms arising from the choice $\mathbb{A} = \mathbb{A}_{\boldsymbol{x}}$ in (4.40). For the time derivative, we write

$$\langle \partial_t \mathbb{B}_\ell, \mathbb{A}_{\boldsymbol{x}} \rangle = \int_0^t \langle \partial_t (b-\omega), (b-\omega)_- \rangle = \tfrac{1}{2} \|(b-\omega)_-(t)\|_2^2, \tag{4.61}$$

where we applied Lemma A.2 below for the Lipschitz function $s \mapsto s_-$ and also (4.44) and (4.4) to eliminate the value at $t = 0$. Furthermore, using integration by parts, $\boldsymbol{v}_\ell \cdot \boldsymbol{n} = 0$ and $\operatorname{div} \boldsymbol{v}_\ell = 0$, we get

$$(\mathbb{B}_\ell \otimes \boldsymbol{v}_\ell, \nabla \mathbb{A}_{\boldsymbol{x}})_Q = \int_0^t ((b-\omega)\boldsymbol{v}_\ell, \nabla (b-\omega)_-) = \frac{1}{2} \int_0^t \int_{\partial\Omega} ((b-\omega)_-)^2 \boldsymbol{v}_\ell \cdot \boldsymbol{n} = 0$$

and also

$$(\lambda(\theta_\ell) \nabla \mathbb{B}_\ell, \nabla \mathbb{A}_{\boldsymbol{x}})_Q = \int_0^t \| \sqrt{\lambda(\theta_\ell)} \nabla (b-\omega)_- \|_2^2 \geq 0.$$

Moreover, we have $b - \omega_P < b - \omega$ and thus, the assumption (3.10) yields

$$\begin{aligned} (\mathbb{P}(\theta_\ell, \mathbb{B}_\ell), \mathbb{A}_{\boldsymbol{x}})_Q &= \int_0^t \int_\Omega (b-\omega)_- \, \mathbb{P}(\theta_\ell, \mathbb{B}_\ell) \boldsymbol{x} \cdot \boldsymbol{x} \\ &= \int_0^t \int\limits_{\{b<\omega\}} (b-\omega) \mathbb{P}(\theta_\ell, (\mathbb{B}_\ell - \omega_P \mathbb{I}) + \omega_P \mathbb{I}) \boldsymbol{x} \cdot \boldsymbol{x} \geq 0. \end{aligned}$$

In addition, the right hand side of (4.40) vanishes due to (4.60). Thus, using the above computation in (4.40), we obtain

$$\|(b-\omega)_-(t)\|_2^2 \leq 0$$

for all $t \in (0,T)$ (recall (4.42)), whence

$$\mathbb{B}_\ell(t) \boldsymbol{x} \cdot \boldsymbol{x} \geq \omega |\boldsymbol{x}|^2 \quad \text{a.e. in } \Omega, \text{ for all } t \in (0,T) \text{ and for every } \boldsymbol{x} \in \mathbb{R}^d. \tag{4.62}$$

Note that this immediately yields $\mathbb{B}_\ell \in \mathbb{R}^{d\times d}_{>0}$, $\mathbb{B}_\ell^{-1} \in \mathbb{R}^{d\times d}_{>0}$ a.e. in Q, and thus

$$|\mathbb{B}_\ell^{-1}| = |\mathbb{B}_\ell^{-\frac{1}{2}} \mathbb{B}_\ell^{-\frac{1}{2}}| \leq |\mathbb{B}_\ell^{-\frac{1}{2}}|^2 = \operatorname{tr} \mathbb{B}_\ell^{-1} \leq \frac{d}{\omega}. \tag{4.63}$$

Also, using the identity

$$\nabla \mathbb{B}_\ell^{-1} = -\mathbb{B}_\ell^{-1} \nabla \mathbb{B}_\ell \mathbb{B}_\ell^{-1},$$

(which is standard for continuously differentiable functions and in general we can approximate $\mathbb{B}_\ell$ by smooth mappings and pass to the limit) and (4.24) we conclude that $\mathbb{B}_\ell^{-1}$ exists a.e. in Q and satisfies

$$\mathbb{B}_\ell^{-1} \in L^\infty(0,T; L^\infty(\Omega; \mathbb{R}^{d\times d}_{>0})) \cap L^2(0,T; W^{1,2}(\Omega; \mathbb{R}^{d\times d}_{>0})). \tag{4.64}$$

Moreover, recalling f from (2.9) and using the simple inequalities

$$\det \mathbb{B}_\ell \geq \omega^d \quad \text{and} \quad |\ln x| \leq x + \frac{1}{x}, \ x > 0,$$

it is easy to see that also

$$f(\mathbb{B}_\ell) \in L^2(0,T; W^{1,2}(\Omega; \mathbb{R}_{\geq 0})) \cap L^{q+2}(Q; \mathbb{R}_{\geq 0}).$$

Next, we prove positivity of θ_ℓ. Since $\theta_\ell \in L^{r+2}(0,T; W^{1,r+2}(\Omega; \mathbb{R}))$, we can use the analogous method as before. Indeed, we start by choosing

$$\tau = \chi_{(0,t)} (\theta_\ell - \omega)_- \in L^{r+2}(0,T; W^{1,r+2}(\Omega; \mathbb{R}))$$

as a test function in (4.59) to get (using $\operatorname{div} \boldsymbol{v}_\ell = 0$)

$$\begin{aligned}&\frac{c_v}{2}\|(\theta_\ell-\omega)_-(t)\|_2^2-\frac{c_v}{2}\|(\theta_\ell-\omega)_-(0)\|_2^2\\&\quad+\int_0^t\|\sqrt{\kappa(\theta_\ell)}\nabla(\theta_\ell-\omega)_-\|_2^2+\int_0^t\|\nabla(\theta_\ell-\omega)_-\|_{r+2}^{r+2}\\&\qquad=\int_0^t\big(\mathbb{S}_\ell^\omega\cdot\mathbb{D}\boldsymbol{v}_\ell,(\theta_\ell-\omega)_-\big)\le 0.\end{aligned}\tag{4.65}$$

Hence, using $\theta_\ell(0)=\theta_0^\omega\ge\omega$ in Ω and (4.51), we obtain that $\|(\theta_\ell(t)-\omega)_-\|_2=0$ for all $t\in(0,T)$, which means

$$\theta_\ell(t)\ge\omega\quad\text{a.e. in }\Omega\text{ and for all }t\in(0,T).\tag{4.66}$$

Consequently, since $\nabla\theta_\ell^{-1}=\theta_\ell^{-2}\nabla\theta_\ell$, we also obtain

$$\theta_\ell^{-1}\in L^\infty(0,T;L^\infty(\Omega;\mathbb{R}_{>0}))\cap L^{r+2}(0,T;W^{1,r+2}(\Omega;\mathbb{R}_{>0})).\tag{4.67}$$

From these findings we also easily read that

$$|\ln\theta_\ell|\le\theta_\ell+\frac{1}{\theta_\ell}\le\theta_\ell+\frac{1}{\omega}\quad\text{and}\quad|\nabla\ln\theta_\ell|=\frac{|\nabla\theta_\ell|}{\theta_\ell}\le\frac{1}{\omega}|\nabla\theta_\ell|,$$

hence also

$$\ln\theta_\ell\in L^{r+2}(0,T;W^{1,r+2}(\Omega;\mathbb{R})).$$

Entropy equation. In order to take the remaining limits $\ell\to\infty$ and $\omega\to0+$, we need to derive the entropy (in)equality from which we then deduce that $\det\mathbb{B}_\ell$ and θ_ℓ remain *strictly* positive a.e. in Q. First, we rewrite (4.59) in the form

$$\langle c_v\partial_t\theta_\ell,\tau\rangle+(c_v\boldsymbol{v}_\ell\cdot\nabla\theta_\ell,\tau)+(\kappa(\theta_\ell)\nabla\theta_\ell+\omega|\nabla\theta_\ell|^r\nabla\theta_\ell,\nabla\tau)=(\mathbb{S}_\ell^\omega\cdot\mathbb{D}\boldsymbol{v}_\ell,\tau)\tag{4.68}$$

for all $\tau\in W^{1,r+2}(\Omega;\mathbb{R})$ and a.e. in $(0,T)$. Then, we take $\phi\in W^{1,\infty}(\Omega;\mathbb{R})$ and note that $\tau=\theta_\ell^{-1}\phi$ can be used as a test function in (4.68) thanks to (4.67). This way, we get

$$\begin{aligned}&\langle c_v\partial_t\theta_\ell,\theta_\ell^{-1}\phi\rangle+(c_v\boldsymbol{v}_\ell\cdot\nabla\ln\theta_\ell,\phi)+(\kappa(\theta_\ell)\nabla\ln\theta_\ell,\nabla\phi)-(\kappa(\theta_\ell)|\nabla\ln\theta_\ell|^2,\phi)\\&\quad+\omega(|\nabla\theta_\ell|^r\nabla\ln\theta_\ell,\nabla\phi)-\omega(|\nabla\theta_\ell|^r|\nabla\ln\theta_\ell|^2,\phi)\\&\qquad=(2\nu(\theta_\ell)\theta_\ell^{-1}|\mathbb{D}\boldsymbol{v}_\ell|^2+2a\mu g_\omega(\mathbb{B}_\ell,\theta_\ell)\mathbb{B}_\ell\cdot\mathbb{D}\boldsymbol{v}_\ell,\phi)\end{aligned}\tag{4.69}$$

a.e. in $(0,T)$. Similarly, we observe that $f'(\mathbb{B}_\ell)\phi=\mu(\mathbb{I}-\mathbb{B}_\ell^{-1})\phi$ (recall (2.14), (2.15)) is a valid test function in (4.40) due to (4.64). Thus, we obtain

$$\begin{aligned}&\langle\partial_t\mathbb{B}_\ell,f'(\mathbb{B}_\ell)\phi\rangle+(\boldsymbol{v}_\ell\cdot\nabla f(\mathbb{B}_\ell),\phi)\\&\quad+(\mu\mathbb{P}(\theta_\ell,\mathbb{B}_\ell)\cdot(\mathbb{I}-\mathbb{B}_\ell^{-1}),\phi)+(\mu\lambda(\theta_\ell)|\mathbb{B}_\ell^{-\frac12}\nabla\mathbb{B}_\ell\mathbb{B}_\ell^{-\frac12}|^2,\phi)\\&\qquad=-(\lambda(\theta_\ell)\nabla f(\mathbb{B}_\ell),\nabla\phi)+(2a\mu g_\omega(\mathbb{B}_\ell,\theta_\ell)\mathbb{B}_\ell\cdot\mathbb{D}\boldsymbol{v}_\ell,\phi)\end{aligned}\tag{4.70}$$

a.e. in $(0,T)$. If we define

$$\eta_\ell\coloneqq c_v\ln\theta_\ell-f(\mathbb{B}_\ell)\tag{4.71}$$

and

$$\begin{aligned}\xi_\ell\coloneqq{}&2\nu(\theta_\ell)\theta_\ell^{-1}|\mathbb{D}\boldsymbol{v}_\ell|^2+\kappa(\theta_\ell)|\nabla\ln\theta_\ell|^2+\omega|\nabla\theta_\ell|^r|\nabla\ln\theta_\ell|^2\\&+\mu\mathbb{P}(\theta_\ell,\mathbb{B}_\ell)\cdot(\mathbb{I}-\mathbb{B}_\ell^{-1})+\mu\lambda(\theta_\ell)|\mathbb{B}_\ell^{-\frac12}\nabla\mathbb{B}_\ell\mathbb{B}_\ell^{-\frac12}|^2\end{aligned}\tag{4.72}$$

and subtract (4.70) from (4.69), we get

$$\begin{aligned}&\langle c_v\partial_t\theta_\ell,\theta_\ell^{-1}\phi\rangle-\langle\partial_t\mathbb{B}_\ell,f'(\mathbb{B}_\ell)\phi\rangle+(\boldsymbol{v}_\ell\cdot\nabla\eta_\ell,\phi)\\&\quad+\big((\kappa(\theta_\ell)+\omega|\nabla\theta_\ell|^r)\nabla\ln\theta_\ell-\lambda(\theta_\ell)\nabla f(\mathbb{B}_\ell),\nabla\phi\big)=(\xi_\ell,\phi)\end{aligned}\tag{4.73}$$

a.e. in $(0,T)$ and for all $\phi\in W^{1,\infty}(\Omega;\mathbb{R})$. It remains to rewrite the time derivative accordingly. Concerning the term containing $\partial_t\theta_\ell$, note that $\psi(s)=\max\{|s|,\omega\}^{-1}$, $s\in\mathbb{R}$, is a bounded Lipschitz function. Since $\theta_\ell\geq\omega$ a.e. in Q by (4.66), we get

$$\int_1^{\theta_\ell}\psi(s)\,\mathrm{d}s=\int_1^{\theta_\ell}\frac{1}{s}\,\mathrm{d}s=\ln\theta_\ell.$$

Thus, Lemma A.2 below yields

$$\langle c_v\partial_t\theta_\ell,\theta_\ell^{-1}\phi\rangle=\frac{\mathrm{d}}{\mathrm{d}t}(c_v\ln\theta_\ell,\phi).$$

If we multiply this by $\varphi\in W^{1,\infty}((0,T);\mathbb{R})$ with $\varphi(T)=0$, integrate over $(0,T)$ and by parts, we are led to

$$\langle c_v\partial_t\theta_\ell,\theta_\ell^{-1}\phi\varphi\rangle=-(c_v\ln\theta_\ell,\phi\partial_t\varphi)_Q-(c_v\ln\theta_0^\omega,\phi\varphi(0)),\tag{4.74}$$

where we also used (4.56). Analogous ideas can be used to rewrite the second term of (4.73). However, since the duality $\langle\partial_t\mathbb{B}_\ell,f'(\mathbb{B}_\ell)\phi\rangle$ cannot be interpreted entry-wise, let us proceed more carefully. We apply Lemma A.1 below to obtain functions $\mathbb{B}_\ell^\varepsilon\in\mathcal{C}^1([0,T];W^{1,2}(\Omega;\mathbb{R}^{d\times d}_{>0})\cap L^{q+2}(\Omega;\mathbb{R}^{d\times d}_{>0}))$, $\varepsilon>0$, such that

$$\|\mathbb{B}_\ell^\varepsilon-\mathbb{B}_\ell\|_{L^2W^{1,2}\cap L^{q+2}L^{q+2}}+\|\partial_t\mathbb{B}_\ell^\varepsilon-\partial_t\mathbb{B}_\ell\|_{(L^2W^{1,2}\cap L^{q+2}L^{q+2})^*}\to0\tag{4.75}$$

as $\varepsilon\to0+$ and also $\Lambda(\mathbb{B}_\ell^\varepsilon)\geq\omega$ a.e. in Q. For such regularization, we have

$$\langle\partial_t\mathbb{B}_\ell^\varepsilon,f'(\mathbb{B}_\ell^\varepsilon)\phi\varphi\rangle=-(f(\mathbb{B}_\ell^\varepsilon(0)),\phi\varphi(0))-(f(\mathbb{B}_\ell^\varepsilon),\phi\partial_t\varphi)_Q\tag{4.76}$$

by the standard calculus and it remains to justify the limit $\varepsilon\to0+$ on both sides of (4.75). Since $\mathbb{B}_\ell\in\mathcal{C}([0,T];L^2(\Omega))$ (cf. (4.42)), we know that

$$\|\mathbb{B}_\ell^\varepsilon-\mathbb{B}_\ell\|_2\rightrightarrows0\quad\text{uniformly in }[0,T].\tag{4.77}$$

Now it is important to observe that since we have $\Lambda(\mathbb{B}_s)\geq\omega$ for all $s\in[0,1]$, where

$$\mathbb{B}_s\coloneqq(1-s)\mathbb{B}_\ell+s\mathbb{B}_\ell^\varepsilon,$$

the convergence (4.77) actually also implies

$$\|f(\mathbb{B}_\ell^\varepsilon)-f(\mathbb{B}_\ell)\|_2+\|(\mathbb{B}_\ell^\varepsilon)^{-1}-\mathbb{B}_\ell^{-1}\|_2\rightrightarrows0\quad\text{uniformly in }[0,T].\tag{4.78}$$

Indeed, this is a simple consequence of the identities

$$\begin{aligned}f(\mathbb{B}_\ell^\varepsilon)-f(\mathbb{B}_\ell)&=\int_0^1\frac{\mathrm{d}}{\mathrm{d}s}f(\mathbb{B}_s)\,\mathrm{d}s=\int_0^1\mu(\mathbb{I}-\mathbb{B}_s^{-1})\cdot(\mathbb{B}_\ell^\varepsilon-\mathbb{B}_\ell)\,\mathrm{d}s,\\(\mathbb{B}_\ell^\varepsilon)^{-1}-\mathbb{B}_\ell^{-1}&=\int_0^1\frac{\mathrm{d}}{\mathrm{d}s}\mathbb{B}_s^{-1}\,\mathrm{d}s=-\int_0^1\mathbb{B}_s^{-1}(\mathbb{B}_\ell^\varepsilon-\mathbb{B}_\ell)\mathbb{B}_s^{-1}\,\mathrm{d}s,\end{aligned}\tag{4.79}$$

(4.77) and the estimate

$$|\mathbb{B}_s^{-1}|\leq\operatorname{tr}\mathbb{B}_s^{-1}\leq\frac{d}{\Lambda(\mathbb{B}_s)}\leq\frac{d}{\omega}.$$

Using the same scheme as in (4.79), we also deduce from (4.42) and (4.44) that

$$f(\mathbb{B}_\ell)\in\mathcal{C}(0,T;L^2(\Omega;\mathbb{R})),\quad f(\mathbb{B}_\ell(0))=f(\mathbb{B}_0^\omega).\tag{4.80}$$

This and $(4.78)_1$ allow us to pass to the desired limit on the right-hand side of (4.76). Next, using (4.63), we can estimate, for any $\phi \in W^{1,\infty}(\Omega;\mathbb{R})$, that

$$\begin{aligned}|\nabla(f'(\mathbb{B}_\ell^\varepsilon)\phi)| &= |(\mathbb{B}_\ell^\varepsilon)^{-1}\nabla\mathbb{B}_\ell^\varepsilon(\mathbb{B}_\ell^\varepsilon)^{-1}\phi + (\mathbb{I}-(\mathbb{B}_\ell^\varepsilon)^{-1})\nabla\phi| \\ &\le C\omega^{-2}|\nabla\mathbb{B}_\ell^\varepsilon||\phi| + (1+C\omega^{-1})|\nabla\phi|.\end{aligned}$$

Using the second line of this estimate to show boundedness and the first line to identify the weak ε-limit using $(4.78)_2$ and (4.75), we eventually obtain

$$f'(\mathbb{B}_\ell^\varepsilon)\phi \rightharpoonup f'(\mathbb{B}_\ell)\phi \text{ weakly in } L^2(0,T;W^{1,2}(\Omega;\mathbb{R}^{d\times d}_{\mathrm{sym}})) \cap L^{q+2}(Q;\mathbb{R}^{d\times d}_{\mathrm{sym}}).$$

If we apply this with (4.75), we get, for all $\varphi \in W^{1,\infty}((0,T);\mathbb{R})$, $\varphi(T)=0$, that

$$\begin{aligned}&|\langle\partial_t\mathbb{B}_\ell^\varepsilon, f'(\mathbb{B}_\ell^\varepsilon)\phi\varphi\rangle - \langle\partial_t\mathbb{B}_\ell, f'(\mathbb{B}_\ell)\phi\varphi\rangle| \\ &\quad\le |\langle\partial_t\mathbb{B}_\ell^\varepsilon - \partial_t\mathbb{B}_\ell, f'(\mathbb{B}_\ell^\varepsilon)\phi\varphi\rangle| + |\langle\partial_t\mathbb{B}_\ell\varphi, f'(\mathbb{B}_\ell^\varepsilon)\phi - f'(\mathbb{B}_\ell)\phi\rangle| \to 0\end{aligned}$$

as $\varepsilon\to 0+$. This validates the limit on the left-hand side of (4.76), and thus

$$\langle\partial_t\mathbb{B}_\ell, f'(\mathbb{B}_\ell)\phi\varphi\rangle = -(f(\mathbb{B}_0^\omega),\phi\varphi(0)) - (f(\mathbb{B}_\ell),\phi\partial_t\varphi)_Q \tag{4.81}$$

for all $\varphi\in W^{1,\infty}(\Omega;\mathbb{R})$, $\varphi(T)=0$, and every $\phi\in W^{1,\infty}(\Omega;\mathbb{R})$. Therefore, after application of (4.74) and (4.81), entropy equation (4.73) becomes

$$\begin{aligned}&-(\eta_\ell,\phi\partial_t\varphi)_Q - (\eta_0^\omega,\phi)\varphi(0) - (\boldsymbol{v}_\ell\eta_\ell,\nabla\phi\varphi)_Q \\ &\quad+\big((\kappa(\theta_\ell)+\omega|\nabla\theta_\ell|^r)\nabla\ln\theta_\ell - \lambda(\theta_\ell)\nabla f(\mathbb{B}_\ell),\nabla\phi\varphi\big)_Q = (\xi_\ell,\phi\varphi)_Q\end{aligned} \tag{4.82}$$

for all $\varphi\in W^{1,\infty}(0,T;\mathbb{R})$, $\varphi(T)=0$, and $\phi\in W^{1,\infty}(\Omega;\mathbb{R})$, where

$$\eta_0^\omega \coloneqq c_v\ln\theta_0^\omega - f(\mathbb{B}_0^\omega).$$

Moreover, since $\ln\theta_\ell\in\mathcal{C}([0,T];L^2(\Omega;\mathbb{R}))$ and (4.80) hold, we easily read

$$\eta_\ell\in\mathcal{C}([0,T];L^2(\Omega;\mathbb{R})),\quad \eta_\ell(0)=\eta_0^\omega. \tag{4.83}$$

Total energy equality. The integrated version of the total energy equality is important in the derivation of the a priori estimates below. We multiply the i-th equation in (4.39) by $(\boldsymbol{v}_\ell,\boldsymbol{w}_i)$, sum up the result over $i=1,\dots,\ell$ and then we add (4.59) with $\tau=1$. This way, after several cancellations using also $(1.14)_1$, we obtain

$$\frac{\mathrm{d}}{\mathrm{d}t}\int_\Omega E_\ell + \alpha\int_{\partial\Omega}|\boldsymbol{v}_\ell|^2 = (\boldsymbol{g},\boldsymbol{v}_\ell)\quad\text{a.e. in } (0,T), \tag{4.84}$$

where $E_\ell\coloneqq\frac12|\boldsymbol{v}_\ell|^2 + c_v\theta_\ell$.

5. Existence of a weak solution: limits $\omega\to 0$, $\ell\to\infty$

This is the most essential part of the paper. Here, we first rigorously derive the estimates independent of ω and ℓ and then let $\omega\to 0+$ and $\ell\to\infty$ (in fact, we take these two limits simultaneously by setting $\omega=\frac{1}{\ell}$). Due to the linearity of the leading differential operators, the limit passage is then relatively straightforward. On the other hand, to obtain the attainment of the initial condition in the strong topology, we need to develop a new technique based on the combination of the entropy inequality and the global energy inequality.

Estimates independent of ℓ, ω based on global energy and entropy. Let us first show that the total energy of the fluid remains bounded. In (4.84), we apply Young's inequality, (3.22) and $\theta_\ell > 0$, to estimate

$$\frac{\mathrm{d}}{\mathrm{d}t}\int_\Omega E_\ell \le \frac{1}{2}\int_\Omega |\boldsymbol{v}_\ell|^2 + \frac{1}{2}\int_\Omega |\boldsymbol{g}|^2 \le \int_\Omega E_\ell + \frac{1}{2}\int_\Omega |\boldsymbol{g}|^2$$

a.e. in $(0,T)$. Hence, by the Gronwall inequality, we get

$$\int_\Omega E_\ell(t) \le e^t\left(\int_\Omega E_\ell(0) + \frac{1}{2}\int_0^t \|\boldsymbol{g}\|_2^2\right) \quad \text{for all } t \in [0,T].$$

Then, we apply (4.45), (4.56) to identify that

$$E_\ell(0) = \frac{1}{2}|P_\ell \boldsymbol{v}_0|^2 + c_v\theta_0^\omega$$

and if we use properties of P_ℓ, (4.6) and (3.22), we arrive at

$$\|\theta_\ell\|_{L^\infty L^1} + \|\boldsymbol{v}_\ell\|_{L^\infty L^2} \le C\|E_\ell\|_{L^\infty L^1} \le C. \tag{5.1}$$

Now we turn our attention to (4.82), which we localize in time by choosing[2] $\varphi = \chi_{(0,t)}$, leading to

$$\int_\Omega \eta_\ell(t)\phi + \int_0^t\int_\Omega \boldsymbol{j}_\ell \cdot \nabla\phi = \int_\Omega \eta_0^\omega\phi + \int_0^t\int_\Omega \xi_\ell\phi \quad \text{for all } \phi \in W^{1,\infty}(\Omega;\mathbb{R}) \tag{5.2}$$

and all $t \in (0,T)$ (in fact, for all $t \in [0,T]$ due to continuity), where

$$\boldsymbol{j}_\ell \coloneqq -\boldsymbol{v}_\ell\eta_\ell + (\kappa(\theta_\ell) + \omega|\nabla\theta_\ell|^r)\nabla\ln\theta_\ell - \lambda(\theta_\ell)\nabla f(\mathbb{B}_\ell) \in L^1(Q;\mathbb{R}^d).$$

In particular, taking $\phi = 1$, we deduce, using $\xi_\ell \ge 0$, that the function $t \mapsto \int_\Omega \eta_\ell(t)$ is non-decreasing, and thus

$$\int_Q \xi_\ell = \max_{t\in[0,T]}\int_0^t\int_\Omega \xi_\ell = \max_{t\in[0,T]}\int_\Omega \eta_\ell(t) - \int_\Omega \eta_0^\omega = \int_\Omega \eta_\ell(T) - \int_\Omega \eta_0^\omega. \tag{5.3}$$

Then, using (4.71), the inequalities

$$\ln x \le x - 1 \quad \text{for all } x > 0 \quad \text{and} \quad f(\mathbb{B}_\ell) \ge 0, \tag{5.4}$$

assumption (3.23) and (5.1) (recall also (4.51)), we obtain

$$\int_Q \xi_\ell \le \int_\Omega (c_v\ln\theta_\ell(T) - f(\mathbb{B}_\ell(T))) + C \le C\int_\Omega (\theta_\ell(T) - 1) + C \le C, \tag{5.5}$$

hence

$$\|\xi_\ell\|_{L^1L^1} \le C. \tag{5.6}$$

Also, it is easy to see using (3.22), (3.23), (4.6), (4.7) and (5.3) that

$$\|\eta_\ell\|_{L^\infty L^1} \le C. \tag{5.7}$$

Estimate (5.6) implies, using (3.2) and (3.9), that

$$\begin{aligned}\|\theta_\ell^{-\frac{1}{2}}\mathbb{D}\boldsymbol{v}_\ell\|_{L^2L^2} + \|\sqrt{\kappa(\theta_\ell)}\nabla\ln\theta_\ell\|_{L^2L^2} + \sqrt{\omega}\||\nabla\theta_\ell|^{\frac{r}{2}}\nabla\ln\theta_\ell\|_{L^2L^2} \\ + \|\mathbb{B}_\ell^{-\frac{1}{2}}\nabla\mathbb{B}_\ell\mathbb{B}_\ell^{-\frac{1}{2}}\|_{L^2L^2} \le C.\end{aligned} \tag{5.8}$$

[2]Strictly speaking, as $\chi_{(0,t)}$ is not Lipschitz, we can not use it directly in (4.82). However, a standard argument using a piecewise linear approximation of $\chi_{(0,t)}$ with the Lebesgue differentiation theorem and absolute continuity of integral shows that $\chi_{(0,t)}$ is a reasonable test function.

Improved ℓ, ω estimates. In what follows, we improve the uniform estimate (5.8) considerably by choosing appropriate test functions in (4.40) and (4.59) and then using (3.11). In fact, we repeat the scheme of estimates presented in (3.16)–(3.21), but now, we prove it fully rigorously.

Our aim is to set $\mathbb{A} := \mathbb{B}_\ell^{q-1}$ in (4.40). To verify that this is a valid test function, we show first that $\mathbb{B}_\ell$ is actually essentially bounded. Indeed, setting first $\mathbb{A} = \chi_{(0,t)}\phi\mathbb{I}$, $t \in (0,T)$, $\phi \in L^{q+2}(0,T;L^{q+2}(\Omega;\mathbb{R})) \cap L^2(0,T;W^{1,2}(\Omega;\mathbb{R}))$, in (4.40) yields

$$\int_0^t \langle \partial_t \operatorname{tr}\mathbb{B}_\ell, \phi\rangle + \int_0^t (\boldsymbol{v}\cdot\nabla \operatorname{tr}\mathbb{B}_\ell, \phi) + \int_0^t (\mathbb{P}(\theta_\ell, \mathbb{B}_\ell)\cdot\mathbb{I}, \phi) + \int_0^t (\lambda(\theta_\ell)\nabla \operatorname{tr}\mathbb{B}_\ell, \nabla\phi) = \int_0^t (2ag_\omega(\mathbb{B}_\ell,\theta_\ell)\mathbb{B}_\ell\cdot\mathbb{D}\boldsymbol{v}_\ell, \phi).$$

Hence, recalling (3.8) to bound the third term on the left hand side and using (4.1) and (4.30) to estimate the right hand side, we see that there exists a constant $C(\ell,\omega) > 0$, such that

$$\int_0^t \langle \partial_t \operatorname{tr}\mathbb{B}_\ell, \phi\rangle + \int_0^t (\boldsymbol{v}\cdot\nabla \operatorname{tr}\mathbb{B}_\ell, \phi) + \int_0^t (\lambda(\theta_\ell)\nabla \operatorname{tr}\mathbb{B}_\ell, \nabla\phi) \le C(\ell,\omega)\int_0^t\int_\Omega |\phi|.$$

Substituting $u(\boldsymbol{x},t) \coloneqq \operatorname{tr}\mathbb{B}_\ell(\boldsymbol{x},t) - C(\ell,\omega)t$ leads to

$$\int_0^t \langle \partial_t u, \phi\rangle + \int_0^t (\boldsymbol{v}\cdot\nabla u, \phi) + \int_0^t (\lambda(\theta_\ell)\nabla u, \nabla\phi) \le C(\ell,\omega)\int_0^t\int_\Omega (|\phi| - \phi).$$

If we choose $\phi = (u-K)_+$ and use (1.8), $(1.14)_1$ to eliminate the convective term, we obtain

$$\frac{1}{2}\|(u(t)-K)_+\|_2^2 + \int_0^t \|\sqrt{\lambda(\theta_\ell)}\nabla(u-K)_+\|_2^2 \le \frac{1}{2}\|(u(0)-K)_+\|_2^2.$$

If we let $K \coloneqq \frac{d}{\omega}$, then (4.44) and (4.5) imply

$$0 \le (u(0)-K)_+ = (\operatorname{tr}\mathbb{B}_0^\omega - \tfrac{d}{\omega})_+ \le (\sqrt{d}|\mathbb{B}_0^\omega| - \tfrac{d}{\omega})_+ = 0 \quad \text{in } \Omega.$$

Thus, we get $\|(u(t) - \frac{d}{\omega})_+\|_2^2 = 0$, hence

$$|\mathbb{B}_\ell| \le \operatorname{tr}\mathbb{B}_\ell \le \tfrac{d}{\omega} + C_0 t \le \tfrac{d}{\omega} + C_0 T$$

and we see that indeed

$$\mathbb{B}_\ell \in L^\infty(0,T;L^\infty(\Omega;\mathbb{R}^{d\times d}_{>0})) \cap L^2(0,T;W^{1,2}(\Omega;\mathbb{R}^{d\times d}_{>0})). \tag{5.9}$$

Due to the fact that $\mathbb{B}_\ell$ is strictly positive definite, we can use the above property to show that the same holds also for $\mathbb{B}_\ell^{q-1}$, which is essential for showing that $\mathbb{A} := \mathbb{B}_\ell^{q-1}$ can be used in (4.40) as a test function. Indeed, the boundedness of $\mathbb{B}_\ell^{q-1}$ is a direct consequence of (5.9) and the spectral decomposition. To show that gradient of $\mathbb{B}_\ell^{q-1}$ is square integrable, we recall the identity

$$\frac{\nabla\mathbb{B}_\ell^{q-1}}{q-1} = \int_0^1\int_0^1 \mathbb{B}_\ell^{(1-s)(q-1)}((1-t)\mathbb{I} + t\mathbb{B}_\ell)^{-1}\nabla\mathbb{B}_\ell((1-t)\mathbb{I} + t\mathbb{B}_\ell)^{-1}\mathbb{B}_\ell^{s(q-1)}\,\mathrm{d}s\,\mathrm{d}t,$$

which is a consequence of the well known identities for $\nabla\exp\mathbb{A}$ and $\nabla\log\mathbb{A}$, see e.g. [57, 58, 3] and references therein for details. Then using also (4.62) to estimate

$$|((1-t)\mathbb{I} + t\mathbb{B}_\ell)^{-1}| \le \frac{\sqrt{d}}{\Lambda((1-t)\mathbb{I} + t\mathbb{B}_\ell)} \le \frac{\sqrt{d}}{\omega}$$

and also (5.9), we see that $\nabla\mathbb{B}_\ell^{q-1} \in L^2(0,T;L^2(\mathbb{R}\times\mathbb{R}^{d\times d}_{\mathrm{sym}}))$ and consequently

$$\mathbb{B}_\ell^{q-1} \in L^\infty(0,T;L^\infty(\Omega;\mathbb{R}^{d\times d}_{>0})) \cap L^2(0,T;W^{1,2}(\Omega;\mathbb{R}^{d\times d}_{>0})).$$

Hence, setting $\mathbb{A} := \chi_{[0,t]}\mathbb{B}_\ell^{q-1}$ in (4.40), using (3.7) and the identities[3]

$$\langle \partial_t \mathbb{B}_\ell, \mathbb{B}_\ell^{q-1}\rangle = \frac{1}{q}\int_\Omega \partial_t \operatorname{tr}\mathbb{B}_\ell^q,$$
$$(\boldsymbol{v}\cdot\nabla\mathbb{B}_\ell, \mathbb{B}_\ell^{q-1}) = \frac{1}{q}\int_\Omega \boldsymbol{v}\cdot\nabla\operatorname{tr}\mathbb{B}_\ell^q = 0$$

and the estimate (see (iv) and (v) in Lemma A.3 below)

$$\nabla\mathbb{B}_\ell\cdot\nabla\mathbb{B}_\ell^{q-1} \geq \frac{4(q-1)}{q^2}|\nabla\mathbb{B}_\ell^{\frac{q}{2}}|^2,$$

we get

$$\begin{aligned}&\frac{1}{q}\int_\Omega(\operatorname{tr}\mathbb{B}_\ell^q(t)-\operatorname{tr}\mathbb{B}_\ell^q(0)) + C_{q-1}\int_0^t\int_\Omega|\mathbb{B}|^{2q} + \frac{4(q-1)}{q^2}\int_0^t\int_\Omega\lambda(\theta_\ell)|\nabla\mathbb{B}_\ell^{\frac{q}{2}}|^2\\ &\quad\leq 2a\int_0^t\int_\Omega g(\mathbb{B}_\ell,\theta_\ell)\mathbb{D}\boldsymbol{v}_\ell\cdot\mathbb{B}_\ell^q + C.\end{aligned}$$

If we apply (4.44), (3.4), $g_\omega\leq 1$ and $|\mathbb{B}_\ell^q|\leq\max\{1,d^{\frac{1-q}{2}}\}|\mathbb{B}_\ell|^q$ (see [3]), we deduce

$$\begin{aligned}&\int_\Omega\operatorname{tr}\mathbb{B}_\ell^q(t) + \int_0^t\int_\Omega|\mathbb{B}_\ell|^{2q} + \int_0^t\int_\Omega|\nabla\mathbb{B}_\ell^{\frac{q}{2}}|^2\\ &\qquad\leq\int_\Omega(\operatorname{tr}\mathbb{B}_0^\omega)^q + C\left(\int_0^t\int_\Omega|\mathbb{D}\boldsymbol{v}_\ell||\mathbb{B}_\ell|^q+1\right).\end{aligned}\tag{5.10}$$

Then, to estimate the term with $\operatorname{tr}\mathbb{B}_0^\omega$, we use (4.6) and (3.22). On the last term on the right hand side, we apply Young's inequality, leading to

$$\|\mathbb{B}_\ell\|^q_{L^\infty L^q} + \|\mathbb{B}_\ell\|^{2q}_{L^{2q}L^{2q}} + \|\nabla\mathbb{B}_\ell^{\frac{q}{2}}\|^2_{L^2L^2} \leq C\left(1+\|\mathbb{D}\boldsymbol{v}_\ell\|^2_{L^2L^2}\right),\tag{5.11}$$

where the right hand side is finite due to (4.30), but we do not have a uniform bound yet. To obtain it, we combine the estimate (5.11) with the temperature equation (4.59) and improve the information about θ_ℓ and $\mathbb{D}\boldsymbol{v}_\ell$.

Let $\beta\in[0,\frac{1}{2}]$ be arbitrary. We define

$$\tau_\beta := -\theta_\ell^{-\beta}.$$

Using Lemma A.2 with $\psi(s) = -\max(s,\omega)^{-\beta}$ to rewrite the time derivative, the a priori bound (5.1) with Young's inequality, $(1.14)_1$ and (3.3), we obtain the estimate

$$\begin{aligned}&\langle c_v\partial_t\theta_\ell,\tau_\beta\rangle + (\kappa(\theta_\ell)\nabla\theta_\ell,\nabla\tau_\beta)_Q + \omega(|\nabla\theta_\ell|^r\nabla\theta_\ell,\nabla\tau_\beta)_Q\\ &\quad\geq\frac{-c_v}{1-\beta}\int_\Omega\theta_\ell^{1-\beta}(T) + \beta\int_Q\theta_\ell^{-1-\beta}\kappa(\theta_\ell)|\nabla\theta_\ell|^2 + \omega\beta\int_Q\theta_\ell^{-1-\beta}|\nabla\theta_\ell|^{r+2}\\ &\quad\geq C\beta\int_Q\left|\nabla\theta_\ell^{\frac{r+1-\beta}{2}}\right|^2 + \omega\beta\int_Q\theta_\ell^{-1-\beta}|\nabla\theta_\ell|^{r+2} - C.\end{aligned}\tag{5.12}$$

The function τ_β evidently satisfies $\tau_\beta\in L^{r+2}(0,T;W^{1,r+2}(\Omega))\cap L^\infty(0,T;L^\infty(\Omega))$ (cf. (4.67)), and is thus an admissible test function in (4.59). This way, noting

[3]To interpret the duality pairing in the first identity, one has to approximate $\mathbb{B}_\ell$ similarly as before when dealing with $\langle\partial_t\mathbb{B}_\ell,(\mathbb{I}-\mathbb{B}_\ell^{-1})\phi\rangle$.

that the convective term $(c_v \boldsymbol{v}_\ell \cdot \nabla \theta_\ell, \tau_\beta)_Q$ disappears since $\operatorname{div} \boldsymbol{v}_\ell = 0$, and having the estimate (5.12) and using also $g_\omega \leq 1$, we deduce

$$\beta \int_Q \Big(\Big| \nabla \theta_\ell^{\frac{r+1-\beta}{2}} \Big|^2 + \omega \theta_\ell^{-1-\beta} |\nabla \theta_\ell|^{r+2} \Big) + \int_Q \theta_\ell^{-\beta} |\mathbb{D} \boldsymbol{v}_\ell|^2 \leq C \Big(\int_Q \theta_\ell^{1-\beta} |\mathbb{B}_\ell| |\mathbb{D} \boldsymbol{v}_\ell| + 1 \Big).$$

Note that since $\beta \in [0, \frac{1}{2}]$ is arbitrary, we can reduce the above inequality to

$$\beta \int_Q \Big(\Big| \nabla \theta_\ell^{\frac{r+1-\beta}{2}} \Big|^2 + \omega \theta_\ell^{-1-\beta} |\nabla \theta_\ell|^{r+2} \Big) + \int_Q |\mathbb{D} \boldsymbol{v}_\ell|^2 \leq C \Big(\int_Q (\theta_\ell + 1) |\mathbb{B}_\ell| |\mathbb{D} \boldsymbol{v}_\ell| + 1 \Big),$$

which is very much similar to (3.17), while the estimate (5.11) mimics (3.16). Hence, applying the Young and the Hölder inequality, and using (5.11), we deduce similarly as in (3.18) that

$$\beta \int_Q \Big(\Big| \nabla \theta_\ell^{\frac{r+1-\beta}{2}} \Big|^2 + \omega \theta_\ell^{-1-\beta} |\nabla \theta_\ell|^{r+2} \Big) + \int_Q |\mathbb{D} \boldsymbol{v}_\ell|^2 \leq C \Big(1 + \int_Q |\theta_\ell|^{2q'} \Big). \tag{5.13}$$

Next, we continue as after (3.18). We recall the interpolation inequality

$$\|\theta_\ell\|_{2q'}^{2q'} \leq C \|\theta_\ell\|_1^{2q' - \frac{d(r-\beta+1)(2q'-1)}{d(r-\beta)+2}} \|\theta_\ell^{\frac{r+1-\beta}{2}}\|_{1,2}^{\frac{2d(2q'-1)}{d(r-\beta)+2}}. \tag{5.14}$$

Thus, using the uniform bound (5.1), the estimate (5.13) and the interpolation inequality (5.14), we deduce

$$\begin{aligned} \beta \int_0^T \|\theta_\ell^{\frac{r+1-\beta}{2}}\|_{1,2}^2 &\leq C\beta \int_0^T \big(\|\nabla \theta_\ell^{\frac{r+1-\beta}{2}}\|_2^2 + \|\theta_\ell\|_1^{r+1-\beta} \big) \\ &\leq C \Big(1 + \int_Q |\theta_\ell|^{2q'} \Big) \leq C + C \int_0^T \|\theta_\ell^{\frac{r+1-\beta}{2}}\|_{1,2}^{\frac{2d(2q'-1)}{d(r-\beta)+2}}. \end{aligned} \tag{5.15}$$

Finally, thanks to (3.11), we can find $\beta_0 > 0$ such that for all $\beta \in (0, \beta_0)$ we have

$$\frac{2d(2q'-1)}{d(r-\beta)+2} < 2.$$

Consequently, we can use the Young inequality in (5.15) and conclude that

$$\|\theta_\ell^{\frac{r+1-\beta}{2}}\|_{L^2 W^{1,2}} \leq C(\beta) \tag{5.16}$$

for all $\beta \in (0, \beta_0)$ (which can be however easily extended via (5.8) to the validity for all $\beta \in (0,1)$). Furthermore, from the interpolation inequality

$$\|\theta_\ell\|_{r+1+\frac{2}{d}-\beta}^{r+1+\frac{2}{d}-\beta} \leq C \|\theta_\ell\|_1^{\frac{2}{d}} \|\theta_\ell^{\frac{r+1-\beta}{2}}\|_{1,2}^2, \tag{5.17}$$

(5.1) and (5.16), we conclude

$$\|\theta_\ell\|_{L^{r+1+\frac{2}{d}-\beta} L^{r+1+\frac{2}{d}-\beta}} \leq C(\beta). \tag{5.18}$$

Summary of all uniform estimates. To summarize the estimates proved up to this point, we recall (5.1), (5.6), (5.7) and (5.8) based on the use of total energy and entropy estimates. Next, having (5.16), we can choose $\beta := \frac{\beta_0}{2}$ and go backward in the computation in the previous part and obtain further a priori estimates. Namely, using (5.16) and (5.14), we see that the right hand side of (5.13) is uniformly

bounded. Then, using (5.13) in (5.11) we deduce also the a priori bound for $\mathbb{B}_\ell$. Thus, we can conclude with the following set of estimates

$$\|\boldsymbol{v}_\ell\|_{L^\infty L^2} + \|\mathbb{D}\boldsymbol{v}_\ell\|_{L^2L^2} \leq C, \tag{5.19}$$

$$\|\mathbb{B}_\ell\|_{L^\infty L^q} + \|\mathbb{B}_\ell\|_{L^{2q}L^{2q}} + \|\nabla \mathbb{B}_\ell^{\frac{q}{2}}\|_{L^2L^2} \leq C, \tag{5.20}$$

$$\|\theta_\ell\|_{L^\infty L^1} + \|\nabla \theta_\ell^{\frac{r+1-\varepsilon}{2}}\|_{L^2L^2} + \|\theta_\ell\|_{L^{r+1+\frac{2}{d}-\varepsilon}L^{r+1+\frac{2}{d}-\varepsilon}} \leq C(\varepsilon) \tag{5.21}$$

for all $\varepsilon \in (0,1)$. Next, in order to obtain estimates on $\nabla\mathbb{B}_\ell$, we separate two cases. If $1 < q < 2$, we use (A.23) and Hölder's inequality, (A.20) and (5.20) to estimate

$$\|\nabla \mathbb{B}_\ell\|_{\frac{4q}{q+2}} \leq 2\|\mathbb{B}_\ell^{1-\frac{q}{2}}\|_{\frac{4q}{2-q}} \|\nabla \mathbb{B}_\ell^{\frac{q}{2}}\|_2 \leq C.$$

On the other hand, if $q \geq 2$, the optimal estimate on $\nabla\mathbb{B}_\ell$ is obtained simply by testing (4.40) with $\mathbb{B}_\ell$ (instead of $\mathbb{B}_\ell^{q-1}$). Indeed, using (5.9), we eventually obtain (5.11), but with $q = 2$. Combination of these two cases leads to

$$\|\nabla \mathbb{B}_\ell\|_{L^m L^m} \leq C. \tag{5.22}$$

Uniform time derivatives estimates. We end this part by derivation of the uniform estimates for the time derivatives. To this end, we need to determine integrability of the non-linear terms in (4.39), (4.40) and (4.82). It follows from an interpolation inequality, Korn's inequality, (5.1) and (5.19) that

$$\|\boldsymbol{v}_\ell\|_{L^2\frac{d+2}{d}L^2\frac{d+2}{d}} \leq C\|\boldsymbol{v}_\ell\|_{L^\infty L^2}^{\frac{2}{d+2}} \|\mathbb{D}\boldsymbol{v}_\ell\|_{L^2L^2}^{\frac{d}{d+2}} \leq C. \tag{5.23}$$

Furthermore, the Hölder inequality, (5.21) and (5.11) yield

$$\|\theta_\ell \mathbb{B}_\ell\|_{L^2L^2} \leq C, \tag{5.24}$$

Hence, as $d \geq 2$, we read from (4.39) that

$$\|\partial_t \boldsymbol{v}_\ell\|_{L^{\frac{d+2}{d}}W^{-1,\frac{d+2}{d}}_{\boldsymbol{n},\mathrm{div}}} \leq C. \tag{5.25}$$

Next, we focus on the non-linear terms in (4.40). Using Hölder's inequality and (5.11), (5.23), we observe that

$$\|\mathbb{B}_\ell \otimes \boldsymbol{v}_\ell\|_{L^{s_1}L^{s_1}} \leq C, \tag{5.26}$$

with

$$s_1 := \Big(\frac{1}{2q} + \frac{d}{2(d+2)}\Big)^{-1} > \Big(\frac{1}{2q} + \frac{1}{2}\Big)^{-1} = \frac{2q}{q+1}. \tag{5.27}$$

Moreover, making use of (5.20) and (3.6), we obtain

$$\|\mathbb{P}(\theta_\ell, \mathbb{B}_\ell)\|_{L^{\frac{2q}{q+1}}L^{\frac{2q}{q+1}}} \leq C. \tag{5.28}$$

Furthermore, using (5.20), (5.19) and Hölder's inequality, we also get

$$\|(a\mathbb{D}\boldsymbol{v}_\ell + \mathbb{W}\boldsymbol{v}_\ell)\mathbb{B}_\ell\|_{L^{\frac{2q}{q+1}}L^{\frac{2q}{q+1}}} \leq C. \tag{5.29}$$

Thus, we read from (4.40) using (5.22) (where note that $m > \frac{2q}{q+1}$), (5.26), (5.27) and (5.28), (5.29) that

$$\|\partial_t \mathbb{B}_\ell\|_{L^{\frac{2q}{q+1}}W^{-1,\frac{2q}{q+1}}} \leq C. \tag{5.30}$$

Next, we examine the non-linearities related to (4.82). Since ξ_ℓ is controlled by (5.6), the problematic terms could be only on the left hand side. To get an appropriate uniform control over the convective term, we estimate

$$\eta_\ell \leq \eta_\ell + f(\mathbb{B}_\ell) = c_v \ln \theta_\ell \leq c_v(\theta_\ell - 1).$$

This, together with (5.1) and (5.7), yields

$$\|\ln \theta_\ell\|_{L^\infty L^1} \leq C. \tag{5.31}$$

Then, since (5.8) and (3.3) give

$$\|\nabla \ln \theta_\ell\|_{L^2 L^2} \leq C, \tag{5.32}$$

we can use Sobolev's inequality, Poincaré's inequality and an interpolation to obtain

$$\|\ln \theta_\ell\|_{L^{2+\frac{2}{d}} L^{2+\frac{2}{d}}} \leq C \|\ln \theta_\ell\|^{\frac{1}{d+1}}_{L^\infty L^1} \|\ln \theta_\ell\|^{\frac{d}{d+1}}_{L^2 W^{1,2}} \leq C. \tag{5.33}$$

Now we observe that a similar reasoning applies also for the quantity $\ln \det \mathbb{B}_\ell$. Indeed, using (5.7), (5.31), (5.11) and (4.71) in the form

$$\ln \det \mathbb{B}_\ell = \frac{1}{\mu}(\eta_\ell - c_v \ln \theta_\ell) + \operatorname{tr} \mathbb{B}_\ell - d,$$

it is clear that

$$\|\ln \det \mathbb{B}_\ell\|_{L^\infty L^1} \leq C. \tag{5.34}$$

Further, the estimate of its derivative follows from a version of Jacobi's formula (see Lemma A.3 below) and (5.8) as

$$\|\nabla \ln \det \mathbb{B}_\ell\|_{L^2 L^2} = \|\operatorname{tr}(\mathbb{B}_\ell^{-\frac{1}{2}} \nabla \mathbb{B}_\ell \mathbb{B}_\ell^{-\frac{1}{2}})\|_{L^2 L^2} \leq C. \tag{5.35}$$

Hence, using again the Sobolev, the Poincaré and interpolation inequalities, we get

$$\|\ln \det \mathbb{B}_\ell\|_{L^{2+\frac{2}{d}} L^{2+\frac{2}{d}}} \leq C. \tag{5.36}$$

From (5.33), (5.36), (5.11) and (4.71), we deduce

$$\|\eta_\ell\|_{L^{s_2} L^{s_2}} \leq C, \quad \text{where} \quad s_2 := \min\{2 + \tfrac{2}{d}, 2q\} > 2, \tag{5.37}$$

and thus

$$\|\boldsymbol{v}_\ell \eta_\ell\|_{L^{s_3} L^{s_3}} \leq C, \quad \text{where} \quad s_3 := \Big(\frac{d}{2(d+2)} + \frac{1}{s_2}\Big)^{-1} > 1. \tag{5.38}$$

We remark that, since

$$\nabla \eta_\ell = c_v \nabla \ln \theta_\ell - \mu(\operatorname{tr} \nabla \mathbb{B}_\ell - \operatorname{tr}(\mathbb{B}_\ell^{-\frac{1}{2}} \nabla \mathbb{B}_\ell \mathbb{B}_\ell^{-\frac{1}{2}})),$$

we also have, using (5.35), (5.32), (5.22), (5.8) and Poincaré's inequality that

$$\|\eta_\ell\|_{L^m W^{1,m}} \leq C. \tag{5.39}$$

Looking at (4.82), we still need to verify that the flux terms are controlled. For the term $\kappa(\theta_\ell)\nabla \ln \theta_\ell$, we first use (3.3) and (5.21) to estimate

$$\|\sqrt{\kappa(\theta_\ell)}\|_{L^{\frac{2d(r+1)+4}{dr}}(Q;\mathbb{R})} \leq C \|1+\theta\|^{\frac{r}{2}}_{L^{r+1+\frac{2}{d}}(Q;\mathbb{R})} \leq C$$

and then, by Hölder's inequality and (5.8), we get

$$\begin{aligned} &\|\kappa(\theta_\ell)\nabla \ln \theta_\ell\|_{L^{\frac{2d(r+1)+4}{2d(r+1)+2-d}}(Q;\mathbb{R}^d)} \\ &\quad \leq \|\sqrt{\kappa(\theta_\ell)}\|_{L^{\frac{2d(r+1)+4}{dr}}(Q;\mathbb{R})} \|\sqrt{\kappa(\theta_\ell)}\nabla \ln \theta_\ell\|_{L^2(Q;\mathbb{R}^d)} \leq C. \end{aligned} \tag{5.40}$$

Further, let us derive an estimate on $\omega|\nabla\theta_\ell|^r\nabla\ln\theta_\ell$, from which it follows that this term vanishes as $\omega\to 0+$. The Hölder inequality, (5.8) and (5.21) yield

$$\begin{aligned}\omega\int_Q|\nabla\theta_\ell|^r|\nabla\ln\theta_\ell| &= \omega^{\frac{1}{r+2}}\int_Q\left(\omega\theta_\ell^{-2}|\nabla\theta_\ell|^{r+2}\right)^{\frac{r+1}{r+2}}\theta_\ell^{2\frac{r+1}{r+2}-1}\\ &\le C\omega^{\frac{1}{r+2}}\Big(\int_Q\theta_\ell^r\Big)^{\frac{1}{r+2}}\le C\omega^{\frac{1}{r+2}}.\end{aligned}\tag{5.41}$$

From this and from (5.40), (5.32), (5.35), (5.38), (5.8), (4.82), we see, using the definition of a weak time derivative, that

$$\|\partial_t\eta_\ell\|_{L^1W^{-M,2}}\le C,\tag{5.42}$$

where M is so large that $W^{M,2}(\Omega;\mathbb{R})\hookrightarrow W^{1,\infty}(\Omega;\mathbb{R})$.

Finally, we focus on terms appearing in the temperature equation (4.68). First, we note that it is a consequence of assumption (3.11), a priori estimates (5.19)–(5.21) and the Hölder inequality, that

$$\int_Q|\theta_\ell\boldsymbol{v}_\ell|+\int_Q|\mathbb{S}^\omega_\ell\cdot\mathbb{D}\boldsymbol{v}_\ell|\le C.\tag{5.43}$$

In the terms involving temperature gradient, we use (3.3), (5.8), (5.21) and the inequality $\max\{2,r+1+\varepsilon\}<r+1+\frac{2}{d}-\varepsilon$ for ε small (recall (3.11)) to estimate

$$\begin{aligned}\int_Q|\kappa(\theta_\ell)\nabla\theta_\ell| &\le C\int_Q(\theta_\ell|\nabla\ln\theta_\ell|+\theta_\ell^{\frac{r+1+\varepsilon}{2}}|\nabla\theta^{\frac{r+1-\varepsilon}{2}}|)\\ &\le C\int_Q(\theta_\ell^2+\theta_\ell^{r+1+\varepsilon})\le C.\end{aligned}\tag{5.44}$$

Proceeding similarly as in (5.41), but using now (5.13) instead of (5.8), we also find

$$\begin{aligned}\int_Q\omega|\nabla\theta_\ell|^{r+1} &= \omega^{\frac{1}{r+2}}\int_Q\left(\omega\theta_\ell^{-1-\beta}|\nabla\theta_\ell|^{r+2}\right)^{\frac{r+1}{r+2}}\theta_\ell^{\frac{(\beta+1)(r+1)}{r+2}}\\ &\le C\omega^{\frac{1}{r+2}}\Big(\int_Q\theta_\ell^{(\beta+1)(r+1)}\Big)^{\frac{1}{r+2}}\le C\omega^{\frac{1}{r+2}},\end{aligned}\tag{5.45}$$

where $\beta>0$ is chosen so small that $(\beta+1)(r+1)<r+1+\frac{2}{d}$. Using the above estimates in (4.59), we deduce

$$\|\partial_t\theta_\ell\|_{L^1W^{-M,2}}\le C,\tag{5.46}$$

for sufficiently large M. Very similarly, choosing $\theta_\ell^{-\frac{1}{2}}\phi$ in (4.59) and repeating the method for estimating $\partial_t\eta_\ell$, we can find that

$$\|\partial_t\theta_\ell^{\frac{1}{2}}\|_{L^1W^{-M,2}}\le C.\tag{5.47}$$

Finally, returning to (4.70) with (4.81) and using the uniform estimates proved so far, it is easy to see that also

$$\|\partial_t f(\mathbb{B}_\ell)\|_{L^1W^{-M,2}}\le C.\tag{5.48}$$

The last two properties will be useful in the initial condition identification.

Limits $\omega \to 0$, $\ell \to \infty$. Let us note that the estimates above are independent not only of ℓ, but also of ω. Hence, we can set $\omega := \ell^{-1}$ and hereby, it remains to take the limit $\ell \to \infty$ only.

By collecting the estimates (5.1), (5.21)–(5.20), (5.22), (5.30), (5.25), (5.37), (5.39), (5.42), (5.44), (5.46), (5.47) and using the Aubin–Lions lemma and Vitali's convergence theorem, we get the following results:

$$\boldsymbol{v}_\ell \rightharpoonup \boldsymbol{v} \qquad \text{weakly in } L^2(0,T;W^{1,2}_{\boldsymbol{n},\mathrm{div}}), \tag{5.49}$$

$$\boldsymbol{v}_\ell \to \boldsymbol{v} \qquad \text{strongly in } L^{2\frac{d+2}{d}-\varepsilon}(Q;\mathbb{R}^d) \text{ and a.e. in } Q, \tag{5.50}$$

$$\partial_t \boldsymbol{v}_\ell \rightharpoonup \partial_t \boldsymbol{v} \qquad \text{weakly in } L^{\frac{d+2}{d}}(0,T;W^{-1,\frac{d+2}{d}}_{\boldsymbol{n},\mathrm{div}}), \tag{5.51}$$

$$\mathbb{B}_\ell \rightharpoonup \mathbb{B} \qquad \text{weakly in } L^m(0,T;W^{1,m}(\Omega;\mathbb{R}^{d\times d}_{\mathrm{sym}})), \tag{5.52}$$

$$\mathbb{B}_\ell \to \mathbb{B} \qquad \text{strongly in } L^{2q-\varepsilon}(Q;\mathbb{R}^{d\times d}_{\mathrm{sym}}) \text{ and a.e. in } Q, \tag{5.53}$$

$$\partial_t \mathbb{B}_\ell \rightharpoonup \partial_t \mathbb{B} \qquad \text{weakly in } L^{\frac{2q}{q+1}}(0,T;W^{-1,\frac{2q}{q+1}}(\Omega;\mathbb{R}^{d\times d}_{\mathrm{sym}})), \tag{5.54}$$

$$\eta_\ell \rightharpoonup \eta \qquad \text{weakly in } L^m(0,T;W^{1,m}(\Omega;\mathbb{R})), \tag{5.55}$$

$$\eta_\ell \to \eta \qquad \text{strongly in } L^{s_2-\varepsilon}(Q;\mathbb{R}) \text{ and a.e. in } Q, \tag{5.56}$$

$$\eta_\ell \overset{*}{\rightharpoonup} \eta \qquad \text{weakly}^* \text{ in } BV(0,T;W^{-M,2}(\Omega;\mathbb{R})), \tag{5.57}$$

$$\theta_\ell^{\frac{r+1-\varepsilon}{2}} \rightharpoonup \theta^{\frac{r+1-\varepsilon}{2}} \qquad \text{weakly in } L^2(0,T;W^{1,2}(\Omega;\mathbb{R})), \tag{5.58}$$

$$\theta_\ell \to \theta \qquad \text{strongly in } L^{r+1+\frac{2}{d}-\varepsilon}(Q;\mathbb{R}), \tag{5.59}$$

$$\theta_\ell^{\frac{1}{2}} \overset{*}{\rightharpoonup} \theta^{\frac{1}{2}} \qquad \text{weakly}^* \text{ in } BV(0,T;W^{-M,2}(\Omega;\mathbb{R})), \tag{5.60}$$

$$\theta_\ell \overset{*}{\rightharpoonup} \theta \qquad \text{weakly}^* \text{ in } BV(0,T;W^{-M,2}(\Omega;\mathbb{R})) \tag{5.61}$$

for any $\varepsilon \in (0,1)$. Using these properties, we shall now explain how to take the limit in equations (4.39), (4.40), (4.82), (4.84) and (4.59).

First, we focus on taking the limit in the function $g_{\frac{1}{\ell}}$. From (4.62), (4.66) and (5.49), (5.52) (or (5.50), (5.53)), we obtain

$$\mathbb{B}\boldsymbol{x}\cdot\boldsymbol{x} \geq 0 \quad \text{for all } \boldsymbol{x}\in\mathbb{R}^d \quad \text{and} \quad \theta \geq 0 \quad \text{a.e. in } Q, \tag{5.62}$$

however, we need these properties with strict inequalities. To this end, we use the Fatou lemma, (5.53) and (5.34) to get

$$\int_\Omega |\ln\det\mathbb{B}| \leq \liminf_{\ell\to\infty}\int_\Omega |\ln\det\mathbb{B}_\ell| \leq C \quad \text{a.e. in } (0,T).$$

Thus, by taking the essential supremum over $(0,T)$, we obtain

$$\|\ln\det\mathbb{B}\|_{L^\infty L^1} < \infty, \tag{5.63}$$

which, together with (5.62) implies

$$\mathbb{B}\boldsymbol{x}\cdot\boldsymbol{x} > 0 \quad \text{for all } \boldsymbol{x}\in\mathbb{R}^d \quad \text{a.e. in } Q. \tag{5.64}$$

An analogous argument, using now (5.59) and (5.31), shows that

$$\theta > 0 \quad \text{a.e. in } Q. \tag{5.65}$$

With this in hand, note that the property (1.6) follows from (4.71) and the pointwise a.e. convergence of η_ℓ, θ_ℓ and $\mathbb{B}_\ell$. Also, from (5.64), (5.65) and the pointwise

convergence we deduce that, at almost every point $(t,x)\in Q$, we can find $M_{t,x}\in\mathbb{N}$ such that for all $\ell > M_{t,x}$ we have

$$\Lambda(\mathbb{B}_\ell(t,x)) > \frac{1}{2}\Lambda(\mathbb{B}(t,x)) > \frac{1}{\ell} \quad \text{and} \quad \theta_\ell(t,x) > \frac{1}{2}\theta(t,x) > \frac{1}{\ell}.$$

Then, looking at the definition of g_λ, we see that at almost every point $(t,x)\in Q$ and for $\ell > M_{t,x}$, the positive parts $\max\{0,\cdot\}$ can be removed and thus, it is clear that $g_{\frac{1}{\ell}}(\mathbb{B}_\ell,\theta_\ell)$ converges pointwise a.e. in Q to 1. Hence, the Vitali theorem and $0\le g_{\frac{1}{\ell}}<1$, imply

$$g_{\frac{1}{\ell}}(\mathbb{B}_\ell,\theta_\ell)\to 1 \quad \text{strongly in} \quad L^p(Q;\mathbb{R}) \quad \text{for any} \quad 1\le p<\infty. \tag{5.66}$$

Therefore, regarding the first two equations (4.39) and (4.40), we can take the limit in the same way as we did in the limit $n\to\infty$. Indeed, the integrability of the resulting non-linear limits was already verified when estimating $\partial_t \boldsymbol{v}_\ell$ and $\partial_t\mathbb{B}_\ell$ ((5.23)–(5.29)). This way, taking (5.66) into account, using the density of $\operatorname{span}\{\boldsymbol{w}_i\}_{i=1}^\infty$ in $W^{1,\frac{d}{2}+1}_{\boldsymbol{n},\mathrm{div}}$ and extending the functional $\partial_t\mathbb{B}$ to the space stated in (3.30) using (5.30), we obtain precisely (3.37) and (3.38).

Next, we show how to take the limit in (4.82). Regarding the initial condition, using (4.7) and (4.6), we estimate

$$|\eta_0^\omega| \le c_v|\ln\theta_0^\omega| + \mu(|\operatorname{tr}\mathbb{B}_0^\omega| + d + |\ln\det\mathbb{B}_0^\omega|) \le C(|\ln\theta_0| + |\mathbb{B}_0| + |\ln\det\mathbb{B}_0| + 1),$$

where the right hand side is integrable by assumptions (3.22) and (3.23). Moreover, the function $\eta_0^{1/\ell}$ converges point-wise a.e. in Ω due to (4.8) and (4.9). Thus, by the dominated convergence theorem, the function $\eta_0^{\frac{1}{\ell}}$ converges to η_0 in $L^1(\Omega;\mathbb{R})$. In order to take the limit in the convective term, we use (5.50), (5.56) and (5.38). Next, the properties (5.33), (5.36), (5.59), (5.53) and (5.35), (5.32) imply

$$\ln\theta_\ell \rightharpoonup \ln\theta \qquad \text{weakly in } L^2(0,T;W^{1,2}(\Omega;\mathbb{R})), \tag{5.67}$$

$$\ln\det\mathbb{B}_\ell \rightharpoonup \ln\det\mathbb{B} \qquad \text{weakly in } L^2(0,T;W^{1,2}(\Omega;\mathbb{R})). \tag{5.68}$$

Further, we use (3.1), (3.3), (5.21) and Vitali's theorem to find that

$$\sqrt{\kappa(\theta_\ell)} \rightharpoonup \sqrt{\kappa(\theta)} \quad \text{strongly in } L^{\frac{2(r+1)}{r}}(Q;\mathbb{R}). \tag{5.69}$$

As a consequence of this, (5.67) and (5.8), we get

$$\sqrt{\kappa(\theta_\ell)}\nabla\ln\theta_\ell \rightharpoonup \sqrt{\kappa(\theta)}\nabla\ln\theta \quad \text{weakly in } L^2(Q;\mathbb{R}^d). \tag{5.70}$$

Therefore, using again (5.69), we obtain

$$\kappa(\theta_\ell)\nabla\ln\theta_\ell \rightharpoonup \kappa(\theta)\nabla\ln\theta \quad \text{weakly in } L^1(Q;\mathbb{R}^d).$$

Next, in the term $\mu\lambda(\theta_\ell)\nabla\operatorname{tr}\mathbb{B}_\ell$, we use (3.1), (3.4), (5.59), Vitali's theorem and (5.52). Analogously, we take the limit in the term $\mu\lambda(\theta_\ell)\nabla\ln\det\mathbb{B}_\ell$, only we use (5.68) instead of (5.52). The term containing $\omega|\nabla\theta_\ell|^r\nabla\ln\theta_\ell$ tends to zero by (5.41).

Now we take the limit in the terms on the right hand side of (4.82), i.e., the function ξ_ℓ defined in (4.72). Note that we just need to pass to the limit with possible inequality sign (selecting non-negative test functions ϕ, φ). To take the limit in the term $\mathbb{P}(\theta_\ell,\mathbb{B}_\ell)\cdot(\mathbb{I}-\mathbb{B}_\ell^{-1})\phi\varphi\ge 0$, we use (5.59), (5.53) and apply Fatou's lemma. Next, in term $\kappa(\theta_\ell)|\nabla\ln\theta_\ell|^2\phi\varphi$, we use (5.70) and the weak lower semi-continuity.

Moreover, the auxiliary term $\omega|\nabla\theta_\ell|^r|\nabla\ln\theta_\ell|^2\phi\varphi$ is simply estimated from below by zero. Thus, in order to let $\ell\to\infty$ in (4.82), it remains to show that

$$\begin{aligned}\liminf_{\ell\to\infty}\int_Q\left(\frac{2\nu(\theta_\ell)}{\theta_\ell}|\mathbb{D}\boldsymbol{v}_\ell|^2+\lambda(\theta_\ell)\left|\mathbb{B}_\ell^{-\frac12}\nabla\mathbb{B}_\ell\mathbb{B}_\ell^{-\frac12}\right|^2\right)\phi\varphi\\ \geq\int_Q\left(\frac{2\nu(\theta)}{\theta}|\mathbb{D}\boldsymbol{v}|^2+\lambda(\theta)\left|\mathbb{B}^{-\frac12}\nabla\mathbb{B}\mathbb{B}^{-\frac12}\right|^2\right)\phi\varphi.\end{aligned}$$

The above inequality is however consequence of the weak lower semicontinuity and the following claim

$$\begin{aligned}\sqrt{\lambda(\theta_\ell)}\mathbb{B}_\ell^{-\frac12}\nabla\mathbb{B}_\ell\mathbb{B}_\ell^{-\frac12}&\rightharpoonup\sqrt{\lambda(\theta)}\mathbb{B}^{-\frac12}\nabla\mathbb{B}\mathbb{B}^{-\frac12}\quad\text{weakly in } L^2(Q;\mathbb{R}^d\times\mathbb{R}^{d\times d}_{\mathrm{sym}}),\\ \sqrt{\frac{2\nu(\theta_\ell)}{\theta_\ell}}\mathbb{D}\boldsymbol{v}_\ell&\rightharpoonup\sqrt{\frac{2\nu(\theta)}{\theta}}\mathbb{D}\boldsymbol{v}\qquad\text{weakly in } L^2(Q;\mathbb{R}^{d\times d}_{\mathrm{sym}}),\end{aligned}\tag{5.71}$$

which we need to obtain. To do so, we start with (5.8) and therefore we have (for a proper subsequence) that

$$\sqrt{\lambda(\theta_\ell)}\mathbb{B}_\ell^{-\frac12}\nabla\mathbb{B}_\ell\mathbb{B}_\ell^{-\frac12}\rightharpoonup G\qquad\text{weakly in } L^2(Q;\mathbb{R}^d\times\mathbb{R}^{d\times d}_{\mathrm{sym}}),\tag{5.72}$$

$$\sqrt{\frac{2\nu(\theta_\ell)}{\theta_\ell}}\mathbb{D}\boldsymbol{v}_\ell\rightharpoonup K\qquad\text{weakly in } L^2(Q;\mathbb{R}^{d\times d}_{\mathrm{sym}}).\tag{5.73}$$

Thus, it remains to show

$$\sqrt{\lambda(\theta)}\mathbb{B}^{-\frac12}\nabla\mathbb{B}\mathbb{B}^{-\frac12}=G,\qquad\sqrt{\frac{2\nu(\theta)}{\theta}}\mathbb{D}\boldsymbol{v}=K.\tag{5.74}$$

First, we use the Egorov theorem and then it follows from (5.34), (5.53), (5.63), (5.65) and (5.59) that for any $\varepsilon>0$ there exists measurable $Q_\varepsilon\subset Q$ fulfilling $|Q\setminus Q_\varepsilon|\leq\varepsilon$ such that

$$\mathbb{B}_\ell^{-\frac12}\rightrightarrows\mathbb{B}^{-\frac12},\qquad\sqrt{\lambda(\theta_\ell)}\rightrightarrows\sqrt{\lambda(\theta)},\qquad\sqrt{\frac{2\nu(\theta_\ell)}{\theta_\ell}}\rightrightarrows\sqrt{\frac{2\nu(\theta)}{\theta}}$$

uniformly in Q_ε. Combining the above uniform convergence results with the weak convergence results (5.49) and (5.52), we deduce

$$\begin{aligned}\sqrt{\lambda(\theta_\ell)}\mathbb{B}_\ell^{-\frac12}\nabla\mathbb{B}_\ell\mathbb{B}_\ell^{-\frac12}&\rightharpoonup\sqrt{\lambda(\theta)}\mathbb{B}^{-\frac12}\nabla\mathbb{B}\mathbb{B}^{-\frac12}\quad\text{weakly in } L^1(Q_\varepsilon;\mathbb{R}^d\times\mathbb{R}^{d\times d}_{\mathrm{sym}}),\\ \sqrt{\frac{2\nu(\theta_\ell)}{\theta_\ell}}\mathbb{D}\boldsymbol{v}_\ell&\rightharpoonup\sqrt{\frac{2\nu(\theta)}{\theta}}\mathbb{D}\boldsymbol{v}\qquad\text{weakly in } L^1(Q_\varepsilon;\mathbb{R}^{d\times d}_{\mathrm{sym}}).\end{aligned}\tag{5.75}$$

Thus, the uniqueness of a weak limit implies that (5.74) is satisfied a.e. in Q_ε. Since $\varepsilon>0$ was arbitrary, we can let $\varepsilon\to0+$ and conclude that (5.74) holds true a.e. in Q. Consequently, we deduced (5.71) and therefore we proved (3.40).

In addition, in very similar manner we can let $\ell\to\infty$ in (4.59) to obtain (3.39). Note that contrary to the entropy inequality, we use here in addition the estimates (5.43), (5.44) and (5.45). Otherwise, the proof is almost identical.

To take the limit in (4.84), we first note, using (4.45) and (4.56), that it implies

$$-(E_\ell,\partial_t\phi)_Q+\alpha(|\boldsymbol{v}_\ell|^2,\phi)_\Sigma=(\tfrac12|P_\ell\boldsymbol{v}_0|^2+c_v\theta_0^{\frac{1}{\ell}},\phi(0))+(\boldsymbol{g},\boldsymbol{v}_\ell\phi)_Q\tag{5.76}$$

for all $\phi \in \mathcal{C}^1([0,T];\mathbb{R})$ with $\phi(T)=0$. Then, recalling (5.50) and (5.59), we see that $E_\ell = \frac{1}{2}|\boldsymbol{v}_\ell|^2 + c_v\theta_\ell$ converges strongly to E and thus, using also properties of P_ℓ and (4.9), we can take the limit in (5.76) to conclude

$$-(E,\partial_t\phi)_Q + \alpha(|\boldsymbol{v}|^2,\phi)_\Sigma = (E_0,\phi(0)) + (\boldsymbol{g},\boldsymbol{v}\phi)_Q, \tag{5.77}$$

where we set $E_0 \coloneqq \frac{1}{2}|\boldsymbol{v}_0|^2 + c_v\theta_0$. In particular, by choosing an appropriate sequence of test functions ϕ, we obtain (3.41).

Attainment of initial conditions. To finish the existence proof, it remains to identify the initial conditions and show that they are attained strongly. Let us start by an observation that $\boldsymbol{v}$ and $\mathbb{B}$ are weakly continuous in time. Indeed, first of all, we recall that

$$\begin{aligned}&\boldsymbol{v} \in L^\infty(0,T;L^2(\Omega;\mathbb{R}^d)), \quad \partial_t\boldsymbol{v} \in L^{\frac{d+2}{d}}(0,T;W^{-1,\frac{d+2}{d}}_{\boldsymbol{n},\mathrm{div}}(\Omega;\mathbb{R}^d)),\\ &\mathbb{B} \in L^\infty(0,T;L^q(\Omega;\mathbb{R}^{d\times d}_{>0})), \quad \partial_t\mathbb{B} \in L^{\frac{2q}{q+1}}(0,T;W^{-1,\frac{2q}{q+1}}(\Omega;\mathbb{R}^{d\times d}_{\mathrm{sym}})),\end{aligned} \tag{5.78}$$

cf. (5.20) and (5.30). From this we obtain, by a standard argument known from the theory of Navier–Stokes equations (see e.g. [43, Sect. 3.8.]), that

$$\boldsymbol{v} \in \mathcal{C}_w([0,T];L^2(\Omega;\mathbb{R}^d)) \quad \text{and} \quad \mathbb{B} \in \mathcal{C}_w([0,T];L^q(\Omega;\mathbb{R}^d)). \tag{5.79}$$

Then, to identify the corresponding weak limits, we can use an analogous idea as in the part where the limit $n\to\infty$ was taken together with (4.8). This way, we obtain

$$\lim_{t\to 0+}(\boldsymbol{v}(t),\boldsymbol{w}) = (\boldsymbol{v}_0,\boldsymbol{w}) \quad \text{for all } \boldsymbol{w}\in L^2(\Omega;\mathbb{R}^d) \tag{5.80}$$

and

$$\lim_{t\to 0+}(\mathbb{B}(t),\mathbb{W}) = (\mathbb{B}_0,\mathbb{W}) \quad \text{for all } \mathbb{W}\in L^{q'}(\Omega;\mathbb{R}^{d\times d}_{\mathrm{sym}}). \tag{5.81}$$

Next, we use a similar procedure for entropy and temperature. Recalling (5.57) and (5.60), we can define for all $t^0\in[0,T]$ the values $\sqrt{\theta}(t^0_\pm)$, $\eta(t^0_\pm)$ such that

$$\lim_{t\to t^0\pm}\big(\big\|\sqrt{\theta(t)} - \sqrt{\theta(t^0_\pm)}\big\|_{W^{-M,2}(\Omega;\mathbb{R})} + \|\eta(t)-\eta(t^0_\pm)\|_{W^{-M,2}(\Omega;\mathbb{R})}\big) = 0. \tag{5.82}$$

Therefore, using the density of $L^w(\Omega;\mathbb{R})$ in $W^{-M,2}(\Omega;\mathbb{R})$, which is valid for all $w\in(1,\infty)$ and M sufficiently large, and recalling the fact that $\theta\in L^\infty(0,T;L^1(\Omega;\mathbb{R}))$, we can deduce that there is non-negative $\theta^*_0\in L^1(\Omega;\mathbb{R})$ fulfilling

$$\lim_{t\to 0+}(\sqrt{\theta(t)},\zeta) = (\sqrt{\theta^*_0},\zeta) \quad \text{for all } \zeta\in L^2(\Omega;\mathbb{R}). \tag{5.83}$$

Our aim is to show that $\theta^*_0=\theta_0$ and that it is attained strongly.

Unlike in the theory of Navier–Stokes(–Fourier) systems, we can not draw information about $\limsup_{t\to0+}\|\boldsymbol{v}(t)\|_2^2$ from the (kinetic) energy estimate directly because of the presence of $\theta\mathbb{B}$ in (3.37). Instead, we need first to combine the total energy and entropy balances to obtain the initial condition for θ. In (5.77) we choose a sequence of test functions ϕ approximating the function $\chi_{[0,t)}$, $t\in(0,T)$. This way, after taking the appropriate limit, we arrive at

$$\int_\Omega E(t) + \alpha\int_0^t\int_\Omega|\boldsymbol{v}|^2 = \int_\Omega E_0 + \int_0^t\int_\Omega \boldsymbol{g}\cdot\boldsymbol{v} \quad \text{for a.a. } t\in(0,T). \tag{5.84}$$

Next, we strengthen the above relation to be valid for all $t \in (0,T)$ with possibly inequality sign. Due to the weak continuity of $\boldsymbol{v}$, see (5.79), we see that $\boldsymbol{v}(\tau)$ is uniquely defined for all $\tau \in (0,T)$ and

$$\lim_{t\to\tau}(\boldsymbol{v}(t),\boldsymbol{w}) = (\boldsymbol{v}(\tau),\boldsymbol{w}) \quad \text{for all } \boldsymbol{w}\in L^2(\Omega;\mathbb{R}^d). \tag{5.85}$$

The same is however not true for θ since it is not weakly continuous w.r.t. $t \in (0,T)$. Nevertheless, we can define one-side values for every $t\in(0,T)$ with the help of (5.60), i.e., using the similar arguments as in (5.83), we have the one-sided uniquely defined weak limit

$$\lim_{t\to\tau\pm}(\sqrt{\theta(t)},\zeta) = (\sqrt{\theta(\tau_\pm)},\zeta) \quad \text{for all } \zeta\in L^2(\Omega;\mathbb{R}). \tag{5.86}$$

Next, we use above weak convergence results in (5.84). Integrating it with respect to $t\in(\tau,\tau+\delta)$, we get

$$\int_\tau^{\tau+\delta}\int_\Omega E(t)\,\mathrm{d}t = \int_\tau^{\tau+\delta}\int_0^t\left(\int_\Omega \boldsymbol{g}\cdot\boldsymbol{v} - \alpha\int_\Omega|\boldsymbol{v}|^2\right)\mathrm{d}t + \delta\int_\Omega E_0.$$

Thus, dividing by δ, letting first $\delta\to 0+$ and then $\tau\to 0+$, we get

$$\lim_{\tau\to0+}\lim_{\delta\to0+}\delta^{-1}\int_\tau^{\tau+\delta}\int_\Omega E(t)\,\mathrm{d}t = \int_\Omega E_0 = \int_\Omega(\tfrac12|\boldsymbol{v}_0|^2 + c_v\theta_0) \tag{5.87}$$

and in a very similar manner, we obtain

$$\lim_{\tau\to0+}\lim_{\delta\to0+}\delta^{-1}\int_{\tau-\delta}^{\tau}\int_\Omega E(t)\,\mathrm{d}t = \int_\Omega E_0 = \int_\Omega(\tfrac12|\boldsymbol{v}_0|^2 + c_v\theta_0). \tag{5.88}$$

We focus on the term on the left hand side. Using the convexity, we have

$$\begin{aligned}
\delta^{-1}\int_\tau^{\tau+\delta}\int_\Omega E(t)\,\mathrm{d}t &= \delta^{-1}\int_\tau^{\tau+\delta}\int_\Omega(\tfrac12|\boldsymbol{v}(t)|^2 + c_v\theta(t))\,\mathrm{d}t\\
&\geq \delta^{-1}\int_\tau^{\tau+\delta}\int_\Omega(\tfrac12|\boldsymbol{v}(\tau)|^2 + c_v\theta(\tau_+)) + \boldsymbol{v}(\tau)\cdot(\boldsymbol{v}(t)-\boldsymbol{v}(\tau))\\
&\qquad + 2c_v\sqrt{\theta(\tau_+)}\big(\sqrt{\theta(t)}-\sqrt{\theta(\tau_+)}\big)\,\mathrm{d}t\\
&\geq \int_\Omega(\tfrac12|\boldsymbol{v}(\tau)|^2 + c_v\theta(\tau_+))\\
&\quad - \sup_{t\in(\tau,\tau+\delta)}\left|\int_\Omega \boldsymbol{v}(\tau)\cdot(\boldsymbol{v}(t)-\boldsymbol{v}(\tau)) + 2c_v\sqrt{\theta(\tau_+)}\big(\sqrt{\theta(t)}-\sqrt{\theta(\tau_+)}\big)\right|.
\end{aligned}$$

Then, it follows from the weak continuity results (5.85) and (5.86) and also from the above inequality that

$$\lim_{\delta\to0+}\delta^{-1}\int_\tau^{\tau+\delta}\int_\Omega E(t)\,\mathrm{d}t \geq \int_\Omega(\tfrac12|\boldsymbol{v}(\tau)|^2 + c_v\theta(\tau_+)).$$

Repeating the same procedure we also get

$$\lim_{\delta\to0+}\delta^{-1}\int_{\tau-\delta}^{\tau}\int_\Omega E(t)\,\mathrm{d}t \geq \int_\Omega(\tfrac12|\boldsymbol{v}(\tau)|^2 + c_v\theta(\tau_-)).$$

Consequently, combining it with (5.87) and (5.88), and also with (3.22), (5.80), (5.83) and weak lower semi-continuity, we get

$$\begin{aligned}\int_\Omega (\tfrac{1}{2}|\boldsymbol{v}_0|^2 + c_v\theta_0) &\geq \limsup_{t\to 0+} \int_\Omega (\tfrac{1}{2}|\boldsymbol{v}(t)|^2 + c_v\theta(t_\pm)) \\ &\geq \liminf_{t\to 0+} \int_\Omega \tfrac{1}{2}|\boldsymbol{v}(t)|^2 + \limsup_{t\to 0+} \int_\Omega c_v\theta(t_\pm) \\ &\geq \int_\Omega \tfrac{1}{2}|\boldsymbol{v}_0|^2 + \limsup_{t\to 0+} \int_\Omega c_v\theta(t_\pm),\end{aligned}$$

hence due to (5.83) and the convexity of the second power, we have

$$\int_\Omega \theta_0^* \leq \limsup_{t\to 0+} \int_\Omega \theta(t_\pm) \leq \int_\Omega \theta_0. \tag{5.89}$$

In what follows we will not distinguish "$\pm$" in $\theta(t_\pm)$ and $\eta(t_\pm)$ and simply write $\theta(t)$ and $\eta(t)$. To obtain also the corresponding lower estimate, we need to extract the available information from the entropy inequality (3.40). To this end, we localize (3.40) in time, using a sequence of non-negative functions approximating $\chi_{[0,t)}$. This way, we eventually obtain

$$\int_\Omega \eta(t)\phi + \int_0^t \int_\Omega \boldsymbol{j}\cdot\nabla\phi \geq \int_\Omega \eta_0\phi + \int_0^t \int_\Omega \xi\phi \tag{5.90}$$

a.e. in $(0,T)$ and for all $\phi \in W^{M,2}(\Omega;\mathbb{R}_{\geq 0})$, where

$$\boldsymbol{j} \coloneqq -\boldsymbol{v}\eta + \kappa(\theta)\nabla\ln\theta - \mu\lambda(\theta)\nabla(\operatorname{tr}\mathbb{B} - d - \ln\det\mathbb{B}) \in L^1(Q;\mathbb{R}^d).$$

Hence, using (5.82) and taking $\liminf_{t\to 0+}$ of (5.90) (which surely exists due to (5.82)), we deduce (3.45). Let us now fix $\varphi \in \mathcal{C}^M(\Omega;\mathbb{R}_{\geq 0})$ such that $\int_\Omega \varphi = 1$. Since f is convex, we get from (3.45) and (5.81) (or (3.43)) that

$$\begin{aligned}\int_\Omega c_v \ln\theta_0\,\varphi = \int_\Omega \eta_0\varphi + \int_\Omega f(\mathbb{B}_0)\varphi &\leq \liminf_{t\to 0+} \int_\Omega \eta(t)\varphi + \liminf_{t\to 0+} \int_\Omega f(\mathbb{B}(t))\varphi \\ &\leq \liminf_{t\to 0+} \int_\Omega c_v \ln\theta(t)\,\varphi.\end{aligned}$$

If we use this information together with Jensen's inequality and the fact that the function $s \mapsto \exp(s/2)$, is increasing and convex in $\mathbb{R}$, we are led to

$$\begin{aligned}\exp\left(\frac{1}{2}\int_\Omega \ln\theta_0\varphi\right) &\leq \exp\left(\frac{1}{2}\liminf_{t\to 0+}\int_\Omega \ln\theta(t)\varphi\right) \\ &= \liminf_{t\to 0+} \exp\left(\int_\Omega \ln\sqrt{\theta(t)}\,\varphi\right) \leq \liminf_{t\to 0+}\int_\Omega \sqrt{\theta(t)}\varphi \\ &= \int_\Omega \sqrt{\theta_0^*}\varphi.\end{aligned} \tag{5.91}$$

In every Lebesgue point $x_0 \in \Omega$ of both $\ln\theta_0$ and θ_0^*, we can localize inequality (5.91) in Ω by choosing a sequence of functions φ that approximates the Dirac delta distribution at $x_0 \in \Omega$. Indeed, appealing to the Lebesgue differentiation theorem, we get this way that

$$\sqrt{\theta_0(x_0)} = \exp\left(\frac{1}{2}\ln\theta(x_0)\right) \leq \sqrt{\theta_0^*(x_0)},$$

hence $\theta_0 \leq \theta_0^*$ a.e. in Ω, which together with (5.89) implies $\theta_0^* = \theta_0$ a.e. in Ω. To show the strong convergence, we use (5.83) with $\zeta := \sqrt{\theta_0}$ and also (5.89), to deduce

$$\limsup_{t\to 0+} \big\| \sqrt{\theta(t)} - \sqrt{\theta_0}\big\|_2^2 = \limsup_{t\to 0+} \int_\Omega \theta(t) + \int_\Omega \theta_0 - 2 \lim_{t\to 0+} \int_\Omega \sqrt{\theta(t)}\sqrt{\theta_0} \leq 0.$$

Hence, the above inequality implies that

$$\sqrt{\theta(t)} \to \sqrt{\theta_0} \quad \text{ strongly in } L^2(\Omega;\mathbb{R}),$$

which implies (3.44).

Using information above, we can now improve the initial condition for $\boldsymbol{v}$ as well. Indeed, from (5.84), (3.44) and (3.22), we obtain

$$\begin{aligned}\limsup_{t\to 0+} \int_\Omega \tfrac{1}{2}|\boldsymbol{v}(t)|^2 &\leq \limsup_{t\to 0+} \int_\Omega E(t) - \liminf_{t\to 0+} \int_\Omega c_v\theta(t)\\ &\leq \int_\Omega E_0 + \lim_{t\to 0+} \int_0^t (\boldsymbol{g},\boldsymbol{v}) - \int_\Omega c_v\theta_0 = \int_\Omega \tfrac{1}{2}|\boldsymbol{v}_0|^2.\end{aligned}$$

Thus, using also (5.80), we conclude that

$$\limsup_{t\to 0+} \|\boldsymbol{v}(t) - \boldsymbol{v}_0\|_2^2 = \limsup_{t\to 0+} \int_\Omega |\boldsymbol{v}(t)|^2 + \int_\Omega |\boldsymbol{v}_0|^2 - 2\lim_{t\to 0+} \int_\Omega \boldsymbol{v}(t)\cdot\boldsymbol{v}_0 \leq 0,$$

which implies (3.42).

Finally, since f is strictly convex on $\mathbb{R}^{d\times d}_{>0}$ as

$$f''(\mathbb{B})\mathbb{A}\cdot\mathbb{A} = \mu\mathbb{B}^{-1}\mathbb{A}\mathbb{B}^{-1}\cdot\mathbb{A} = \mu|\mathbb{B}^{-\frac{1}{2}}\mathbb{A}\mathbb{B}^{-\frac{1}{2}}|^2, \quad \mathbb{B}\in\mathbb{R}^{d\times d}_{>0}, \quad \mathbb{A}\in\mathbb{R}^{d\times d},$$

the strong attainment of the initial condition for $\mathbb{B}$ (3.43) follows readily from (5.81), the classical result [55, Theorem 3 (i)] and the Vitali's theorem once we show the property

$$\limsup_{t\to 0+} \int_\Omega f(\mathbb{B}(t)) \leq \int_\Omega f(\mathbb{B}_0). \tag{5.92}$$

To this end, we make an observation that in (4.70) (with (4.81) in place), we can choose $\phi = 1$, drop the non-negative terms, integrate over $(0,t)$ and then estimate the right-hand side using Hölder inequality, (5.20) and (5.19) to obtain

$$\int_\Omega f(\mathbb{B}_\ell(t)) - \int_\Omega f(\mathbb{B}_0^{\frac{1}{\ell}}) \leq \int_0^t \int_\Omega 2a\mu g_{\frac{1}{\ell}}(\mathbb{B}_\ell,\theta_\ell)\mathbb{B}_\ell\cdot\mathbb{D}\boldsymbol{v}_\ell \leq Ct^{\frac{1}{2q'}}. \tag{5.93}$$

Note that again we rely on (5.48) to give a proper meaning to the left-hand side of (5.93) for all $t\in(0,T)$. Utilizing now the convexity and continuity of f on $\mathbb{R}^{d\times d}_{>0}$ and (4.8), taking the limit $\ell\to\infty$ in (5.93) leads to

$$\int_\Omega f(\mathbb{B}(t)) - \int_\Omega f(\mathbb{B}_0) \leq \liminf_{\ell\to\infty} \Big(\int_\Omega f(\mathbb{B}_\ell(t)) - \int_\Omega f(\mathbb{B}_0^{\frac{1}{\ell}})\Big) \leq Ct^{\frac{1}{2q'}},$$

from which (5.92) immediately follows.

6. Global energy equality for $d \leq 3$

To derive (3.46) (which is a weak version of (1.12)), we need to construct the pressure p and ensure that every term appearing (3.46) is integrable. To this end, we apply the conditions (3.11). Moreover, we need to be able to test the momentum equation with $\boldsymbol{v}\phi$, where ϕ is some smooth function on Q. Unfortunately, we can not do this operation in (3.37) nor at any stage of our approximation scheme.

The remedy is to truncate the convection term in the balance of momentum. However, then we are just mimicking the existence proof that is done in [7] for a different non-linear fluid. Thus, let us only verify the weak compactness of weak solutions $(\boldsymbol{v}_\delta, \mathrm{p}_\delta, \mathbb{B}_\delta, \theta_\delta, \eta_\delta)$ to the system $\operatorname{div} \boldsymbol{v}_\delta = 0$, (3.38), (3.40),

$$\langle \partial_t \boldsymbol{v}_\delta, \boldsymbol{\varphi}\rangle - (T_\delta \boldsymbol{v}_\delta \otimes \boldsymbol{v}_\delta, \nabla\boldsymbol{\varphi})_Q + (\mathbb{S}_\delta, \nabla\boldsymbol{\varphi})_Q + \alpha(\boldsymbol{v}_\delta \boldsymbol{\varphi})_\Sigma = (\mathrm{p}_\delta, \operatorname{div}\boldsymbol{\varphi})_Q + (\boldsymbol{g}, \boldsymbol{\varphi})_Q \quad (6.1)$$

for all $\boldsymbol{\varphi} \in L^\infty(0,T; W^{1,\infty}_{\boldsymbol{n}})$, with $T_\delta \boldsymbol{v}_\delta = ((\boldsymbol{v}_\delta s_\delta) * r_\delta)_{\mathrm{div}}$, where s_δ is a truncation near $\partial\Omega$, r_δ is a standard mollifier and $(\cdot)_{\mathrm{div}}$ is a Helmholtz projection onto divergence-free functions, and

$$\begin{aligned} &- (E_0, \phi)\varphi(0) - (E_\delta, \phi\partial_t\varphi)_Q + \alpha(|\boldsymbol{v}_\delta|^2, \phi\varphi)_\Sigma + (\kappa(\theta_\delta)\nabla\theta_\delta, \nabla\phi\varphi)_Q \\ &\qquad = (E_\delta \boldsymbol{v}_\delta + \mathrm{p}_\delta \boldsymbol{v}_\delta - \mathbb{S}_\delta \boldsymbol{v}_\delta, \nabla\phi\varphi)_Q \end{aligned} \quad (6.2)$$

for all $\varphi \in W^{1,\infty}((0,T);\mathbb{R})$, $\varphi(T) = 0$, and every $\phi \in W^{1,\infty}(\Omega;\mathbb{R})$. The existence of such solutions follows by combining the approximation scheme from Section 4 together with the one in [7]. In view of the uniform estimates derived in Sections 4–5, we may suppose that the sequence $\{(\boldsymbol{v}_\delta, \mathrm{p}_\delta, \mathbb{B}_\delta, \theta_\delta, \eta_\delta)\}_{\delta>0}$ is uniformly bounded in the spaces depicted in (3.25)–(3.36) and that we have the same convergence results as in (5.49)–(5.59) and so forth (with ℓ replaced by δ). We may also suppose that, say $\mathrm{p}_\delta \in L^2(Q;\mathbb{R})$ with $\int_\Omega \mathrm{p}_\delta = 0$. Then, since we have $\nu(\theta_\delta)\mathbb{D}\boldsymbol{v}_\delta, \theta_\delta\mathbb{B}_\delta \in L^2(Q;\mathbb{R}^{d\times d}_{\mathrm{sym}})$ and the convection term is truncated, equation (6.1) is valid for all $\boldsymbol{\varphi} \in L^2(0,T;W^{1,2}_{\boldsymbol{n}})$, in fact. What is missing is the uniform estimate of the pressure. By localizing (6.1) in time, choosing $\boldsymbol{\varphi} = \nabla u$ and using $\operatorname{div}\boldsymbol{v}_\delta = 0$, we obtain

$$-(\mathrm{p}_\delta, \Delta u) = (T_\delta \boldsymbol{v}_\delta \otimes \boldsymbol{v}_\delta - \mathbb{S}_\delta, \nabla\nabla u) - \alpha(\boldsymbol{v}_\delta, \nabla u)_{\partial\Omega} + (\boldsymbol{g}, \nabla u)$$

a.e. in $(0,T)$. There the convective term, if not truncated, is the most irregular one (recall that $\|\boldsymbol{v}_\delta \otimes \boldsymbol{v}_\delta\|_{L^{\frac{d+2}{d}} L^{\frac{d+2}{d}}} \le C$). Thus, expecting p_δ to have the same integrability, we may choose $u \in W^{2,(\frac{d+2}{d})'}(\Omega;\mathbb{R})$ to be the solution to the Neumann problem

$$\begin{aligned} -\Delta u &= |\mathrm{p}_0|^{\frac{d+2}{d}-2}\mathrm{p}_0 - \frac{1}{|\Omega|}\int_\Omega |\mathrm{p}_0|^{\frac{d+2}{d}-2}\mathrm{p}_0 \quad \text{in } \Omega, \\ \nabla u \cdot \boldsymbol{n} &= 0 \quad \text{on } \partial\Omega \end{aligned}$$

a.e. in $(0,T)$, where $\mathrm{p}_0 = \mathrm{p}_\delta - \frac{1}{|\Omega|}\int_\Omega \mathrm{p}_\delta$. Since $\|u\|_{2,(\frac{d+2}{d})'} \le C\|\mathrm{p}_0\|_{\frac{d+2}{d}}$ by the corresponding L^q-theory (here we used $\Omega \in \mathcal{C}^{1,1}$), the test function u eventually leads to

$$\|\mathrm{p}_\delta\|_{L^{\frac{d+2}{d}} L^{\frac{d+2}{d}}} \le C,$$

see [7] for details.

Taking the limit $\delta \to 0+$ in (6.1), (3.38) and (3.40) can be done analogously as when we considered the limit $\ell \to \infty$. Indeed, in the additional term $\int_0^T (\mathrm{p}_\delta, \operatorname{div}\boldsymbol{\varphi})$, we simply use the fact that $\mathrm{p}_\delta \rightharpoonup \mathrm{p}$ weakly in $L^{\frac{d+2}{d}}(Q;\mathbb{R})$. It remains to take the limit $\delta \to 0+$ in (6.2). Since $\boldsymbol{v}_\delta$ converges strongly in $L^{2\frac{d+2}{d}-\varepsilon}(Q;\mathbb{R}^d)$ and $d \le 3$, we deduce that the terms $p_\delta \boldsymbol{v}_\delta$ and $|\boldsymbol{v}_\delta|^2\boldsymbol{v}_\delta$ converge weakly to their limits. The limits in other terms were already discussed and we omit it here. Thus, the proof of Theorem 3.1 is complete.

Appendix A. Auxiliary results

In this additional section, we prove those auxiliary results which were used above but are not completely standard in the existing literature. On the other hand, they are not new and serve only to clarify some arguments used in the proof.

For the purposes of this section, we replace the interval $(0,T)$ (or $[0,T]$) by an arbitrary bounded interval $I \subset \mathbb{R}$ and set $Q = I \times \Omega$. The set Ω is always assumed to be a bounded Lipschitz domain in $\mathbb{R}^d$, $d \in \mathbb{N}$.

Intersections of Sobolev-Bochner spaces. If $X \overset{\text{dense}}{\hookrightarrow} H \overset{\text{dense}}{\hookrightarrow} X^*$ is a Gelfand triple, it is well known that

$$\mathcal{C}^1(I;X) \overset{\text{dense}}{\hookrightarrow} \mathcal{W}^p_X \hookrightarrow \mathcal{C}(I;H), \tag{A.1}$$

where

$$\mathcal{W}^p_X \coloneqq \big(\{u \in L^p(I;X);\ \partial_t u \in (L^p(I;X))^*\}, \|\cdot\|_{L^pX} + \|\partial_t \cdot\|_{L^{p'}X^*}\big), \quad 1<p<\infty.$$

The first embedding in (A.1) is useful to manipulate certain duality pairings involving time derivatives, while the second embedding is important for the identification of boundary values (i.e. initial conditions) and the corresponding integration by parts formulas. We would like to generalize (A.1) for the space

$$\mathcal{W}^{p,q}_{X,Y} \coloneqq \big(\{u \in L^p(I;X)\cap L^q(I;Y);\ \partial_t u \in (L^p(I;X)\cap L^q(I;Y))^*\},$$
$$\|\cdot\|_{L^pX\cap L^qY} + \|\partial_t\cdot\|_{(L^pX\cap L^qY)^*}\big), \quad 1<p,q<\infty,$$

The primary application which we have in mind is the case where $X = W^{1,2}(\Omega)$, $Y = L^\omega(\Omega)$ and $\omega > \frac{2d}{d-2}$ (i.e., we know better integrability than what follows from the Sobolev embedding, recall the function $\mathbb{B}_\ell$). Thus, we may assume that both X and Y admit the Gelfand triplet structure with a common Hilbert space H.

Lemma A.1. *Let $1 < p,q < \infty$ and suppose that X, Y are separable reflexive Banach spaces and H is separable Hilbert space forming Gelfand triples in the sense that*

$$X \overset{\text{dense}}{\hookrightarrow} H \overset{\text{dense}}{\hookrightarrow} X^* \quad \text{and} \quad Y \overset{\text{dense}}{\hookrightarrow} H \overset{\text{dense}}{\hookrightarrow} Y^*. \tag{A.2}$$

Then, we have the embeddings

$$\mathcal{C}^1(I;X\cap Y) \overset{\text{dense}}{\hookrightarrow} \mathcal{W}^{p,q}_{X,Y} \hookrightarrow \mathcal{C}(I;H). \tag{A.3}$$

Moreover, the integration by parts formula

$$(u(t_2),v(t_2))_H - (u(t_1),v(t_1))_H = \int_{t_1}^{t_2} \langle \partial_t u, v\rangle + \int_{t_1}^{t_2} \langle \partial_t v, u\rangle \tag{A.4}$$

holds for any $u,v \in \mathcal{W}^{p,q}_{X,Y}$ and any $t_1,t_2 \in I$.

Proof. The proof of the first embedding in (A.3) can be done in a standard way by extending u outside I evenly, taking the convolution with a smooth kernel and then estimating the difference from u and $\partial_t u$ in the respective norms. See [27] or [60] for details.

If $u,v \in \mathcal{C}^1(I;X\cap Y) \hookrightarrow \mathcal{C}(I;H)$, then $\partial_t u, \partial_t v \in \mathcal{C}(I;X\cap Y) \hookrightarrow \mathcal{C}(I;H)$ and, using density of the embeddings in (A.2), the duality in (A.4) can be represented as

$$\langle \partial_t u, v\rangle + \langle \partial_t v, u\rangle = (\partial_t u, v)_H + (\partial_t v, u)_H = \partial_t (u,v)_H \quad \text{a.e. in } I,$$

hence (A.4) is obvious in that case. Next, we can proceed as in [53, Lemma 7.3.] to prove that

$$\|u(t)\|_H \le C(\|u\|_{L^1 H} + \|u\|_{\mathcal{W}^{p,q}_{X,Y}}) \tag{A.5}$$

for all $t \in I$ and every $u \in \mathcal{C}^1(I; X \cap Y)$. Moreover, by (A.2), we have

$$\mathcal{W}^{p,q}_{X,Y} \hookrightarrow L^p(I;X) \cap L^q(I;Y) \hookrightarrow L^1(I;X) \cap L^1(I;Y) \hookrightarrow L^1(I; X+Y) \hookrightarrow L^1(I;H),$$

and thus (A.5) yields

$$\|u\|_{\mathcal{C}(I;H)} \le C\|u\|_{\mathcal{W}^{p,q}_{X,Y}}. \tag{A.6}$$

Since $\mathcal{C}^1(I; X \cap Y)$ is dense in $\mathcal{W}^{p,q}_{X,Y}$, the estimate (A.6) and identity (A.4) remain valid for all $u \in \mathcal{W}^{p,q}_{X,Y}$. Moreover, if $u \in \mathcal{W}^{p,q}_{X,Y}$, then we can take $v = u$ and $t_2 \to t_1$ in (A.4) to deduce that $u \in \mathcal{C}(I;H)$. Thus, the embedding $\mathcal{W}^{p,q}_{X,Y} \hookrightarrow \mathcal{C}(I;H)$ holds and the proof is finished. □

Since $\mathcal{W}^{p,p}_{X,X} = \mathcal{W}^p_X$, the classical result (A.1) can be seen as an obvious corollary.

Fundamental theorem of calculus in the Sobolev-Bochner setting. Let $H = L^2(\Omega)$. The formula (A.4) can be used to identify that

$$\langle \partial_t u, u\rangle = \frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}t}\int_\Omega u^2 \tag{A.7}$$

a.e. in I. However, in certain situations we would like to generalize (A.7) to

$$\langle \partial_t u, \psi(u)\rangle = \frac{\mathrm{d}}{\mathrm{d}t}\int_\Omega \int_w^u \psi(s)\,\mathrm{d}s.$$

Whether this is possible depends on what kind of function ψ is and also on the choice of X. The next lemma characterizes one such situation.

Lemma A.2. *Let $1 < p, q < \infty$. Suppose that $\psi : \mathbb{R} \to \mathbb{R}$ is a Lipschitz function. For $w \in \mathbb{R}$, we define*

$$\Psi(x) = \int_w^x \psi(s)\,\mathrm{d}s, \quad x \in \mathbb{R}.$$

Then, for any $u \in \mathcal{W}^p_{W^{1,q}(\Omega)}$, there holds

$$\Psi(u) \in \mathcal{C}(I; L^1(\Omega)) \tag{A.8}$$

and

$$\int_{t_1}^{t_2} \langle \partial_t u, \psi(u)\rangle = \int_\Omega \Psi(u(t_2)) - \int_\Omega \Psi(u(t_1)) \quad \text{for all } t_1, t_2 \in I. \tag{A.9}$$

Moreover, if ψ is bounded, then

$$\Psi(u) \in \mathcal{C}(I; L^2(\Omega)).$$

Proof. First of all, we remark that $\psi(u) \in W^{1,q}(\Omega)$ a.e. in I, by a classical result (see e.g. [61, Theorem 2.1.11.]), and thus the duality in (A.9) is well defined. Next, we apply Theorem A.1 to find $u_\varepsilon \in \mathcal{C}^1(I; W^{1,q}(\Omega))$ satisfying

$$\|u_\varepsilon - u\|_{L^p W^{1,q}} + \|\partial_t u_\varepsilon - \partial_t u\|_{L^{p'} W^{-1,q'}} \to 0 \quad \text{as } \varepsilon \to 0+. \tag{A.10}$$

Then, using the standard calculus, it is easy to see that the identity

$$\begin{aligned}\int_{t_1}^{t_2}\langle\partial_t u_\varepsilon,\psi(u_\varepsilon)\rangle &= \int_{t_1}^{t_2}\int_\Omega \psi(u_\varepsilon)\partial_t u_\varepsilon \\ &= \int_{t_1}^{t_2}\int_\Omega \partial_t\Psi(u_\varepsilon) = \int_\Omega \Psi(u_\varepsilon(t_2)) - \int_\Omega \Psi(u_\varepsilon(t_1))\end{aligned} \tag{A.11}$$

holds for any $t_1, t_2 \in I$. Denoting the Lipschitz constant of ψ by $L \geq 0$, we estimate

$$|\psi(u_\varepsilon)| \leq |\psi(u_\varepsilon) - \psi(0)| + |\psi(0)| \leq L|u_\varepsilon| + |\psi(0)|$$

and

$$|\nabla\psi(u_\varepsilon)| \leq |\psi'(u_\varepsilon)||\nabla u_\varepsilon| \leq L|\nabla u_\varepsilon|.$$

Hence, the sequence $\psi(u_\varepsilon)$ is bounded in $L^p(I;W^{1,q}(\Omega))$. As $1 < p, q < \infty$, this is a separable reflexive space, and thus, there exist a subsequence and its limit $\overline{\psi(u)} \in L^p(I;W^{1,q}(\Omega))$ such that

$$\psi(u_\varepsilon) \rightharpoonup \overline{\psi(u)} \quad \text{weakly in } L^p(I;W^{1,q}(\Omega)). \tag{A.12}$$

Since $p > 1$, a subsequence of u_ε converges point-wise a.e. in Q to u, and thus $\overline{\psi(u)} = \psi(u)$ using the continuity of ψ. Hence, by (A.10) and (A.12), we obtain

$$\begin{aligned}\int_{t_1}^{t_2}\langle\partial_t u_\varepsilon,\psi(u_\varepsilon)\rangle &= \int_{t_1}^{t_2}\langle\partial_t u_\varepsilon - \partial_t u,\psi(u_\varepsilon)\rangle + \int_{t_1}^{t_2}\langle\partial_t u,\psi(u_\varepsilon)\rangle \\ &\to \int_{t_1}^{t_2}\langle\partial_t u,\psi(u)\rangle\end{aligned} \tag{A.13}$$

as $\varepsilon \to 0+$. Next, using the embedding $\mathcal{W}^p_{W^{1,q}(\Omega)} \hookrightarrow \mathcal{C}(I;L^2(\Omega))$ and (A.10), we get, for any $t_0 \in I$, that

$$\|u(t) - u(t_0)\|_2 \to 0 \quad \text{as} \quad t \to t_0 \tag{A.14}$$

and

$$\|u_\varepsilon(t_0) - u(t_0)\|_2 \to 0 \quad \text{as} \quad \varepsilon \to 0+. \tag{A.15}$$

Then, the Lipschitz continuity of ψ, Hölder's inequality and (A.14) yield

$$\begin{aligned}\int_\Omega |\Psi(u(t)) - \Psi(u(t_0))| &= \int_\Omega \left|\int_{u(t_0)}^{u(t)} \psi(s)\,\mathrm{d}s\right| \leq \int_\Omega\int_{u(t_0)}^{u(t)} (|\psi(0)| + L|s|)\,\mathrm{d}s \\ &\leq \int_\Omega\int_{u(t_0)}^{u(t)} C(1 + |u(t_0)| + |u(t)|) \leq C\int_\Omega (1 + |u(t_0)| + |u(t)|)|u(t) - u(t_0)| \\ &\leq C\|1 + |u(t_0)| + |u(t)|\|_2\|u(t) - u(t_0)\|_2 \leq C\|u(t) - u(t_0)\|_2 \to 0\end{aligned} \tag{A.16}$$

as $t \to t_0$, which proves (A.8) (and thus, the values $\Phi(u(t))$, $t \in I$, are well defined). By an analogous estimate, using (A.15) instead of (A.14), we can prove that

$$\int_\Omega |\Psi(u_\varepsilon(t_0)) - \Psi(u(t_0))| \to 0 \quad \text{as } \varepsilon \to 0+$$

for any $t \in I$. This and (A.13) used in (A.11) to take the limit $\varepsilon \to 0+$ proves (A.9).

If ψ is bounded, we replace (A.16) by

$$\int_\Omega |\Psi(u(t)) - \Psi(u(t_0))|^2 = \int_\Omega \left|\int_{u(t_0)}^{u(t)} \psi(s)\,\mathrm{d}s\right|^2 \leq C\int_\Omega |u(t) - u(t_0)|^2$$

and the rest of the proof remains the same. □

Clearly, we can also replace ψ by $\psi\phi$, where $\phi \in W^{1,\infty}(\Omega;\mathbb{R})$, leading to

$$\int_0^t \langle \partial_t u, \psi(u)\phi\rangle = \int_\Omega \int_w^{u(t)} \psi(s)\,\mathrm{d}s\,\phi - \int_\Omega \int_w^{u(0)} \psi(s)\,\mathrm{d}s\,\phi \quad \text{for all } t \in I. \quad \text{(A.17)}$$

Then, since ϕ is a Lipschitz (time independent) function, the proof is basically the same as the one presented above.

Calculus for positive definite matrices. We recall that the operations "$\cdot$" and $|\cdot|$ on matrices are defined by

$$\mathbb{A}_1 \cdot \mathbb{A}_2 = \sum_{i=1}^d \sum_{j=1}^d (\mathbb{A}_1)_{ij}(\mathbb{A}_2)_{ij} \quad \text{and} \quad |\mathbb{A}| = \sqrt{\mathbb{A}\cdot\mathbb{A}},$$

respectively. Then, the object $|\mathbb{A}|$ coincides, in fact, with the Frobenius matrix norm of $\mathbb{A}$.

The next lemma is formulated for a function $\mathbb{A} : Q \to \mathbb{R}^{d\times d}_{>0}$ and for simplicity, we shall assume that $\mathbb{A}$ is continuously differentiable with respect to all variables, i.e., $\mathbb{A} \in \mathcal{C}^1(Q;\mathbb{R}^{d\times d}_{>0})$. In particular situations, this assumption can be of course removed by an appropriate approximation (convolution smoothing) and the assertions of the following lemma hereby extend to the setting of weakly differentiable functions. Let us also denote any of the space-time derivatives by a generic symbol ∂.

Lemma A.3. *Let $\mathbb{A} \in \mathcal{C}^1(Q;\mathbb{R}^{d\times d}_{>0})$. Then*

(i) $$0 \le \operatorname{tr}\mathbb{A} - d - \ln\det\mathbb{A}, \quad \text{(A.18)}$$

(ii) $$|\mathbb{A}| \le \operatorname{tr}\mathbb{A} \le \sqrt{d}|\mathbb{A}|, \quad \text{(A.19)}$$

(iii) $$\min\{1, d^{\frac{1-\alpha}{2}}\}|\mathbb{A}|^\alpha \le |\mathbb{A}^\alpha| \le \max\{1, d^{\frac{1-\alpha}{2}}\}|\mathbb{A}|^\alpha \quad \textit{for any} \quad \alpha \ge 0, \quad \text{(A.20)}$$

(iv) $$\partial\mathbb{A}\cdot\mathbb{A}^\alpha = \begin{cases} \frac{1}{\alpha+1}\partial\operatorname{tr}\mathbb{A}^{\alpha+1} & \textit{if} \quad \alpha \ne -1; \\ \partial\ln\det\mathbb{A} = \partial\operatorname{tr}\log\mathbb{A} & \textit{if} \quad \alpha = -1, \end{cases} \quad \text{(A.21)}$$

(v) $$(\operatorname{sign}\alpha)\partial\mathbb{A}\cdot\partial\mathbb{A}^\alpha \ge \begin{cases} \frac{4|\alpha|}{(\alpha+1)^2}|\partial\mathbb{A}^{\frac{\alpha+1}{2}}|^2 & \textit{if} \quad \alpha \ne -1; \\ |\partial\log\mathbb{A}|^2 & \textit{if} \quad \alpha = -1, \end{cases} \quad \text{(A.22)}$$

(vi) $$|\partial\mathbb{A}| \le 2|\mathbb{A}^{1-\alpha}\partial\mathbb{A}^\alpha| \quad \textit{for all} \quad \alpha \in [\tfrac12, 1). \quad \text{(A.23)}$$

Proof. Property (i) follows by passing to the spectral decomposition of $\mathbb{A}$ and from the fact that $x \mapsto x - 1 - \ln x$ attains its minimum at $x = 1$. Estimate (ii) is a consequence of the Cauchy–Schwarz inequality since

$$|\mathbb{A}| = |(\mathbb{A}^{\frac12})^T\mathbb{A}^{\frac12}| \le |\mathbb{A}^{\frac12}|^2 = \operatorname{tr}\mathbb{A} = \mathbb{I}\cdot\mathbb{A} \le |\mathbb{I}||\mathbb{A}| = \sqrt{d}|\mathbb{A}|.$$

For (iii), we refer to [3, Proposition 1] and for (iv), (v) to [3, Theorem 1]. The relation (iv) with $\alpha = -1$ is also known as the Jacobi identity.

Finally, property (A.23) can be shown using the idea from the proof of [3, Theorem 3], which we briefly sketch here. For any natural numbers p, q satisfying $q - p \le p < q$ (so that $\alpha = \frac{p}{q} \in [\frac12, 1)$), we may use the Young inequality to write

$$|\partial\mathbb{B}^q|^2 = |\partial\mathbb{B}^{q-p}\mathbb{B}^p + \mathbb{B}^{q-p}\partial\mathbb{B}^p|^2 \le 2|\partial\mathbb{B}^{q-p}\mathbb{B}^p|^2 + 2|\mathbb{B}^{q-p}\partial\mathbb{B}^p|^2 =: 2A + 2B. \quad \text{(A.24)}$$

Now we simply expand the derivative and rearrange the terms to get

$$A = \Big| \sum_{i=0}^{q-p-1} \mathbb{B}^i \partial\mathbb{B}\mathbb{B}^{q-1-i} \Big|^2 = \sum_{i=0}^{q-p-1} \sum_{j=0}^{q-p-1} \big| \mathbb{B}^{\frac{i+j}{2}} \partial\mathbb{B}\mathbb{B}^{q-1-\frac{i+j}{2}} \big|^2$$
$$= \sum_{s=0}^{2(q-p-1)} (1+\min\{s, 2(q-p-1)-s\}) \big| \mathbb{B}^{\frac{s}{2}} \partial\mathbb{B}\mathbb{B}^{q-1-\frac{s}{2}} \big|^2, \tag{A.25}$$

whereas for B, a completely analogous computation yields

$$B = \sum_{s=0}^{2(p-1)} (1+\min\{s, 2(p-1)-s\}) \big| \mathbb{B}^{\frac{s}{2}} \partial\mathbb{B}\mathbb{B}^{q-1-\frac{s}{2}} \big|^2.$$

Then, using $q-p \le p$ first inside the minimum in (A.25) and then in the number of terms of the sum (relying on the non-negativity of each term), we see that $A \le B$. Returning with this information to (A.24) and setting $\mathbb{B} = \mathbb{A}^{\frac{1}{q}}$, we easily conclude (A.23) for rational powers α. The general case follows by a density argument (the continuity of the mapping $\alpha \mapsto \mathbb{A}^{1-\alpha}$ follows immediately from the spectral decomposition, while continuity of $\alpha \mapsto \partial\mathbb{A}^\alpha$ is a consequence of the integral representation formula for $\partial \exp \mathbb{X}$, see [57] or [3] for more details).

□

References

[1] N. Arada and A. Sequeira, *Strong Steady Solutions for a Generalized Oldroyd-B Model with Shear-Dependent Viscosity in a Bounded Domain*, Mathematical Models and Methods in Applied Sciences, 13 (2003), pp. 1303–1323.

[2] J. W. Barrett and S. Boyaval, *Existence and approximation of a (regularized) Oldroyd-B model*, Mathematical Models and Methods in Applied Sciences, 21 (2011), pp. 1783–1837.

[3] M. Bathory, *Sharp nonlinear estimates for multiplying derivatives of positive definite tensor fields*, Mathematical Inequalities & Applications, 25 (2022), pp. 751–769.

[4] M. Bathory, M. Bulíček, and J. Málek, *Large data existence theory for three-dimensional unsteady flows of rate-type viscoelastic fluids with stress diffusion*, Advances in Nonlinear Analysis, 10 (2020), pp. 501–521.

[5] J. Blechta, J. Málek, and K. R. Rajagopal, *On the Classification of Incompressible Fluids and a Mathematical Analysis of the Equations That Govern Their Motion*, SIAM J. Math. Anal., 52 (2020), pp. 1232–1289.

[6] M. Bulíček, E. Feireisl, and J. Málek, *On a class of compressible viscoelastic rate-type fluids with stress-diffusion*, Nonlinearity, 32 (2019), pp. 4665–4681.

[7] M. Bulíček, J. Málek, and K. R. Rajagopal, *Mathematical analysis of unsteady flows of fluids with pressure, shear-rate, and temperature dependent material moduli that slip at solid boundaries*, SIAM Journal on Mathematical Analysis, 41 (2009), pp. 665–707.

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

[9] M. Bulíček and J. Havrda, *On existence of weak solution to a model describing incompressible mixtures with thermal diffusion cross effects*, ZAMM Z. Angew. Math. Mech., 95 (2015), pp. 589–619.

[10] M. Bulíček, T. Los, Y. Lu, and J. Málek, *On planar flows of viscoelastic fluids of Giesekus type*, Nonlinearity, 35 (2022), pp. 6557–6604.

[11] M. Bulíček and J. Málek, *Internal flows of incompressible fluids subject to stick-slip boundary conditions*, Vietnam J. Math., 45 (2017), pp. 207–220.

[12] ———, *Large data analysis for Kolmogorov's two-equation model of turbulence*, Nonlinear Anal. Real World Appl., 50 (2019), pp. 104–143.

[13] M. Bulíček, J. Málek, V. Průša, and E. Süli, *PDE analysis of a class of thermodynamically compatible viscoelastic rate-type fluids with stress-diffusion*, in Mathematical analysis in fluid mechanics—selected recent results, vol. 710 of Contemp. Math., Amer. Math. Soc., Providence, RI, 2018, pp. 25–51.

[14] M. Bulíček, J. Málek, V. Průša, and E. Süli, *On incompressible heat-conducting viscoelastic rate-type fluids with stress-diffusion and purely spherical elastic response*, SIAM J. Math. Anal., 53 (2021), pp. 3985–4030.

[15] M. Bulíček, J. Málek, and K. R. Rajagopal, *Navier's slip and evolutionary Navier-Stokes-like systems with pressure and shear-rate dependent viscosity*, Indiana University Mathematics Journal, 56 (2007), pp. 51–85.

[16] M. Bulíček, J. Málek, and C. Rodriguez, *Global well-posedness for two-dimensional flows of viscoelastic rate-type fluids with stress diffusion*, J. Math. Fluid Mech., 24 (2022), pp. Paper No. 61, 19.

[17] H. B. Callen, *Thermodynamics and an introduction to thermostatistics*, Wiley, New York, second ed., 1985.

[18] L. Chupin, *Global Strong Solutions for Some Differential Viscoelastic Models*, SIAM Journal on Applied Mathematics, 78 (2018), pp. 2919–2949.

[19] E. A. Coddington and N. Levinson, *Theory of ordinary differential equations*, McGraw-Hill Book Company, Inc., New York-Toronto-London, 1955.

[20] P. Constantin and M. Kliegl, *Note on Global Regularity for Two-Dimensional Oldroyd-B Fluids with Diffusive Stress*, Archive for Rational Mechanics and Analysis, 206 (2012), pp. 725–740.

[21] M. Dostalík, V. Průša, and T. Skřivan, *On diffusive variants of some classical viscoelastic rate-type models*, AIP Conference Proceedings, 2107 (2019).

[22] M. Dressler, B. J. Edwards, and H. C. Öttinger, *Macroscopic thermodynamics of flowing polymeric liquids*, Rheologica Acta, 38 (1999), p. 117 – 136.

[23] E. Feireisl, *Relative entropies in thermodynamics of complete fluid systems*, Discrete & Continuous Dynamical Systems - A, 32 (2012), pp. 3059–3080.

[24] E. Feireisl, M. Frémond, E. Rocca, and G. Schimperna, *A new approach to non-isothermal models for nematic liquid crystals*, Archive for Rational Mechanics and Analysis, 205 (2012), pp. 651–672.

[25] E. Feireisl and J. Málek, *On the Navier-Stokes equations with temperature-dependent transport coefficients*, Differ. Equ. Nonlinear Mech., (2006), pp. Art. ID 90616, 14.

[26] E. Feireisl, A. Novotný, and Y. Sun, *On the motion of viscous, compressible, and heat-conducting liquids*, Journal of Mathematical Physics, 57 (2016), p. 083101.

[27] H. Gajewski, K. Gröger, and K. Zacharias, *Nichtlineare Operatorgleichungen und Operatordifferentialgleichungen*, Akademie-Verlag, Berlin, 1974. Mathematische Lehrbücher und Monographien, II. Abteilung, Mathematische Monographien, Band 38.

[28] R. J. Gordon and W. R. Schowalter, *Anisotropic fluid theory: A different approach to the dumbbell theory of dilute polymer solutions*, Transactions of the Society of Rheology, 16 (1972), pp. 79–97.

[29] C. Guillopé and J.-C. Saut, *Existence results for the flow of viscoelastic fluids with a differential constitutive law*, Nonlinear Analysis. Theory, Methods & Applications. An International Multidisciplinary Journal, 15 (1990), pp. 849–869.

[30] M. Heida, R. I. A. Patterson, and D. R. M. Renger, *Topologies and measures on the space of functions of bounded variation taking values in a Banach or metric space*, J. Evol. Equ., 19 (2019), pp. 111–152.

[31] J. Hron, V. Miloš, V. Průša, O. Souček, and K. Tůma, *On thermodynamics of incompressible viscoelastic rate type fluids with temperature dependent material coefficients*, International Journal of Non-Linear Mechanics, 95 (2017), pp. 193–208.

[32] I. Ireka and T. Chinyoka, *Non-isothermal flow of a johnson-segalman liquid in a lubricated pipe with wall slip*, Journal of Non-Newtonian Fluid Mechanics, 192 (2013), p. 20 – 28.

[33] M. Johnson and D. Segalman, *A model for viscoelastic fluid behavior which allows non-affine deformation*, Journal of Non-Newtonian Fluid Mechanics, 2 (1977), pp. 255–270.

[34] O. Kreml, *Mathematical Analysis of Models for Viscoelastic Fluids*, Ph.D. thesis, Charles University, 2010.

[35] O. Kreml, M. Pokorný, and P. Šalom, *On the global existence for a regularized model of viscoelastic non-Newtonian fluid*, Colloquium Mathematicum, 139 (2015), pp. 149–163.

[36] A. Leonov, *Nonequilibrium thermodynamics and rheology of viscoelastic polymer media*, Rheologica Acta, 15 (1976), pp. 85–98.
[37] P. L. Lions and N. Masmoudi, *Global solutions for some oldroyd models of non-newtonian flows*, Chinese Annals of Mathematics, 21 (2000), pp. 131–146.
[38] Y. Lu and M. Pokorný, *Global existence of large data weak solutions for a simplified compressible Oldroyd-B model without stress diffusion*, Analysis in Theory and Applications, 36 (2020), pp. 348–372.
[39] M. Lukáčová-Medviďová, H. Mizerová, Š. Nečasová, and M. Renardy, *Global existence result for the generalized Peterlin viscoelastic model*, SIAM Journal on Mathematical Analysis, 49 (2017), pp. 2950–2964.
[40] J. Málek, J. Nečas, M. Rokyta, and M. Růžička, *Weak and Measure-valued Solutions to Evolutionary PDEs*, Chapman & Hall, 1996.
[41] J. Málek and V. Průša, *Derivation of equations for continuum mechanics and thermodynamics of fluids*, in Handbook of mathematical analysis in mechanics of viscous fluids, Springer, Cham, 2018, pp. 3–72.
[42] J. Málek, V. Průša, T. Skřivan, and E. Süli, *Thermodynamics of viscoelastic rate-type fluids with stress diffusion*, Physics of Fluids, 30 (2018).
[43] J. Málek and K. R. Rajagopal, *Mathematical issues concerning the Navier-Stokes equations and some of its generalizations*, in Evolutionary equations. Vol. II, Handb. Differ. Equ., Elsevier/North-Holland, Amsterdam, 2005, pp. 371–459.
[44] J. Málek, K. R. Rajagopal, and K. Tůma, *On a variant of the Maxwell and Oldroyd-B models within the context of a thermodynamic basis*, International Journal of Non-Linear Mechanics, 76 (2015), pp. 42 – 47.
[45] ———, *Derivation of the variants of the Burgers model using a thermodynamic approach and appealing to the concept of evolving natural configurations*, Fluids, 3 (2018).
[46] N. Masmoudi, *Global existence of weak solutions to macroscopic models of polymeric flows*, Journal de Mathématiques Pures et Appliquées, 96 (2011), pp. 502–520.
[47] J. Nečas, *Direct methods in the theory of elliptic equations*, Springer Monographs in Mathematics, Springer, Heidelberg, 2012. Translated from the 1967 French original by Gerard Tronel and Alois Kufner, Editorial coordination and preface by Šárka Nečasová and a contribution by Christian G. Simader.
[48] P. D. Olmsted, O. Radulescu, and C.-Y. D. Lu, *Johnson–segalman model with a diffusion term in cylindrical couette flow*, Journal of Rheology, 44 (2000), pp. 257–275.
[49] R. Pivokonský, P. Filip, and J. Zelenková, *The role of the gordon–schowalter derivative term in the constitutive models—improved flexibility of the modified XPP model*, Colloid and Polymer Science, 293 (2015), pp. 1227–1236.
[50] K. R. Rajagopal and A. R. Srinivasa, *A thermodynamic frame work for rate type fluid models*, J. Non-Newton. Fluid Mech., 88 (2000), pp. 207–227.
[51] ———, *On thermomechanical restrictions of continua*, Proc. R. Soc. Lond., Ser. A, Math. Phys. Eng. Sci., 460 (2004), pp. 631–651.
[52] I. J. Rao and K. R. Rajagopal, *A thermodynamic framework for the study of crystallization in polymers*, Zeitschrift für angewandte Mathematik und Physik ZAMP, 53 (2002), pp. 365–406.
[53] T. Roubíček, *Nonlinear partial differential equations with applications*, vol. 153 of International Series of Numerical Mathematics, Birkhäuser/Springer Basel AG, Basel, second ed., 2013.
[54] R. I. Tanner, *The changing face of rheology*, Journal of Non-Newtonian Fluid Mechanics, 157 (2009), p. 141 – 144.
[55] A. Visintin, *Strong convergence results related to strict convexity*, Communications in Partial Differential Equations, 9 (1984), pp. 439–466.
[56] P. Wapperom and M. Hulsen, *Thermodynamics of viscoelastic fluids: the temperature equation*, Journal of Rheology, 42 (1998), pp. 999–1019.
[57] R. M. Wilcox, *Exponential operators and parameter differentiation in quantum physics*, Journal of Mathematical Physics, 8 (1967), pp. 962–982.
[58] A. Wouk, *Integral representation of the logarithm of matrices and operators*, Journal of Mathematical Analysis and Applications, 11 (1965), pp. 131–138.
[59] S. Zaremba, *Sur une forme perfectionee de la theorie de la relaxation*, Bull. Int. Acad. Sci. Cracovie, (1903), pp. 594–614.

[60] E. Zeidler, *Nonlinear functional analysis and its applications. II/A*, Springer-Verlag, New York, 1990. Linear monotone operators, Translated from the German by the author and Leo F. Boron.

[61] W. P. Ziemer, *Weakly differentiable functions*, vol. 120 of Graduate Texts in Mathematics, Springer-Verlag, New York, 1989. Sobolev spaces and functions of bounded variation.

Michal Bathory, University of Vienna, Faculty of Mathematics, Oskar-Morgenstern-Platz 1, 1090 Wien, Austria

Email address: michal.bathory@univie.ac.at

Miroslav Bulíček, Charles University, Faculty of Mathematics and Physics, Mathematical Institute, Sokolovská 83, 186 75 Praha 8, Czech Republic

Email address: mbul8060@karlin.mff.cuni.cz

Josef Málek, Charles University, Faculty of Mathematics and Physics, Mathematical Institute, Sokolovská 83, 186 75 Praha 8, Czech Republic

Email address: malek@karlin.mff.cuni.cz