FACULTY
OF MATHEMATICS
AND PHYSICS
Charles University

DOCTORAL THESIS

Michal Bathory

Analysis of unsteady flows of incompressible heat-conducting rate-type viscoelastic fluids with stress-diffusion

Mathematical Institute of Charles University

Supervisor of the doctoral thesis: RNDr. Miroslav Bulíček, Ph.D.
Study programme: Physics
Study branch: Mathematical and computer modeling

Prague 2020

I declare that I carried out this doctoral thesis independently, and only with the cited sources, literature and other professional sources. It has not been used to obtain another or the same degree.

I understand that my work relates to the rights and obligations under the Act No. 121/2000 Sb., the Copyright Act, as amended, in particular the fact that the Charles University has the right to conclude a license agreement on the use of this work as a school work pursuant to Section 60 subsection 1 of the Copyright Act.

In Prague on May 20, 2020 . . . .

Author's signature

I thank Miroslav Bulíček for his insightful advice that kept me on the right path to my goals.

I also thank Josef Málek for support and motivation during my studies.

Finally, I thank my family and friends for the support and for providing me with the right distractions.

Title: Analysis of unsteady flows of incompressible heat-conducting rate-type viscoelastic fluids with stress-diffusion

Author: Michal Bathory

Institute: Mathematical Institute of Charles University

Supervisor: RNDr. Miroslav Bulíček, Ph.D., Mathematical Institute of Charles University

Abstract: We prove a global-in-time and large-data existence of a suitable weak solution to a system of partial differential equations describing an unsteady flow of homogeneous incompressible viscoelastic rate-type fluid. The material parameters are continuous functions of temperature and, in particular, the dependence of the shear modulus is assumed to be linear. It is shown that studied models obey the fundamental laws of thermodynamics. The key step towards the existence proof is derivation of the balance of entropy. This inequality is paramount in the analysis and as its consequence, we obtain sufficient a priori estimates, positivity of temperature and also regularity of the elastic deformation. The second part of the thesis deals with the existence analysis for the isothermal case, however using a completely different method, which is of independent interest.

Keywords: viscoelastic fluid, heat conducting fluid, Navier-Stokes equations, Oldroyd-B model, rate type fluid

Contents

1. Introduction

This work is devoted to a fast developing field of viscoelastic fluids and their mathematical analysis. It seems that the growing interest in viscoelastic fluids is related to a recent progress in various engineering areas, where these fluids find many applications. Although there are plenty of studies devoted to the numerical analysis of these fluids in some particular settings, a rigorous treatment of the models which describe the motion of these fluids is hard to find. Actually, the corresponding analysis in a reasonably general setting (including the temperature evolution) is non-existent at the present time. This is due to intrinsic difficulties hidden in the equations describing the evolution of the elastic component of the fluid and also of the temperature.

The thesis consists of two independent parts. The first part is the main part, which has four chapters and concerns an analysis of heat conducting viscoelastic fluids, where the material coefficients are allowed to depend on the temperature. As such, the presented existence theory for these fluids is the first of its kind. Then, the second part is included in Chapter 5 and provides an analysis in a simpler isothermal case. This part coincides with an author's upcoming journal article and thus, it is completely stand-alone and can be read independently. The existence analysis of each part rests upon different mathematical ideas. Throughout the thesis, we sometimes refer to the Chapter 5 in order to explain why ideas from the isothermal case do not carry over to the general case. A reader who is unfamiliar with the mathematical and/or physical theory for viscoelastic fluids might want to start reading Chapter 5, which is simpler, and thus more lucid. On the other hand, the main contribution of the thesis is the existence result for heat-conducting viscoelastic fluids. To be more precise, we prove that, without any restriction on the size of data or dimension of the space, there exists a suitable weak solution to a system of nonlinear partial differential equations with initial and boundary conditions, which describes an unsteady flow of a homogeneous, incompressible, heat-conducting, rate-type viscoelastic fluid (with stress diffusion), that fills up a mechanically and thermally isolated container. Throughout the thesis, we explain the meaning of all quantities mentioned in the previous sentence. Also, in this chapter, we describe the results of both parts of the thesis in more detail. We start the explanatory part with the most important term, the notion of *viscoelastic fluid.*

1.1 What is a viscoelastic fluid?

The adjective "viscoelastic" means that the fluid has both viscous and elastic properties. In other words, compared to standard viscous fluids such as water or oil, viscoelastic fluids behave partly as a rubber band, for example. More precisely, they can store and release mechanical energy under compression, stretching or twisting of the material. Examples of materials that fits well to such concepts are: synthetic polymers, rubbers, molten glasses or metals. Furthermore, many biomaterials are viscoelastic. Let us name at least two examples: a tendon and blood. Although these two materials seemingly have nothing in common, they can be both regarded as viscoelastic. Indeed, the blood is not a standard viscous

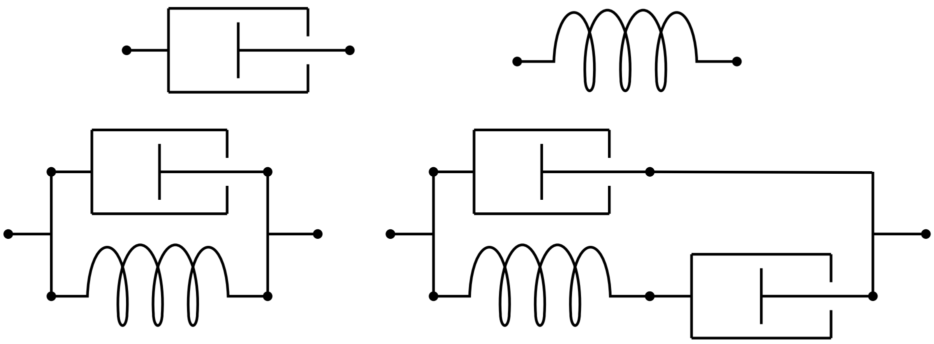

Figure 1.1: On the first row there are 1D analogues of viscous and elastic materials: dashpot (damper) and spring. On the second row we see two of many possible combinations of the basic elements. These are examples of 1D models of a viscoelastic fluid.

fluid as it contains red blood cells which have the tendency to always return to their original shape. Thus, these cells are responsible for the elastic behaviour of blood. Vice versa, the fact that an elastic tendon may be stretched during an exercise proves that a viscous component must be present. In fact, it turns out that many real-world materials are not purely viscous, nor purely elastic, and therefore should be regarded and modelled as viscoelastic (see e.g. [2], [22], [44], [46], [7]). However, it seems that a fluid flow and an elastic deformation can not be truthfully described within a single physical theory. As the title suggests, we focus only on the fluid flow. Thus, in the biological analogy, this thesis is about blood and not about tendons. To say it more elaborately, we assume that the material and the time frame in which we consider its motion fits better to a fluid-like movement than to a solid-like deformation. Furthermore, on a theoretical side, this means that the material characteristics are studied in an infinitesimal neighbourhood of its current state, rather than relating them to some initial state.

The idea that a fluid possesses an elastic part can be nicely visualized using the one-dimensional elements, which we can think of as of the fluid "molecules", see Figure 1.1. While the energy which we use to stretch or compress a dashpot partially dissipates to heat, a spring can store this energy and release it in the same form. Thus, it is also useful to think of the dashpot and spring as representatives of irreversible and reversible processes, respectively. However, these simple ideas are very idealized for several reasons. First, it is not clear how the multi-dimensional analogues of the one-dimensional models depicted on the second line of Figure 1.1 should look like. This issue can be partially resolved by appealing to certain physical principles, to be seen in Chapter 2. The major issue, however, is that in reality properties of a viscoelastic fluid (such as resistance to deformation) change dramatically with temperature changes. Thus, the temperature evolution and its effects on the flow characteristics should not be omitted for viscoelastic fluids. The main contribution of this thesis is that we provide an existence analysis for a variant of a viscoelastic fluid model that incorporates the full evolution of the temperature and also of the whole extra stress tensor.

The concept of the extra stress tensor is common for all viscoelastic fluids. In Chapter 2 we show that the Cauchy stress tensor $\mathbb{T}$, which represents the forces appearing in the balance of momentum, decomposes as

$$\mathbb{T} = \mathbb{T}_{\text{viscous}} + \mathbb{T}_{\text{elastic}} \tag{1.1}$$

for viscoelastic fluids. The first part $\mathbb{T}_{\text{viscous}}$ is the usual Cauchy stress tensor as in the theory of Newtonian fluids (with possibly a temperature dependent viscosity, however). The other part $\mathbb{T}_{\text{elastic}}$ is the one which is usually called the extra stress tensor in the literature and it corresponds to the forces arising from the elastic response of the fluid. Thus, viscoelastic fluids are non-Newtonian by definition and they always exhibit the effects known as non-linear creep and stress relaxation (see e.g. [68]).

In what comes next, we specify precisely what kind of viscoelastic fluids we are interested in.

1.2 Brief description of the model

We show in Chapter 2 that a flow of a viscoelastic fluid is governed by a system of non-linear partial differential equations completed by initial and boundary conditions. To fix ideas, we sketch this system here, although we refer the reader to Chapter 3 for a rigorous formulation.

Let $\Omega \subset \mathbb{R}^d$, $d \in \mathbb{N}$, be a domain. Let $T > 0$ be the final time and set $Q = (0,T) \times \Omega$. Furthermore, let $\boldsymbol{v}_0 : \Omega \to \mathbb{R}^d$, $\mathbb{B}_0 : \Omega \to \mathbb{R}^{d\times d}_{>0}$, $\theta_0 : \Omega \to (0,\infty)$ be some given initial data and let $\boldsymbol{f} : Q \to \mathbb{R}^d$ be a density of external body forces (such as the gravity). Suppose also that a, c_v, μ are some constants and $\nu, \gamma, \kappa, \delta, \lambda$ are some positive continuous functions (these quantities are the material parameters, or coefficients, of the model). Finally, let us denote the symmetric and antisymmetric parts of the velocity gradient $\nabla\boldsymbol{v}$ as

$$\mathbb{D}\boldsymbol{v} = \frac{1}{2}(\nabla\boldsymbol{v} + (\nabla\boldsymbol{v})^T), \qquad \mathbb{W}\boldsymbol{v} = \frac{1}{2}(\nabla\boldsymbol{v} - (\nabla\boldsymbol{v})^T).$$

Our goal is to show that under some reasonable assumptions, there exist a velocity $\boldsymbol{v} : Q \to \mathbb{R}^d$, a “pressure” $\mathrm{p} : Q \to \mathbb{R}$, an extra stress tensor $\mathbb{B} : Q \to \mathbb{R}^{d\times d}_{>0}$ and a temperature $\theta : Q \to (0,\infty)$ solving the following initial-boundary value problem (in some appropriate sense):

$$\operatorname{div}\boldsymbol{v} = 0, \tag{1.2}$$

$$\partial_t\boldsymbol{v} + \boldsymbol{v}\cdot\nabla\boldsymbol{v} - \operatorname{div}(2\nu(\theta)\mathbb{D}\boldsymbol{v}) + \nabla\mathrm{p} = 2a\mu\operatorname{div}(\theta\,\mathbb{B}) + \boldsymbol{f}, \tag{1.3}$$

$$\begin{aligned}\partial_t\mathbb{B} + \boldsymbol{v}\cdot\nabla\mathbb{B} + \delta(\theta)\gamma(|\mathbb{B}-\mathbb{I}|)(\mathbb{B}-\mathbb{I}) - \operatorname{div}(\lambda(\theta)\nabla\mathbb{B})\\ = \mathbb{W}\boldsymbol{v}\mathbb{B} - \mathbb{B}\mathbb{W}\boldsymbol{v} + a(\mathbb{D}\boldsymbol{v}\mathbb{B} + \mathbb{B}\mathbb{D}\boldsymbol{v}),\end{aligned} \tag{1.4}$$

$$c_v\partial_t\theta + c_v\boldsymbol{v}\cdot\nabla\theta - \operatorname{div}(\kappa(\theta)\nabla\theta) = 2\nu(\theta)|\mathbb{D}\boldsymbol{v}|^2 + 2a\mu\,\theta\,\mathbb{B}\cdot\mathbb{D}\boldsymbol{v}, \tag{1.5}$$

$$\begin{aligned}\partial_t E + \boldsymbol{v}\cdot\nabla E - \operatorname{div}(\kappa(\theta)\nabla\theta)\\ = \operatorname{div}(-\mathrm{p}\boldsymbol{v} + 2\nu(\mathbb{D}\boldsymbol{v})\boldsymbol{v} + 2a\mu\theta\mathbb{B}\boldsymbol{v}) + \boldsymbol{f}\cdot\boldsymbol{v},\end{aligned} \tag{1.6}$$

$$\partial_t\eta + \boldsymbol{v}\cdot\nabla\eta - \operatorname{div}(\kappa(\theta)\nabla\ln\theta - \lambda(\theta)\mu\nabla(\operatorname{tr}\mathbb{B} - d - \ln\det\mathbb{B})) = \xi \tag{1.7}$$

in Q with boundary conditions ($\boldsymbol{n}$ is the outward unit normal vector to $\partial\Omega$)

$$\boldsymbol{v} = 0, \qquad \boldsymbol{n}\cdot\nabla\mathbb{B} = 0, \qquad \boldsymbol{n}\cdot\nabla\theta = 0 \qquad \text{on } (0,T)\times\partial\Omega$$

and with initial conditions

$$\boldsymbol{v} = \boldsymbol{v}_0, \qquad \mathbb{B} = \mathbb{B}_0, \qquad \theta = \theta_0 \qquad \text{in } \{0\} \times \Omega.$$

Here, the total energy E is defined as

$$E = \tfrac{1}{2}|\boldsymbol{v}|^2 + c_v\theta,$$

the entropy is related to θ and $\mathbb{B}$ via the formula

$$\eta = c_v \ln\theta - \mu(\operatorname{tr}\mathbb{B} - d - \ln\det\mathbb{B}) \tag{1.8}$$

and the rate of entropy production ξ is given by

$$\xi = \frac{2\nu}{\theta}|\mathbb{D}\boldsymbol{v}|^2 + \kappa(\theta)|\nabla\ln\theta|^2 + \delta(\theta)\gamma(|\mathbb{B} - \mathbb{I}|)\left|\mathbb{B}^{\frac{1}{2}} - \mathbb{B}^{-\frac{1}{2}}\right|^2 + \mu\lambda(\theta)\left|\mathbb{B}^{-\frac{1}{2}}\nabla\mathbb{B}\mathbb{B}^{-\frac{1}{2}}\right|^2$$

However, in such a form, this goal seems to be too ambitious. In fact, we need to relieve from this solution concept in several different aspects. Namely, equation (1.5) is to be abandoned completely, (1.6) is replaced by its global version inequality and (1.7) is also weakened to just an inequality. The precise reasons for such reductions are discussed in Chapter 3. On top of that, we work only with weak solutions, of course. Nevertheless, even then we are be able to show that the solution which we construct obeys basic physical principles.

Equations (1.2), (1.3) and (1.5) form an incompressible Navier-Stokes-Fourier system with viscosity depending on temperature and with two additional terms $2a\mu\operatorname{div}(\theta\mathbb{B})$ and $2a\mu\theta\mathbb{B}\cdot\mathbb{D}\boldsymbol{v}$ that both depend on $\mathbb{B}$. Since the fluid is assumed to be homogeneous, its density is taken to be equal to one for simplicity. The unknown p, which was labelled as "pressure" has no connection to any physical pressure (see [62] for various notions of pressure), except for some simple situations, such as a pipe flow. In general, it is merely a Lagrange multiplier corresponding to the incompressibility constraint (1.2) and it is determined only up to a constant (see [57, Sect. 4.2.3.] for a detailed discussion). In fact, in our analysis, we do not care about the construction of a "pressure" at all since it is eliminated by taking the Leray projection $\boldsymbol{u} \mapsto \boldsymbol{u} - \nabla\Delta^{-1}\operatorname{div}\boldsymbol{u}$ of (1.3). This is indeed possible since

1) we do not consider any dependence of material coefficients on the pressure,

2) we avoid the local form of (1.6), where p appears explicitly,

3) we consider Dirichlet boundary conditions for the velocity.

If we wished to relax any of these assumptions, we would need to construct the pressure using an additional approximation (see e.g. [12]). This, however, is not our goal here.

The quantity $\mathbb{B}$ that appears in the system above models the elastic part of the total deformation of the fluid. In our case, it coincides with $\mathbb{T}_{\text{elastic}}$ from (1.1), up to a temperature-dependent multiplicative coefficient. It is possible to give $\mathbb{B}$ a specific physical meaning by appealing to the concept of an evolving natural (or stress-free) configuration. Then, the quantity $\mathbb{B}$ can be interpreted as the left Cauchy-Green tensor with respect to the stress free configuration. We refer the reader to [63] for details. For our purposes, the quantity $\mathbb{B}$ is merely an unknown of the system, which must be a positive definite tensor in order to ensure that the elastic deformation is regular.

The evolution of $\mathbb{B}$ is governed by (1.4), which is a diffusive variant of Oldroyd-B (or rather Johnson-Segalman, see [41]) model, but with a non-standard growth given by γ. We want to point out that none of the equations (1.3), (1.4), (1.5) can be decoupled from the other two, even if the material parameters were independent of the temperature. Indeed, this is caused by the term $2a\mu \operatorname{div}(\theta \mathbb{B})$ on the right hand side of (1.3). In fact, this was one of our motivations to study the system above, to illustrate that the existence theory for such systems is achievable even if the equations for θ and $\mathbb{B}$ can not be decoupled from the Navier-Stokes system.

The term $D = -\operatorname{div}(\lambda(\theta)\nabla\mathbb{B})$ indicates that we are dealing with the diffusive variant of a viscoelastic model. This term is physically well motivated, see [27], [30] and references therein. Although it is generally believed that presence of a diffusion term improves mathematical properties of any system, it is not so clear here. On one hand, the diffusion should lead to an estimate of $\nabla\mathbb{B}$ in some reasonable function space that should, eventually, lead to certain compactness property of $\mathbb{B}$. On the other hand, obtaining such an estimate is a very delicate (if possible) procedure. Indeed, in general, we can not test (1.4) by $\mathbb{B}$ since the other terms of the type $\nabla\boldsymbol{v}\mathbb{B}$ have very poor integrability. In our case however, we are able to test (1.4) by a small power of $\operatorname{tr}\mathbb{B}$, which improves the integrability of $\mathbb{B}$ and, consequently, also differentiability. We remark that the form of D is by no means unique, one can derive similar models with diffusion terms different from D simply by modifying the entropy production mechanism (see Chapter 2 for details). However, such a modification can have a large (usually negative) impact on the mathematical analysis of the model. For example if the diffusion term took the form $-\mathbb{B}(\Delta\mathbb{B}) - (\Delta\mathbb{B})\mathbb{B}$ (as suggested e.g. in [53]) then it would be unclear whether it would be even possible to define such a term in a reasonable way ($\mathbb{B}$ is not twice differentiable, in general). Another possibility is to consider $-\operatorname{div}(\beta(\nabla\mathbb{B})\nabla\mathbb{B})$, where β satisfies certain growth restrictions, similarly as for the p-Laplacian problems (see [42]). On the other hand, the obvious advantage of D is that it is linear, which is preferred both physically (for easy interpretation) and mathematically: we can always assign a weak sense to a Laplacian and we can easily identify the weak limit, which is not the case for non-linear problems.

Presence of products of the type $\nabla\boldsymbol{v}\mathbb{B}$ in (1.4) suggests that we are considering a rate-type model, where an objective derivative of $\mathbb{B}$, defined by

$$\overset{\circ}{\mathbb{B}} = \partial_t\mathbb{B} + \boldsymbol{v}\cdot\nabla\mathbb{B} - (\mathbb{W}\boldsymbol{v}\mathbb{B} - \mathbb{B}\mathbb{W}\boldsymbol{v}) - a(\mathbb{D}\boldsymbol{v}\mathbb{B} + \mathbb{B}\mathbb{D}\boldsymbol{v}),$$

is used. Here on the right hand side, the first two terms correspond to the usual convected time derivative, while the third term is needed to ensure the material frame indifference of $\overset{\circ}{\mathbb{B}}$ (see Chapter 2 for more details). The case $a = 1$ corresponds to the upper convected Oldroyd derivative (cf. [61]), the case $a = 0$ yields the corrotational Jaumann-Zaremba derivative (see [73]) and the case $a \in [-1, 1]$ gives the class of Gordon-Schowalter derivatives (cf. [36]). The case $a = 0$ is much easier to handle mathematically (since the products of the type $\nabla\boldsymbol{v}\mathbb{B}$ disappear upon testing (1.4) with $\mathbb{B}$), while the case $a = 1$ seems to be the correct choice from the physical point of view. We wish to point out that in our analysis we place no restriction on parameter a.

The equation (1.5) governs the evolution of temperature in the moving fluid. Its particular form (1.5) which is considered in this work deserves some clarification. If we compare it to the temperature equations derived in [40] or [53], our form (1.5) is much simpler. This is a consequence of the fact that the underlying Helmholtz free energy of the system (1.2)–(1.7) is chosen as

$$\psi(\theta, \mathbb{B}) = -c_v\theta(\ln\theta - 1) + \mu\theta(\operatorname{tr}\mathbb{B} - d - \ln\det\mathbb{B}). \tag{1.9}$$

In particular, the shear modulus takes the linear form $\mu\theta$, where $\mu > 0$ is constant. Hence, we also obtain a linear relation between the internal energy and temperature: $e = c_v\theta$, where the constant $c_v > 0$ is the heat capacity. On the other hand, the constitutive equation (1.8) for entropy η is slightly more complicated than it would be in the case of constant shear modulus. A similar model including the assumption (1.9) is used in [65, p. 379] (and in follow-up works) to model certain polymer melts. Nevertheless, we are aware that the assumption on the linearity of shear modulus with respect to the temperature should be relaxed in general (cf. [40, Sect. 2]). This however, would lead to many other complications in the mathematical analysis of the model that would obscure the other ideas in this thesis. Thus, we avoid this generalization for now, although in Chapter 2 we derive the model for as general ψ as possible, which opens the door for future studies. We also remark that the existence analysis for a similar model with general shear modulus is appearing soon in [20], but with a different simplification. The reduction, which occurs if $e = c_v\theta$, is not the only reason why we choose (1.9) to work with. The other reason is that in the balance equations (1.3) and (1.5), the terms $\operatorname{div}(\theta\mathbb{B})$ and $\theta\mathbb{B}\cdot\mathbb{D}\boldsymbol{v}$ appear, respectively. These terms are quite irregular from the point of view of a priori estimates and subsequent analysis. Moreover, the term $\operatorname{div}(\theta\mathbb{B})$ on the right hand side of (1.3) introduces an unavoidable coupling to the rest of the system. These are the main mathematical issues which occur also in the case of general shear modulus (cf. [20]). Therefore, the model (1.2)–(1.7), while simpler in structure, really seems as an appropriate "toy-problem" for the analysis of models with general shear moduli.

The expression $\theta\mathbb{B}\cdot\mathbb{D}\boldsymbol{v}$ appears in (1.5) as a consequence of the laws of thermodynamics. It is thus surprising that this term is sometimes omitted even in some recent works on viscoelastic fluids (see [40] for more details). It can be seen in Chapter 3 that this term is by far the most difficult to define by using only the a priori estimates that can be derived for the system (1.2)–(1.7). For this reason, at some point, we abandon (1.5) completely and replace it by two inequalities. First one is the global version of the total energy balance (1.6). Second one is the balance of entropy (1.7). This way, we avoid mathematical difficulties connected with (1.5), while still obtaining physically relevant solution. We thus see that the actual form of the temperature equation is not so important for the analysis. While we use (1.5) in the construction of our weak solution (which might be impossible in the case of non-linear shear modulus), we may choose not to do so. Indeed, it is possible to use the balance of internal energy instead, which always has a simpler structure (see e.g. [20]). However, this brings other technical difficulties, which we want to avoid here.

1.3 Isothermal case

In relation to the model introduced above, let us now describe the result in Chapter 5 which concerns only the isothermal case and in a slightly different setting. If the temperature is constant, then the equations (1.5), (1.7) are trivially satisfied and it is enough to work with the system (1.2)–(1.4). Moreover, motivated by the structure of the original Oldroyd-B and Giesekus models, we forbid the non-standard growth of the term

$$\delta(\theta)\gamma(|\mathbb{B}-\mathbb{I}|)(\mathbb{B}-\mathbb{I}) \quad \text{and replace it by} \quad \delta_1(\mathbb{B}-\mathbb{I})+\delta_2(\mathbb{B}^2-\mathbb{B}), \quad \delta_1,\delta_2\geq 0.$$

Furthermore, the coefficient γ (which, in a sense, improves the mathematical properties of the system) now plays a different role: it is merely a constant satisfying $\gamma\in(0,1)$. Finally, we also consider Navier-slip boundary conditions for the velocity (which seem to be physically appropriate and also can approximate the no-slip boundary condition). Thus, the final form of system studied in Chapter 5 is:

$$\operatorname{div}\boldsymbol{v}=0, \tag{1.10}$$

$$\begin{aligned}\partial_t\boldsymbol{v}+\boldsymbol{v}\cdot\nabla\boldsymbol{v}-\nu\Delta\boldsymbol{v}+\nabla p&\\ =2\mu a\operatorname{div}((1-\gamma)(\mathbb{B}-\mathbb{I})&+\gamma(\mathbb{B}^2-\mathbb{B}))+\boldsymbol{f},\end{aligned} \tag{1.11}$$

$$\begin{aligned}\partial_t\mathbb{B}+\boldsymbol{v}\cdot\nabla\mathbb{B}+\delta_1(\mathbb{B}-\mathbb{I})+\delta_2(\mathbb{B}^2-\mathbb{B})-\lambda\Delta\mathbb{B}&\\ =\mathbb{W}\mathbb{B}-\mathbb{B}\mathbb{W}+a(\mathbb{D}\boldsymbol{v}\mathbb{B}+\mathbb{B}\mathbb{D}\boldsymbol{v}&),\end{aligned} \tag{1.12}$$

in Q with boundary conditions[1]

$$\begin{aligned}\boldsymbol{v}\cdot\boldsymbol{n}=0, \qquad & -\sigma\boldsymbol{v}_\tau=\Big(\big(2\nu\mathbb{D}\boldsymbol{v}+2\mu a(1-\gamma)(\mathbb{B}-\mathbb{I})+2\mu a\gamma(\mathbb{B}^2-\mathbb{B})\big)\,\boldsymbol{n}\Big)_\tau,\\ \boldsymbol{n}\cdot\nabla\mathbb{B}=0,\end{aligned}$$

on $(0,T)\times\partial\Omega$ and with the initial conditions

$$\boldsymbol{v}(0,\cdot)=\boldsymbol{v}_0, \quad \mathbb{B}(0,\cdot)=\mathbb{B}_0 \qquad \text{in } \Omega.$$

Note that compared to (1.4), the equation (1.11) has a different right hand side. In terms of the elastic part of the Cauchy stress tensor, we replaced

$$\mathbb{T}_{\text{elastic}}=2\mu a\theta\mathbb{B}$$

by

$$\mathbb{T}_{\text{elastic}}=2\mu a(1-\gamma)(\mathbb{B}-\mathbb{I})+\gamma(\mathbb{B}^2-\mathbb{B}).$$

It is shown in Chapter 5 that this modification is a direct consequence of the following definition of Helmholtz free energy:

$$\psi=\mu((1-\gamma)(\operatorname{tr}\mathbb{B}-d-\ln\det\mathbb{B})+\tfrac{1}{2}\gamma|\mathbb{B}-\mathbb{I}|^2).$$

If we compare this to (1.9), the crucial difference is that for $\gamma>0$, the term $\frac{1}{2}|\mathbb{B}-\mathbb{I}|^2$ appears. It turns out that this term can be physically well justified and improves mathematical properties of the system (1.10)–(1.12) significantly, compared to the case $\gamma=0$.

[1] The subscript denotes the tangential part of a vector, i.e., $\boldsymbol{u}_\tau=\boldsymbol{u}-(\boldsymbol{u}\cdot\boldsymbol{n})\boldsymbol{n}=\boldsymbol{n}\times\boldsymbol{u}\times\boldsymbol{n}$.

To summarize: In both isothermal and non-isothermal models introduced above, the main obstacle for an successful analysis of these models are terms coming from the objective derivative, such as $\mathbb{B}\nabla \boldsymbol{v}$. In both cases, to overcome the lack of regularity to define $\mathbb{B}\nabla \boldsymbol{v}$, we use the (material) parameter γ. In the non-isothermal case, the coefficient γ improves the energy dissipation of the model. On the other hand, in the isothermal case, the parameter γ is used to improve the energy storage mechanism (far away from the rest state).

1.4 Our goal

We aim to prove that the system of equations (1.2)–(1.7) has a solution. Thus, we want to make clear that this thesis is concerned primarily with the mathematical analysis of the corresponding model and we do not venture too deep into physical discussions, nor we adhere to some particular physical notation. For this we refer the reader to classical works by Truesdell, Noll and Rajagopal [69], [70]. There are also many articles on the derivation of the models of viscoelastic fluids even in the non-isothermal case already. It seems however, that any kind of analysis of these models falls very much behind. This thesis should at least partially fill this gap. Nevertheless, some physical considerations are applied in the Chapter 2 to derive the studied model properly. This means that we want our model to obey all the natural physical and thermodynamical principles as there is no reason why not to do so. Also, it turns out that it is actually useful in the analysis to know the way in which the model was derived. In particular, in connection with viscoelastic fluids, it is advantageous to determine the Helmholtz free energy corresponding to the system. Indeed, this functional characterizes the way in which the fluid stores the mechanical energy.

The mathematical analysis in this work is solely the existence analysis, i.e., finding a suitable notion of a solution to our problem and proving that such a solution exists. This is a fundamental question which should be always resolved first, before any kind of further (numerical) analysis is executed (in practice, it is usually the other way around). An appropriate existence theory indicates what is the minimal regularity that we can expect from the solution. On the other hand, we do not attempt to answer the question of uniqueness of our solution as this problem is notoriously difficult already for the Navier-Stokes (or Euler) equations (see [72]). In fact, the recent result [10] for the Navier-Stokes equations implies that the uniqueness fails if the class of solutions where it is studied is too broad, broader than the Leray-Hopf class (cf. [48]). This suggests that one should rather aim to construct a weak solution satisfying as many physical principles as possible and hence, try to narrow down the class of uniqueness. This is precisely our philosophy in this work.

1.5 State-of-the-art existence results

Regarding the existence analysis of a viscoelastic fluid model including the full temperature evolution, there is an upcoming study [20]. There the authors show global and large-data existence of a weak solution to 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 author's best knowledge, there is no existence theory in a setting that would be of similar generality as considered in this thesis. Thus, here, for the first time, we provide an existence analysis for a viscoelastic fluid model that is coupled with a full evolution of temperature and that takes into account all components of the extra stress tensor. Moreover, the equation for the temperature that we consider is derived from fundamental thermodynamical laws (similarly as in [20], [40], [53]) and consequently, the heating originates from both the viscous and elastic forces. Also, we would like to point out that the majority of material coefficients of the model are allowed to be temperature dependent here. Although we place some restrictions on the growth of these coefficients, these are only asymptotic and therefore irrelevant from the physical point of view. Furthermore, the model considered here has the property that the evolution of the temperature can not 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 viscoelastic models, 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 some way. For example, in [49] 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 Oldroyd derivative (see [57], [55], [56], [63] or [64]). Then, the follow-up of this is [59], where the author claims to prove existence of a weak solution to FENE-P, Giesekus and PTT viscoelastic models. However, in these works it is only shown that certain defect measures of the non-linear terms are compact. Furthermore, in the scalar case, that is if $\mathbb{B} = b\mathbb{I}$, we refer to [19] (and [11] in the compressible case) for an analysis of such models. In the two-dimensional case, existence and regularity results can be found in [25]. An existence theory for related viscoelastic models (Peterlin class) was developed, e.g., in [50]. 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 [43], where the author obtains a global weak solution to an Oldroyd-like diffusive model under certain growth assumptions on the material coefficients. Furthermore, there are existence results for viscoelastic models involving various approximations that improve properties of the system, see e.g. [5] or [42]. Finally, the forthcoming article [6] (in Chapter 5) contains the existence theory for viscoelastic diffusive Oldroyd-B or Giesekus models. This result, however, relies on a certain physical correction of the energy storage mechanism away from the stress-free state and thus improving the a priori estimates of the system. There are also existence results that are of local nature or for small (initial) data. Again, we stress out that all these results concern only the isothermal case. Local-in-time existence of regular solutions to a viscoelastic Oldroyd-B model without diffusion was shown in [38]. 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. [3].

1.6 Basic notation

In this section, we introduce the basic notation which is used throughout the thesis, but might be regarded as non-standard.

We distinguish scalar, vector and tensor quantities by using different fonts: a for scalars, $\boldsymbol{a}$ for vectors and $\mathbb{A}$ for tensors (matrices). The symbol "$\otimes$" denotes an outer product while "$\cdot$" is used to denote an inner product. We do not use any additional notation for vector or tensor inner or outer products as this is clear from context. If there is any doubt, we use a general rule that the lower rank of the two objects in the inner product determines how many indices are contracted. For example, the product $(\mathbb{B}\otimes\boldsymbol{v})\cdot\nabla\mathbb{A}$ is understood as $\sum_{i,j,k}\mathbb{B}_{ij}\boldsymbol{v}_k\partial_k\mathbb{A}_{ij}$, while the product $\boldsymbol{v}\cdot\nabla\mathbb{B}$ means $\sum_i \boldsymbol{v}_i\partial_i\mathbb{B}$ and the product $\nabla\mathbb{B}\cdot\mathbb{A}$ translates as $\sum_{jk}\nabla\mathbb{B}_{jk}\mathbb{A}_{jk}$ (every rank-3 tensor which we encounter decomposes naturally in a vectorial and matrix part).

By $\mathbb{R}^{d\times d}_{\mathrm{sym}}$ we denote the set of real symmetric $d\times d$ matrices. Furthermore, by $\mathbb{R}^{d\times d}_{>0}$ we denote the subset of $\mathbb{R}^{d\times d}_{\mathrm{sym}}$ which consists of positive definite matrices, i.e., those matrices $\mathbb{A}$ satisfying

$$\mathbb{A}\boldsymbol{z}\cdot\boldsymbol{z} > 0 \quad \text{for all } \boldsymbol{z}\in\mathbb{R}^d\setminus\{0\}.$$

The symbol $|\mathbb{A}|$ is used to denote the Euclidian (or Frobenius) matrix norm, defined by

$$|\mathbb{A}| = \sqrt{\sum_{i=1}^{d}\sum_{j=1}^{d}(\mathbb{A})^2_{ij}}.$$

Such a particular choice (all matrix norms are equivalent since $\mathbb{R}^{d\times d}$ is a finite dimensional space) is convenient in our computations. Note also that this norm satisfies the frequently used sub-multiplicative property in the form

$$|\mathbb{A}\mathbb{B}| \le |\mathbb{A}||\mathbb{B}| \quad \text{for all } \mathbb{A},\mathbb{B}\in\mathbb{R}^{d\times d}$$

due to the Cauchy-Schwarz inequality, see [39, Sect. 5.6, 2nd example] for details. Further, by $\operatorname{tr}\mathbb{A}$, we denote the trace of $\mathbb{A}$, defined as

$$\operatorname{tr}\mathbb{A} = \sum_{i=1}^{d}(\mathbb{A})_{ii}$$

and $\det\mathbb{A}$ stands for the determinant of $\mathbb{A}$.

The symbols C, C_i, $i\in\mathbb{N}$, are systematically used in estimates to denote generic positive constants. Their value may change during computations, but whenever such a constant depends on some important quantity ω, we indicate it as $C(\omega)$. The symbol $\hookrightarrow$ always denotes a continuous embedding: If $X\hookrightarrow Y$ for some normed vector spaces X, Y, then this means precisely that $X\subset Y$ and there exists a constant $C>0$, such that

$$\|u\|_Y \le C\|u\|_X \quad \text{for all } u\in X.$$

Moreover, we write $X \overset{\mathrm{d}}{\hookrightarrow} Y$ if $X\hookrightarrow Y$ and X is dense in Y.

2. Physical meaning of the model

In this chapter, we derive system (1.2)–(1.7) from physical principles. The starting point are the balance laws of: mass, linear & angular momenta, energy and entropy. In fact, these laws can be recognized directly in (1.2), (1.3), (1.5) and (1.7) (see below). These general laws do not carry any information about properties of the material, which we are modelling. This is the point where the constitutive theory comes into the game. This theory brings certain restrictions on how the stress can depend on strain (or vice versa), how the temperature changes may induce the flux of heat and so on. In fact, whole equation (1.4) can be perceived as a constitutive relation. In a broader sense, one could also include the choice of boundary conditions into the constitutive theory, but we do not discuss this possibility as such a generality is not the primary goal of this work. Generally speaking, it is a difficult task to determine the right constitutive relations for a given material, especially for viscoelastic fluids. Though one may be able to simply fit some experimental data, this still might give no clue about characteristic properties of the studied material if it is very complex. On the other hand, if one tries to derive the constitutive laws from some additional information such as a fluid microstructure, then the transition back to a macrostructure is usually very difficult or impossible, again due to the complexity of the material. However, for us, these are not relevant issues since we are interested only in the resulting (macroscopic) viscoelastic models and their mathematical analysis. Thus, we actually want to derive as general class of viscoelastic models as possible, which contains the well established physical models as special cases. Then, the question is, what are the minimal requirements which every reasonable model should meet. This is the point, where the laws of thermodynamics can be applied. Namely, the second law of thermodynamics tells us that whatever happens in our isolated container filled with fluid, the total entropy can not decrease. This principle proves very useful in nailing down the right constitutive relations, but is not sufficient by itself. Returning to the one-dimensional analogues from Figure 1.1, one can realize that mechanical energy in a viscoelastic fluid is managed in two different ways: it can be either stored for later as a spring does, or it can be dissipated (and thus converted to different forms) as in a dashpot. Thus, these two mechanisms should be carefully specified: we do so by writing explicit formulas for the Helmholtz free energy ψ and for the production of entropy ξ, respectively. We refer to [63] and [64] for a detailed treatment and further ideas involved in this approach.

Strictly speaking, this chapter contains no new results as there are already many studies in which even more general models are derived (see, e.g., [27], [40] or [53]). Nevertheless, the computations made in this chapter should provide an useful insight into the analysis that is done in Chapter 3. Indeed, one can recognize quite easily the quantities that should be a priori under control (in our case this is the entropy production and the total energy). Moreover, one can also see very clearly which test functions should be used to draw basic information from the system. On the other hand, nothing in this chapter is really necessary to understand the analysis in Chapter 3, therefore it is possible to skip it.

First, by keeping ψ in an implicit form, we derive viscoelastic models that are considerably more general than the model studied in Chapters 1 and 3. The author hopes to provide an existence analysis for these models as well in forthcoming articles. Then we obtain the model (1.2)–(1.7) by an appropriate choice of ψ.

2.1 Balance equations

In Chapter 1 we stated the system (1.2)–(1.7) without explaining the physical meaning of the quantities involved. Therefore, let us first make an overview of all quantities that have clear physical interpretation:

$\boldsymbol{v}$	flow velocity,
p	pressure,
$\mathbb{T}$	Cauchy stress tensor,
$\mathbb{B}$	elastic stress tensor,
θ	temperature,
e	internal energy,
$\boldsymbol{j}_e$	flux of internal energy,
E	total energy (sum of e and $\frac{1}{2}\lvert\boldsymbol{v}\rvert^2$),
ψ	Helmholtz free energy,
η	entropy,
$\boldsymbol{j}_\eta$	entropy flux,
ξ	production of entropy,
μ	coefficient of shear modulus,
c_v	heat capacity,
ν	kinematic viscosity,
κ	thermal conductivity,
λ	stress diffusion coefficient,
$\boldsymbol{f}$	external body forces.

Note that we did not include the density ρ of the fluid in the list. Since we are considering only a homogeneous incompressible fluid, the density ρ is constant and thus we can renormalize the other parameters in a way that $\rho = 1$.

In this chapter, to simplify the notation, we omit the dependence of the material parameters on the temperature although it is always assumed. Furthermore, we denote the material time derivative by overset " • ", i.e., we set

$$\overset{\bullet}{u} = \frac{\mathrm{d}}{\mathrm{d}t} u = \partial_t u + \boldsymbol{v} \cdot \nabla u.$$

Then, the local balance equations take the form:

$$0 = \operatorname{div} \boldsymbol{v} \qquad \text{balance of mass,} \tag{2.1}$$

$$\overset{\bullet}{\boldsymbol{v}} = \operatorname{div} \mathbb{T} + \boldsymbol{f}, \quad \mathbb{T} = \mathbb{T}^T \qquad \text{balance of lin. \& ang. momenta,} \tag{2.2}$$

$$\overset{\bullet}{e} = \mathbb{T} \cdot \mathbb{D}\boldsymbol{v} - \operatorname{div} \boldsymbol{j}_e \qquad \text{balance of internal energy,} \tag{2.3}$$

$$\overset{\bullet}{E} = \operatorname{div}(\mathbb{T}\boldsymbol{v} - \boldsymbol{j}_e) + \boldsymbol{f} \cdot \boldsymbol{v} \qquad \text{balance of total energy,}$$

$$\overset{\bullet}{\eta} = \xi - \operatorname{div} \boldsymbol{j}_\eta \qquad \text{balance of entropy.} \tag{2.4}$$

We may notice that the system (2.1)–(2.4) contains some quantities that are not explicitly seen in the system (1.2)–(1.7), namely, these are $\mathbb{T}$, e, $\boldsymbol{j}_e$, $\boldsymbol{j}_\eta$. Vice versa, the system (2.1)–(2.4) does not explicitly mention $\mathbb{B}$ or p. To see the connection between these systems, we need to prescribe constitutive relations.

2.2 Constitutive relations

Let us start by specifying the Helmholtz free energy ψ, which is a function of θ and $\mathbb{B}$. In specific physical settings, the function ψ will depend only on certain quantities formed from $\mathbb{B}$ (due to material frame indifference and also due to symmetry of $\mathbb{B}$). For example, if the fluid is isotropic, then it is known that ψ is a function of θ and only of eigenvalues of $\mathbb{B}$, or, equivalently, there exists a function $\overline{\psi}$ such that $\psi = \overline{\psi}(\theta, \operatorname{tr}\mathbb{B}, \operatorname{tr}\mathbb{B}^2, \ldots, \operatorname{tr}\mathbb{B}^d)$. Nevertheless, this information is not needed in what follows and thus, we write simply $\psi = \psi(\theta, \mathbb{B})$. We suppose that $\psi = \psi(\theta, \mathbb{B}) : (0, \infty) \times \mathbb{R}^{d\times d}_{>0} \to \mathbb{R}$ is a twice differentiable function which, for all $\theta > 0$, satisfies:

$$
\begin{aligned}
\partial^2_{\theta\theta}\psi(\theta,\mathbb{B}) < 0 &\quad \text{for all } \mathbb{B} \in \mathbb{R}^{d\times d}_{>0}, && (2.5)\\
\partial^2_{\mathbb{B}\mathbb{B}}\psi(\theta,\mathbb{B})\mathbb{A}\cdot\mathbb{A} \geq 0 &\quad \text{for all } \mathbb{B} \in \mathbb{R}^{d\times d}_{>0} \text{ and every } \mathbb{A} \in \mathbb{R}^{d\times d}_{\mathrm{sym}}, && (2.6)\\
\partial_{\mathbb{B}}\psi(\theta,\mathbb{B})\cdot(\mathbb{B}-\mathbb{I}) \geq 0 &\quad \text{for all } \mathbb{B} \in \mathbb{R}^{d\times d}_{>0}, && (2.7)\\
\partial_{\mathbb{B}}\psi(\theta,\mathbb{B})\mathbb{B} - \mathbb{B}\partial_{\mathbb{B}}\psi(\theta,\mathbb{B}) = 0 &\quad \text{for all } \mathbb{B} \in \mathbb{R}^{d\times d}_{>0}, && (2.8)\\
\psi(\theta,\mathbb{B}) \to \infty &\quad \text{as } \det\mathbb{B} \to 0_+ \text{ or } |\mathbb{B}| \to \infty. && (2.9)
\end{aligned}
$$

The assumptions (2.5) and (2.6) concern strict concavity and convexity with respect to θ and $\mathbb{B}$, respectively. Then, properties (2.6) and (2.7) together imply that, for all $\theta > 0$, $\psi(\theta, \cdot)$ has a local minimum at $\mathbb{I}$. Indeed, since $\mathbb{R}^{d\times d}_{>0}$ is an open set, if we take any $\mathbb{A} \in \mathbb{R}^{d\times d}_{\mathrm{sym}}$, then we can choose $\varepsilon > 0$ so small that $\mathbb{I} + \varepsilon\mathbb{A} \in \mathbb{R}^{d\times d}_{>0}$. Then, on setting $\mathbb{B} = \mathbb{I} + \varepsilon\mathbb{A}$ in (2.7) and dividing by ε, we get

$$\partial_{\mathbb{B}}\psi(\theta, \mathbb{I} + \varepsilon\mathbb{A}) \cdot \mathbb{A} \geq 0.$$

Thus, using the continuity of $\partial_{\mathbb{B}}\psi(\theta, \cdot)$, we can take the limit $\varepsilon \to 0_+$ and obtain $\partial_{\mathbb{B}}\psi(\theta, \mathbb{I}) \cdot \mathbb{A} \geq 0$. But since $\mathbb{A} \in \mathbb{R}^{d\times d}_{\mathrm{sym}}$ was arbitrary, we obviously get

$$\partial_{\mathbb{B}}\psi(\theta, \mathbb{I}) = 0,$$

which together with (2.6) proves the claim. Thus, no mechanical energy is stored if $\mathbb{B} = \mathbb{I}$. Actually, the whole term

$$\partial_{\mathbb{B}}\psi(\theta, \mathbb{B}) \cdot (\mathbb{B} - \mathbb{I})$$

is precisely one of the entropy producing mechanisms, as we shall see below. Thus, from this point of view, the assumption (2.7) is natural and we also see that if $\mathbb{B} = \mathbb{I}$, there is no production of entropy due to the elastic effects. Next, the assumption (2.8) means simply that the matrices $\partial_{\mathbb{B}}\psi(\theta, \mathbb{B})$ and $\mathbb{B}$ commute. Finally, assumption (2.9) penalizes certain unphysical deformations such as a compression of the material to a point or an infinite expansion. The reasons why all these assumptions are needed become clearer throughout the thesis. Once we agree that the function ψ is known, we can define the entropy as

$$\eta = \eta(\theta, \mathbb{B}) = -\partial_\theta \psi(\theta, \mathbb{B}), \tag{2.10}$$

according to the first law of thermodynamics. Then, the internal energy e is defined by

$$e = e(\theta, \mathbb{B}) = \psi(\theta, \mathbb{B}) + \theta\eta(\theta, \mathbb{B}), \tag{2.11}$$

using another fundamental identity of thermodynamics. In the literature dealing with the derivation of temperature dependent viscoelastic models it is common to assign some particular function to ψ and only after that proceed with the derivation of the model. Definitely one of the reasons for this is that a particular form of ψ helps to identify an explicit form of the equation for the temperature, which, otherwise, is only implicitly encoded in (2.4) and (2.10) as

$$-\frac{\mathrm{d}}{\mathrm{d}t}(\partial_\theta \psi) = \xi - \operatorname{div} \boldsymbol{j}_\eta. \tag{2.12}$$

For us, this approach makes no sense since in the analysis we need to avoid the temperature equation anyway (due to its possibly ill-posed terms). Thus, it turns out that the particular form of the temperature equation has very little importance for the mathematical analysis of the model. Moreover, by keeping the general form of ψ, one obtains an entire class of models, where it is easy to trace the effect of ψ on the resulting system of equations.

We proceed with the constitutive equations for quantities other than ψ. We make the following choice:

$$\mathbb{T} = -\mathrm{p}\mathbb{I} + 2\nu\mathbb{D}\boldsymbol{v} + 2a\mu\mathbb{B}\partial_{\mathbb{B}}\psi, \tag{2.13}$$

$$0 = \overset{\circ}{\mathbb{B}} + \delta\gamma(\mathbb{B} - \mathbb{I}) - \operatorname{div}(\lambda\nabla\mathbb{B}), \tag{2.14}$$

$$\overset{\circ}{\mathbb{B}} = \overset{\bullet}{\mathbb{B}} - (\mathbb{W}\boldsymbol{v}\mathbb{B} - \mathbb{B}\mathbb{W}\boldsymbol{v}) - a(\mathbb{D}\boldsymbol{v}\mathbb{B} + \mathbb{B}\mathbb{D}\boldsymbol{v}), \tag{2.15}$$

$$\boldsymbol{j}_e = -\kappa\nabla\theta + \lambda(\theta\partial^2_{\theta\mathbb{B}}\psi - \partial_{\mathbb{B}}\psi) \cdot \nabla\mathbb{B}, \tag{2.16}$$

$$\boldsymbol{j}_\eta = -\kappa\nabla\ln\theta + \lambda\partial^2_{\theta\mathbb{B}}\psi \cdot \nabla\mathbb{B}. \tag{2.17}$$

Let us now comment on these relations. In (2.13) we see clearly how the Cauchy stress tensor decomposes into a "viscous" and an "elastic" part, as was mentioned in the introduction. The equation (2.14) is a variant of diffusive Oldroyd-B model with enhanced growth given by γ and with a generalized objective derivative $\overset{\circ}{\mathbb{B}}$ defined by (2.15). We recall that we do not place any restriction on $a \in \mathbb{R}$ and that

$$\overset{\circ}{\mathbb{B}} = \overset{\bullet}{\mathbb{B}} - \nabla\boldsymbol{v}\mathbb{B} - \mathbb{B}(\nabla\boldsymbol{v})^T \quad \text{if } a = 1,$$

which coincides with the upper convected Oldroyd derivative, that seems to have a special physical meaning, as mentioned before. The need for such a kind of derivative stems from the fact that the material time derivative $\overset{\bullet}{\mathbb{B}}$ alone is not an objective physical quantity. Here, by objectivity, we mean invariance with respect to a time dependent rotation (and translation) of an observer. While the quantities $\mathbb{B}$, $\mathbb{D}\boldsymbol{v}$ are objective, the quantity $\overset{\bullet}{\mathbb{B}}$ must be corrected by $\mathbb{W}\boldsymbol{v}\mathbb{B} - \mathbb{B}\mathbb{W}\boldsymbol{v}$, for example. Next, recent studies (such as [30]) suggest that the presence of the diffusion term in (2.14) is based on solid physical grounds. We remark that it appears in (2.14) in a favourable linear form, although it is theoretically possible to consider also non-linear dependencies. Finally, the equation (2.16) represents a kind of a generalized Fourier's law.

2.3 Fulfilment of the laws of thermodynamics

Note that in the previous section we did not specify ξ. The reason is that this quantity can now be computed from (2.1) and (2.13)–(2.17) using the basic thermodynamical identities (2.10) and (2.11). Indeed, taking the material time derivative of (2.11) and using (2.10), we obtain

$$\dot{e} = \partial_{\mathbb{B}}\psi \cdot \dot{\mathbb{B}} + \partial_\theta \psi \dot{\theta} + \dot{\theta}\eta + \theta\dot{\eta} = \partial_{\mathbb{B}}\psi \cdot \dot{\mathbb{B}} + \theta\dot{\eta}.$$

Then, we use the above identity in (2.3) and further use the identities (2.15), (2.4) to evaluate the material time derivatives of $\mathbb{B}$, η, and use the commutativity property (2.8) to deduce

$$\begin{aligned}
\mathbb{T} \cdot \mathbb{D}\boldsymbol{v} - \operatorname{div} \boldsymbol{j}_e &= \dot{e} = \dot{\mathbb{B}} \cdot \partial_{\mathbb{B}}\psi + \theta\dot{\eta} \\
&= \mathring{\mathbb{B}} \cdot \partial_{\mathbb{B}}\psi + ((\mathbb{W}\boldsymbol{v}\mathbb{B} - \mathbb{B}\mathbb{W}\boldsymbol{v}) + a(\mathbb{D}\boldsymbol{v}\mathbb{B} + \mathbb{B}\mathbb{D}\boldsymbol{v})) \cdot \partial_{\mathbb{B}}\psi + \theta\xi - \theta \operatorname{div} \boldsymbol{j}_\eta \\
&= \mathring{\mathbb{B}} \cdot \partial_{\mathbb{B}}\psi + \mathbb{W}\boldsymbol{v} \cdot (\partial_{\mathbb{B}}\psi\mathbb{B} - \mathbb{B}\partial_{\mathbb{B}}\psi) + a\mathbb{D}\boldsymbol{v} \cdot (\partial_{\mathbb{B}}\psi\mathbb{B} + \mathbb{B}\partial_{\mathbb{B}}\psi) + \theta\xi - \theta \operatorname{div} \boldsymbol{j}_\eta \\
&= \mathring{\mathbb{B}} \cdot \partial_{\mathbb{B}}\psi + 2a\mathbb{B}\partial_{\mathbb{B}}\psi \cdot \mathbb{D}\boldsymbol{v} - \theta \operatorname{div} \boldsymbol{j}_\eta + \theta\xi.
\end{aligned}$$

From this, we easily compute

$$\theta\xi = (\mathbb{T} - 2a\mathbb{B}\partial_{\mathbb{B}}\psi) \cdot \mathbb{D}\boldsymbol{v} - \mathring{\mathbb{B}} \cdot \partial_{\mathbb{B}}\psi + \operatorname{div}(\theta\boldsymbol{j}_\eta - \boldsymbol{j}_e) - \nabla\theta \cdot \boldsymbol{j}_\eta, \tag{2.18}$$

or alternatively

$$\xi = \frac{\mathbb{T} - 2a\mathbb{B}\partial_{\mathbb{B}}\psi}{\theta} \cdot \mathbb{D}\boldsymbol{v} - \mathring{\mathbb{B}} \cdot \frac{\partial_{\mathbb{B}}\psi}{\theta} + \operatorname{div}\left(\boldsymbol{j}_\eta - \frac{\boldsymbol{j}_e}{\theta}\right) - \nabla \ln\theta \cdot \frac{\boldsymbol{j}_e}{\theta}$$

after dividing by θ. Finally, in (2.18) we use the constitutive relations (2.13), (2.14), (2.16) and (2.17). This way, using also the constraint $0 = \operatorname{div} \boldsymbol{v} = \mathbb{I} \cdot \mathbb{D}\boldsymbol{v}$, we obtain

$$\begin{aligned}
\theta\xi &= 2\nu|\mathbb{D}\boldsymbol{v}|^2 + \delta\gamma(\mathbb{B} - \mathbb{I}) \cdot \partial_{\mathbb{B}}\psi - \operatorname{div}(\lambda\nabla\mathbb{B}) \cdot \partial_{\mathbb{B}}\psi \\
&\qquad + \operatorname{div}(\lambda\partial_{\mathbb{B}}\psi \cdot \nabla\mathbb{B}) + \kappa\theta|\nabla\ln\theta|^2 - \lambda\partial^2_{\theta\mathbb{B}}\psi \cdot (\nabla\theta \cdot \nabla\mathbb{B}) \\
&= 2\nu|\mathbb{D}\boldsymbol{v}|^2 + \kappa\theta|\nabla\ln\theta|^2 + \delta\gamma(\mathbb{B} - \mathbb{I}) \cdot \partial_{\mathbb{B}}\psi + \lambda\partial^2_{\mathbb{B}\mathbb{B}}\psi\nabla\mathbb{B} \cdot \nabla\mathbb{B}.
\end{aligned} \tag{2.19}$$

Now it is clear that the assumptions (2.6) and (2.7) are needed to get

$$\xi \geq 0, \tag{2.20}$$

which encodes local version of the second law of thermodynamics. Indeed, integration of (2.4) over Ω together with the boundary conditions

$$\boldsymbol{n} \cdot \nabla\theta = 0, \quad \text{and} \quad \boldsymbol{n} \cdot \nabla\mathbb{B} = 0$$

and (2.20) yield

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

Thus, in this section, we have verified that the fundamental laws of thermodynamics (2.11), (2.10) and (2.20) are satisfied provided that the participating physical quantities obey the constitutive relations (2.6), (2.7), (2.8) and (2.13)–(2.17).

Based on the method developed in [63], the computations above may serve as a justification of the selected constitutive relations. Indeed, this is best seen if we start the derivation again, but now we replace the information (2.13)–(2.17) by the single assumption

$$\theta\xi = 2\nu|\mathbb{D}\boldsymbol{v}|^2 + \kappa\theta|\nabla \ln\theta|^2 + \delta\gamma(\mathbb{B}-\mathbb{I})\cdot\partial_{\mathbb{B}}\psi + \lambda\partial^2_{\mathbb{B}\mathbb{B}}\psi\nabla\mathbb{B}\cdot\nabla\mathbb{B}. \tag{2.21}$$

With some effort, such an assumption can be justified: The first two terms correspond to an usual Navier-Stokes-Fourier fluid, the third one is common for all Oldroyd-B type models, while the last one generates the stress diffusion effect. We remark that it is also possible to interpret the stress diffusion as a consequence of a non-standard energy storage mechanism; however, this would yield a different class of models (see [53] for derivation and also [19] for a corresponding analysis in a simplified setting). Note that, so far, we only specified two scalar functions: ψ (though implicitly) and ξ. This should be, in principle, enough as ψ tells us how the fluid stores the energy, while ξ describes its dissipation. Then, we can again use the balance equations (2.1)–(2.4) to derive (2.18). Thus, we have two different equations for ξ. The idea now is to separate the independent mechanisms hidden in our system and to deduce the constitutive relations by comparing the equations for ξ. To this end, we first trace back the computation in (2.19) to find that (2.21) is equivalent to

$$\begin{aligned}\theta\xi &= 2\nu|\mathbb{D}\boldsymbol{v}|^2 + \delta\gamma(\mathbb{B}-\mathbb{I})\cdot\partial_{\mathbb{B}}\psi - \operatorname{div}(\lambda\nabla\mathbb{B})\cdot\partial_{\mathbb{B}}\psi + \operatorname{div}(\lambda\partial_{\mathbb{B}}\psi\cdot\nabla\mathbb{B})\\ &\qquad + \kappa\theta|\nabla\ln\theta|^2 - \lambda\partial^2_{\theta\mathbb{B}}\psi\cdot(\nabla\theta\cdot\nabla\mathbb{B})\\ &= 2\nu\mathbb{D}\boldsymbol{v}\cdot\mathbb{D}\boldsymbol{v} + (\delta\gamma(\mathbb{B}-\mathbb{I}) - \operatorname{div}(\lambda\nabla\mathbb{B}))\cdot\partial_{\mathbb{B}}\psi + \operatorname{div}(\lambda\partial_{\mathbb{B}}\psi\cdot\nabla\mathbb{B})\\ &\qquad + \nabla\theta\cdot(\kappa\nabla\ln\theta - \lambda\partial^2_{\theta\mathbb{B}}\psi\cdot\nabla\mathbb{B}).\end{aligned}$$

Now by comparison of this with (2.18), which is

$$\theta\xi = (\mathbb{T} - 2a\mathbb{B}\partial_{\mathbb{B}}\psi)\cdot\mathbb{D}\boldsymbol{v} - \overset{\circ}{\mathbb{B}}\cdot\partial_{\mathbb{B}}\psi + \operatorname{div}(\theta\boldsymbol{j}_\eta - \boldsymbol{j}_e) - \nabla\theta\cdot\boldsymbol{j}_\eta,$$

we can precisely read the constitutive relations (2.13)–(2.17). However, this final step is, of course, ambiguous and must be seen only as a motivation for the constitutive relations. First of all, one has to assume that the terms in ξ can really be separated so that they represent independent entropy producing mechanisms. Then, even if we assume this and obtain, for example, the identity

$$2\nu|\mathbb{D}\boldsymbol{v}|^2 = (\mathbb{T} - 2a\mathbb{B}\partial_{\mathbb{B}}\psi)\cdot\mathbb{D}\boldsymbol{v},$$

then this only implies that

$$\mathbb{T} = -\mathrm{p}\mathbb{I} + 2\nu\mathbb{D}\boldsymbol{v} + 2a\mathbb{B}\partial_{\mathbb{B}}\psi + \mathbb{A} - \frac{\mathbb{A}\cdot\mathbb{D}\boldsymbol{v}}{|\mathbb{D}\boldsymbol{v}|^2}\mathbb{D}\boldsymbol{v},$$

where $\mathbb{A}$ can be any matrix. Thus, such a derivation can be unambiguous only if we introduce further physical assumptions (such as the material frame indifference or the maximisation of the entropy production principle, see [64]).

Finally, we remark that this is, of course, not the only way how to derive (non-isothermal) viscoelastic models. For different or earlier approach, we refer to [47] or [71] and references therein.

2.4 Complete system of equations

Here we put together the constitutive assumptions with the balance laws and deduce the final appearance of the system of partial differential equations. We still keep the general form of ψ, which could be useful for future studies of similar models.

By simply plugging (2.13)–(2.17) into (2.1)–(2.4) and rearranging the terms, we obtain the system

$$\left.\begin{aligned}
&\operatorname{div}\boldsymbol{v} = 0,\\
&\partial_t\boldsymbol{v} + \boldsymbol{v}\cdot\nabla\boldsymbol{v} - \operatorname{div}(2\nu\mathbb{D}\boldsymbol{v}) + \nabla \mathrm{p} = \operatorname{div}(2a\mathbb{B}\partial_{\mathbb{B}}\psi) + \boldsymbol{f},\\
&\partial_t\mathbb{B} + \boldsymbol{v}\cdot\nabla\mathbb{B} + \delta\gamma(\mathbb{B}-\mathbb{I}) - \operatorname{div}(\lambda\nabla\mathbb{B})\\
&\qquad\qquad = \mathbb{W}\boldsymbol{v}\mathbb{B} - \mathbb{B}\mathbb{W}\boldsymbol{v} + a(\mathbb{D}\boldsymbol{v}\mathbb{B} + \mathbb{B}\mathbb{D}\boldsymbol{v}),\\
&\partial_t e + \boldsymbol{v}\cdot\nabla e - \operatorname{div}(\kappa\nabla\theta) + \operatorname{div}(\lambda(\theta\partial^2_{\theta\mathbb{B}}\psi - \partial_{\mathbb{B}}\psi)\cdot\nabla\mathbb{B})\\
&\qquad\qquad = 2\nu|\mathbb{D}\boldsymbol{v}|^2 + 2a\mathbb{B}\partial_{\mathbb{B}}\psi\cdot\mathbb{D}\boldsymbol{v},\\
&\partial_t E + \boldsymbol{v}\cdot\nabla E - \operatorname{div}(\kappa\nabla\theta) + \operatorname{div}(\lambda(\theta\partial^2_{\theta\mathbb{B}}\psi - \partial_{\mathbb{B}}\psi)\cdot\nabla\mathbb{B})\\
&\qquad\qquad = \operatorname{div}(-\mathrm{p}\boldsymbol{v} + 2\nu(\mathbb{D}\boldsymbol{v})\boldsymbol{v} + 2a(\mathbb{B}\partial_{\mathbb{B}}\psi)\boldsymbol{v}) + \boldsymbol{f}\cdot\boldsymbol{v},\\
&\partial_t\eta + \boldsymbol{v}\cdot\nabla\eta - \operatorname{div}(\kappa\nabla\ln\theta) + \operatorname{div}(\lambda\partial^2_{\theta\mathbb{B}}\psi\cdot\nabla\mathbb{B})\\
&\qquad = \frac{2\nu}{\theta}|\mathbb{D}\boldsymbol{v}|^2 + \kappa|\nabla\ln\theta|^2 + \frac{\delta\gamma}{\theta}(\mathbb{B}-\mathbb{I})\cdot\partial_{\mathbb{B}}\psi + \frac{\lambda}{\theta}\partial^2_{\mathbb{B}\mathbb{B}}\psi\nabla\mathbb{B}\cdot\nabla\mathbb{B},
\end{aligned}\right\}\tag{2.22}$$

where e, E, η are given by

$$\begin{aligned}
e &= \psi(\theta,\mathbb{B}) - \theta\partial_\theta\psi(\theta,\mathbb{B}),\\
E &= \tfrac{1}{2}|\boldsymbol{v}|^2 + c_v\theta,\\
\eta &= -\partial_\theta\psi(\theta,\mathbb{B})
\end{aligned}$$

and the functions $\boldsymbol{v}$, p, $\mathbb{B}$, θ are the unknowns of (2.22). An explicit equation for the temperature can be derived from (2.12) by using the chain rule and (2.14) to substitute for $\dot{\mathbb{B}}$.

Though it would be tempting to try to do an existence analysis for the system (2.22) with a general ψ satisfying (2.5)–(2.9), it is not done in this work. Instead, we choose a canonical representative for ψ, which makes the existence analysis of Chapter 3 more illuminating. This way, we do not have to refer to the assumptions (2.5)–(2.9) all the time as they are implicitly encoded in the system (2.22). Moreover, the assumptions (2.5)–(2.9) might not be sufficient for the existence theory, but they are definitely necessary, should the second law of thermodynamics hold.

A common choice for the Helmholtz free energy (see e.g. [53]) is

$$\psi(\theta,\mathbb{B}) = \psi_0(\theta) + \mu\psi_2(\mathbb{B}),\tag{2.23}$$

where $\mu > 0$ is a constant. In this setting, the coupling in the system (2.22) is very weak. Indeed, one can get the a priori estimates for $\boldsymbol{v}$ and $\mathbb{B}$ without using the equation for internal energy. Also, since $\partial^2_{\theta\mathbb{B}}\psi = 0$, we see that the entropy is just a function of temperature and also that the fourth term in the entropy equation vanishes. Consequently, the estimates for temperature are then rather standard and one can follow the classical theory for Navier-Stokes-Fourier system, compare also with [17]. Also, the model (2.23) is oversimplified from the point of view of physics since one would like to allow a non-constant coefficient μ depending on the temperature, i.e.,

$$\psi(\theta, \mathbb{B}) = \psi_0(\theta) + \mu(\theta)\psi_2(\mathbb{B}). \tag{2.24}$$

Then the analysis of (2.22) indeed becomes a challenging problem for several reasons, as can be seen in [20], where the authors treated such system under the additional hypothesis $\mathbb{B} = b\mathbb{I}$. One of the difficulties is that if one starts with construction of the approximate solutions using the equation for e (which seems to be the only way if we want to use a constructive approximative scheme), then at some point it is needed to invert the relation between the temperature and the internal energy (as the material coefficients are temperature dependent). To do so, one can introduce an approximation and make the dependence of e on θ linear near zero as it is done in [20]. Inspired by this idea, but not going into such technicalities now, let us make a compromise between (2.23) and (2.24) and set

$$\psi(\theta, \mathbb{B}) = \psi_0(\theta) + \mu\theta\psi_2(\mathbb{B}),$$

where $\mu > 0$ is a constant. We recall that such a model was considered in [65], for example. Further, we consider the common choice of the functions ψ_0, ψ_2 (see e.g. [27]). Thus, our final choice of ψ for analysis of the next chapter is

$$\psi(\theta, \mathbb{B}) = -c_v\theta(\ln\theta - 1) + \mu\theta(\operatorname{tr}\mathbb{B} - d - k\ln\det\mathbb{B}), \tag{2.25}$$

where, however, we set $k = 1$ to simplify the notation (as long as $k > 0$, this parameter does not affect the subsequent analysis). Note that in this case, we have

$$\eta(\theta, \mathbb{B}) = -\partial_\theta\psi(\theta, \mathbb{B}) = c_v\ln\theta - \mu(\operatorname{tr}\mathbb{B} - d - \ln\det\mathbb{B}),$$

and thus

$$e(\theta, \mathbb{B}) = \psi(\theta, \mathbb{B}) + \theta\eta(\theta, \mathbb{B}) = c_v\theta. \tag{2.26}$$

Also, we can easily verify that ψ satisfies (2.5)–(2.9). To this end, we apply (4.40) and observe that, for all $\theta > 0$ and $\mathbb{B} \in \mathbb{R}^{d\times d}_{>0}$, we have

$$\partial^2_{\theta\theta}\psi(\theta, \mathbb{B}) = -\frac{c_v}{\theta} < 0, \qquad \partial_{\mathbb{B}}\psi(\theta, \mathbb{B}) = \mu\theta(\mathbb{I} - \mathbb{B}^{-1})$$

and, using also (4.38), (4.39), we obtain

$$\begin{aligned}\partial^2_{\mathbb{B}\mathbb{B}}\psi(\theta, \mathbb{B})\mathbb{A} \cdot \mathbb{A} = -\mu\theta\partial_{\mathbb{B}}(\mathbb{B}^{-1})\mathbb{A} \cdot \mathbb{A} &= \mu\theta\mathbb{B}^{-1}\mathbb{A}\mathbb{B}^{-1} \cdot \mathbb{A} \\ &= \mu\theta|\mathbb{B}^{-\frac{1}{2}}\mathbb{A}\mathbb{B}^{-\frac{1}{2}}|^2 \geq 0.\end{aligned} \tag{2.27}$$

Finally, upon inserting (2.25), (2.26) into (2.22), redefining the pressure p by

$$\mathrm{p}_1 = \mathrm{p} + 2a\mu\theta,$$

noting that

$$\theta\partial^2_{\theta\mathbb{B}}\psi - \partial_{\mathbb{B}}\psi = 0,$$

using

$$
\begin{aligned}
(\mathbb{B}-\mathbb{I})\cdot\partial_{\mathbb{B}}\psi(\theta,\mathbb{B}) &= \mu\theta(\mathbb{B}-\mathbb{I})\cdot(\mathbb{I}-\mathbb{B}^{-1})\\
&= \mu\theta\mathbb{B}^{\frac{1}{2}}(\mathbb{B}^{\frac{1}{2}}-\mathbb{B}^{-\frac{1}{2}})\cdot\mathbb{B}^{-\frac{1}{2}}(\mathbb{B}^{\frac{1}{2}}-\mathbb{B}^{-\frac{1}{2}})\\
&= \mu\theta|\mathbb{B}^{\frac{1}{2}}-\mathbb{B}^{-\frac{1}{2}}|^2
\end{aligned}
$$

and also

$$
\partial^2_{\mathbb{B}\mathbb{B}}\psi(\theta,\mathbb{B})\nabla\mathbb{B}\cdot\nabla\mathbb{B} = \mu\theta\mathbb{B}^{-1}\nabla\mathbb{B}\mathbb{B}^{-1}\cdot\nabla\mathbb{B} = \mu\theta|\mathbb{B}^{-\frac{1}{2}}\nabla\mathbb{B}\mathbb{B}^{-\frac{1}{2}}|^2,
$$

we obtain (1.2)–(1.7) in the form

$$
\begin{aligned}
&\operatorname{div}\boldsymbol{v} = 0,\\
&\partial_t\boldsymbol{v} + \boldsymbol{v}\cdot\nabla\boldsymbol{v} - \operatorname{div}(2\nu\mathbb{D}\boldsymbol{v}) + \nabla \mathrm{p}_1 = \operatorname{div}(2a\mu\theta\mathbb{B}) + \boldsymbol{f},\\
&\partial_t\mathbb{B} + \boldsymbol{v}\cdot\nabla\mathbb{B} + \delta\gamma(\mathbb{B}-\mathbb{I}) - \operatorname{div}(\lambda\nabla\mathbb{B}) = \mathbb{W}\boldsymbol{v}\mathbb{B} - \mathbb{B}\mathbb{W}\boldsymbol{v} + a(\mathbb{D}\boldsymbol{v}\mathbb{B} + \mathbb{B}\mathbb{D}\boldsymbol{v}),\\
&c_v\partial_t\theta + c_v\boldsymbol{v}\cdot\nabla\theta - \operatorname{div}(\kappa\nabla\theta) = 2\nu|\mathbb{D}\boldsymbol{v}|^2 + 2a\mu\theta\mathbb{B}\cdot\mathbb{D}\boldsymbol{v},\\
&\partial_t E + \boldsymbol{v}\cdot\nabla E - \operatorname{div}(\kappa\nabla\theta) = \operatorname{div}(-\mathrm{p}\boldsymbol{v} + 2\nu(\mathbb{D}\boldsymbol{v})\boldsymbol{v} + 2a\mu\theta\mathbb{B}\boldsymbol{v}) + \boldsymbol{f}\cdot\boldsymbol{v},\\
&\partial_t\eta + \boldsymbol{v}\cdot\nabla\eta - \operatorname{div}(\kappa\nabla\ln\theta - \mu\lambda\nabla(\operatorname{tr}\mathbb{B} - d - \ln\det\mathbb{B}))\\
&\qquad\qquad = \tfrac{2\nu}{\theta}|\mathbb{D}\boldsymbol{v}|^2 + \kappa|\nabla\ln\theta|^2 + \mu\delta\gamma|\mathbb{B}^{\frac{1}{2}} - \mathbb{B}^{-\frac{1}{2}}|^2 + \mu\lambda|\mathbb{B}^{-\frac{1}{2}}\nabla\mathbb{B}\mathbb{B}^{-\frac{1}{2}}|^2.
\end{aligned}
$$

3. Mathematical theory

In this chapter, we investigate the system (1.2)–(1.7), that was introduced in Chapter 1, from the analytical point of view. In particular, we discuss several possibilities how to capture the temperature evolution and choose the one that is most convenient for an existence analysis. Then we define the notion of suitable weak solution and verify that such a definition is meaningful. After that, we state and prove our main result about existence of a suitable weak solution to (1.2)–(1.7). Finally, we also determine what additional conditions are needed for the fulfilment of local balances of total and internal energy.

3.1 Function spaces

First, let us introduce the function spaces used in the thesis and the corresponding notation. If not stated otherwise, the set $\Omega \subset \mathbb{R}^d$ is always an open bounded set with a Lipschitz boundary (in the sense of [60, Sect. 2.1.1]). By $L^p(\Omega)$ and $W^{n,p}(\Omega)$, $1 \le p \le \infty$, $n \in \mathbb{N}$, we denote the usual Lebesgue and Sobolev spaces, with their usual norms denoted as $\|\cdot\|_p$ and $\|\cdot\|_{n,p}$, respectively. In certain situations, we use the notation $\|\cdot\|_{p;\Omega}$ instead to clarify which domain is considered in the norm. Further, if $p > 1$, we set $W^{-n,p}(\Omega) = (W^{n,p'}(\Omega))^*$, where $p' = p/(p-1)$, $n \in \mathbb{N}$, and the star symbol "$*$" denotes the topological (continuous) dual space. These Banach spaces are separable if $1 \le p < \infty$ and reflexive if $1 < p < \infty$ (see e.g. [45, Ch. 2, 5] or [29, Corollary IV.8.2.]). Occasionally, we take advantage of the uniform convexity of the spaces $L^p(\Omega)$, $1 < p < \infty$, for this, we refer to [26, Ch. 3]. Every uniformly convex Banach space X has the following property:

If $\{u_k\}_{k=1}^{\infty} \subset X$ converges weakly to $u \in X$ and if $\|u_k\|_X \to \|u\|_X$, then $u_k \to u$ strongly in X. We refer to this property as to the Radon-Riesz property of X in the thesis.

We use the same notation for the function spaces of scalar-, vector-, or tensor-valued functions. We do not specify the meaning of the duality pairing $\langle \cdot, \cdot \rangle$, since it is always clear from the context. For certain subspaces of vector valued functions, we use the following notation:

$$
\begin{aligned}
\mathcal{D}(\Omega) &= \Big\{ \boldsymbol{w} \in \mathcal{C}^\infty(\Omega),\ \overline{\{\boldsymbol{w} \neq 0\}} \text{ is a compact set in } \Omega \Big\}, \\
\mathcal{D}_{\mathrm{div}} &= \{ \boldsymbol{w} \in \mathcal{D}(\Omega) : \operatorname{div} \boldsymbol{w} = 0 \text{ in } \Omega \}, \\
W^{1,p}_{0,\mathrm{div}} &= \overline{\mathcal{D}_{\mathrm{div}}}^{\|\cdot\|_{1,p}}, \quad W^{-1,p}_{0,\mathrm{div}} = \left(W^{1,p'}_{0,\mathrm{div}}(\Omega) \right)^*, \quad 1 < p < \infty.
\end{aligned}
$$

The standard inner product in $L^2(\Omega)$ is denoted as $(\cdot, \cdot)$.

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

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

is finite. If $X = L^q(\Omega)$ or $X = W^{k,q}(\Omega)$, with $1 \le q \le \infty$, we shorten the notation and use the symbols $\|\cdot\|_{L^pL^q}$ or $\|\cdot\|_{L^pW^{k,q}}$, respectively, for corresponding norms.

The space $\mathcal{C}([0,T];X)$ contains continuous X-valued functions on $[0,T]$, i.e., such mappings $u : [0,T] \to X$, for which

$$\lim_{t\to t_0} \|u(t) - u(t_0)\|_X = 0, \quad \text{for any } t_0 \in [0,T].$$

This space is equipped with the norm

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

Furthermore, the space $\mathcal{C}_w([0,T];X) \subset L^\infty(0,T;X)$ denotes a space of weakly continuous functions on $[0,T]$, i.e., for every $u \in \mathcal{C}_w([0,T];X)$ and every $g \in X^*$ there holds

$$\lim_{t\to t_0} \langle u(t), g\rangle = \langle u(t_0), g\rangle \quad \text{for any } t_0 \in [0,T].$$

We use certain Bochner spaces also for weakly differentiable mappings from $[0,T] \to X$. For $u \in L^1(0,T;X)$, the default meaning of the symbols $u' = \frac{\mathrm{d}}{\mathrm{d}t}u = \partial_t u$ is always the distributional derivative of u that coincides with an integrable function. In other words, if there exists $w \in L^1(0,T;Y)$, where Y is a Banach space satisfying $X \hookrightarrow Y$, such that

$$\int_0^T w\varphi = -\int_0^T u\partial_t\varphi \quad \text{for all } \varphi \in \mathcal{D}(0,T), \tag{3.1}$$

then we say that $\partial_t u = w$. In case that $Y = X^*$, it is assumed that X admits the Gelfand triplet structure, i.e., that X is separable, reflexive and there exists a separable Hilbert space H such that $X \overset{\mathrm{d}}{\hookrightarrow} H$. Then, any element $u \in X$ belongs to H and thus defines a continuous linear functional f_u on H by virtue of the isomorphism

$$\Phi : H \to H^*, \quad \Phi(u) \mapsto f_u = (u,\cdot)_H.$$

Moreover, using continuity and density of the embedding $X \overset{\mathrm{d}}{\hookrightarrow} H$, one can show that

$$X \overset{\mathrm{d}}{\hookrightarrow} H = H^* \overset{\mathrm{d}}{\hookrightarrow} X^*$$

and that the map $\Psi : X \to X^*$, $u \mapsto f_u$ is continuous and injective. In this sense we interpret the embedding $X \overset{\mathrm{d}}{\hookrightarrow} X^*$. Then, we define the space

$$W^{1,p}(0,T;X) = \Big\{u \in L^p(0,T;X);\ \partial_t u \in L^{p'}(0,T;X^*)\Big\} \tag{3.2}$$

and equip it with the norm

$$\|u\|_{W^{1,p}X} = \|u\|_{L^pX} + \|\partial_t u\|_{L^{p'}X^*}.$$

We recall that, for any $u \in W^{1,p}(0,T;X)$, identity (3.1) can be rewritten (using only the classical Lebesgue integral) as

$$\int_0^T \langle \partial_t u, g\rangle\varphi = -\int_0^T (u,g)_H \partial_t\varphi \quad \text{for all } \varphi \in \mathcal{D}(0,T) \quad \text{and every } g \in X,$$

see [74, Proposition 23.20]. Moreover, we also define the space $\mathcal{C}^1([0,T];X)$ as the space of functions u such that $u, \partial_t u \in \mathcal{C}([0,T];X)$, where $\partial_t u$ now coincides with the classical derivative of u. For more details regarding the definition of $W^{1,p}(0,T;X)$, we refer to [74, Ch. 23].

Suppose that X is a separable reflexive Banach space. The spaces $L^p(0,T;X)$, $W^{1,p}(0,T;X)$ are separable if $1 \le p < \infty$ and reflexive if $1 < p < \infty$, see [74, Propositions 23.2, 23.7]. These properties are used in Lemma 4.1, which is then applied several times in the proof of our main result Theorem 3.2 below. Moreover, for $1 < p < \infty$ and X separable, reflexive, we can make the identification

$$(L^p(0,T;X))^* = L^{p'}(0,T;X^*), \tag{3.3}$$

see [74, Convention 23.8].

3.2 System of equations and its variants

We recall that $T > 0$ and Ω is a domain in $\mathbb{R}^d$ with a Lipschitz boundary $\partial\Omega$. Then, we set $Q = (0,T) \times \Omega$. Assume that ν, γ, κ, δ, λ are real functions that are continuous, positive, and with an appropriate growth near 0 and ∞ (specified in the next section). Finally, let $\boldsymbol{f}$, $\boldsymbol{v}_0$, $\mathbb{B}_0$ and θ_0 be some appropriate data.

In the best case scenario, we would like to find a sufficiently regular triple $(\boldsymbol{v}, \mathbb{B}, \theta) : Q \to \mathbb{R}^d \times \mathbb{R}^{d\times d}_{>0} \times (0,\infty)$ that solves the system

$$\operatorname{div} \boldsymbol{v} = 0, \tag{3.4}$$

$$\partial_t \boldsymbol{v} + \boldsymbol{v} \cdot \nabla \boldsymbol{v} - \operatorname{div}(2\nu(\theta)\mathbb{D}\boldsymbol{v}) + \nabla \mathrm{p} = 2a\mu \operatorname{div}(\theta\mathbb{B}) + \boldsymbol{f}, \tag{3.5}$$

$$\begin{aligned}\partial_t \mathbb{B} + \boldsymbol{v} \cdot \nabla \mathbb{B} + \delta(\theta)\gamma(|\mathbb{B} - \mathbb{I}|)(\mathbb{B} - \mathbb{I}) - \operatorname{div}(\lambda(\theta)\nabla\mathbb{B}) \\ = \mathbb{W}\boldsymbol{v}\mathbb{B} - \mathbb{B}\mathbb{W}\boldsymbol{v} + a(\mathbb{D}\boldsymbol{v}\mathbb{B} + \mathbb{B}\mathbb{D}\boldsymbol{v}),\end{aligned} \tag{3.6}$$

$$c_v\partial_t\theta + c_v\boldsymbol{v} \cdot \nabla\theta - \operatorname{div}(\kappa(\theta)\nabla\theta) = 2\nu(\theta)|\mathbb{D}\boldsymbol{v}|^2 + 2a\mu\,\theta\mathbb{B} \cdot \mathbb{D}\boldsymbol{v} \tag{3.7}$$

everywhere in $Q = (0,T) \times \Omega$, $T > 0$, fulfils the boundary conditions

$$\boldsymbol{v} = 0, \quad \boldsymbol{n} \cdot \nabla\theta = 0, \quad \boldsymbol{n} \cdot \nabla\mathbb{B} = 0 \quad \text{on } (0,T) \times \partial\Omega \tag{3.8}$$

and satisfies the initial conditions

$$\boldsymbol{v}(0) = \boldsymbol{v}_0, \quad \mathbb{B}(0) = \mathbb{B}_0, \quad \theta(0) = \theta_0 \quad \text{in } \Omega.$$

Moreover, we require that such a solution satisfies basic physical principles, namely the conservation of total energy and the second law of thermodynamics (at least in some weakened sense). As long as $(\boldsymbol{v}, \mathbb{B}, \theta)$ are so smooth that every term in (3.4)–(3.7) is well defined and that (3.5) can be tested by $\boldsymbol{v}$, one can derive the local form of the total energy balance. Indeed, if we take the scalar product of (3.5) with $\boldsymbol{v}$ and add the result to (3.7), we obtain

$$\partial_t E + \boldsymbol{v} \cdot \nabla E - \operatorname{div}(\kappa(\theta)\nabla\theta) = \operatorname{div}(-\mathrm{p}\boldsymbol{v} + 2\nu(\theta)(\mathbb{D}\boldsymbol{v})\boldsymbol{v} + 2a\mu\theta\mathbb{B}\boldsymbol{v}) + \boldsymbol{f} \cdot \boldsymbol{v}. \tag{3.9}$$

However, unless in some rather unrealistic setting, it is not known whether the multiplication of (3.5) by $\boldsymbol{v}$ can be justified rigorously due to low regularity of $\boldsymbol{v}$, as can be seen already in the case of Navier-Stokes equations in three dimensions. This issue can be overcome by enforcing validity of (3.9), i.e., simply by replacing (3.7) with (3.9). This has several positive side effects. Firstly, the term $2\nu(\theta)|\mathbb{D}\boldsymbol{v}|^2$ (usually only an integrable quantity) is replaced by the term $2\nu(\theta)(\mathbb{D}\boldsymbol{v})\boldsymbol{v}$, whose integrability is always better. On the other hand, it is then necessary to take care of the pressure, which appears explicitly in (3.9). This procedure was fully exploited in [12] in the case of the Navier-Stokes-Fourier system with temperature, pressure and shear rate dependent coefficients. The second benefit of (3.9) over (3.7) applies only in our case as it concerns the additional term $2a\mu\theta\mathbb{B}\cdot\mathbb{D}\boldsymbol{v}$. We shall see later in this chapter that it may be impossible to give a meaning to this term on the level of weak solutions (depending on the growth of κ, γ and on $\mathbb{B}_0$). This is also the term which is very often omitted in the "naive" approach to the temperature dependent viscoelastic models (see the discussion in [40] and references therein). But we have seen in Chapter 2 that such a simplification is physically incorrect. Thus, the equation (3.9) really seems as a suitable replacement for (3.7).

However, our solution concept defined below is so weak that even the term $2a\mu\theta\mathbb{B}\boldsymbol{v}$ (or $\kappa(\theta)\nabla\theta$) may not be integrable. Thus, in general, we need to avoid (3.9) in its local form (as well as (3.7)). Instead, we only require that the total energy of the whole fluid is conserved, i.e., we integrate (3.9) over Ω and, using the integration by parts and boundary conditions (3.8), it reduces to

$$\frac{\mathrm{d}}{\mathrm{d}t}\int_\Omega E = \int_\Omega \boldsymbol{f}\cdot\boldsymbol{v}. \tag{3.10}$$

The equation (3.10) always makes sense for our weak solution and it plays an important role in the a priori estimates.

Since (3.10) is only an ordinary differential equation in time, we clearly need to supplement our system with another information. This is the point where the laws of thermodynamics come into play. In our case, these can be expressed by

$$\eta = c_v \ln\theta - \mu(\operatorname{tr}\mathbb{B} - d - \ln\det\mathbb{B}), \tag{3.11}$$

$$\begin{aligned}\xi = \frac{2\nu(\theta)}{\theta}|\mathbb{D}\boldsymbol{v}|^2 + \kappa(\theta)|\nabla\ln\theta|^2 + \mu\delta(\theta)\gamma(|\mathbb{B}-\mathbb{I}|)|\mathbb{B}^{\frac12} - \mathbb{B}^{-\frac12}|^2 \\ + \mu\lambda(\theta)|\mathbb{B}^{-\frac12}\nabla\mathbb{B}\mathbb{B}^{-\frac12}|^2 \geq 0,\end{aligned} \tag{3.12}$$

where (3.11) is the definition of entropy in our model, while (3.12) expresses the fact that the entropy of a closed system never decreases over time. These quantities are related by the balance equation

$$\partial_t\eta + \boldsymbol{v}\cdot\nabla\eta - \operatorname{div}(\kappa(\theta)\nabla\ln\theta - \mu\lambda(\theta)\nabla(\operatorname{tr}\mathbb{B} - d - \ln\det\mathbb{B})) = \xi, \tag{3.13}$$

which is added to our system and which can be derived from (3.6) and (3.7) as follows. First, we multiply (3.6) by $\mu(\mathbb{I}-\mathbb{B}^{-1})$. To this end let us denote

$$\psi_e(\mathbb{B}) = \mu(\operatorname{tr}\mathbb{B} - d - \ln\det\mathbb{B})$$

and compute the resulting terms separately. By (4.40) from Chapter 4, we have

$$\partial\psi_e(\mathbb{B}) = \mu(\mathbb{I}-\mathbb{B}^{-1}),$$

hence we can rewrite the first two terms as

$$\mu(\mathbb{I}-\mathbb{B}^{-1})\cdot(\partial_t\mathbb{B} + \boldsymbol{v}\cdot\nabla\mathbb{B}) = \partial_t\psi_e + \boldsymbol{v}\cdot\nabla\psi_e. \tag{3.14}$$

Next, relying on (4.34) and positive definiteness of $\mathbb{B}$, we may write that

$$\begin{aligned}\mu(\mathbb{I}-\mathbb{B}^{-1})\cdot\delta(\theta)\gamma(|\mathbb{B}-\mathbb{I}|)(\mathbb{B}-\mathbb{I})&\\ &=\mu\delta(\theta)\gamma(|\mathbb{B}-\mathbb{I}|)\mathbb{B}^{-\frac{1}{2}}(\mathbb{B}^{\frac{1}{2}}-\mathbb{B}^{-\frac{1}{2}})\cdot\mathbb{B}^{\frac{1}{2}}(\mathbb{B}^{\frac{1}{2}}-\mathbb{B}^{-\frac{1}{2}})\\ &=\mu\delta(\theta)\gamma(|\mathbb{B}-\mathbb{I}|)|\mathbb{B}^{\frac{1}{2}}-\mathbb{B}^{-\frac{1}{2}}|^2.\end{aligned}$$

Further, utilizing (4.39), we get

$$\begin{aligned}-\mu(\mathbb{I}-\mathbb{B}^{-1})\cdot\operatorname{div}(\lambda(\theta)\nabla\mathbb{B})=&\\ &-\operatorname{div}(\mu\lambda(\theta)(\mathbb{I}-\mathbb{B}^{-1})\cdot\nabla\mathbb{B})-\mu\lambda(\theta)\nabla\mathbb{B}^{-1}\cdot\nabla\mathbb{B}\\ &-\operatorname{div}(\lambda(\theta)\nabla\psi_e)+\mu\lambda(\theta)|\mathbb{B}^{-\frac{1}{2}}\nabla\mathbb{B}\mathbb{B}^{-\frac{1}{2}}|^2.\end{aligned}$$

Finally, using (4.34) and (3.4), we deduce that

$$\begin{aligned}\mu(\mathbb{I}-\mathbb{B}^{-1})\cdot(\mathbb{W}\boldsymbol{v}\mathbb{B}-\mathbb{B}\mathbb{W}\boldsymbol{v}+a(\mathbb{D}\boldsymbol{v}\mathbb{B}+\mathbb{B}\mathbb{D}\boldsymbol{v}))&\\ &=\mu\mathbb{I}\cdot(\mathbb{W}\boldsymbol{v}\mathbb{B}-\mathbb{B}\mathbb{W}\boldsymbol{v}+a(\mathbb{D}\boldsymbol{v}\mathbb{B}+\mathbb{B}\mathbb{D}\boldsymbol{v}))\\ &=a\mu\mathbb{I}\cdot(\mathbb{D}\boldsymbol{v}\mathbb{B}+\mathbb{B}\mathbb{D}\boldsymbol{v})\\ &=2a\mu\mathbb{B}\cdot\mathbb{D}\boldsymbol{v}.\end{aligned}\tag{3.15}$$

Therefore, applying (3.14)–(3.15) in (3.6) multiplied by $\mu(\mathbb{I}-\mathbb{B}^{-1})$, we obtain

$$\begin{aligned}&\partial_t\psi_e(\mathbb{B})+\boldsymbol{v}\cdot\nabla\psi_e(\mathbb{B})-\operatorname{div}(\lambda(\theta)\nabla\psi_e(\mathbb{B}))\\ &\quad=-\mu\delta(\theta)\gamma(|\mathbb{B}-\mathbb{I}|)|\mathbb{B}^{\frac{1}{2}}-\mathbb{B}^{-\frac{1}{2}}|^2-\mu\lambda(\theta)|\mathbb{B}^{-\frac{1}{2}}\nabla\mathbb{B}\mathbb{B}^{-\frac{1}{2}}|^2+2a\mu\mathbb{B}\cdot\mathbb{D}\boldsymbol{v}.\end{aligned}\tag{3.16}$$

Secondly, we multiply (3.6) by $\frac{1}{\theta}$ to get

$$\begin{aligned}&c_v\partial_t\ln\theta+c_v\boldsymbol{v}\cdot\nabla\ln\theta-\operatorname{div}(\kappa(\theta)\nabla\ln\theta)\\ &\qquad=\kappa(\theta)|\nabla\ln\theta|^2+\frac{2\nu(\theta)}{\theta}|\mathbb{D}\boldsymbol{v}|^2+2a\mu\mathbb{B}\cdot\mathbb{D}\boldsymbol{v}.\end{aligned}\tag{3.17}$$

If we subtract (3.16) from (3.17), we obtain precisely the entropy equality (3.13).

Note that, from the point of view of analysis, the equation (3.13) should really be preferred over (3.7) or (3.9) since the terms $2a\mu\theta\mathbb{B}\cdot\mathbb{D}\boldsymbol{v}$, $2a\mu\theta\mathbb{B}\boldsymbol{v}$ disappeared and, most importantly, every term in (3.13) is either a derivative of something, or non-negative. Thus, by integrating (3.13) over Ω, we find

$$\frac{\mathrm{d}}{\mathrm{d}t}\int_\Omega\eta=\int_\Omega\xi\geq 0,\tag{3.18}$$

which is the second law of thermodynamics. This relation is the cornerstone of our a priori estimates since it yields integrability of ξ.

Here it is interesting to make another comparison with the work [12]. There, an analogous version of entropy equality (actually an inequality, see below) is derived merely as an additional property of the constructed weak solution and is not needed for its existence. The key uniform estimates are deduced by testing some approximating equations for $\boldsymbol{v}$ and θ separately by $\boldsymbol{v}$ and θ^λ, $\lambda\in(-1,0)$. On the other hand, in our case, testing (3.6) by $\boldsymbol{v}$ gives us nothing as the right hand side now contains $\mathbb{B}$. Instead, we draw most of the information from (3.18) (3.10) and from (3.7) (in its approximated form). Thus, unlike in [12], the fulfilment of the entropy inequality is absolutely crucial in our analysis. A similar remark actually applies also for analysis of "scalar" viscoelastic models as in [19], [11] or [20]. There one obtains uniform estimates on $\nabla \boldsymbol{v}$ simply by testing the momentum equation with $\boldsymbol{v}$ and using the Young inequality to absorb the term $\boldsymbol{f} \cdot \boldsymbol{v}$. This is not possible in our situation as we need first to add the equation for internal energy (to eliminate the terms with $\mathbb{B}$), which then annihilates the viscous dissipation term and leads only to (3.10). We, on the other hand derive the uniform estimates on $\nabla \boldsymbol{v}$ from the entropy inequality and from the approximated temperature equation. These issues suggest that the fundamental information about the solution stems from the entropy inequality rather than from the energy balances.

Another major difference compared to the theory of Navier-Stokes-Fourier fluids can be seen in (3.7). If $\mathbb{B} = \mathbb{I}$, then the term $2a\mu\theta\mathbb{B} \cdot \mathbb{D}\boldsymbol{v}$ vanishes and the right hand side of (3.7) is positive. Thus, one obtains a minimum principle for the temperature and the analysis in [12] is very much built on this fact. However, in our case, the term $2a\mu\theta\mathbb{B} \cdot \mathbb{D}\boldsymbol{v}$ does not have a sign, nor can it be estimated and hence it is impossible to prove a minimum principle for the temperature. Instead, we again have to rely solely on the entropy inequality to show that $\theta > 0$ (and that $\mathbb{B}$ is positive definite) almost everywhere in Q.

The reason why we speak about entropy inequality (instead of equality) is that ξ is a priori only an integrable quantity. Hence, when constructing a solution as a weak limit of some approximations, it is not clear if one can pass to the limit in (3.13) without using measures. Therefore we take advantage of $\xi \geq 0$ and impose (3.13) only with inequality sign (relying on the Fatou lemma or weak lower semi-continuity in the limiting processes). Thus, we actually construct a weak solution whose dissipation is at least ξ. Hereby we are admitting that there might be further entropy producing mechanisms that we do not know. For similar reasons, (3.10) is going to be only an inequality as well. On the physical side, this means that the energy can not be spontaneously created from nothing, but it is allowed to transmute into some forms that are not modelled. A precise nature of this transformation of energy is not yet known even in the much simpler case of (three-dimensional) Navier-Stokes equations, even though a formula for it was found in [28] (actually, this issue is intimately connected with the open problem of regularity of solution to Navier-Stokes equations).

The importance of the entropy inequality in the mathematical analysis of fluids has been observed in several other works treating fairly complex fluids. For example, we can mention its use in the theory of compressible Navier-Stokes-Fourier equations (see [32] or [33]). There, the so-called relative entropy/energy inequalities play a key role in proving, e.g., the weak-strong uniqueness property of solutions. Furthermore, in [18], the entropy inequality is a crucial tool for proving existence of a weak solution to a model describing an unsteady flow of an incompressible heat-conducting mixture of several fluids.

An interesting comparison can be made with the theory in Chapter 5, where the isothermal case of a similar viscoelastic model is studied. There, the analysis relies on the uniform control of the quantity $\frac{\mathrm{d}}{\mathrm{d}t}(\frac{1}{2}|\boldsymbol{v}|^2 + \psi)$. If we go back to (2.11), take the material time derivative and use the balance equations (2.2), (2.3), (2.4), we (eventually) obtain

$$\frac{\mathrm{d}}{\mathrm{d}t}(\tfrac{1}{2}|\boldsymbol{v}|^2+\psi)+\theta\xi+\overset{\bullet}{\theta}\eta+\nabla\theta\cdot\boldsymbol{j}_\eta=\operatorname{div}(\mathbb{T}\boldsymbol{v}-\boldsymbol{j}_e+\theta\boldsymbol{j}_\eta)+\boldsymbol{f}\cdot\boldsymbol{v}. \tag{3.19}$$

From this we clearly see that in the isothermal case one obtains, after integration over Ω and application of boundary conditions, that

$$\frac{\mathrm{d}}{\mathrm{d}t}\int_\Omega(\tfrac{1}{2}|\boldsymbol{v}|^2+\psi)+\int_\Omega\theta\xi=\int_\Omega\boldsymbol{f}\cdot\boldsymbol{v}. \tag{3.20}$$

As $\xi\geq 0$, this is a very powerful identity, which yields all the necessary a priori estimates. This approach completely fails in our case, as we can not eliminate the term $\overset{\bullet}{\theta}\eta$ in (3.19). Although we can use that

$$\overset{\bullet}{\psi}+\overset{\bullet}{\theta}\eta=\partial_\theta\psi\overset{\bullet}{\theta}+\partial_{\mathbb{B}}\psi\cdot\overset{\bullet}{\mathbb{B}}-\overset{\bullet}{\theta}\partial_\theta\psi=\partial_{\mathbb{B}}\psi\cdot\overset{\bullet}{\mathbb{B}},$$

this does not help either, since the right hand side is not in the form of a total derivative (since ψ now depends also on θ). Intuitively, this phenomenon can be understood as follows. In (3.20), the term $\theta\xi$ represents energy which is dissipated into some unknown quantities that are not modelled in the isothermal case and thus have no influence on $\boldsymbol{v}$ and $\mathbb{B}$. On the other hand, in (3.19), we clearly see that this dissipation has an immediate impact on the evolution of temperature, which in turn affects the evolution of $\boldsymbol{v}$ and $\mathbb{B}$ via corresponding changes in the material coefficients. Therefore, the identity (3.19) is merely a tautology and gives us no information at all. Finally, in relation with the analysis in [6], we remark that modifying the Helmholtz free energy with the term $\frac{1}{2}|\mathbb{B}-\mathbb{I}|^2$ is possible in our situation as well. However, the major drawback of this approach in our setting is that (3.5) newly includes the term $2a\mu\operatorname{div}(\theta\mathbb{B}^2)$, which has very poor regularity due to the presence of θ (unlike in [6]).

Now we state the final form of the model (3.4)–(3.7), which we consider in the subsequent analysis. For the reasons explained above, instead of (3.4)–(3.7), we actually study the system

$$\operatorname{div}\boldsymbol{v}=0, \tag{3.21}$$

$$\partial_t\boldsymbol{v}+\boldsymbol{v}\cdot\nabla\boldsymbol{v}-\operatorname{div}(2\nu(\theta)\mathbb{D}\boldsymbol{v})+\nabla\mathrm{p}=2a\mu\operatorname{div}(\theta\,\mathbb{B})+\boldsymbol{f}, \tag{3.22}$$

$$\begin{aligned}\partial_t\mathbb{B}+\boldsymbol{v}\cdot\nabla\mathbb{B}+\delta(\theta)\gamma(|\mathbb{B}-\mathbb{I}|)(\mathbb{B}-\mathbb{I})-\operatorname{div}(\lambda(\theta)\nabla\mathbb{B})&\\ =\mathbb{W}\boldsymbol{v}\mathbb{B}-\mathbb{B}\mathbb{W}\boldsymbol{v}+a(\mathbb{D}\boldsymbol{v}\mathbb{B}+\mathbb{B}\mathbb{D}\boldsymbol{v}),&\end{aligned} \tag{3.23}$$

$$\partial_t\eta+\boldsymbol{v}\cdot\nabla\eta-\operatorname{div}(\kappa(\theta)\nabla\ln\theta-\lambda(\theta)\mu\nabla(\operatorname{tr}\mathbb{B}-d-\ln\det\mathbb{B}))\geq\xi, \tag{3.24}$$

$$\begin{aligned}\eta&=c_v\ln\theta-\mu(\operatorname{tr}\mathbb{B}-d-\ln\det\mathbb{B}),\\ \xi&=\frac{2\nu}{\theta}|\mathbb{D}\boldsymbol{v}|^2+\kappa(\theta)|\nabla\ln\theta|^2+\delta(\theta)\gamma(|\mathbb{B}-\mathbb{I}|)|\mathbb{B}^{\frac{1}{2}}-\mathbb{B}^{-\frac{1}{2}}|^2\\ &\qquad+\mu\lambda(\theta)|\mathbb{B}^{-\frac{1}{2}}\nabla\mathbb{B}\mathbb{B}^{-\frac{1}{2}}|^2\end{aligned} \tag{3.25}$$

in Q together with the inequality

$$\frac{\mathrm{d}}{\mathrm{d}t}\int_\Omega(\tfrac{1}{2}|\boldsymbol{v}|^2+c_v\theta)\leq\int_\Omega\boldsymbol{f}\cdot\boldsymbol{v},\qquad\text{in }(0,T), \tag{3.26}$$

with boundary conditions

$$\boldsymbol{v}=0,\quad\boldsymbol{n}\cdot\nabla\theta=0,\quad\boldsymbol{n}\cdot\nabla\mathbb{B}=0\qquad\text{on }(0,T)\times\partial\Omega$$

and with initial conditions

$$\boldsymbol{v}(0) = \boldsymbol{v}_0, \quad \mathbb{B}(0) = \mathbb{B}_0, \quad \theta(0) = \theta_0 \qquad \text{in } \Omega. \tag{3.27}$$

Although the system (3.21)–(3.25) is stated in the classical way (for better readability), our existence analysis concerns only its weak formulation, of course. This seems to be more natural even from the physical point of view, as the primary form of balance laws is global, i.e., integrated over some volume of the fluid (cf. [17]). In the literature, it is common to refer to weak solutions of Navier-Stokes(-Fourier) equations as to suitable weak solutions if they fulfil the second law of thermodynamics. We stick to this convention, however we do not claim that our notion of suitable weak solution coincides with, for example, the similar notion in [12], [17] or [21], even if $\mathbb{B} = \mathbb{I}$. The crucial difference is that we consider the balance of the total energy only globally and thus, there is no local relation between $\partial_t\theta$ and $\partial_t|\boldsymbol{v}|^2$. However, we can observe that our suitable weak solution is also a suitable weak solution in the sense of [17] if it is sufficiently regular. Indeed, if we were able to multiply (3.22) by $\boldsymbol{v}$, (3.23) by $\mu(\mathbb{I} - \mathbb{B}^{-1})$ and (3.24) by θ, then, after summing everything together, we would obtain

$$\partial_t E + \boldsymbol{v} \cdot \nabla E - \operatorname{div}(\kappa(\theta)\nabla\theta) \geq \operatorname{div}(-\mathrm{p}\boldsymbol{v} + 2\nu(\theta)(\mathbb{D}\boldsymbol{v})\boldsymbol{v} + 2a\mu\theta\mathbb{B}\boldsymbol{v}) + \boldsymbol{f} \cdot \boldsymbol{v}.$$

However, this inequality must actually be an equality since the strict inequality (on any subset of Q of positive measure) would violate the global conservation of total energy (3.26). Thus, we recover the local balance of total energy (3.9) and the equivalence to the concept of weak solutions in [17] easily follows (in the case $\mathbb{B} = \mathbb{I}$). Of course, the requirement that (3.22) can be tested by $\boldsymbol{v}$ (and so on) is overly optimistic (unless $d \leq 2$). However, since our solution can be constructed as a limit of suitable approximations, the decisive criterion for the validity of (3.9) is merely that every term can be defined in a distributional sense. We discuss in Section 3.6 below when this actually happens.

3.3 Assumptions on the material coefficients

In this section we specify the growth properties of the material coefficients ν, γ, κ, δ, λ quantitatively. Assume that for some numbers $r > 0, q > 1$ (that are further specified below) the parameters of the model fulfil

$$a \in \mathbb{R}, \tag{3.28}$$
$$c_v, \mu > 0, \tag{3.29}$$
$$0 < \nu, \gamma, \kappa, \delta, \lambda \in \mathcal{C}(0, \infty), \tag{3.30}$$
$$C^{-1} \leq \nu(s) \leq C \quad \text{for all } s > 0, \tag{3.31}$$
$$C^{-1}s^q \leq \gamma(s) \leq Cs^q \quad \text{for all } s > s_0, \tag{3.32}$$
$$C^{-1}(1 + s^r) \leq \kappa(s) \leq C(1 + s^r) \quad \text{for all } s > 0, \tag{3.33}$$
$$C^{-1} \leq \delta(s) \leq C \quad \text{for all } s > 0, \tag{3.34}$$
$$C^{-1} \leq \lambda(s) \leq C \quad \text{for all } s > 0. \tag{3.35}$$

for some constants $C > 0$, $s_0 > 0$. Let us make a few remarks concerning these assumptions. The property (3.31) is quite relevant for fluids as there seems to be an experimental evidence that the viscosity of fluids (unlike gases) is not growing above all bounds with increasing temperature. Thus, we decided to stick with the assumption (3.31) even though a possible growth of ν for large temperatures (which is typically the case for compressible gases) leads to less restrictive assumptions on other parameters (for example, if ν was linear, then one could choose $r = 2$ to proceed with the analysis). The assumption (3.32) is purely theoretical; we need it to show sufficient integrability of $\mathbb{B}$, which is otherwise not available as the a priori estimates on $\nabla\mathbb{B}$ are too weak. However, as this term is related to the dissipation, we do not feel it is a serious drawback. Indeed, assumption (3.32) gives a restriction only for extreme values of $\mathbb{B}$, which are never attained in reality. In addition, there are also models (the so-called FENE models), where the a priori estimates require that $\mathbb{B}$ remains bounded. We remark that (3.30) and (3.32) imply that there exist $C, C_1, C_2 > 0$ such that

$$C_1 s^q - C_2 \leq \gamma(s) \leq C(1 + s^q) \quad \text{for all } s > 0, \tag{3.36}$$

which is a convenient form of (3.32) for the analysis below. The assumption (3.33) governs the growth of the thermal conductivity coefficient. In the existence theory below, we require that either r, or q, is fairly large. For example, we assume that

$$r \to \infty \quad \text{if} \quad q \to 1_+$$

(cf. (A1) below). Finally, the assumptions (3.33) and (3.35) are chosen just for simplicity. On one hand we could easily assume some growth with respect to the temperature, on the other hand, the experimental data do not show any rapid growth of these parameters neither for large and for small temperatures.

In addition to q and r, our model also contains the parameter $\varrho > 1$, which is the integrability exponent of the initial datum for the unknown $\mathbb{B}$. The parameter ϱ has a direct impact not only on the integrability of the unknown $\mathbb{B}$, but also on the other unknowns $\boldsymbol{v}$ and θ. The larger ϱ is, the better information can be drawn from the equation (3.23). Nevertheless, the case $\varrho > q$, gives no further advantage over the case $\varrho = q$ since we cannot expect that the quantity $\mathbb{D}\boldsymbol{v}$ (which is present on the right hand side of (3.23)) is better than square integrable. Vice versa, in the excluded[1] limiting case $\varrho = 1$, the equation (3.23) does not seem to provide any additional information than what is encoded in the energy and entropy balances.

Next, we place the restrictions on r, q and ϱ that are needed for:

1) existence of suitable weak solution,
2) fulfilment of the local balance of the total energy,
3) fulfilment of the local balance of the internal energy (or the temperature inequality in our case).

To this end, let us first define

$$\sigma = \min\{q, \varrho\} \qquad \text{and} \qquad r_d = r + \frac{2}{d}. \tag{3.37}$$

[1] The case $\varrho = 1$ is excluded since then the third term of (3.23) becomes only a L^1 quantity, which causes difficulties in the construction of a solution.

Conditions needed for existence of a suitable weak solution

We always suppose that r, q and ϱ satisfy

$$r > 1 - \frac{2}{d}, \qquad q > 1, \qquad \varrho > 1. \tag{A0}$$

Then, the conditions

$$(r_d - 1)(q - 1) > 2, \tag{A1}$$

$$(r_d - 1)\left(q + \sigma - \frac{2d}{d+2}\right) > \frac{4d}{d+2} \tag{A2}$$

are the minimal requirements for which we can prove existence of a suitable weak solution, as we explain in the next section.

Conditions needed for the local balance of E

Further, the conditions

$$(r_d - 1)(q + \sigma - 2) > 4 - \frac{2}{d}, \tag{B1}$$

$$(r_d - 1)\left(q + \sigma - \frac{3d}{d+2}\right) > \frac{6d}{d+2} \tag{B2}$$

seem necessary should the local balance of the total energy (3.9) be fulfilled if $d = 2$ or $d = 3$. The case $d = 1$ necessitates another restriction on r, q, ϱ and is omitted for simplicity. On the other hand, in the cases $d \geq 4$, it is unclear if $\boldsymbol{v} \in L^3(Q)$, for any choice of r, q and ϱ.

Condition needed for the local balance of e

Finally, the condition

$$(r_d - 1)(q + \sigma - 2) > 4 \tag{C}$$

is needed for the validity of the temperature inequality

$$c_v \partial_t \theta + c_v \boldsymbol{v} \cdot \nabla \theta - \operatorname{div}(\kappa(\theta)\nabla\theta) \geq 2\nu(\theta)|\mathbb{D}\boldsymbol{v}|^2 + 2a\mu\, \theta \mathbb{B} \cdot \mathbb{D}\boldsymbol{v}$$

(i.e., for the local balance of the internal energy).

Since $\sigma \leq q$, it is obvious that (C) implies (A1) and (A2), written symbolically as

$$\text{(C)} \Rightarrow \text{(A1)} \wedge \text{(A2)}.$$

Moreover, if $d \leq 3$, it is easy to see that the relation

$$\text{(C)} \Rightarrow \text{(A1)} \wedge \text{(B1)} \wedge \text{(B2)} \Rightarrow \text{(A1)} \wedge \text{(A2)}$$

is valid. Instead of (A1), we often use one of its equivalent versions:

$$r_d > \frac{q+1}{q-1}, \quad r_d > 2q' - 1, \quad \text{or} \quad r_d' < \frac{q+1}{2}.$$

Likewise, it is useful to note that (A2) is equivalent to

$$r_d > 1 + \frac{\frac{4d}{d+2}}{q + \sigma - \frac{2d}{d+2}},$$

hence also to

$$r_d' < \frac{1 + \frac{\frac{4d}{d+2}}{q+\sigma-\frac{2d}{d+2}}}{\frac{\frac{4d}{d+2}}{q+\sigma-\frac{2d}{d+2}}} = \frac{d+2}{4d}\left(q + \sigma + \frac{2d}{d+2}\right). \tag{3.38}$$

Similarly, condition (B2) is equivalent to

$$r_d' < \frac{d+2}{6d}\left(q + \sigma + \frac{3d}{d+2}\right). \tag{3.39}$$

Furthermore, defining

$$r_0 = \frac{q+1}{q-1} \quad \text{and} \quad r_1 = \frac{q+\sigma+2}{q+\sigma-2},$$

we observe that

$$\text{(A1)} \quad \text{is equivalent to} \quad r_d > r_0 \quad \text{and} \quad \text{(C)} \quad \text{is equivalent to} \quad r_d > r_1.$$

Let us make one important remark on the assumptions above. By imposing (A0)–(A2) and (3.31)–(3.35), we place some restrictions on the coefficients of the model which may not agree with experimental measurements. Thus, one could think that this renders our analysis useless in the actual applications. However, from the physical point of view, these assumptions are completely irrelevant. Indeed, note that (3.31)–(3.35) 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 (3.33) 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 (1.2)–(1.7). 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 coefficient γ. If $|\mathbb{B} - \mathbb{I}|$ is too large, any realistic material eventually breaks down. Thus, we may set $\gamma(s) = 1$, $s \in [0, s_1)$, where s_1 is large, to mimic the Oldroyd-B model, for example.

3.4 Definition of suitable weak solution

To state the definition of the weak solution conveniently, let us first define certain quantities that depend only on the given numbers q, r, ϱ and d. These play an important role in the existence theory below. We set

$$p = \begin{cases} \dfrac{q+\sigma}{2r_d' - 1} & \text{if} \quad r_d < r_1, \\ 2) & \text{if} \quad r_d = r_1, \\ 2 & \text{if} \quad r_d > r_1, \end{cases} \tag{3.40}$$

where the symbol $x_0)$, $x_0 \in \mathbb{R}$, is an abbreviation for any number from a (sufficiently small) left neighbourhood of x_0, excluding x_0. Furthermore, we set

$$R_d = R + \frac{2}{d}, \quad \text{where} \quad R = \begin{cases} \dfrac{(r_d - 1)(q+\sigma)}{2} - \dfrac{2}{d} & \text{if} \quad r_d < r_1, \\ r+1) & \text{if} \quad r_d \geq r_1. \end{cases} \tag{3.41}$$

Next, we define

$$p_1 = \min\left\{p, p\frac{d+2}{2d}\right\} \tag{3.42}$$

and

$$\sigma_1 = \begin{cases} \sigma & \text{if} \quad \sigma < q, \\ \sigma) & \text{if} \quad \sigma = q. \end{cases} \tag{3.43}$$

Finally, we set

$$s_0 = \frac{2(q+\sigma)}{q+\sigma+2} \quad \text{and} \quad s_4 = \min\left\{2 + \frac{2}{d}, q+\sigma\right\}. \tag{3.44}$$

Definition 3.1. *Let $T > 0$ and let $\Omega \subset \mathbb{R}^d$, $d \in \mathbb{N}$ be a Lipschitz domain. Assume that the constants a, c_v, μ and the functions $\nu, \delta, \gamma, \kappa, \lambda$ fulfil the assumptions* (3.28)–(3.35) *with the numbers q, r, ϱ satisfying* (A0)–(A2). *Let the numbers σ, p, R, R_d, p_1, σ_1, s_0 and s_4 be defined by* (3.37), (3.40), (3.41), (3.42), (3.43) *and* (3.44), *respectively. Suppose that the data satisfy*

$$\boldsymbol{v}_0 \in L^2(\Omega), \quad \mathbb{B}_0 \in L^{\varrho}(\Omega), \quad \theta_0 \in L^1(\Omega), \quad \boldsymbol{f} \in L^2(0,T;L^2(\Omega)), \tag{3.45}$$

where the function $\mathbb{B}_0$ is positive definite a.e. in Ω and the function θ_0 is positive a.e. in Ω. Moreover, assume that the function η_0 defined by

$$\eta_0 = c_v \ln\theta_0 - \mu(\operatorname{tr}\mathbb{B}_0 - d - \ln\det\mathbb{B}_0)$$

fulfils

$$\eta_0 \in L^1(\Omega). \tag{3.46}$$

Then, we say that a function $(\boldsymbol{v}, \mathbb{B}, \theta, \eta) : Q \to \mathbb{R}^d \times \mathbb{R}^{d\times d}_{>0} \times (0,\infty) \times \mathbb{R}$ is a suitable weak solution of the initial-boundary value problem (3.21)–(3.27) *if the following properties are satisfied:*

$$\begin{aligned}
\boldsymbol{v} &\in L^p(0,T;W^{1,p}_{0,\mathrm{div}}) \cap \mathcal{C}_w([0,T];L^2(\Omega)), && (3.47)\\
\partial_t\boldsymbol{v} &\in L^{p_1}(0,T;W^{-1,p_1}_{0,\mathrm{div}}), && (3.48)\\
\mathbb{B} &\in L^{s_0}(0,T;W^{1,s_0}(\Omega)) \cap \mathcal{C}_w([0,T];L^{\sigma}(\Omega)) \cap L^{q+\sigma}(Q), && (3.49)\\
\partial_t\mathbb{B} &\in \left(L^{s_0'}(0,T;W^{1,s_0'}(\Omega)) \cap L^{\frac{q+\sigma}{\sigma-1}}(Q)\right)^*, && (3.50)\\
\mathbb{B}^{-\frac{1}{2}}\nabla\mathbb{B}\mathbb{B}^{-\frac{1}{2}} &\in L^2(Q), && \\
\ln\det\mathbb{B} &\in L^2(0,T;W^{1,2}(\Omega)) \cap L^\infty(0,T;L^1(\Omega)), && (3.51)\\
\theta^{\frac{R}{2}} &\in L^2(0,T;W^{1,2}(\Omega)), && (3.52)\\
\theta &\in L^\infty(0,T;L^1(\Omega)) \cap L^{R_d}(Q), && (3.53)\\
\ln\theta &\in L^2(0,T;W^{1,2}(\Omega)) \cap L^\infty(0,T;L^1(\Omega)), && (3.54)\\
\eta &\in L^{s_0}(0,T;W^{1,s_0}(\Omega)) \cap L^\infty(0,T;L^1(\Omega)) \cap L^{s_4}(Q); && (3.55)
\end{aligned}$$

the identity

$$\eta = c_v \ln\theta - \mu(\operatorname{tr}\mathbb{B} - d - \ln\det\mathbb{B}), \tag{3.56}$$

holds almost everywhere in Q; equations (3.22)*,* (3.23)*,* (3.24) *and* (3.26) *are satisfied in the following sense:*

$$\left.\begin{aligned} &\int_0^T (\langle\partial_t \boldsymbol{v}, \boldsymbol{\varphi}\rangle - (\boldsymbol{v}\otimes\boldsymbol{v}, \nabla\boldsymbol{\varphi}) + (2\nu(\theta)\mathbb{D}\boldsymbol{v}, \nabla\boldsymbol{\varphi})) \\ &\qquad\qquad = -\int_0^T (2a\mu\theta\mathbb{B}, \nabla\boldsymbol{\varphi}) + \int_0^T (\boldsymbol{f}, \boldsymbol{\varphi}) \\ &\qquad\qquad \textit{for all} \quad \boldsymbol{\varphi} \in L^{p_1'}(0,T; W^{1,p_1'}_{0,\mathrm{div}}), \end{aligned}\right\} \tag{3.57}$$

$$\left.\begin{aligned} &\int_0^T \langle\partial_t \mathbb{B}, \mathbb{A}\rangle - \int_0^T (\mathbb{B}\otimes\boldsymbol{v}, \nabla\mathbb{A}) \\ &\qquad + \int_0^T (\delta(\theta)\gamma(|\mathbb{B}-\mathbb{I}|)(\mathbb{B}-\mathbb{I}), \mathbb{A}) + \int_0^T (\lambda(\theta)\nabla\mathbb{B}, \nabla\mathbb{A}) \\ &\qquad\qquad = \int_0^T ((a\mathbb{D}\boldsymbol{v} + \mathbb{W}\boldsymbol{v})\mathbb{B}, \mathbb{A} + \mathbb{A}^T) \\ &\qquad \textit{for all} \quad \mathbb{A} \in L^{s_0'}(0,T; W^{1,s_0'}(\Omega)) \cap L^{\frac{q+\sigma}{\sigma-1}}(Q), \end{aligned}\right\} \tag{3.58}$$

$$\left.\begin{aligned} &(\eta_0, \phi)\varphi(0) - \int_0^T (\eta, \phi)\partial_t\varphi \\ &\qquad + \int_0^T \Big(\kappa(\theta)\nabla\ln\theta - \mu\lambda(\theta)\nabla(\operatorname{tr}\mathbb{B} - d - \ln\det\mathbb{B}) - \eta\boldsymbol{v}, \nabla\phi\Big)\varphi \\ &\geq \int_0^T \Big(\frac{2\nu(\theta)}{\theta}|\mathbb{D}\boldsymbol{v}|^2 + \kappa(\theta)|\nabla\ln\theta|^2 \\ &\qquad + \mu\delta(\theta)\gamma(|\mathbb{B}-\mathbb{I}|)|\mathbb{B}^{\frac12} - \mathbb{B}^{-\frac12}|^2 + \mu\lambda(\theta)|\mathbb{B}^{-\frac12}\nabla\mathbb{B}\mathbb{B}^{-\frac12}|^2, \phi\Big)\varphi \\ &\textit{for all } 0 \le \varphi \in W^{1,\infty}(0,T),\ \varphi(T) = 0, \ \textit{ and every } 0 \le \phi \in W^{1,\infty}(\Omega), \end{aligned}\right\} \tag{3.59}$$

$$\int_\Omega (\tfrac12|\boldsymbol{v}|^2 + c_v\theta)(t) \le \int_\Omega (\tfrac12|\boldsymbol{v}_0|^2 + c_v\theta_0) + \int_0^t (\boldsymbol{f}, \boldsymbol{v}) \ \textit{ for a.a. } t \in (0,T); \tag{3.60}$$

and the initial data are attained in the following way:

$$\lim_{t\to 0_+} \|\boldsymbol{v}(t) - \boldsymbol{v}_0\|_2 = 0, \tag{3.61}$$

$$\lim_{t\to 0_+} \|\mathbb{B}(t) - \mathbb{B}_0\|_{\sigma_1} = 0, \tag{3.62}$$

$$\operatorname*{ess\,lim}_{t\to 0_+} \|\theta(t) - \theta_0\|_1 = 0, \tag{3.63}$$

$$\operatorname*{ess\,lim\,inf}_{t\to 0_+} \int_\Omega \eta(t)\phi \ge \int_\Omega \eta_0\phi \quad \textit{for all} \quad 0 \le \phi \in W^{1,\infty}(\Omega). \tag{3.64}$$

The bounds (A1)–(A2) are needed to define all the terms appearing in (3.57)–(3.59). Let us now quickly verify this fact using Hölder's inequality and properties (3.47)–(3.55). Also, for this purpose, let us assume that the test functions are smooth, the precise computations are carried out in Section 3.5 below. From an interpolation inequality, the Sobolev inequality and (3.47), we get

$$\boldsymbol{v} \in L^{(1+\frac{2}{d})p}(0,T; L^{(1+\frac{2}{d})p}(\Omega)). \tag{3.65}$$

Now we observe that (A2) ensures that the exponent $(1+\frac{2}{d})p$ is greater than two. Indeed, recalling (3.40), this is obvious if $r_d \geq r_1$, while if $r_0 < r_d < r_1$ we use (A2) in the form (3.38) to estimate

$$p = \frac{q+\sigma}{2r_d' - 1} > \frac{q+\sigma}{\frac{d+2}{2d}(q+\sigma)} = \frac{2d}{d+2}.$$

Thus, by the virtue of the Cauchy-Schwarz inequality, the convective term $\boldsymbol{v} \otimes \boldsymbol{v}$ in (3.57) is integrable. Next, using (3.41) and (A1), we get

$$\frac{1}{R_d} + \frac{1}{q+\sigma} = \frac{2}{(r_d-1)(q+\sigma)} + \frac{1}{q+\sigma} = \frac{r_d+1}{(r_d-1)(q+\sigma)} = \frac{1}{p} < 1$$

if $r_d < r_1$ and the same inequality also holds in the (subcritical) case $r_d \geq r_1$ since

$$\frac{1}{r_d+1} + \frac{1}{q+\sigma} < \frac{1}{2q'} + \frac{1}{q+\sigma} < \frac{1}{q'} + \frac{1}{q} = 1.$$

Thus, appealing to (3.49), (3.53) and Hölder's inequality, the term $\theta\mathbb{B}$ appearing in (3.57) is also integrable. In (3.58), the expression $\delta(\theta)\gamma(|\mathbb{B}-\mathbb{I}|)(\mathbb{B}-\mathbb{I})$ is well defined due to (3.36), (3.34), (3.49) and $q+1 < q+\sigma$. In certain sense the worst term is $\nabla\boldsymbol{v}\mathbb{B}$, which nevertheless, is integrable by (3.49) and (3.47) since (A1) implies

$$\frac{1}{p} + \frac{1}{q+\sigma} = \frac{2r_d'-1}{q+\sigma} + \frac{1}{q+\sigma} < \frac{q+1}{q+\sigma} < 1 \quad \text{if } r_d < r_1,$$

while in the case $r_d \geq r_1$, it is enough to use $q+\sigma > 2$. Moreover, the convective term $\mathbb{B} \otimes \boldsymbol{v}$ is clearly more regular than $\nabla\boldsymbol{v}\mathbb{B}$, recall (3.65). In equation (3.59), the convective term $\eta\boldsymbol{v}$ is also well defined due to (3.55): we already know that $\boldsymbol{v}$ is better than square integrable and η has the same property since

$$s_2 = \min\left\{2 + \frac{2}{d}, q+\sigma\right\} > 2.$$

The term $\kappa(\theta)\nabla\ln\theta$ is integrable provided that $\sqrt{\kappa(\theta)}$ is square integrable (using (3.54)), which is true since $\theta^{\frac{r}{2}} \in L^{\frac{2R_d}{r}}(Q)$ by (3.53) and

$$R_d > r_d > r.$$

Indeed, this inequality is obvious if $r_d \geq r_1$, while if $r_0 < r_d < r_1$, it follows from (A1) as

$$\begin{aligned} R_d = \frac{(r_d-1)(q+\sigma)}{2} &= r_d + \frac{r_d(q+\sigma-2) - q - \sigma}{2} \\ &> r_d + \frac{q^2 + q\sigma - 2q + q + \sigma - 2 - q^2 - q\sigma + q + \sigma}{2(q-1)} \\ &= r_d + \frac{\sigma-1}{q-1} > r_d. \end{aligned}$$

The term $\nabla \ln\det\mathbb{B}$ is square integrable due to (3.51). Finally, each term on the right hand side of (3.59) is integrable since every one of them is non-negative and $\eta \in L^\infty(0,T;L^1(\Omega))$.

Thus, we have verified that Definition 3.1 is reasonable if (A0)–(A2) hold and also that the requirements (3.47)–(3.55) are natural. The purpose of the further assumptions (B1), (B2) and (C) is explained later in Section 3.6 and Section 3.7, respectively.

It is easy to show that every smooth suitable weak solution satisfies the weak form of local balance of total energy, using the same manipulations as in Section 3.2. Then, using integration by parts and the fundamental lemma of variational calculus, we obtain precisely the system (3.21)–(3.27), (3.9). This is the the weak-strong compatibility of the suitable weak solution.

In our main result below we show that under the assumptions of Definition 3.1 a suitable weak solution exists. It may be interesting to note that $p < 2$ if $r_d < r_1$. In other words, we obtain a weak solution of our system even without knowing whether the viscous dissipation $2\nu(\theta)|\mathbb{D}\boldsymbol{v}|^2$ is an integrable quantity or not. A similar phenomenon is also observed in the existence theory for heat-conducting compressible Navier-Stokes equations (see [33], [34, Ch. 2]) or for certain fluid mixtures.

It remains to show that a suitable weak solution to (3.21)–(3.27) exists.

3.5 Existence of a suitable weak solution

In this section we state and prove the main result of this thesis, which is the following theorem.

Theorem 3.2. *Suppose that all the assumptions of Definition 3.1 are fulfilled. Then, there exists a suitable weak solution to the system* (3.21)–(3.27).

Proof. Although the Theorem 3.2 is stated for $d \in \mathbb{N}$, the proof is done for $d \geq 3$. The cases $d = 1$ and $d = 2$ are, of course, simpler and can be obtained by obvious modifications of the proof below.

The general strategy is the following: We approximate the system (3.4)–(3.7) (the one with temperature equation) using several parameters to obtain a proper Galerkin approximation and we 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. After taking the limit in these, we can extend the space of test functions in the equations for $\mathbb{B}$ and θ. Using this, we prove invertibility of θ and $\mathbb{B}$, which, in turn, enables us to derive the entropy equation. From this equation we read the fundamental uniform estimates. Some of these estimates are then improved by considering appropriate test functions in the equations for θ and $\mathbb{B}$. Finally, we pass to the final limit, identify the non-linear terms and initial conditions, hereby obtaining a solution of the original problem.

3.5.1 Approximation scheme

Here we introduce an approximation, which is essential for the proof. It is constructed in a way that one can prove a minimum principle for the spectrum of $\mathbb{B}$ and for θ. This is a key step in obtaining the entropy inequality, from which we then read a powerful a priori estimate that is uniform with respect to all parameters. We also prepare some simple estimates corresponding to this approximation, that are used later in the proof.

For any $\omega \in (0,1)$, let us define the "cut-off" function

$$g_\omega(\mathbb{A},\tau) = \frac{\max\{0,\Lambda(\mathbb{A})-\omega\}\max\{0,\tau-\omega\}}{(|\Lambda(\mathbb{A})|+\omega)(1+\omega|\mathbb{A}|^3)(|\tau|+\omega)(1+\omega\tau^2)}, \quad \mathbb{A} \in \mathbb{R}^{d\times d}_{\mathrm{sym}}, \quad \tau \in \mathbb{R},$$

where

$$\Lambda(\mathbb{A}) = \text{the smallest eigenvalue of } \mathbb{A}.$$

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

$$\lim_{\omega\to 0_+} g_\omega(\mathbb{A},\tau) = 1.$$

Furthermore, we remark that

$$g_\omega(\mathbb{A},\tau)(1+|\mathbb{A}|+|\mathbb{A}|^2+|\mathbb{A}|^3)(1+\tau+\tau^2) \le C(\omega). \tag{3.66}$$

Then, we consider the following ω-approximated system of equations in Q:

$$\partial_t \boldsymbol{v} + \operatorname{div}(\boldsymbol{v}\otimes\boldsymbol{v}) - \operatorname{div}(2\nu(\theta)\mathbb{D}\boldsymbol{v}) + \nabla \mathrm{p} = \operatorname{div}(2a\mu g_\omega(\mathbb{B},\theta)\theta\mathbb{B}) + \boldsymbol{f}, \tag{3.67}$$

$$\begin{aligned} &\partial_t\mathbb{B} + \boldsymbol{v}\cdot\nabla\mathbb{B} + \delta(\theta)\gamma(|\mathbb{B}-\mathbb{I}|)(\mathbb{B}-\mathbb{I}) - \operatorname{div}(\lambda(\theta)\nabla\mathbb{B}) \\ &\qquad\qquad = g_\omega(\mathbb{B},\theta)(a(\mathbb{D}\boldsymbol{v}\mathbb{B}+\mathbb{B}\mathbb{D}\boldsymbol{v}) + \mathbb{W}\boldsymbol{v}\mathbb{B} - \mathbb{B}\mathbb{W}\boldsymbol{v}), \end{aligned}$$

$$\begin{aligned} &c_v\partial_t\theta + c_v\boldsymbol{v}\cdot\nabla\theta - \operatorname{div}(\kappa(\theta)\nabla\theta) - \omega\operatorname{div}(|\nabla\theta|^r\nabla\theta) \\ &\qquad\qquad = 2\nu(\theta)|\mathbb{D}\boldsymbol{v}|^2 + 2a\mu g_\omega(\mathbb{B},\theta)\theta\mathbb{B}\cdot\mathbb{D}\boldsymbol{v} \end{aligned} \tag{3.68}$$

with the boundary conditions

$$\boldsymbol{v} = 0, \quad \boldsymbol{n}\cdot\nabla\mathbb{B} = 0, \quad \boldsymbol{n}\cdot\nabla\theta = 0 \qquad \text{on } (0,T)\times\partial\Omega \tag{3.69}$$

and the initial conditions

$$\boldsymbol{v}(0) = \boldsymbol{v}_0, \quad \mathbb{B}(0) = \mathbb{B}_0^\omega, \quad \theta(0) = \theta_0^\omega \qquad \text{in } \Omega,$$

where $\mathbb{B}_0^\omega$ and θ_0^ω are defined by

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

$$\theta_0^\omega(x) = \begin{cases} \theta_0(x) & \text{if } \omega < \theta_0(x) < \frac{1}{\omega}, \\ 1 & \text{elsewhere.} \end{cases}$$

With such a definition, these functions clearly satisfy

$$\Lambda(\mathbb{B}_0^\omega) > \omega, \qquad \theta_0^\omega > \omega \tag{3.70}$$

and

$$|\mathbb{B}_0^\omega| < \frac{\sqrt{d}}{\omega}, \qquad |\theta_0^\omega| < \frac{1}{\omega} \tag{3.71}$$

in Ω. Moreover, it is evident that

$$|\mathbb{B}_0^\omega| \le \sqrt{d} + |\mathbb{B}_0|, \qquad \theta_0^\omega \le 1 + \theta_0, \tag{3.72}$$

and, since $\ln 1 = 0$, also that

$$|\ln \det \mathbb{B}_0^\omega| \le |\ln \det \mathbb{B}_0|, \qquad |\ln \theta_0^\omega| \le |\ln \theta_0| \tag{3.73}$$

a.e. in Ω. Let us further remark that, since $\mathbb{B}_0 \in L^\sigma(\Omega)$ is positive definite a.e. in Ω, the Lebesgue measure of the sets $\{\Lambda(\mathbb{B}_0) \le \omega\}$ and $\{|\mathbb{B}_0| \ge \frac{1}{\omega}\}$ tends to zero as $\omega \to 0_+$, and thus

$$\|\mathbb{B}_0^\omega - \mathbb{B}_0\|_\sigma^\sigma = \int_{\Lambda(\mathbb{B}_0)\le\omega} |\mathbb{I} - \mathbb{B}_0|^\sigma + \int_{|\mathbb{B}_0|\ge\frac{1}{\omega}} |\mathbb{I} - \mathbb{B}_0|^\sigma \to 0. \tag{3.74}$$

Using a completely analogous argument for θ_0^ω and relying on the assumptions $\theta_0 \in L^1(\Omega)$, $\theta > 0$ a.e. in Ω, we also obtain

$$\|\theta_0^\omega - \theta_0\|_1 \to 0, \quad \omega \to 0_+. \tag{3.75}$$

In (3.68) we also included the term $-\omega \operatorname{div}(|\nabla\theta|^r \nabla\theta)$ which does not appear in the original temperature equation. This term vanishes in the limit $\omega \to 0_+$ and is used only to avoid the construction of a weighted Sobolev space of the type

$$\left\{ u \in W^{1,2}(\Omega) : \int_\Omega \kappa(\theta)|\nabla u|^2 < \infty \right\}$$

(where the density of smooth functions is not available in general).

Galerkin approximation

Next, we discretize the ω-approximated system in space by the Galerkin method. Following the results from [52], we find the bases $\{\boldsymbol{w}_i\}_{i=1}^\infty$, $\{\mathbb{W}_j\}_{j=1}^\infty$ and $\{w_k\}_{k=1}^\infty$ of $W^{N,2}(\Omega) \cap W^{1,2}_{0,\mathrm{div}}$, $W^{N,2}(\Omega)$ and $W^{N,2}(\Omega)$, respectively, with the following property: The bases are orthonormal in $L^2(\Omega)$ and orthogonal in $W^{N,2}(\Omega)$ with $N \in \mathbb{N}$ so large that $W^{N,2}(\Omega) \hookrightarrow W^{1,\infty}(\Omega)$. Let us also assume, without loss of generality, that $w_1 = |\Omega|^{-\frac{1}{2}}$ in Ω and that $\mathbb{W}_j^T = \mathbb{W}_j$ for every $j \in \mathbb{N}$. Moreover, for any $\ell, n \in \mathbb{N}$, there exist L^2-orthogonal projections

$$\begin{aligned} P_\ell &: L^2(\Omega) \to \operatorname{span}\{\boldsymbol{w}_i\}_{i=1}^\ell, \\ Q_n &: L^2(\Omega) \to \operatorname{span}\{\mathbb{W}_j\}_{j=1}^n, \\ R_n &: L^2(\Omega) \to \operatorname{span}\{w_k\}_{k=1}^n \end{aligned}$$

and

$$P_\ell, Q_n, R_n \text{ are continuous in } L^2(\Omega) \text{ and } W^{N,2}(\Omega), \text{ independently of } \ell, n. \tag{3.76}$$

We fix $\ell, n \in \mathbb{N}$ and consider the problem of finding the functions $\alpha_{\ell n}^i$, $\beta_{\ell n}^j$, $\gamma_{\ell n}^k$ of time, where $i = 1, \dots, \ell$ and $j, k = 1, \dots, n$, such that the functions $\boldsymbol{v}_{\ell n}$, $\mathbb{B}_{\ell n}$, $\theta_{\ell n}$ defined as

$$\begin{gathered} \boldsymbol{v}_{\ell n}(t,x) = \sum_{i=1}^{\ell} \alpha_{\ell n}^i(t) \boldsymbol{w}_i(x), \\ \mathbb{B}_{\ell n}(t,x) = \sum_{j=1}^{n} \beta_{\ell n}^j(t) \mathbb{W}_j(x) \quad \text{and} \quad \theta_{\ell n} = \sum_{k=1}^{n} \gamma_{\ell n}^k(t) w_k(x) \end{gathered} \tag{3.77}$$

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

$$\begin{aligned}(\partial_t\boldsymbol{v}_{\ell n},\boldsymbol{w}_i)-(\boldsymbol{v}_{\ell n}\otimes\boldsymbol{v}_{\ell n},\nabla\boldsymbol{w}_i)+(2\nu(\theta_{\ell n})\mathbb{D}\boldsymbol{v}_{\ell n},\nabla\boldsymbol{w}_i)\\ =-(2a\mu g_\omega(\mathbb{B}_{\ell n},\theta_{\ell n})\theta_{\ell n}\mathbb{B}_{\ell n},\nabla\boldsymbol{w}_i)+(\boldsymbol{f},\boldsymbol{w}_i),\end{aligned}\tag{3.78}$$

$$\begin{aligned}&(\partial_t\mathbb{B}_{\ell n},\mathbb{W}_j)+(\boldsymbol{v}_{\ell n}\cdot\nabla\mathbb{B}_{\ell n},\mathbb{W}_j)\\ &\quad+(\delta(\theta_{\ell n})\gamma(|\mathbb{B}_{\ell n}-\mathbb{I}|)(\mathbb{B}_{\ell n}-\mathbb{I}),\mathbb{W}_j)+(\lambda(\theta_{\ell n})\nabla\mathbb{B}_{\ell n},\nabla\mathbb{W}_j)\\ &\qquad=(2g_\omega(\mathbb{B}_{\ell n},\theta_{\ell n})(a\mathbb{D}\boldsymbol{v}_{\ell n}+\mathbb{W}\boldsymbol{v}_{\ell n})\mathbb{B}_{\ell n},\mathbb{W}_j),\end{aligned}\tag{3.79}$$

$$\begin{aligned}(c_v\partial_t\theta_{\ell n},w_k)+(c_v\boldsymbol{v}_{\ell n}\cdot\nabla\theta_{\ell n},w_k)+((\kappa(\theta_{\ell n})+\omega|\nabla\theta_{\ell n}|^r)\nabla\theta_{\ell n},\nabla w_k)\\ =(2\nu(\theta_{\ell n})|\mathbb{D}\boldsymbol{v}_{\ell n}|^2+2a\mu g_\omega(\mathbb{B}_{\ell n},\theta_{\ell n})\theta_{\ell n}\mathbb{B}_{\ell n}\cdot\mathbb{D}\boldsymbol{v}_{\ell n},w_k),\end{aligned}\tag{3.80}$$

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

$$\boldsymbol{v}_{\ell n}(0)=P_\ell\boldsymbol{v}_0,\quad \mathbb{B}_{\ell n}(0)=Q_n\mathbb{B}_0^\omega,\quad \theta_{\ell n}(0)=R_n\theta_0^\omega\quad\text{in }\Omega.\tag{3.81}$$

Since, 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

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

the system (3.78)–(3.80) can be rewritten as

$$\left.\begin{aligned}(\alpha_{\ell n}^i)'&=F_1(t,\alpha_{\ell n}^1,\dots,\alpha_{\ell n}^\ell), && i=1,\dots,\ell,\\ (\beta_{\ell n}^j)'&=F_2(\beta_{\ell n}^1,\dots,\beta_{\ell n}^n), && j=1,\dots,n,\\ (\gamma_{\ell n}^k)'&=F_3(\gamma_{\ell n}^1,\dots,\gamma_{\ell n}^n), && k=1,\dots,n.\end{aligned}\right\}\tag{3.82}$$

This is a system of $\ell+2n$ ordinary differential equations. Though it contains many non-linearities, it is easy to see, using (3.30), that F_1,F_2 and F_3 are continuous with respect to the variables $\alpha_{\ell n}^i$, $\beta_{\ell n}^j$ and $\gamma_{\ell n}^k$, respectively. Moreover, the explicit dependence of F_1 on time is controlled by

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

Thus, we can apply the Caratheodory existence theorem (see [24, Chapter 2, Theorem 1]) 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 (3.82) on $(0,T_0)$, where $T_0<T$ is the time of the first blow-up, i.e., the time, for which

$$\lim_{t\to(T_0)_-}\left(\sum_{i=1}^{\ell}|\alpha_{\ell n}^i(t)|+\sum_{j=1}^{n}|\beta_{\ell n}^j(t)|+\sum_{k=1}^{n}|\gamma_{\ell n}^k(t)|\right)=\infty.\tag{3.83}$$

If we use the a priori estimates derived in the next section (see e.g. (3.85)) and apply them onto the interval $(0,T_0)$, we can prove that

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

which contradicts (3.83). 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 $[0,T]$.

3.5.2 Limit $n \to \infty$

By multiplying the i-th equation in (3.78) by $\alpha^i_{\ell n}$, summing the result over all $i = 1, \dots, \ell$ and using (3.69), integration by parts, and the fact that $\operatorname{div} \boldsymbol{v} = 0$ (so that the convective term vanishes), we obtain

$$
\begin{aligned}
\frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}t}\|\boldsymbol{v}_{\ell n}\|_2^2 + \left\|\sqrt{2\nu(\theta_{\ell n})}\mathbb{D}\boldsymbol{v}_{\ell n}\right\|_2^2 \\
&= -(2a\mu g_\omega(\mathbb{B}_{\ell n}, \theta_{\ell n})\theta_{\ell n}\mathbb{B}_{\ell n}, \mathbb{D}\boldsymbol{v}_{\ell n}) + (\boldsymbol{f}, \boldsymbol{v}_{\ell n})
\end{aligned}
$$

a.e. in $(0, T)$. Then we use (3.31), (3.66), (3.81), Korn's and Young's inequality, and deduce

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

a.e. in $(0, T)$. Integration with respect to time and the use of (3.76) and (3.45) directly leads to

$$
\sup_{t\in(0,T)} \|\boldsymbol{v}_{\ell n}(t)\|_2^2 + \int_0^T \|\nabla\boldsymbol{v}_{\ell n}\|_2^2 \le C(\omega). \tag{3.84}
$$

(we do not need to trace the dependence of constants C on the data $\boldsymbol{f}$, $\boldsymbol{v}_0$, θ_0, or $\mathbb{B}_0$ as these are fixed functions in our setting). Recalling the construction of $\boldsymbol{v}_{\ell n}$ in (3.77) and L^2-orthonormality of the basis vectors $\{\boldsymbol{w}_i\}_{i=1}^\ell$, we note that

$$
\|\boldsymbol{v}_{\ell n}(t)\|_2^2 = \sum_{i=1}^{\ell}(\alpha^i_{\ell n}(t))^2.
$$

Hence, the estimate (3.84) yields

$$
\sup_{t\in(0,T)} \sum_{i=1}^{\ell}(\alpha^i_{\ell n}(t))^2 \le C(\omega), \tag{3.85}
$$

which, together with $\boldsymbol{w}_i \in W^{1,\infty}(\Omega)$, $i = 1, \dots, \ell$, implies

$$
\|\boldsymbol{v}_{\ell n}\|_{L^\infty W^{1,\infty}} \le C(\omega, \ell). \tag{3.86}
$$

Using (3.85) together with (3.66) and (3.31) in (3.78), we see that

$$
\begin{aligned}
&\left\|(\alpha^i_{\ell n})'\right\|_{2;(0,T)} = \|(\partial_t\boldsymbol{v}_{\ell n}, \boldsymbol{w}_i)\|_{2;(0,T)} \\
&= \|(\boldsymbol{v}_{\ell n}\otimes\boldsymbol{v}_{\ell n} - 2\nu(\theta_{\ell n})\mathbb{D}\boldsymbol{v}_{\ell n} - 2a\mu g_\omega(\mathbb{B}_{\ell n}, \theta_{\ell n})\theta_{\ell n}\mathbb{B}_{\ell n}, \nabla\boldsymbol{w}_i) + (\boldsymbol{f}, \boldsymbol{w}_i)\|_{2;(0,T)} \\
&\le C(\ell)\left\|\sum_{i=1}^{\ell}\left((\alpha^i_{\ell n})^2 + |\alpha^i_{\ell n}| + 1\right)\right\|_{2;(0,T)} + C(\ell)\|\boldsymbol{f}\|_{L^2L^2} \\
&\le C(\omega, \ell).
\end{aligned}
$$

Thus, we get

$$
\|\partial_t\boldsymbol{v}_{\ell n}\|_{L^2W^{1,\infty}} = \left\|\sum_{i=1}^{\ell}(\alpha^i_{\ell n})'\boldsymbol{w}_i\right\|_{L^2W^{1,\infty}} \le C(\omega, \ell) \tag{3.87}
$$

and, using the fundamental theorem of calculus (see [67, Theorem 7.20]) and Hölder's inequality, also that

$$|\alpha^i_{\ell n}(t) - \alpha^i_{\ell n}(s)| \leq \int_s^t |(\alpha^i_{\ell n})'| \leq C(\omega,\ell)|t-s|^{\frac{1}{2}} \quad \text{for every } t,s \in [0,T] \tag{3.88}$$

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

Next, we multiply the j-th equation in (3.79) by $\beta^j_{\ell n}$ and sum the result over $j = 1,\dots,n$. Note that the convective term vanishes after integration by parts and use of (3.69) and $\operatorname{div} \boldsymbol{v} = 0$. Also the term including $\mathbb{W}\boldsymbol{v}_{\ell n}$ vanishes due to (4.34) and symmetry of $\mathbb{B}^2_{\ell n}$. Thus, we obtain

$$\begin{aligned}\frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}t}\|\mathbb{B}_{\ell n}\|_2^2 + (\delta(\theta_{\ell n})\gamma(|\mathbb{B}_{\ell n} - \mathbb{I}|)(\mathbb{B}_{\ell n} - \mathbb{I}), \mathbb{B}_{\ell n}) + \left\|\sqrt{\lambda(\theta_{\ell n})}\nabla\mathbb{B}_{\ell n}\right\|_2^2 \\ = (2ag_\omega(\mathbb{B}_{\ell n},\theta_{\ell n})\mathbb{D}\boldsymbol{v}_{\ell n}\mathbb{B}_{\ell n}, \mathbb{B}_{\ell n})\end{aligned} \tag{3.89}$$

a.e. in $(0,T)$. To estimate the second term in (3.89) from below, first we remark, using the Young inequality, that

$$(\mathbb{B}_{\ell n} - \mathbb{I})\cdot\mathbb{B}_{\ell n} = |\mathbb{B}_{\ell n} - \mathbb{I}|^2 + (\mathbb{B}_{\ell n} - \mathbb{I})\cdot\mathbb{I} \geq \tfrac{1}{2}|\mathbb{B}_{\ell n} - \mathbb{I}|^2 - \tfrac{d}{2},$$

and then we apply (3.36), (3.34) and use the Young inequality again to get

$$\begin{aligned}&\delta(\theta_{\ell n})\gamma(|\mathbb{B}_{\ell n} - \mathbb{I}|)(\mathbb{B}_{\ell n} - \mathbb{I})\cdot\mathbb{B}_{\ell n} \\ &\qquad \geq \tfrac{1}{2}\delta(\theta_{\ell n})\gamma(|\mathbb{B}_{\ell n} - \mathbb{I}|)|\mathbb{B}_{\ell n} - \mathbb{I}|^2 - \delta(\theta_{\ell n})\gamma(|\mathbb{B}_{\ell n} - \mathbb{I}|)\tfrac{d}{2} \\ &\qquad \geq C_1|\mathbb{B}_{\ell n} - \mathbb{I}|^{2+q} - C_2|\mathbb{B}_{\ell n} - \mathbb{I}|^2 - C|\mathbb{B}_{\ell n} - \mathbb{I}|^q - C \\ &\qquad \geq C_1|\mathbb{B}_{\ell n} - \mathbb{I}|^{2+q} - C.\end{aligned}$$

If we use this estimate in (3.89), together with (3.35), (3.66), (3.81), we obtain, after integration over $(0,t)$, $t \in (0,T)$, that

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

From this, using (3.76) and (3.71), we easily read that

$$\|\mathbb{B}_{\ell n}\|_{L^\infty L^2} + \|\mathbb{B}_{\ell n}\|_{L^{2+q}L^{2+q}} + \|\nabla\mathbb{B}_{\ell n}\|_{L^2L^2} \leq C(\omega,\ell). \tag{3.90}$$

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 $\|\mathbb{A}\|_{L^{q+2}W^{N,2}} \leq 1$ and use (3.79), Hölder's inequality, (3.90), (3.86), (3.34), (3.36), (3.35), (3.66), (3.76) and

$$\left(\min\left\{2, \frac{q+2}{q+1}\right\}\right)' = \left(\frac{q+2}{q+1}\right)' = q+2$$

to get

$$
\begin{aligned}
\int_0^T \langle \partial_t \mathbb{B}_{\ell n}, \mathbb{A}\rangle &= \int_0^T (\partial_t \mathbb{B}_{\ell n}, Q_n\mathbb{A}) \\
&= -\int_0^T (\boldsymbol{v}_{\ell n}\cdot\nabla \mathbb{B}_{\ell n}, Q_n\mathbb{A}) - \int_0^T (\delta(\theta_{\ell n})\gamma(|\mathbb{B}_{\ell n}-\mathbb{I}|)(\mathbb{B}_{\ell n}-\mathbb{I}), Q_n\mathbb{A}) \\
&\quad - \int_0^T (\lambda(\theta_{\ell n})\nabla \mathbb{B}_{\ell n}, \nabla Q_n\mathbb{A}) + \int_0^T (2g_\omega(\mathbb{B}_{\ell n},\theta_{\ell n})(a\mathbb{D}\boldsymbol{v}_{\ell n} + \mathbb{W}\boldsymbol{v}_{\ell n})\mathbb{B}_{\ell n}, Q_n\mathbb{A}) \\
&\le C(\omega,\ell)\int_0^T\int_\Omega \Big(|\nabla \mathbb{B}_{\ell n}||Q_n\mathbb{A}| + |\mathbb{B}_{\ell n}-\mathbb{I}|^{q+1}|Q_n\mathbb{A}| + |\nabla\mathbb{B}_{\ell n}||\nabla Q_n\mathbb{A}| + |Q_n\mathbb{A}|\Big) \\
&\le C(\omega,\ell)\int_0^T (\|\nabla\mathbb{B}_{\ell n}\|_1 + \|\mathbb{B}_{\ell n}-\mathbb{I}\|_{q+1}^{q+1} + \|\mathbb{B}_{\ell n}\|_1)\|Q_n\mathbb{A}\|_{1,\infty} \\
&\le C(\omega,\ell)\int_0^T (\|\nabla\mathbb{B}_{\ell n}\|_2 + \|\mathbb{B}_{\ell n}\|_{q+2}^{q+1} + 1)\|Q_n\mathbb{A}\|_{N,2} \\
&\le C(\omega,\ell)\|\mathbb{A}\|_{L^{q+2}W^{N,2}} \le C(\omega,\ell),
\end{aligned}
$$

hence

$$
\|\partial_t \mathbb{B}_{\ell n}\|_{L^{\frac{q+2}{q+1}}W^{-N,2}} \le C(\omega,\ell). \tag{3.91}
$$

Finally, we multiply the k-th equation in (3.80) by $\gamma^k_{\ell n}$, sum the result over $k = 1,\dots,n$, use (3.69) and integration by parts in the convective term to get

$$
\begin{aligned}
\frac{c_v}{2}\frac{\mathrm{d}}{\mathrm{d}t}\|\theta_{\ell n}\|_2^2 + \left\|\sqrt{\kappa(\theta_{\ell n})}\nabla\theta_{\ell n}\right\|_2^2 + \omega\|\nabla\theta_{\ell n}\|_{r+2}^{r+2} \\
= (2\nu(\theta_{\ell n})|\mathbb{D}\boldsymbol{v}_{\ell n}|^2 + 2a\mu g_\omega(\mathbb{B}_{\ell n},\theta_{\ell n})\theta_{\ell n}\mathbb{B}_{\ell n}\cdot\mathbb{D}\boldsymbol{v}_{\ell n}, \theta_{\ell n})
\end{aligned} \tag{3.92}
$$

a.e. in $(0,T)$. Therefore, integrating this inequality over $(0,t)$, $t\in(0,T)$, using (3.29), (3.31), (3.33), (3.66), (3.86), (3.90), Young's inequality, (3.76) and (3.71), we deduce

$$
\|\theta_{\ell n}(t)\|_2^2 + \int_0^t \left\|\nabla\theta_{\ell n}^{\frac{r}{2}+1}\right\|_2^2 + \int_0^t \left\|\sqrt{\kappa(\theta_{\ell n})}\nabla\theta_{\ell n}\right\|_2^2 + \int_0^t \|\nabla\theta_{\ell n}\|_{r+2}^{r+2} \le C(\omega,\ell).
$$

This, with the help of the interpolation inequality[2]

$$
\|\theta_{\ell n}\|_{L^{r+2+\frac{4}{d}}L^{r+2+\frac{4}{d}}} \le \|\theta_{\ell n}\|_{L^\infty L^2}^{\frac{4}{(r+2)d+4}}\|\theta_{\ell n}\|_{L^{r+2}L^{\frac{d}{d-2}(r+2)}}^{\frac{(r+2)d}{(r+2)d+4}} = \|\theta_{\ell n}\|_{L^\infty L^2}^{\frac{4}{(r+2)d+4}}\left\|\theta_{\ell n}^{\frac{r}{2}+1}\right\|_{L^2L^{\frac{2d}{d-2}}}^{\frac{2d}{(r+2)d+4}},
$$

Sobolev's inequality and also Poincaré's inequality yields

$$
\begin{aligned}
\|\theta_{\ell n}\|_{L^\infty L^2} + \left\|\sqrt{\kappa(\theta_{\ell n})}\nabla\theta_{\ell n}\right\|_{L^2L^2} + \|\theta_{\ell n}\|_{L^{r+2+\frac{4}{d}}L^{r+2+\frac{4}{d}}} \\
+ \|\nabla\theta_{\ell n}\|_{L^{r+2}L^{r+2}} \le C(\omega,\ell).
\end{aligned} \tag{3.93}
$$

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 (3.80), Young's inequality, Hölder's inequality, the inequality

$$
(r+1)\frac{r+2}{r+1} = \frac{r+2}{r+1} + b + \frac{b}{r+1} \le \frac{r+2}{r+1} + b + \frac{r}{r+1} = r+2,
$$

[2] A better estimate could be derived using $\nabla\theta_\ell \in L^{r+2}L^{r+2}$ instead. However, at this moment we do not need it, and later we shall need ω-uniform estimates only.

(3.31), (3.33) (3.86), (3.93), (3.66) and (3.76), we obtain

$$
\begin{aligned}
\int_0^T \langle \partial_t \theta_{\ell n}, \tau \rangle &= \int_0^T (\partial_t \theta_{\ell n}, R_n \tau) \\
&= - \int_0^T (c_v \boldsymbol{v}_{\ell n} \cdot \nabla \theta_{\ell n}, R_n \tau) - \int_0^T (\kappa(\theta_{\ell n}) \nabla \theta_{\ell n} + \omega |\nabla \theta_{\ell n}|^r \nabla \theta_{\ell n}, \nabla R_n \tau) \\
&\quad + \int_0^T (2\nu(\theta_{\ell n}) |\mathbb{D} \boldsymbol{v}_{\ell n}|^2 + 2a\mu g_\omega(\mathbb{B}_{\ell n}, \theta_{\ell n}) \theta_{\ell n} \mathbb{B}_{\ell n} \cdot \mathbb{D} \boldsymbol{v}_{\ell n}, R_n \tau) \\
&\le C(\omega, \ell) \int_0^T \int_\Omega \Big(|\nabla \theta_{\ell n}| |R_n \tau| + |\theta_{\ell n}|^{\frac{r}{2}} \left| \sqrt{\kappa(\theta_{\ell n})} \nabla \theta_{\ell n} \right| |\nabla R_n \tau| \\
&\qquad\qquad + |\nabla \theta_{\ell n}|^{r+1} |\nabla R_n \tau| + |R_n \tau| \Big) \\
&\le C(\omega, \ell) \int_0^T \int_\Omega \Big(|\nabla \theta_{\ell n}| + |\theta_{\ell n}|^{r+1} + \left| \sqrt{\kappa(\theta_{\ell n})} \nabla \theta_{\ell n} \right|^{\frac{2r+2}{r+2}} \\
&\qquad\qquad + |\nabla \theta_{\ell n}|^{r+1} + 1 \Big) \|R_n \tau\|_{1,\infty} \\
&\le C(\omega, \ell) \int_0^T \Big(\|\nabla \theta_{\ell n}\|_{\frac{r+2}{r+1}} + \|\theta_{\ell n}\|_{r+2}^{r+1} + \left\| \sqrt{\kappa(\theta_{\ell n})} \nabla \theta_{\ell n} \right\|_2^{\frac{2r+2}{r+2}} \\
&\qquad\qquad + \|\nabla \theta_{\ell n}\|_{r+1}^{r+1} + 1 \Big) \|R_n \tau\|_{N,2} \\
&\le C(\omega, \ell) \|\tau\|_{L^{r+2} W^{N,2}} \le C(\omega, \ell),
\end{aligned}
$$

hence

$$
\|\partial_t \theta_{\ell n}\|_{L^{\frac{r+2}{r+1}} W^{-N,2}} \le C(\omega, \ell). \tag{3.94}
$$

At this point, we want to apply Lemma 4.1 to obtain weakly converging subsequences and their limits. Before we do that, let us make two conventions. First, we never relabel subsequences obtained from Lemma 4.1 or by similar arguments. Second, instead of

$$
u_k \to u \quad \text{strongly in } L^s(Q) \quad \text{for any } 1 \le s < S,
$$

we write just

$$
u_k \to u \quad \text{strongly in } L^{S)}(Q)
$$

(and analogously for other spaces) in order to avoid cumulation of unimportant parameters.

Taking the limit $n \to \infty$

For every $i = 1, \ldots, \ell$, the sequence $\{\alpha^i_{\ell n}\}_{n=1}^\infty \subset \mathcal{C}([0,T])$ is bounded due to (3.85) and uniformly equicontinuous by (3.88). Hence, using the Arzelà-Ascoli theorem (see [45, Theorem 1.5.3] or [29, Theorem IV.6.7]), for every $i = 1, \ldots, \ell$, we obtain $\alpha^i_\ell \in \mathcal{C}([0,T])$ and a subsequence (not relabelled) such that

$$
\alpha^i_{\ell n} \to \alpha^i_\ell \quad \text{strongly in } \mathcal{C}([0,T])
$$

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

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

and note that

$$\boldsymbol{v}_{\ell n} \to \boldsymbol{v}_\ell \quad \text{strongly in } \mathcal{C}([0,T]; W^{1,\infty}(\Omega)). \tag{3.95}$$

According to estimates (3.87), (3.90), (3.91), (3.93), (3.94) and Lemma 4.1, 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

$$\begin{aligned}
\partial_t \boldsymbol{v}_{\ell n} &\overset{*}{\rightharpoonup} \partial_t \boldsymbol{v}_\ell && \text{weakly* in } L^2(0,T; W^{1,\infty}(\Omega)), \\
\mathbb{B}_{\ell n} &\rightharpoonup \mathbb{B}_\ell && \text{weakly in } L^2(0,T,W^{1,2}(\Omega)), && (3.96)\\
\mathbb{B}_{\ell n} &\to \mathbb{B}_\ell && \text{strongly in } L^{2+q)}(Q) \text{ and a.e. in } Q, \\
\partial_t \mathbb{B}_{\ell n} &\rightharpoonup \partial_t \mathbb{B}_\ell && \text{weakly in } L^{\frac{q+2}{q+1}}(0,T; W^{-N,2}(\Omega)), \\
\theta_{\ell n} &\rightharpoonup \theta_\ell && \text{weakly in } L^{r+2}(0,T,W^{1,r+2}(\Omega)), && (3.97)\\
\theta_{\ell n} &\to \theta_\ell && \text{strongly in } L^{r+2+\frac{4}{d})}(Q) \text{ and a.e. in } Q, && (3.98)\\
\partial_t \theta_{\ell n} &\rightharpoonup \partial_t \theta_\ell && \text{weakly in } L^{\frac{r+2}{r+1}}(0,T; W^{-N,2}(\Omega)). && (3.99)
\end{aligned}$$

Now we explain how to take the limit in the non-linear terms appearing in (3.78), (3.79) and (3.80). To handle most of the terms, namely

$$\begin{aligned}
T = \{&\boldsymbol{v}_{\ell n} \otimes \boldsymbol{v}_{\ell n}, \quad \nu(\theta_{\ell n})\mathbb{D}\boldsymbol{v}_{\ell n}, \quad g_\omega(\mathbb{B}_{\ell n}, \theta_{\ell n})\theta_{\ell n}\mathbb{B}_{\ell n}, \quad \delta(\theta_{\ell n})\gamma(|\mathbb{B}_{\ell n} - \mathbb{I}|)(\mathbb{B}_{\ell n} - \mathbb{I}), \\
&\lambda(\theta_{\ell n})\nabla\mathbb{B}_{\ell n}, g_\omega(\mathbb{B}_{\ell n}, \theta_{\ell n})(a\mathbb{D}\boldsymbol{v}_{\ell n} + \mathbb{W}\boldsymbol{v}_{\ell n})\mathbb{B}_{\ell n}, \quad \boldsymbol{v}_{\ell n} \cdot \nabla\theta_{\ell n}, \\
&g_\omega(\mathbb{B}_{\ell n}, \theta_{\ell n})\theta_{\ell n}\mathbb{B}_{\ell n} \cdot \mathbb{D}\boldsymbol{v}_{\ell n}\},
\end{aligned}$$

we use the following generic scheme. Any term $U \in T$ can be written as $U = WS$, where

$$W \quad \text{is linear and converges weakly in } L^{q_1} \tag{3.100}$$

and

$$S \quad \text{converges strongly in } L^{q_2}, \tag{3.101}$$

where

$$\frac{1}{q_1} + \frac{1}{q_2} \le 1, \tag{3.102}$$

Indeed, to deduce a weak convergence of W, we either set it to 1, or use (3.95), (3.96), or (3.97). Further, the strong convergence of S follows from Vitali's theorem if we use pointwise convergence of $\boldsymbol{v}_{\ell n}$, $\theta_{\ell n}$, $\mathbb{B}_{\ell n}$ together with continuity of ν, δ, γ, λ and g_ω and their growth properties (3.31)–(3.35). That q_1, q_2 can be chosen so as to satisfy (3.102) is obvious from the spaces in which the convergence results (3.95)–(3.99) hold. Then, by (3.100), (3.101) and (3.102), we obtain the weak convergence of U to its appropriate limit.

At this point, we have every information needed to take the limit $n \to \infty$ in the equations (3.78) and (3.79). In (3.79), we first multiply the equation by a function $\varphi \in \mathcal{C}^1([0,T])$, integrate over $(0,T)$, then take the limit and finally use the density of functions of the form $\varphi\mathbb{A}$, $\mathbb{A} \in \operatorname{span}\{\mathbb{W}_j\}_{j=1}^\infty$, in the space $L^{(q+2)'}(0,T; W^{N,2}(\Omega))$. This way, we obtain

$$\begin{aligned}
(\partial_t \boldsymbol{v}_\ell, \boldsymbol{w}_i) - (\boldsymbol{v}_\ell \otimes \boldsymbol{v}_\ell, \nabla \boldsymbol{w}_i) &+ (2\nu(\theta_\ell)\mathbb{D}\boldsymbol{v}_\ell, \nabla \boldsymbol{w}_i) \\
&= -(2a\mu g_\omega(\mathbb{B}_\ell, \theta_\ell)\theta_\ell\mathbb{B}_\ell, \nabla \boldsymbol{w}_i) + (\boldsymbol{f}, \boldsymbol{w}_i) \qquad (3.103)\\
&\qquad \text{for every } i = 1, \dots, \ell
\end{aligned}$$

a.e. in $(0,T)$ and

$$\begin{aligned}\int_0^T \langle \partial_t \mathbb{B}_\ell, \mathbb{A}\rangle + \int_0^T (\boldsymbol{v}_\ell \cdot \nabla \mathbb{B}_\ell, \mathbb{A}) + \int_0^T (\delta(\theta_\ell)\gamma(|\mathbb{B}_\ell - \mathbb{I}|)(\mathbb{B}_\ell - \mathbb{I}), \mathbb{A}) \\ + \int_0^T (\lambda(\theta_\ell)\nabla \mathbb{B}_\ell, \nabla \mathbb{A}) = \int_0^T (2g_\omega(\mathbb{B}_\ell, \theta_\ell)(a\mathbb{D}\boldsymbol{v}_\ell + \mathbb{W}\boldsymbol{v}_\ell)\mathbb{B}_\ell, \mathbb{A}) \\ \text{for all } \mathbb{A} \in L^{(q+2)'}(0,T; W^{N,2}(\Omega)),\ \mathbb{A} = \mathbb{A}^T,\end{aligned} \tag{3.104}$$

almost everywhere in $(0,T)$. However, the space of test functions in (3.104) can be enlarged using a standard density argument. Indeed, using Hölder's inequality, it is easy to see that every term of (3.104) (taking aside the time derivative) is well defined provided that

$$\mathbb{A} \in L^2(0,T; W^{1,2}(\Omega)) \cap L^{q+2}(Q)$$

and thus, we can read from (3.104) that

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

Since we also have that $\mathbb{B}_\ell \in L^2(0,T; W^{1,2}(\Omega)) \cap L^{q+2}(Q)$, it follows from Theorem 4.2 below (with $X = W^{1,2}(\Omega)$, $Y = L^{q+2}(\Omega)$ and $H = L^2(\Omega)$) that

$$\mathbb{B}_\ell \in \mathcal{C}([0,T]; L^2(\Omega)). \tag{3.105}$$

Now let us identify $\mathbb{B}_\ell(0)$. Clearly, we can use $\mathbb{A}(t,x) = \psi(t)\mathbb{P}(x)$ in (3.104), where $\psi \in \mathcal{C}^1([0,T])$, $\psi(0) = 1$, $\psi(T) = 0$, and $\mathbb{P} \in W^{N,2}(\Omega)$, to get, after integration by parts, that

$$\begin{aligned}(\mathbb{B}_\ell(0), \mathbb{P}) = -\int_0^T \Big((\mathbb{B}_\ell, \mathbb{P})\partial_t\psi + (\boldsymbol{v}_\ell \cdot \nabla \mathbb{B}_\ell, \mathbb{P})\psi \\ + (\delta(\theta_\ell)\gamma(|\mathbb{B}_\ell - \mathbb{I}|)(\mathbb{B}_\ell - \mathbb{I}), \mathbb{P})\psi \\ - (\lambda(\theta_\ell)\nabla \mathbb{B}_\ell, \nabla \mathbb{P})\psi - (2g_\omega(\mathbb{B}_\ell, \theta_\ell)(a\mathbb{D}\boldsymbol{v}_\ell + \mathbb{W}\boldsymbol{v}_\ell)\mathbb{B}_\ell, \mathbb{P})\psi\Big).\end{aligned} \tag{3.106}$$

On the other hand, if we multiply (3.79) by ψ, integrate over $(0,T)$ and by parts in the time derivative using (3.81), we obtain

$$\begin{aligned}(Q_n\mathbb{B}_0^\omega, \mathbb{W}_j) = -\int_0^T \Big((\mathbb{B}_{\ell n}, \mathbb{W}_j)\partial_t\psi + (\boldsymbol{v}_{\ell n} \cdot \nabla \mathbb{B}_{\ell n}, \mathbb{W}_j)\psi \\ + (\delta(\theta_{\ell n})\gamma(|\mathbb{B}_{\ell n} - \mathbb{I}|)(\mathbb{B}_{\ell n} - \mathbb{I}), \mathbb{W}_j)\psi \\ - (\lambda(\theta_{\ell n})\nabla \mathbb{B}_{\ell n}, \nabla \mathbb{W}_j)\psi - (2g_\omega(\mathbb{B}_{\ell n}, \theta_{\ell n})(a\mathbb{D}\boldsymbol{v}_{\ell n} + \mathbb{W}\boldsymbol{v}_{\ell n})\mathbb{B}_{\ell n}, \mathbb{W}_j)\psi\Big).\end{aligned} \tag{3.107}$$

for every $j = 1, \ldots, n$. Then, we use completeness of $\{\mathbb{W}_j\}_{j=1}^\infty$ in $L^2(\Omega)$ and the same arguments as before to take the limit $n \to \infty$ in (3.107). This way, using also density of $\operatorname{span}\{\mathbb{W}_j\}_{j=1}^\infty$ in $W^{N,2}(\Omega)$, we get, for all $\mathbb{P} \in W^{N,2}(\Omega)$, that

$$\begin{aligned}(\mathbb{B}_0^\omega, \mathbb{P}) = -\int_0^T \Big((\mathbb{B}_\ell, \mathbb{P})\partial_t\psi + (\boldsymbol{v}_\ell \cdot \nabla \mathbb{B}_\ell, \mathbb{P})\psi + (\delta(\theta_\ell)\gamma(|\mathbb{B}_\ell - \mathbb{I}|)(\mathbb{B}_\ell - \mathbb{I}), \mathbb{P})\psi \\ - (\lambda(\theta_\ell)\nabla \mathbb{B}_\ell, \nabla \mathbb{P})\psi - (2g_\omega(\mathbb{B}_\ell, \theta_\ell)(a\mathbb{D}\boldsymbol{v}_\ell + \mathbb{W}\boldsymbol{v}_\ell)\mathbb{B}_\ell, \mathbb{P})\psi\Big).\end{aligned}$$

If we compare this with (3.106) and use density of $W^{N,2}(\Omega)$ in $L^2(\Omega)$, we deduce

$$\mathbb{B}_\ell(0) = \mathbb{B}_0^\omega \quad \text{a.e. in } \Omega. \tag{3.108}$$

We can use an analogous procedure to identify $\boldsymbol{v}_\ell(0)$. Indeed, here the situation is even simpler since (3.95) directly implies $\boldsymbol{v}_\ell \in \mathcal{C}([0,T]; W^{1,\infty}(\Omega))$ and thus, we obtain

$$\boldsymbol{v}_\ell(0) = P_\ell \boldsymbol{v}_0. \tag{3.109}$$

Our aim is now to take the limit in equation (3.80), where we need to justify the limit in the terms $\kappa(\theta_{\ell n})\nabla\theta_{\ell n}$, $|\nabla\theta_{\ell n}|^r\nabla\theta_{\ell n}$ and $2\nu(\theta_{\ell n})|\mathbb{D}\boldsymbol{v}_{\ell n}|^2$ (the other terms can be easily handled according to the scheme (3.100)–(3.102)). For the first one, we use (3.30), (3.33), (3.98) and Vitali's theorem to get

$$\sqrt{\kappa(\theta_{\ell n})} \to \sqrt{\kappa(\theta_\ell)} \quad \text{strongly in } L^{2+\frac{4}{r}}(Q) \tag{3.110}$$

and then we combine this with (3.97), to obtain

$$\sqrt{\kappa(\theta_{\ell n})}\nabla\theta_{\ell n} \rightharpoonup \sqrt{\kappa(\theta_\ell)}\nabla\theta_\ell \quad \text{weakly in } L^1(Q). \tag{3.111}$$

However, by the estimate (3.93) we know that (3.111) is valid also in $L^2(Q)$ up to a subsequence, and hence, using again (3.110), we obtain

$$\kappa(\theta_{\ell n})\nabla\theta_{\ell n} = \sqrt{\kappa(\theta_{\ell n})}\sqrt{\kappa(\theta_{\ell n})}\nabla\theta_{\ell n} \rightharpoonup \kappa(\theta_\ell)\nabla\theta_\ell \quad \text{weakly in } L^{\frac{r+2}{r+1}}(Q). \tag{3.112}$$

Next, to take the limit of the term $2\nu(\theta_{\ell n})|\mathbb{D}\boldsymbol{v}_{\ell n}|^2$, we first remark, using (3.30), (3.31), (3.98) and Vitali's theorem that

$$\nu(\theta_{\ell n}) \to \nu(\theta_\ell) \quad \text{strongly in } L^{\infty)}(Q).$$

This and

$$\mathbb{D}\boldsymbol{v}_{\ell n} \to \mathbb{D}\boldsymbol{v}_\ell \quad \text{strongly in } \mathcal{C}([0,T]; L^\infty(\Omega))$$

(cf. (3.95)) clearly proves that

$$2\nu(\theta_{\ell n})|\mathbb{D}\boldsymbol{v}_{\ell n}|^2 \to 2\nu(\theta_\ell)|\mathbb{D}\boldsymbol{v}_\ell|^2 \quad \text{strongly in } L^{\infty)}(Q). \tag{3.113}$$

Finally, due to (3.93) and reflexivity of the space $L^{(r+2)'}(Q)$, there exists $K \in L^{(r+2)'}(Q)$ such that

$$|\nabla\theta_{\ell n}|^r\nabla\theta_{\ell n} \rightharpoonup K \quad \text{weakly in } L^{(r+2)'}(Q). \tag{3.114}$$

Then, using also (3.112), (3.113) and previous convergence results, we can take the limit in (3.80) and obtain, for all $\tau \in L^{r+2}(0,T; W^{N,2}(\Omega))$, that

$$\begin{aligned} \int_0^T \langle c_v \partial_t\theta_\ell, \tau\rangle + \int_0^T (c_v \boldsymbol{v}_\ell \cdot \nabla\theta_\ell, \tau) + \int_0^T (\kappa(\theta_\ell)\nabla\theta_\ell, \nabla\tau) + \omega\int_0^T (K, \nabla\tau) \\ = \int_0^T (2\nu(\theta_\ell)|\mathbb{D}\boldsymbol{v}_\ell|^2 + 2a\mu g_\omega(\mathbb{B}_\ell, \theta_\ell)\theta_\ell\mathbb{B}_\ell \cdot \mathbb{D}\boldsymbol{v}_\ell, \tau). \end{aligned} \tag{3.115}$$

Recalling (3.99), (3.112) and (3.114), we easily conclude, using a density argument, that (3.115) is valid for all $\tau \in L^{r+2}(0,T; W^{1,r+2}(\Omega))$ and that the time derivative extends to the functional $\partial_t\theta_\ell \in L^{(r+2)'}(0,T; W^{-1,(r+2)'})$. Thus, using Theorem 4.2, we also see that

$$\theta_\ell \in \mathcal{C}([0,T]; L^2(\Omega)). \tag{3.116}$$

Furthermore, choosing $\tau = \theta_\ell$ in (3.115), rewriting the time derivative term and integrating by parts in the convective term leads to

$$\begin{aligned}\omega \int_Q K \cdot \nabla\theta_\ell = -\frac{c_v}{2}\|\theta_\ell(T)\|_2^2 + \frac{c_v}{2}\|\theta_\ell(0)\|_2^2 - \int_Q \kappa(\theta_\ell)|\nabla\theta_\ell|^2 \\ + \int_0^T (2\nu(\theta_\ell)|\mathbb{D}\boldsymbol{v}_\ell|^2 + 2a\mu g_\omega(\mathbb{B}_\ell, \theta_\ell)\theta_\ell\mathbb{B}_\ell \cdot \mathbb{D}\boldsymbol{v}_\ell, \theta_\ell).\end{aligned} \tag{3.117}$$

We use this information to identify K as follows. We note that weak lower semi-continuity and (3.111) (which is valid in $L^2(Q)$) imply

$$\int_Q \kappa(\theta_\ell)|\nabla\theta_\ell|^2 \le \liminf_{n\to\infty} \int_Q \kappa(\theta_{\ell n})|\nabla\theta_{\ell n}|^2. \tag{3.118}$$

Thus, if we integrate (3.92) over $(0, T)$ and use (3.118), (3.113), weak lower semi-continuity of $\|\cdot\|_2$ and the convergence results above to take the limes superior $n \to \infty$ and then apply (3.117), we get

$$\begin{aligned}&\omega \limsup_{n\to\infty} \int_Q |\nabla\theta_{\ell n}|^{r+2} \\ &= -\liminf_{n\to\infty} \frac{c_v}{2}\|\theta_{\ell n}(T)\|_2^2 + \frac{c_v}{2}\|\theta_0^\omega\|_2^2 - \liminf_{n\to\infty} \int_Q \kappa(\theta_{\ell n})|\nabla\theta_{\ell n}|^2 \\ &\quad + \lim_{n\to\infty} \int_0^T (2\nu(\theta_{\ell n})|\mathbb{D}\boldsymbol{v}_{\ell n}|^2 + 2a\mu g_\omega(\mathbb{B}_{\ell n}, \theta_{\ell n})\theta_{\ell n}\mathbb{B}_{\ell n} \cdot \mathbb{D}\boldsymbol{v}_{\ell n}, \theta_{\ell n}) \\ &\le -\frac{c_v}{2}\|\theta_\ell(T)\|_2^2 + \frac{c_v}{2}\|\theta_0^\omega\|_2^2 - \int_Q \kappa(\theta_\ell)|\nabla\theta_\ell|^2 \\ &\quad + \int_0^T (2\nu(\theta_\ell)|\mathbb{D}\boldsymbol{v}_\ell|^2 + 2a\mu g_\omega(\mathbb{B}_\ell, \theta_\ell)\theta_\ell\mathbb{B}_\ell \cdot \mathbb{D}\boldsymbol{v}_\ell, \theta_\ell) \\ &= \frac{c_v}{2}\|\theta_0^\omega\|_2^2 - \frac{c_v}{2}\|\theta_\ell(0)\|_2^2 + \omega \int_Q K \cdot \nabla\theta_\ell.\end{aligned} \tag{3.119}$$

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

$$\theta_\ell(t) \rightharpoonup \theta_0^\omega \quad \text{weakly in } L^2(\Omega) \tag{3.120}$$

as $t \to 0_+$ since then we can use (3.116) to conclude

$$\theta_\ell(0) = \theta_0^\omega \quad \text{a.e. in } \Omega \tag{3.121}$$

by the uniqueness of a (weak) limit. To prove (3.120), we return to (3.80), which we multiply by $\varphi \in W^{1,\infty}(0, T)$ fulfilling $\varphi(0) = 1$, $\varphi(T) = 0$ and integrate the result over $(0, T)$ to get

$$-(c_v\theta_0^\omega, w_k) - \int_0^T (c_v\theta_{\ell n}, w_k)\partial_t\varphi = \int_0^T f_n\varphi. \tag{3.122}$$

for all $k = 1, \dots, n$, where we integrated by parts in the time derivative and used the abbreviation

$$\begin{aligned}f_n = -(c_v\boldsymbol{v}_{\ell n} \cdot \nabla\theta_{\ell n}, w_k) - (\kappa(\theta_{\ell n})\nabla\theta_{\ell n} + \omega|\nabla\theta_{\ell n}|^r\nabla\theta_{\ell n}, \nabla w_k) \\ + (2\nu(\theta_{\ell n})|\mathbb{D}\boldsymbol{v}_{\ell n}|^2 + 2a\mu g_\omega(\mathbb{B}_{\ell n}, \theta_{\ell n})\theta_{\ell n}\mathbb{B}_{\ell n} \cdot \mathbb{D}\boldsymbol{v}_{\ell n}, w_k).\end{aligned}$$

It follows from the results above (cf. the derivation of (3.115)) that

$$f_n \rightharpoonup f \quad \text{weakly in } L^{(r+2)'}(0, T),$$

where
$$f = -(c_v \boldsymbol{v}_\ell \cdot \nabla\theta_\ell, w_k) - (\kappa(\theta_\ell)\nabla\theta_\ell, \nabla w_k) - \omega(K, \nabla w_k) + (2\nu(\theta_\ell)|\mathbb{D}\boldsymbol{v}_\ell|^2 + 2a\mu g_\omega(\mathbb{B}_\ell, \theta_\ell)\theta_\ell\mathbb{B}_\ell \cdot \mathbb{D}\boldsymbol{v}_\ell, w_k).$$
Thus, by taking the limit $n \to \infty$ in (3.122), we arrive at
$$-(c_v\theta_0^\omega, w_k) - \int_0^T (c_v\theta_\ell, w_k)\partial_t\varphi = \int_0^T f\varphi.$$
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
$$-(c_v\theta_0^\omega, w_k) + \frac{1}{\varepsilon}\int_t^{t+\varepsilon} (c_v\theta_\ell, w_k) = \int_0^{t+\varepsilon} f\varphi_\varepsilon.$$
Furthermore, we can take the limit $\varepsilon \to 0_+$ in this equation using (3.116) 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 (3.120) follows by exploiting the density of the set $\operatorname{span}\{w_k\}_{k=1}^\infty$ in $L^2(\Omega)$. Hence, the identity (3.121) is proved and (3.119) hereby simplifies to
$$\limsup_{n\to\infty} \int_Q |\nabla\theta_{\ell n}|^{r+2} \le \int_Q K \cdot \nabla\theta_\ell. \tag{3.123}$$
Since the operator $M : \boldsymbol{u} \mapsto |\boldsymbol{u}|^r\boldsymbol{u}$ is monotone, we have that
$$0 \le \int_Q (|\nabla\theta_{\ell n}|^r\nabla\theta_{\ell n} - |\boldsymbol{u}|^r\boldsymbol{u}) \cdot (\nabla\theta_{\ell n} - \boldsymbol{u}) \quad \text{for all } \boldsymbol{u} \in L^{r+2}(Q).$$
Thus, taking the limes superior in this inequality and using (3.123), (3.114) and (3.97), we obtain
$$\begin{aligned} 0 &\le \limsup_{n\to\infty} \int_Q |\theta_{\ell n}|^{r+2} - \lim_{n\to\infty} \int_Q |\nabla\theta_{\ell n}|^r\nabla\theta_{\ell n} \cdot \boldsymbol{u} \\ &\qquad - \lim_{n\to\infty} \int_Q |\boldsymbol{u}|^r\boldsymbol{u} \cdot \nabla\theta_{\ell n} + \int_Q |\boldsymbol{u}|^{r+2} \\ &\le \int_Q K \cdot \nabla\theta_\ell - \int_Q K \cdot \boldsymbol{u} - \int_Q |\boldsymbol{u}|^r\boldsymbol{u} \cdot \nabla\theta_\ell + \int_Q |\boldsymbol{u}|^{r+2} \\ &= \int_Q (K - |\boldsymbol{u}|^r\boldsymbol{u}) \cdot (\nabla\theta_\ell - \boldsymbol{u}). \end{aligned}$$
Then, if we choose $\boldsymbol{u} = \nabla\theta_\ell - \varepsilon\boldsymbol{\varphi}$, where $\varepsilon > 0$ and $\boldsymbol{\varphi} \in L^{r+2}(Q)$ (following the Minty method), we get
$$0 \le \int_Q (K - |\nabla\theta_\ell - \varepsilon\boldsymbol{\varphi}|^r(\nabla\theta_\ell - \varepsilon\boldsymbol{\varphi})) \cdot \boldsymbol{\varphi}$$

after dividing by ε. Using continuity of the operator M and the dominated convergence theorem, we arrive at

$$0 \le \int_Q (K - |\nabla\theta_\ell|^r \nabla\theta_\ell) \cdot \boldsymbol{\varphi}.$$

Since this holds for arbitrary $\boldsymbol{\varphi} \in L^{r+2}(Q)$ (and thus even with equality), we find

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

Hence, we proved that

$$\begin{aligned} \int_0^T \langle c_v \partial_t \theta_\ell, \tau\rangle + \int_0^T (c_v \boldsymbol{v}_\ell \cdot \nabla\theta_\ell, \tau) + \int_0^T (\kappa(\theta_\ell)\nabla\theta_\ell + \omega|\nabla\theta_\ell|^r \nabla\theta_\ell, \nabla\tau) \\ = \int_0^T (2\nu(\theta_\ell)|\mathbb{D}\boldsymbol{v}_\ell|^2 + 2a\mu g_\omega(\mathbb{B}_\ell, \theta_\ell)\theta_\ell \mathbb{B}_\ell \cdot \mathbb{D}\boldsymbol{v}_\ell, \tau) \end{aligned} \tag{3.124}$$

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

3.5.3 Positive definiteness of $\mathbb{B}_\ell$ and positivity of θ_ℓ

Here we closely follow the method developed in [6] (cf. (5.52) in Chapter 5), i.e., we use

$$\mathbb{A}_{\boldsymbol{x}} = \chi_{(0,t)}(\mathbb{B}_\ell \boldsymbol{x} \cdot \boldsymbol{x} - \omega|\boldsymbol{x}|^2)_- \, \boldsymbol{x} \otimes \boldsymbol{x},$$

in (3.104), where $\boldsymbol{x} \in \mathbb{R}^d$, $t \in (0,T)$ and

$$f_+ = \max\{0, f\}, \qquad f_- = \min\{0, f\}.$$

Note that, since $\mathbb{B}_\ell \in L^2(0,T; W^{1,2}(\Omega)) \cap L^{q+2}(Q)$ and $\boldsymbol{x}$ is a constant vector, the function $\mathbb{A}_{\boldsymbol{x}}$ belongs to the same space, and is thus a valid test function in (3.104). The key property of $\mathbb{A}_{\boldsymbol{x}}$ is that it vanishes whenever the smallest eigenvalue of $\mathbb{B}_\ell$ is greater than ω (since $\mathbb{B}_\ell \boldsymbol{y} \cdot \boldsymbol{y} \ge \omega|\boldsymbol{y}|^2$ for all $\boldsymbol{y} \in \mathbb{R}^d$ in such a case). Thus, we have

$$(\Lambda(\mathbb{B}_\ell) - \omega)_+ (\mathbb{B}_\ell \boldsymbol{x} \cdot \boldsymbol{x} - \omega|\boldsymbol{x}|^2)_- = 0,$$

which implies

$$g_\omega(\mathbb{B}_\ell, \theta_\ell)\mathbb{A}_x = 0 \quad \text{a.e. in } Q. \tag{3.125}$$

Let us now evaluate separately the terms arising from the choice $\mathbb{A} = \mathbb{A}_{\boldsymbol{x}}$ in (3.104). For the time derivative, we write

$$\begin{aligned} \int_0^T \langle \partial_t \mathbb{B}_\ell, \mathbb{A}_{\boldsymbol{x}}\rangle &= \int_0^t \Big\langle \partial_t(\mathbb{B}_\ell \boldsymbol{x} \cdot \boldsymbol{x} - \omega|\boldsymbol{x}|^2), (\mathbb{B}_\ell \boldsymbol{x} \cdot \boldsymbol{x} - \omega|\boldsymbol{x}|^2)_- \Big\rangle \\ &= \frac{1}{2}\Big\|(\mathbb{B}_\ell \boldsymbol{x} \cdot \boldsymbol{x} - \omega|\boldsymbol{x}|^2)_-(t)\Big\|_2^2 - \frac{1}{2}\Big\|(\mathbb{B}_\ell \boldsymbol{x} \cdot \boldsymbol{x} - \omega|\boldsymbol{x}|^2)_-(0)\Big\|_2^2 \\ &= \frac{1}{2}\Big\|(\mathbb{B}_\ell \boldsymbol{x} \cdot \boldsymbol{x} - \omega|\boldsymbol{x}|^2)_-(t)\Big\|_2^2, \end{aligned}$$

where the first equality is a consequence of the linearity of the weak time derivative, the second one follows from Lemma 4.4 since the function $s \mapsto s_-$ is Lipschitz and

$$\int_0^v s_- \, \mathrm{d}s = \frac{1}{2}(v_-)^2, \quad v \in \mathbb{R},$$

and the third equality follows from (3.108) and (3.70). Furthermore, using integration by parts and the facts that $\boldsymbol{v}_\ell \cdot \boldsymbol{n} = 0$ on $\partial\Omega$ and $\operatorname{div} \boldsymbol{v}_\ell = 0$ in Q, we get

$$\begin{aligned}\int_0^T (\boldsymbol{v}_\ell \cdot \nabla \mathbb{B}_\ell, \mathbb{A}_{\boldsymbol{x}}) &= \int_0^t (\boldsymbol{v}_\ell \cdot \nabla (\mathbb{B}_\ell \boldsymbol{x}\cdot\boldsymbol{x} - \omega|\boldsymbol{x}|^2), (\mathbb{B}_\ell \boldsymbol{x}\cdot\boldsymbol{x} - \omega|\boldsymbol{x}|^2)_-) \\ &= \frac{1}{2}\int_0^t \int_{\partial\Omega} ((\mathbb{B}_\ell \boldsymbol{x}\cdot\boldsymbol{x} - \omega|\boldsymbol{x}|^2)_-)^2 \boldsymbol{v}_\ell \cdot \boldsymbol{n} = 0\end{aligned}$$

and also

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

Moreover, since $\omega < 1$ and the functions δ, γ are non-negative, we also obtain

$$\begin{aligned}&\int_0^T (\delta(\theta_\ell)\gamma(|\mathbb{B}_\ell - \mathbb{I}|)(\mathbb{B}_\ell - \mathbb{I}), \mathbb{A}_{\boldsymbol{x}}) \\ &\qquad = \int_0^t \int_\Omega \delta(\theta_\ell)\gamma(|\mathbb{B}_\ell - \mathbb{I}|)(\mathbb{B}_\ell \boldsymbol{x}\cdot\boldsymbol{x} - |\boldsymbol{x}|^2)(\mathbb{B}_\ell \boldsymbol{x}\cdot\boldsymbol{x} - \omega|\boldsymbol{x}|^2)_- \geq 0.\end{aligned}$$

In addition, the right hand side of (3.104) vanishes due to (3.125). Thus, using the above computation in (3.104), we obtain

$$\left\| (\mathbb{B}_\ell \boldsymbol{x}\cdot\boldsymbol{x} - \omega|\boldsymbol{x}|^2)_-(t) \right\|_2^2 \leq 0$$

for all $t \in (0,T)$ (recall (3.105)), whence

$$\mathbb{B}_\ell(t)\boldsymbol{x}\cdot\boldsymbol{x} \geq \omega|\boldsymbol{x}|^2 \quad \text{a.e. in } \Omega, \text{ for all } t \in (0,T) \text{ and for every } \boldsymbol{x} \in \mathbb{R}^d. \tag{3.126}$$

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

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

Also, using the identity

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

(see (4.38)) and (3.90) we conclude that $\mathbb{B}_\ell^{-1}$ exists a.e. in Q and satisfies

$$\mathbb{B}_\ell^{-1} \in L^\infty(0,T; L^\infty(\Omega)) \cap L^2(0,T; W^{1,2}(\Omega)). \tag{3.127}$$

Moreover, we define

$$\psi_2(\mathbb{B}_\ell) = \operatorname{tr} \mathbb{B}_\ell - d - \ln \det \mathbb{B}_\ell$$

and observe, using (4.36), (4.37) and simple inequalities

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

that

$$\begin{aligned}0 \leq \psi_2(\mathbb{B}_\ell) &\leq \sqrt{d}|\mathbb{B}_\ell| - d + d\left(|\det \mathbb{B}_\ell|^{\frac{1}{d}} + |\det \mathbb{B}_\ell|^{-\frac{1}{d}}\right) \\ &\leq \sqrt{d}|\mathbb{B}_\ell| - d + C\left(|\mathbb{B}_\ell| + \frac{1}{\omega}\right) \\ &\leq C|\mathbb{B}_\ell| + \frac{C}{\omega}\end{aligned}$$

and also, using (4.40), that

$$|\nabla\psi_2(\mathbb{B}_\ell)| = |(\mathbb{I}-\mathbb{B}_\ell^{-1})\cdot\nabla\mathbb{B}_\ell| \le C\left(1+\frac{1}{\omega}\right)|\nabla\mathbb{B}_\ell|.$$

Hence, we conclude

$$\psi_2(\mathbb{B}_\ell) \in L^2(0,T;W^{1,2}(\Omega)) \cap L^{q+2}(Q).$$

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

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

as a test function in (3.124) to get

$$\begin{aligned}
&\frac{c_v}{2}\|(\theta_\ell-\omega)_-(t)\|_2^2 - \frac{c_v}{2}\|(\theta_\ell-\omega)_-(0)\|_2^2 \\
&\quad + \int_0^t \left\|\sqrt{\kappa(\theta_\ell)}\nabla(\theta_\ell-\omega)_-\right\|_2^2 + \int_0^t \|\nabla(\theta_\ell-\omega)_-\|_{r+2}^{r+2} \\
&\qquad = \int_0^t \Big(2\nu(\theta_\ell)|\mathbb{D}\boldsymbol{v}_\ell|^2 + 2a\mu g_\omega(\mathbb{B}_\ell,\theta_\ell)\theta_\ell\mathbb{B}_\ell\cdot\mathbb{D}\boldsymbol{v}_\ell, (\theta_\ell-\omega)_-\Big). \\
&\qquad \le 0
\end{aligned}$$

Hence, using $\theta_\ell(0) = \theta_0^\omega \ge \omega$ in Ω and (3.116), we obtain that

$$\|(\theta_\ell(t)-\omega)_-\|_2 = 0 \quad \text{for all } t\in(0,T),$$

which means

$$\theta_\ell(t) \ge \omega \quad \text{a.e. in } \Omega \text{ and for all } t\in(0,T). \tag{3.128}$$

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

$$\theta_\ell^{-1} \in L^\infty(0,T;L^\infty(\Omega)) \cap L^{r+2}(0,T;W^{1,r+2}(\Omega)). \tag{3.129}$$

From these findings we also easily read that

$$|\ln\theta_\ell| \le \theta_\ell + \frac{1}{\theta_\ell} \le \theta_\ell + \frac{1}{\omega}$$

and

$$|\nabla\ln\theta_\ell| = \frac{|\nabla\theta_\ell|}{\theta_\ell} \le \frac{1}{\omega}|\nabla\theta_\ell|,$$

hence

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

3.5.4 Entropy equation

In order to take the remaining limits $\ell\to\infty$ and $\omega\to 0_+$, we need to replace (3.124) by the equation for entropy, whose terms are easier to handle. From this equation, we then deduce that $\det\mathbb{B}_\ell$ and θ_ℓ remain *strictly* positive a.e. in Q.

First, we rewrite (3.124) in the form

$$\begin{aligned}
\langle c_v\partial_t\theta_\ell,\tau\rangle + (c_v\boldsymbol{v}_\ell\cdot\nabla\theta_\ell,\tau) + (\kappa(\theta_\ell)\nabla\theta_\ell + \omega|\nabla\theta_\ell|^r\nabla\theta_\ell,\nabla\tau)& \\
= (2\nu(\theta_\ell)|\mathbb{D}\boldsymbol{v}_\ell|^2 + 2a\mu g_\omega(\mathbb{B}_\ell,\theta_\ell)\theta_\ell\mathbb{B}_\ell\cdot\mathbb{D}\boldsymbol{v}_\ell,\tau)&
\end{aligned} \tag{3.130}$$

for all $\tau \in W^{1,r}(\Omega)$ and a.e. in $(0,T)$. Then, we take $\phi \in W^{1,\infty}(\Omega)$ and note that $\tau = \theta_\ell^{-1}\phi$ can be used as a test function in (3.130) thanks to (3.129). This way, we get

$$\begin{aligned}\Big\langle c_v\partial_t\theta_\ell, \frac{\phi}{\theta_\ell}\Big\rangle &+ (c_v\boldsymbol{v}_\ell \cdot \nabla \ln\theta_\ell, \phi)\\ &+ (\kappa(\theta_\ell)\nabla\ln\theta_\ell, \nabla\phi) - (\kappa(\theta_\ell)|\nabla\ln\theta_\ell|^2, \phi)\\ &+ \omega(|\nabla\theta_\ell|^r\nabla\ln\theta_\ell, \nabla\phi) - \omega(|\nabla\theta_\ell|^r|\nabla\ln\theta_\ell|^2, \phi)\\ &\quad = \Big(\frac{2\nu(\theta_\ell)}{\theta_\ell}|\mathbb{D}\boldsymbol{v}_\ell|^2 + 2a\mu g_\omega(\mathbb{B}_\ell,\theta_\ell)\mathbb{B}_\ell \cdot \mathbb{D}\boldsymbol{v}_\ell, \phi\Big)\end{aligned} \tag{3.131}$$

a.e. in $(0,T)$. Similarly, we observe that $\mu(\mathbb{I} - \mathbb{B}_\ell^{-1})\phi$ is a valid test function in (the localized version of) (3.104) due to (3.127). Thus, using the same algebraic manipulations as those leading to (3.13) in Section 3.2, we obtain

$$\begin{aligned}&\Big\langle \partial_t\mathbb{B}_\ell, \mu(\mathbb{I} - \mathbb{B}_\ell^{-1})\phi\Big\rangle + (\mu\boldsymbol{v}_\ell \cdot \nabla\psi_2(\mathbb{B}_\ell), \phi)\\ &\quad + (\mu\delta(\theta_\ell)\gamma(|\mathbb{B}_\ell - \mathbb{I}|)|\mathbb{B}_\ell^{\frac12} - \mathbb{B}_\ell^{-\frac12}|^2, \phi) + (\mu\lambda(\theta_\ell)|\mathbb{B}_\ell^{-\frac12}\nabla\mathbb{B}_\ell\mathbb{B}_\ell^{-\frac12}|^2, \phi)\\ &\quad\quad = -(\mu\lambda(\theta_\ell)\nabla\psi_2(\mathbb{B}_\ell), \nabla\phi) + (2a\mu g_\omega(\mathbb{B}_\ell,\theta_\ell)\mathbb{B}_\ell \cdot \mathbb{D}\boldsymbol{v}_\ell, \phi)\end{aligned} \tag{3.132}$$

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

$$\eta_\ell = c_v \ln\theta_\ell - \mu\psi_2(\mathbb{B}_\ell) \tag{3.133}$$

and

$$\begin{aligned}\xi_\ell = \frac{2\nu(\theta_\ell)}{\theta_\ell}|\mathbb{D}\boldsymbol{v}_\ell|^2 &+ \kappa(\theta_\ell)|\nabla\ln\theta_\ell|^2 + \omega|\nabla\theta_\ell|^r|\nabla\ln\theta_\ell|^2\\ &+ \mu\delta(\theta_\ell)\gamma(|\mathbb{B}_\ell - \mathbb{I}|)|\mathbb{B}_\ell^{\frac12} - \mathbb{B}_\ell^{-\frac12}|^2 + \mu\lambda(\theta_\ell)|\mathbb{B}_\ell^{-\frac12}\nabla\mathbb{B}_\ell\mathbb{B}_\ell^{-\frac12}|^2\end{aligned}$$

and subtract (3.132) from (3.131), we get

$$\begin{aligned}&\Big\langle c_v\partial_t\theta_\ell, \frac{\phi}{\theta_\ell}\Big\rangle - \Big\langle \partial_t\mathbb{B}_\ell, \mu(\mathbb{I} - \mathbb{B}_\ell^{-1})\phi\Big\rangle + (\boldsymbol{v}_\ell \cdot \nabla\eta_\ell, \phi)\\ &\quad + \Big((\kappa(\theta_\ell) + \omega|\nabla\theta_\ell|^r)\nabla\ln\theta_\ell - \mu\lambda(\theta_\ell)\nabla\psi_2(\mathbb{B}_\ell), \nabla\phi\Big) = (\xi_\ell, \phi)\end{aligned} \tag{3.134}$$

a.e. in $(0,T)$ and for all $\phi \in W^{1,\infty}(\Omega)$.

Obviously, we need 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 (3.128) and $\omega < 1$, 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, if we apply Lemma 4.4, we get

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

by taking the (weak) time derivative of (4.32). Hence, if we multiply this by $\varphi \in W^{1,\infty}(0,T)$ with $\varphi(T) = 0$, integrate over $(0,T)$ and by parts, we are led to

$$\int_0^T \Big\langle c_v\partial_t\theta_\ell, \frac{\phi}{\theta_\ell}\Big\rangle\varphi = -\int_0^T\int_\Omega c_v\ln\theta_\ell\phi\partial_t\varphi - \int_\Omega c_v\ln\theta_0^\omega\phi\varphi(0), \tag{3.135}$$

where we also used (3.121).

Analogous ideas can be used to rewrite the second term of (3.134). However, we can not apply Lemma 4.4 directly since the duality $\left\langle \partial_t \mathbb{B}_\ell, (\mathbb{I} - \mathbb{B}_\ell^{-1})\phi \right\rangle$ can not be interpreted entry-wise. Thus, let us proceed more explicitly. We apply Theorem 4.2 to obtain functions $\mathbb{B}_\ell^\varepsilon \in \mathcal{C}^1([0,T]; W^{1,2}(\Omega) \cap L^{q+2}(\Omega))$ such that

$$\|\mathbb{B}_\ell^\varepsilon - \mathbb{B}_\ell\|_{L^2W^{1,2}\cap L^{q+2}L^{q+2}} + \|\partial_t \mathbb{B}_\ell^\varepsilon - \partial_t \mathbb{B}_\ell\|_{L^2W^{-1,2}+L^{\frac{q+2}{q+1}}L^{\frac{q+2}{q+1}}} \to 0 \tag{3.136}$$

as $\varepsilon \to 0_+$ and also

$$\Lambda(\mathbb{B}_\ell^\varepsilon) \geq \omega \quad \text{a.e. in } Q.$$

Since $\mathbb{B}_\ell \in \mathcal{C}([0,T]; L^2(\Omega))$ (cf. (3.105)), we know that

$$\|\mathbb{B}_\ell^\varepsilon - \mathbb{B}_\ell\|_2 \rightrightarrows 0 \quad \text{uniformly in } [0,T]. \tag{3.137}$$

Furthermore, using (4.38) and (3.127), we can write, for any $\phi \in W^{1,\infty}(\Omega)$, that

$$\begin{aligned}\left|\nabla((\mathbb{I} - (\mathbb{B}_\ell^\varepsilon)^{-1})\phi)\right| &= \left|(\mathbb{B}_\ell^\varepsilon)^{-1}\nabla\mathbb{B}_\ell^\varepsilon(\mathbb{B}_\ell^\varepsilon)^{-1}\phi + (\mathbb{I} - (\mathbb{B}_\ell^\varepsilon)^{-1})\nabla\phi\right| \\ &\leq \frac{C}{\omega^2}|\nabla\mathbb{B}_\ell^\varepsilon||\phi| + \left(1 + \frac{C}{\omega}\right)|\nabla\phi|\end{aligned}$$

and thus, we eventually obtain that

$$(\mathbb{I} - (\mathbb{B}_\ell^\varepsilon)^{-1})\phi \rightharpoonup (\mathbb{I} - \mathbb{B}_\ell^{-1})\phi \quad \text{weakly in } L^2(0,T; W^{1,2}(\Omega)) \cap L^{q+2}(0,T; L^{q+2}(\Omega)).$$

By applying this with (3.136), we get, for all $\varphi \in W^{1,\infty}(0,T)$, $\varphi(T) = 0$, that

$$\begin{aligned}&\left|\int_0^T \left\langle \partial_t\mathbb{B}_\ell^\varepsilon, \mu(\mathbb{I} - (\mathbb{B}_\ell^\varepsilon)^{-1})\phi\right\rangle\varphi - \int_0^T \left\langle \partial_t\mathbb{B}_\ell, \mu(\mathbb{I} - \mathbb{B}_\ell^{-1})\phi\right\rangle\varphi\right| \\ &\quad\leq \int_0^T \left|\left\langle \partial_t\mathbb{B}_\ell^\varepsilon - \partial_t\mathbb{B}_\ell, \mu(\mathbb{I} - (\mathbb{B}_\ell^\varepsilon)^{-1})\phi\right\rangle\right||\varphi| \\ &\qquad + \left|\int_0^T \left\langle \partial_t\mathbb{B}_\ell\varphi, \mu(\mathbb{I} - (\mathbb{B}_\ell^\varepsilon)^{-1})\phi - \mu(\mathbb{I} - \mathbb{B}_\ell^{-1})\phi\right\rangle\right| \\ &\quad\to 0 \quad \text{as } \varepsilon \to 0_+.\end{aligned} \tag{3.138}$$

On the other hand, using $\partial_t\mathbb{B}_\ell^\varepsilon \in \mathcal{C}([0,T]; W^{1,2}(\Omega) \cap L^{q+2}(\Omega))$ and (4.40), we find

$$\begin{aligned}\int_0^T \left\langle \partial_t\mathbb{B}_\ell^\varepsilon, \mu(\mathbb{I} - (\mathbb{B}_\ell^\varepsilon)^{-1})\phi\right\rangle\varphi &= \int_0^T (\mu\partial_t\psi_2(\mathbb{B}_\ell^\varepsilon), \phi)\varphi \\ &= -\int_\Omega \mu\psi_2(\mathbb{B}_\ell^\varepsilon(0))\phi\varphi(0) - \int_0^T\int_\Omega \mu\psi_2(\mathbb{B}_\ell^\varepsilon)\phi\partial_t\varphi\end{aligned} \tag{3.139}$$

To take the limit in the last two terms, let us first remark that the set

$$\left\{\mathbb{A} \in \mathbb{R}^{d\times d} : \ \mathbb{A}\boldsymbol{x}\cdot\boldsymbol{x} \geq \omega \quad \text{for all } \boldsymbol{x} \in \mathbb{R}^d\right\}$$

is convex in $\mathbb{R}^{d\times d}$ (as it is defined by a linear constraint). This, (3.126) and (4.37) imply, for any $s \in (0,1)$, that

$$\begin{aligned}\left|(\mathbb{B}_\ell + s(\mathbb{B}_\ell^\varepsilon - \mathbb{B}_\ell))^{-1}\right| &\leq \operatorname{tr}(\mathbb{B}_\ell + s(\mathbb{B}_\ell^\varepsilon - \mathbb{B}_\ell))^{-1} \\ &\leq \frac{1}{d\Lambda(\mathbb{B}_\ell + s(\mathbb{B}_\ell^\varepsilon - \mathbb{B}_\ell))} \leq \frac{1}{d\omega}\end{aligned}$$

a.e. in Ω. Thus, by the mean value theorem, (4.40) and (3.137), we get

$$\begin{aligned}\int_\Omega |\psi_2(\mathbb{B}_\ell^\varepsilon) - \psi_2(\mathbb{B}_\ell)|^2 \\ &= \int_\Omega \left| \int_0^1 (\mathbb{I} - (\mathbb{B}_\ell + s(\mathbb{B}_\ell^\varepsilon - \mathbb{B}_\ell))^{-1}) \cdot (\mathbb{B}_\ell^\varepsilon - \mathbb{B}_\ell) \,\mathrm{d}s \right|^2 \\ &\le C\left(1 + \frac{1}{\omega}\right) \|\mathbb{B}_\ell^\varepsilon - \mathbb{B}_\ell\|_2^2 \rightrightarrows 0 \quad \text{uniformly in } [0,T]\end{aligned} \tag{3.140}$$

as $\varepsilon \to 0_+$. Similarly, using (3.105) and (3.108), we can also show that

$$\psi_2(\mathbb{B}_\ell) \in \mathcal{C}(0,T; L^2(\Omega)), \quad \psi_2(\mathbb{B}_\ell(0)) = \psi_2(\mathbb{B}_0^\omega). \tag{3.141}$$

Using this and (3.140), we take the limit in (3.139) and then compare to (3.138) to obtain

$$\int_0^T \left\langle \partial_t \mathbb{B}_\ell, \mu(\mathbb{I} - \mathbb{B}_\ell^{-1})\phi \right\rangle \varphi = -\int_\Omega \mu\psi_2(\mathbb{B}_0^\omega)\phi\varphi(0) - \int_0^T \int_\Omega \mu\psi_2(\mathbb{B}_\ell)\phi\partial_t\varphi \tag{3.142}$$

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

Finally, if we subtract (3.142) from (3.135) and use (3.133), we can rewrite (3.134) as

$$\begin{aligned}&-\int_0^T (\eta_\ell, \phi)\partial_t\varphi - (\eta_0^\omega, \phi)\varphi(0) - \int_0^T (\boldsymbol{v}_\ell \eta_\ell, \nabla\phi)\varphi \\ &+ \int_0^T \Big((\kappa(\theta_\ell) + \omega|\nabla\theta_\ell|^r)\nabla \ln\theta_\ell - \mu\lambda(\theta_\ell)\nabla\psi_2(\mathbb{B}_\ell), \nabla\phi\Big)\varphi = \int_0^T (\xi_\ell, \phi)\varphi\end{aligned} \tag{3.143}$$

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

$$\eta_0^\omega = c_v \ln\theta_0^\omega - \mu\psi_2(\mathbb{B}_0^\omega).$$

Moreover, since $\ln\theta_\ell \in \mathcal{C}([0,T]; L^2(\Omega))$ and (3.141) hold, we easily read

$$\eta_\ell \in \mathcal{C}([0,T]; L^2(\Omega)), \quad \eta_\ell(0) = \eta_0^\omega. \tag{3.144}$$

3.5.5 Total energy equality

The integrated version of the total energy equality is important in the derivation of the apriori estimates below.

Let

$$E_\ell = \frac{1}{2}|\boldsymbol{v}_\ell|^2 + c_v\theta_\ell.$$

We multiply the i-th equation in (3.103) by $(\boldsymbol{v}_\ell, \boldsymbol{w}_i)$, sum up the result over $i = 1, \ldots, \ell$ and then we add (3.124) with $\tau = 1$. This way, after several cancellations using also (3.69), we obtain

$$\frac{\mathrm{d}}{\mathrm{d}t}\int_\Omega E_\ell = (\boldsymbol{f}, \boldsymbol{v}_\ell) \tag{3.145}$$

a.e. in $(0,T)$.

3.5.6 Final limit

We start by derivation of an estimate that is uniform with respect to ℓ. In fact, since we derive this estimate without appealing to the properties of g_ω (except for $0 \leq g_\omega \leq 1$), it is uniform also with respect to ω. This saves us some work in exchange for the fact that this estimate is not optimal for the ω-approximated system.

Uniform estimates

As we explained several times before, the key to the uniform estimates is the entropy (in)equality used together with the energy (in)equality. Let us first show that the total energy of the fluid remains bounded. In (3.145), we apply Young's inequality, (3.45) and $\theta_\ell > 0$, to estimate

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

a.e. in $(0,T)$. Hence, by the Gronwall inequality (see e.g. [31, B.2.]), we get

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

Then, we apply (3.109), (3.121) to identify that

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

and if we use (3.76), (3.72), (3.45), we arrive at

$$\|\theta_\ell\|_{L^\infty L^1} + \|\boldsymbol{v}_\ell\|_{L^\infty L^2} \leq C\|E_\ell\|_{L^\infty L^1} \leq C. \tag{3.146}$$

Now we turn our attention to (3.143), which we need to localize in time. To this end, we want to multiply (3.143) by $\varphi = \chi_{(0,t)}$ with $t \in (0,T)$. However, such a test function is not admissible in (3.143) and therefore, we approximate it in a standard way as follows. Fix $\phi \in W^{1,\infty}(\Omega)$, and let us define

$$\begin{aligned}
u &= \int_\Omega \eta_\ell\phi \in \mathcal{C}([0,T]) \quad (\text{cf. } (3.144)) \\
\boldsymbol{j}_\ell &= -\boldsymbol{v}_\ell\eta_\ell + (\kappa(\theta_\ell) + \omega|\nabla\theta_\ell|^r)\nabla\ln\theta_\ell - \lambda(\theta_\ell)\nabla\psi_2(\mathbb{B}_\ell) \in L^1(Q) \\
v &= \int_\Omega (\boldsymbol{j}_\ell\cdot\nabla\phi - \xi_\ell\phi) \in L^1(0,T).
\end{aligned} \tag{3.147}$$

Fix $t \in (0,T)$ and let

$$\varphi_k(s) = \begin{cases} 1 & s \leq t \\ (t-s)k+1 & s \in (t, t+\frac{1}{k}) \\ 0 & s \geq t+\frac{1}{k} \end{cases}$$

Then, for any $k > \frac{1}{T-t}$, we have $\varphi_k \in W^{1,\infty}(\Omega)$ and $\varphi_k(T) = 0$. Hence, by considering φ_k in (3.143), we obtain

$$
\begin{aligned}
0 &= -\int_Q \eta_\ell \partial_t \varphi_k \phi - \int_\Omega \eta_0^\omega \varphi_k(0)\phi + \int_Q \varphi_k \boldsymbol{j}_\ell \cdot \nabla\phi - \int_Q \xi_\ell \varphi_k \phi \\
&= k\int_t^{t+\frac{1}{k}} \int_\Omega \eta_\ell \phi - \int_\Omega \eta_0^\omega \phi + \int_Q (\boldsymbol{j}_\ell \cdot \nabla\phi - \xi_\ell \phi)\varphi_k \\
&= k\int_t^{t+\frac{1}{k}} u - \int_\Omega \eta_0^\omega \phi + \int_0^{t+\frac{1}{k}} v\varphi_k \\
&= u(t) - \int_\Omega \eta_0^\omega \phi + \int_0^t v + k\int_t^{t+\frac{1}{k}} (u - u(t)) + \int_t^{t+\frac{1}{k}} v\varphi_k .
\end{aligned}
\tag{3.148}
$$

Since

$$
\left| k\int_t^{t+\frac{1}{k}} (u-u(t)) \right| \le k\int_t^{t+\frac{1}{k}} |u-u(t)| \to 0 \quad \text{as } k\to\infty \tag{3.149}
$$

by the continuity of u and

$$
\left| \int_t^{t+\frac{1}{k}} v\varphi_k \right| \le \int_t^{t+\frac{1}{k}} |v| \to 0 \quad \text{as } k\to\infty
$$

by the absolute continuity of Lebesgue integral, we obtain from (3.148) that

$$
\int_\Omega \eta_\ell(t)\phi + \int_0^t \int_\Omega \boldsymbol{j}_\ell \cdot \nabla\phi = \int_\Omega \eta_0^\omega \phi + \int_0^t \int_\Omega \xi_\ell \phi \quad \text{for all } \phi \in W^{1,\infty}(\Omega) \tag{3.150}
$$

and all $t \in (0,T)$ (in fact, for all $t \in [0,T]$ due to continuity of both sides of (3.150)). In particular, taking $\phi = 1$,[3] we deduce, using $\xi_\ell \ge 0$, that the function $t \mapsto \int_\Omega \eta_\ell(t)$ is non-decreasing, and thus

$$
\int_Q \xi_\ell = \max_{t\in[0,T]} \int_0^t \int_\Omega \xi_\ell = \max_{t\in[0,T]} \int_\Omega \eta_\ell(t) - \int_\Omega \eta_0^\omega = \int_\Omega \eta_\ell(T) - \int_\Omega \eta_0^\omega . \tag{3.151}
$$

Then, using (3.133), the inequalities

$$
\ln x \le x - 1 \quad \text{for all } x > 0 \tag{3.152}
$$

and

$$
\mu\psi_2(\mathbb{B}_\ell) = \mu(\operatorname{tr}\mathbb{B}_\ell - d - \ln\det\mathbb{B}_\ell) \ge 0 \tag{3.153}
$$

(see (4.36)), assumption (3.46) and (3.146) (recall also (3.116)), we obtain

$$
\int_Q \xi_\ell \le \int_\Omega (c_v \ln\theta_\ell(T) - \mu\psi_2(\mathbb{B}_\ell(T))) + C \le C\int_\Omega (\theta_\ell(T) - 1) + C \le C,
$$

hence

$$
\|\xi_\ell\|_{L^1L^1} \le C. \tag{3.154}
$$

Also, it is easy to see using (3.151), (3.73), (3.72), (3.45) and (3.46) that

$$
\|\eta_\ell\|_{L^\infty L^1} \le C. \tag{3.155}
$$

[3] The case with $\phi \neq 1$ is used for the identification of initial conditions below

Estimate (3.154) implies, using (3.31), (3.34) and (3.35), that

$$\begin{aligned}\left\|\frac{1}{\sqrt{\theta_\ell}}\mathbb{D}\boldsymbol{v}_\ell\right\|_{L^2L^2} + \left\|\sqrt{\kappa(\theta_\ell)}\nabla\ln\theta_\ell\right\|_{L^2L^2} + \omega\left\||\nabla\theta_\ell|^{\frac{r}{2}}\nabla\ln\theta_\ell\right\|_{L^2L^2} \\ + \left\|\sqrt{\gamma(|\mathbb{B}_\ell-\mathbb{I}|)}(\mathbb{B}_\ell^{\frac{1}{2}}-\mathbb{B}_\ell^{-\frac{1}{2}})\right\|_{L^2L^2} + \left\|\mathbb{B}_\ell^{-\frac{1}{2}}\nabla\mathbb{B}_\ell\mathbb{B}_\ell^{-\frac{1}{2}}\right\|_{L^2L^2} \le C.\end{aligned} \tag{3.156}$$

In what follows, we improve the uniform estimate (3.156) considerably by choosing appropriate test functions in (3.104) and (3.124) and then using (A1) and the definitions of p, R, σ to estimate the right hand sides.

From (3.156) and (3.36) one can deduce, using Young's inequality that

$$\|\mathbb{B}_\ell\|_{L^{q+1}L^{q+1}} \le C.$$

However, we can obtain better information (cf. (3.49)) as follows. For any $K > d$ let us define the truncation function T_K as

$$T_K(s) = \min\{\max\{s,\omega\},K\}, \quad s \ge 0,$$

and then we also define

$$\phi(s) = (T_K(s))^{\sigma-1}, \quad s \ge 0.$$

As T_K is a Lipschitz function bounded from below by ω, from above by K and

$$\phi'(s) = (\sigma-1)(T_K(s))^{\sigma-2}T_K'(s),$$

the function ϕ is Lipschitz as well. Then, since $\mathbb{B}_\ell \in L^2(0,T;W^{1,2}(\Omega)) \cap L^{q+2}(Q)$, we obtain

$$\phi(\operatorname{tr}\mathbb{B}_\ell) \in L^2(0,T;W^{1,2}(\Omega)) \cap L^\infty(0,T;L^\infty(\Omega))$$

by a standard result (see e.g. [75, Theorem 2.1.11.]). Thus, we see that

$$\mathbb{A}_K = \phi(\operatorname{tr}\mathbb{B}_\ell)\mathbb{I} = (\min\{\operatorname{tr}\mathbb{B}_\ell, K\})^{\sigma-1}\mathbb{I} \tag{3.157}$$

is a valid test function in (3.104) (recall that $\operatorname{tr}\mathbb{B}_\ell \ge \Lambda(\mathbb{B}_\ell) \ge \omega$). Moreover, it follows from Lemma 4.4 that

$$\langle\partial_t\mathbb{B}_\ell, \mathbb{A}_K\rangle = \left\langle\partial_t\operatorname{tr}\mathbb{B}_\ell, T_K(\operatorname{tr}\mathbb{B}_\ell)^{\sigma-1}\right\rangle = \frac{\mathrm{d}}{\mathrm{d}t}\int_\Omega \Phi(\operatorname{tr}\mathbb{B}_\ell), \tag{3.158}$$

where

$$\Phi(s) = \int_\omega^s \phi(s).$$

Furthermore, noting that ϕ is non-decreasing as $\sigma > 1$, we have

$$\begin{aligned}(\nabla\mathbb{B}_\ell, \nabla\mathbb{A}_K) = (\nabla\operatorname{tr}\mathbb{B}_\ell, \nabla\phi(\operatorname{tr}\mathbb{B}_\ell)) = (\nabla\operatorname{tr}\mathbb{B}_\ell, \phi'(\operatorname{tr}\mathbb{B}_\ell)\nabla\operatorname{tr}\mathbb{B}_\ell) \\ = \left\|\sqrt{\phi'(\operatorname{tr}\mathbb{B}_\ell)}\nabla T_K(\operatorname{tr}\mathbb{B}_\ell)\right\|_2^2 \ge 0.\end{aligned} \tag{3.159}$$

Next, using $\operatorname{div}\boldsymbol{v}_\ell = 0$ in Ω, $\boldsymbol{v}_\ell\cdot\boldsymbol{n} = 0$ on $\partial\Omega$ and integration by parts, we can write

$$(\boldsymbol{v}_\ell\cdot\nabla\mathbb{B}_\ell, \mathbb{A}_K) = (\boldsymbol{v}_\ell\cdot\nabla\operatorname{tr}\mathbb{B}_\ell, \phi(\operatorname{tr}\mathbb{B}_\ell)) = \int_\Omega \operatorname{div}(\Phi(\operatorname{tr}\mathbb{B}_\ell)\boldsymbol{v}_\ell) = 0. \tag{3.160}$$

Furthermore, we use (3.36), (3.34), (4.37) and Young's inequality to obtain

$$
\begin{aligned}
\delta(\theta_\ell)\gamma(|\mathbb{B}_\ell-\mathbb{I}|)(\operatorname{tr}\mathbb{B}_\ell-d)(\operatorname{tr}\mathbb{B}_\ell)^{\sigma-1}&\\
&\geq (C_1|\mathbb{B}_\ell|^q-C_2)(\operatorname{tr}\mathbb{B}_\ell)^\sigma - C(|\mathbb{B}_\ell|^q+1)(\operatorname{tr}\mathbb{B}_\ell)^{\sigma-1}\\
&\geq C_1|\mathbb{B}_\ell|^{q+\sigma}-C_2|\mathbb{B}_\ell|^\sigma - C|\mathbb{B}|^{q+\sigma-1}-C|\mathbb{B}_\ell|^{\sigma-1}\\
&\geq C_1|\mathbb{B}_\ell|^{q+\sigma}-C.
\end{aligned}
$$

Since $K>d$, this yields the estimate

$$
\begin{aligned}
(\delta(\theta_\ell)\gamma(|\mathbb{B}_\ell-\mathbb{I}|)(\mathbb{B}_\ell-\mathbb{I}),\mathbb{A}_K)&\\
&=(\delta(\theta_\ell)\gamma(|\mathbb{B}_\ell-\mathbb{I}|)(\operatorname{tr}\mathbb{B}_\ell-d),T_K(\operatorname{tr}\mathbb{B}_\ell)^{\sigma-1})\\
&\geq \int_\Omega (C_1|\mathbb{B}_\ell^{q+\sigma}|-C)\chi_{\{\operatorname{tr}\mathbb{B}_\ell<K\}}.
\end{aligned}
\tag{3.161}
$$

Hence, if we use (3.158), (3.159), (3.160) and (3.161) in (3.104) multiplied by $\mathbb{A}_K$, we obtain

$$
\frac{\mathrm{d}}{\mathrm{d}t}\int_\Omega \Phi(\operatorname{tr}\mathbb{B}_\ell)+\int_\Omega |\mathbb{B}_\ell|^{q+\sigma}\,\chi_{\{\operatorname{tr}\mathbb{B}_\ell<K\}}\leq C\int_\Omega \mathbb{B}_\ell\cdot\mathbb{D}\boldsymbol{v}_\ell\,T_K(\operatorname{tr}\mathbb{B}_\ell)^{\sigma-1}+C.
$$

By integrating this inequality over $(0,t)$, $t\in(0,T)$, using (3.108), (4.37) and $T_K(\operatorname{tr}\mathbb{B}_\ell)\leq\operatorname{tr}\mathbb{B}_\ell$ in Q, we get

$$
\int_\Omega \Phi(\operatorname{tr}\mathbb{B}_\ell(t))+\int_0^t\int_\Omega |\mathbb{B}_\ell|^{q+\sigma}\chi_{\{\operatorname{tr}\mathbb{B}_\ell<K\}}\leq C\int_0^t\int_\Omega |\mathbb{B}_\ell|^\sigma|\mathbb{D}\boldsymbol{v}_\ell|+\int_\Omega\Phi(\operatorname{tr}\mathbb{B}_0^\omega)+Ct.
$$

Since

$$
\begin{aligned}
\Phi(\operatorname{tr}\mathbb{B}_\ell(t))=\int_\omega^{\operatorname{tr}\mathbb{B}_\ell(t)}\min\{s,K\}^{\sigma-1}\,\mathrm{d}s&\geq\chi_{\{\operatorname{tr}\mathbb{B}_\ell(t)<K\}}\int_\omega^{\operatorname{tr}\mathbb{B}_\ell(t)} s^{\sigma-1}\,\mathrm{d}s\\
&\geq\frac{1}{\sigma}\Big((\operatorname{tr}\mathbb{B}_\ell(t))^\sigma\chi_{\{\operatorname{tr}\mathbb{B}_\ell(t)<K\}}-\omega^\sigma\Big)
\end{aligned}
$$

and

$$
\Phi(\operatorname{tr}\mathbb{B}_0^\omega)=\int_\omega^{\operatorname{tr}\mathbb{B}_0^\omega}\min\{s,K\}^{\sigma-1}\,\mathrm{d}s\leq\int_\omega^{\operatorname{tr}\mathbb{B}_0^\omega}s^{\sigma-1}=\frac{1}{\sigma}\left((\operatorname{tr}\mathbb{B}_0^\omega)^\sigma-\omega^\sigma\right)
$$

a.e. in Ω, we are led to

$$
\begin{aligned}
\int_\Omega(\operatorname{tr}\mathbb{B}_\ell(t))^\sigma\chi_{\{\operatorname{tr}\mathbb{B}_\ell(t)<K\}}+\int_0^t\int_\Omega|\mathbb{B}_\ell|^{q+\sigma}\chi_{\{\operatorname{tr}\mathbb{B}_\ell<K\}}&\\
\leq C\int_0^t\int_\Omega|\mathbb{B}_\ell|^\sigma|\mathbb{D}\boldsymbol{v}_\ell|+\int_\Omega(\operatorname{tr}\mathbb{B}_0^\omega)^\sigma+Ct.&
\end{aligned}
$$

As the integrands on the left hand side are non-negative and converge point-wise as $K\to\infty$ (due to $\operatorname{tr}\mathbb{B}_\ell(t)\in L^1(\Omega)$ and $\operatorname{tr}\mathbb{B}_\ell\in L^1(Q)$), the application of the limes inferior and the Fatou lemma gives

$$
\int_\Omega(\operatorname{tr}\mathbb{B}_\ell(t))^\sigma+\int_0^t\int_\Omega|\mathbb{B}_\ell|^{q+\sigma}\leq C\int_0^t\int_\Omega|\mathbb{B}_\ell|^\sigma|\mathbb{D}\boldsymbol{v}_\ell|+\int_\Omega(\operatorname{tr}\mathbb{B}_0^\omega)^\sigma+Ct. \tag{3.162}
$$

Thus, taking the essential supremum over $(0,T)$, using Young's inequality on the right hand side (with exponents $\frac{q+\sigma}{\sigma}$ and $\frac{q+\sigma}{q}$), (3.72), (3.45) and (4.37), we arrive at

$$
\|\mathbb{B}_\ell\|_{L^\infty L^\sigma}+\|\mathbb{B}_\ell\|_{L^{q+\sigma}L^{q+\sigma}}\leq C\|\mathbb{D}\boldsymbol{v}_\ell\|^{\frac{1}{q}}_{L^{\frac{q+\sigma}{q}}L^{\frac{q+\sigma}{q}}}+C, \tag{3.163}
$$

where note that the right hand side is finite due to $\boldsymbol{v}_\ell \in L^\infty(0,T;W^{1,\infty}(\Omega))$.

Next, we use (3.163) and (3.124) to improve the information about θ_ℓ and $\mathbb{D}\boldsymbol{v}_\ell$. Here, we want to make clear that the term $\omega|\nabla\theta_\ell|^r\nabla\theta_\ell$ appearing in (3.124) is not used any more to deduce estimates on θ_ℓ since we actually want that the resulting estimate is uniform with respect to ω. For any $\beta \in [0,1)$, we can show that $\theta_\ell^{-\beta} \in L^{r+2}(0,T;W^{1,r+2}(\Omega)) \cap L^\infty(0,T;L^\infty(\Omega))$ similarly as in (3.129), and thus $\tau_\beta = -\theta_\ell^{-\beta}$ is an admissible test function in (3.124). Using Lemma 4.4 with $\psi(s) = -\max(s,\omega)^{-\beta}$ to rewrite the time derivative, (3.146) with Young's inequality, integration by parts, $\boldsymbol{v}_\ell \in W^{1,2}_{0,\mathrm{div}}$ and (3.33), we obtain

$$
\begin{aligned}
&\int_0^T \langle c_v\partial_t\theta_\ell, \tau_\beta\rangle + \int_0^T (c_v\boldsymbol{v}_\ell\cdot\nabla\theta_\ell, \tau_\beta) + \int_0^T (\kappa(\theta_\ell)\nabla\theta_\ell, \nabla\tau_\beta) + \omega\int_0^T (|\nabla\theta_\ell|^r\nabla\theta_\ell, \nabla\tau_\beta)\\
&\quad\geq \frac{-c_v}{1-\beta}\frac{\mathrm{d}}{\mathrm{d}t}\int_0^T\int_\Omega \theta_\ell^{1-\beta} - \frac{c_v}{1-\beta}\int_0^T\int_{\partial\Omega}\theta_\ell^{1-\beta}\boldsymbol{v}_\ell\cdot\boldsymbol{n} + \beta\int_Q \theta_\ell^{-1-\beta}\kappa(\theta_\ell)|\nabla\theta_\ell|^2\\
&\quad= \frac{c_v}{1-\beta}\int_\Omega ((\theta_0^{\frac{1}{\omega}})^{1-\beta} - \theta_\ell^{1-\beta}(T)) + \beta\int_Q \theta_\ell^{-1-\beta}\kappa(\theta_\ell)|\nabla\theta_\ell|^2\\
&\quad\geq C\beta\int_Q \left|\nabla\theta_\ell^{\frac{r+1-\beta}{2}}\right|^2 - C.
\end{aligned}
$$

We use this estimate in (3.124) with $\tau = \tau_\beta$ to deduce, using also (3.31), $g_\omega \leq 1$, Hölder's inequality and (3.163) that, in the case $\sigma < q$, we have

$$
\begin{aligned}
\beta\int_Q \left|\nabla\theta_\ell^{\frac{r+1-\beta}{2}}\right|^2 + \int_Q \theta_\ell^{-\beta}|\mathbb{D}\boldsymbol{v}_\ell|^2 &\leq C\int_Q \theta_\ell^{1-\beta}|\mathbb{B}_\ell||\mathbb{D}\boldsymbol{v}_\ell| + C\\
&\leq C\left\|\theta_\ell^{1-\frac{\beta}{2}}\right\|_{\frac{2(q+\sigma)}{q+\sigma-2};Q}\|\mathbb{B}_\ell\|_{q+\sigma;Q}\left\|\theta_\ell^{-\frac{\beta}{2}}\mathbb{D}\boldsymbol{v}_\ell\right\|_{2;Q} + C\\
&\leq C\|\theta_\ell\|^{1-\frac{\beta}{2}}_{\frac{(2-\beta)(q+\sigma)}{q+\sigma-2};Q}\|\mathbb{D}\boldsymbol{v}_\ell\|^{\frac{1}{q}}_{\frac{q+\sigma}{q};Q}\left\|\theta_\ell^{-\frac{\beta}{2}}\mathbb{D}\boldsymbol{v}_\ell\right\|_{2;Q}\\
&\qquad + C\|\theta_\ell\|^{1-\frac{\beta}{2}}_{\frac{(2-\beta)(q+\sigma)}{q+\sigma-2};Q}\left\|\theta_\ell^{-\frac{\beta}{2}}\mathbb{D}\boldsymbol{v}_\ell\right\|_{2;Q} + C\\
&\leq C\|\theta_\ell\|^{1-\frac{\beta}{2}}_{\frac{(2-\beta)(q+\sigma)}{q+\sigma-2};Q}\|\theta_\ell\|^{\frac{\beta}{2q}}_{\frac{\beta(q+\sigma)}{q-\sigma};Q}\left\|\theta_\ell^{-\frac{\beta}{2}}\mathbb{D}\boldsymbol{v}_\ell\right\|^{1+\frac{1}{q}}_{2;Q}\\
&\qquad + C\|\theta_\ell\|^{1-\frac{\beta}{2}}_{\frac{(2-\beta)(q+\sigma)}{q+\sigma-2};Q}\left\|\theta_\ell^{-\frac{\beta}{2}}\mathbb{D}\boldsymbol{v}_\ell\right\|_{2;Q} + C
\end{aligned}
$$

while if $\sigma = q$, we omit the final step to get

$$
\begin{aligned}
\beta\int_Q \left|\nabla\theta_\ell^{\frac{r+1-\beta}{2}}\right|^2 + \int_Q \theta_\ell^{-\beta}|\mathbb{D}\boldsymbol{v}_\ell|^2 &\leq C\|\theta_\ell\|^{1-\frac{\beta}{2}}_{(2-\beta)q';Q}\|\mathbb{D}\boldsymbol{v}_\ell\|^{\frac{1}{q}}_{2;Q}\left\|\theta_\ell^{-\frac{\beta}{2}}\mathbb{D}\boldsymbol{v}_\ell\right\|_{2;Q}\\
&\qquad + C\|\theta_\ell\|^{1-\frac{\beta}{2}}_{(2-\beta)q';Q}\left\|\theta_\ell^{-\frac{\beta}{2}}\mathbb{D}\boldsymbol{v}_\ell\right\|_{2;Q} + C.
\end{aligned}
$$

Thus, using $q > 1$, $(\frac{2q}{q+1})' = 2q'$ and the Young inequality, we arrive at

$$
\begin{aligned}
\beta\int_Q \left|\nabla\theta_\ell^{\frac{r+1-\beta}{2}}\right|^2 + \int_Q \theta_\ell^{-\beta}|\mathbb{D}\boldsymbol{v}_\ell|^2 &\leq C\|\theta_\ell\|^{(2-\beta)q'}_{\frac{(2-\beta)(q+\sigma)}{q+\sigma-2};Q}\|\theta_\ell\|^{\frac{\beta q'}{q}}_{\frac{\beta(q+\sigma)}{q-\sigma};Q}\\
&\qquad + C\|\theta_\ell\|^{2-\beta}_{\frac{(2-\beta)(q+\sigma)}{q+\sigma-2};Q} + C
\end{aligned}
\tag{3.164}
$$

if $\sigma < q$ and

$$\beta \int_Q \left| \nabla \theta_\ell^{\frac{r+1-\beta}{2}} \right|^2 + \int_Q \theta_\ell^{-\beta} |\mathbb{D}\boldsymbol{v}_\ell|^2 \le C \|\theta_\ell\|_{(2-\beta)q';Q}^{2-\beta} \|\mathbb{D}\boldsymbol{v}_\ell\|_{2;Q}^{\frac{2}{q}} + C\|\theta_\ell\|_{(2-\beta)q';Q}^{2-\beta} + C \tag{3.165}$$

if $\sigma = q$, respectively. Next we focus on the case $\sigma < q$. Let

$$\beta_0 = \max\left\{0, r_d + 1 - \frac{(r_d-1)(q+\sigma)}{2}\right\}, \qquad \beta_1 = \min\left\{1, \frac{(r_d+1)(q-\sigma)}{2q}\right\}$$

and note that

$$0 \le \beta_0 < \beta_1 \le 1$$

since (recall $\sigma > 1$, $r_d > r_0 = \frac{q+1}{q-1}$)

$$\begin{aligned} r_d + 1 - \frac{(r_d-1)(q+\sigma)}{2} &< r_d + 1 - \frac{(r_d-1)(q+1)}{2} = 1 - \frac{r_d q - r_d - q - 1}{2} \\ &= 1 - (q-1)\frac{r_d - r_1}{2} < 1 \end{aligned}$$

and

$$\begin{aligned} r_d + 1 - \frac{(r_d-1)(q+\sigma)}{2} &= r_d + 1 - \left(1 - \frac{2}{r_d+1}\right) q \frac{(r_d+1)(q+\sigma)}{2q} \\ &< r_d + 1 - \frac{(r_d+1)(q+\sigma)}{2q} = \frac{(r_d+1)(q-\sigma)}{2q}. \end{aligned}$$

Then, we observe that the inequality

$$\max\left\{(2-\beta)\frac{q+\sigma}{q+\sigma-2}, \beta\frac{q+\sigma}{q-\sigma}\right\} \le r_d + 1 - \beta \tag{3.166}$$

holds if

$$r_0 < r_d < r_1 \quad \text{and} \quad \beta_0 \le \beta \le \beta_1, \tag{3.167}$$

or

$$r_d \ge r_1 \quad \text{and} \quad 0 \le \beta \le \beta_1. \tag{3.168}$$

Indeed, in the first case, we write

$$\begin{aligned} (2-\beta)\frac{q+\sigma}{q+\sigma-2} &= 2 - \beta + \frac{2(2-\beta)}{q+\sigma-2} \\ &\le 2 - \beta + \frac{2 - 2r_d + (r_d-1)(q+\sigma)}{q+\sigma-2} = r_d + 1 - \beta \end{aligned}$$

and in the second case we have

$$\begin{aligned} (2-\beta)\frac{q+\sigma}{q+\sigma-2} &= r_1 + \frac{(1-\beta)(q+\sigma+2) - 2(2-\beta)}{q+\sigma-2} \\ &\le r_d + 1 - \beta\frac{q+\sigma}{q+\sigma-2} \le r_d + 1 - \beta, \end{aligned}$$

while the bound $\beta \leq \beta_1$ is needed for

$$\beta \frac{q+\sigma}{q-\sigma} = -\beta + \beta \frac{2q}{q-\sigma} \leq r_d + 1 - \beta.$$

Thus, by application of (3.166), the Hölder inequality, an interpolation inequality, the Sobolev inequality, the Poincaré inequality and (3.146), we get

$$\begin{aligned} \|\theta_\ell\|_{\frac{(2-\beta)(q+\sigma)}{q+\sigma-2};Q} + \|\theta_\ell\|_{\frac{\beta(q+\sigma)}{q-\sigma};Q} &\leq C\|\theta_\ell\|_{r_d+1-\beta;Q} \\ &\leq C\|\theta_\ell\|_{L^\infty L^1}^{\frac{2}{d(r+1-\beta)+2}} \|\theta_\ell\|_{L^{r+1-\beta}L^{\frac{d}{d-2}(r+1-\beta)}}^{\frac{d(r+1-\beta)}{d(r+1-\beta)+2}} \\ &\leq C\left\|\theta_\ell^{\frac{r+1-\beta}{2}}\right\|_{L^2L^{\frac{2d}{d-2}}}^{\frac{2d}{d(r+1-\beta)+2}} \\ &\leq C\left\|\nabla\theta_\ell^{\frac{r+1-\beta}{2}}\right\|_{L^2L^2}^{\frac{2}{r_d+1-\beta}} + C. \end{aligned} \tag{3.169}$$

If we use (3.166) and Hölder's inequality in (3.164), and then also (3.169), we obtain

$$\begin{aligned} \beta\int_Q \left|\nabla\theta_\ell^{\frac{r+1-\beta}{2}}\right|^2 + \int_Q \theta_\ell^{-\beta}|\mathbb{D}\boldsymbol{v}_\ell|^2 &\leq C\|\theta_\ell\|_{r_d+1-\beta;Q}^{2q'-\beta} + C\|\theta_\ell\|_{r_d+1-\beta;Q}^{2-\beta} + C \\ &\leq C\left\|\nabla\theta_\ell^{\frac{r+1-\beta}{2}}\right\|_{2;Q}^{\omega_1} + C\left\|\nabla\theta_\ell^{\frac{r+1-\beta}{2}}\right\|_{2;Q}^{\omega_2} + C, \end{aligned} \tag{3.170}$$

where

$$0 < \omega_2 < \omega_1 = \frac{2(2q'-\beta)}{r_d+1-\beta} < 2$$

due to $2q' < r_d + 1$ (which is equivalent to $r_d > r_0$). Hence, we can apply the Young inequality in (3.170) to finally get

$$\beta\int_Q \left|\nabla\theta_\ell^{\frac{r+1-\beta}{2}}\right|^2 + \int_Q \theta_\ell^{-\beta}|\mathbb{D}\boldsymbol{v}_\ell|^2 \leq C(\beta) \tag{3.171}$$

for any r, β fulfilling (3.167) or (3.168). In case that (3.167) holds, we make the optimal choice

$$\beta = \beta_0 \tag{3.172}$$

and note that

$$\frac{r+1-\beta}{2} = -\frac{1}{d} + \frac{(r_d-1)(q+\sigma)}{4} = \frac{R}{2},$$

cf. (3.41). Then, from Hölder's inequality and (3.169), we deduce

$$\|\mathbb{D}\boldsymbol{v}_\ell\|_{p;Q} \leq \left\|\theta_\ell^{-\frac{\beta}{2}}\mathbb{D}\boldsymbol{v}_\ell\right\|_{2;Q}\left\|\theta_\ell^{\frac{\beta}{2}}\right\|_{2\frac{r_d+1-\beta}{\beta};Q} \leq C,$$

where

$$p = 2\frac{r_d+1-\beta}{r_d+1} = 2 - \frac{2\beta}{r_d+1} < 2.$$

Note again that this definition of p agrees with the one given in (3.40) since

$$2\frac{r_d+1-\beta}{r_d+1} = \frac{r_d-1}{r_d+1}(q+\sigma)$$

by (3.172). In the special case $r_d = r_1$, we can repeat the above estimates without (3.172), choosing instead $\beta > 0$ arbitrarily small. Finally, if $r_d > r_1$, we can improve the information on $\mathbb{D}\boldsymbol{v}_\ell$ simply by taking $\tau = -1$ in (3.124). Then, using similar computation as above with β chosen as to satisfy

$$0 < \beta < \min\left\{ r_d + 1 - 2\frac{q+\sigma}{q+\sigma-2}, r_d - 1 \right\} = r_d + 1 - 2\frac{q+\sigma}{q+\sigma-2} \tag{3.173}$$

and using (3.163), (3.169), $\sigma \le q$, we obtain

$$\begin{aligned}\int_Q |\mathbb{D}\boldsymbol{v}_\ell|^2 &\le C\int_Q \theta_\ell |\mathbb{B}_\ell||\mathbb{D}\boldsymbol{v}_\ell| \le C\|\mathbb{B}_\ell\|_{q+\sigma;Q}\|\mathbb{D}\boldsymbol{v}_\ell\|_{2;Q}\|\theta_\ell\|_{\frac{2(q+\sigma)}{q+\sigma-2};Q} + C \\ &\le C\left(\|\mathbb{D}\boldsymbol{v}_\ell\|^{\frac{1}{q}}_{\frac{q+\sigma}{q};Q} + 1\right)\|\mathbb{D}\boldsymbol{v}_\ell\|_{2;Q}\|\theta_\ell\|_{r_d+1-\beta;Q} + C \\ &\le C\left(\|\mathbb{D}\boldsymbol{v}_\ell\|^{1+\frac{1}{q}}_{2;Q} + \|\mathbb{D}\boldsymbol{v}_\ell\|_{2;Q}\right)\left(\left\|\nabla\theta_\ell^{\frac{r+1-\beta}{2}}\right\|^{\frac{2}{r_d+1-\beta}}_{2;Q} + 1\right) + C.\end{aligned} \tag{3.174}$$

Hence, using $1 + \frac{1}{q} < 2$, (3.173), Young's inequality and (3.171), we get

$$\beta\int_Q \left|\nabla\theta_\ell^{\frac{r+1-\beta}{2}}\right|^2 + \int_Q |\mathbb{D}\boldsymbol{v}_\ell|^2 \le C(\beta) \tag{3.175}$$

for any β satisfying (3.173). Finally, it remains to consider the excluded case $\sigma = q$. However, in this situation, we have $r_1 = r_0 < r_d$, and thus we can take $\tau = -1$ in (3.124) as before. This way, adding also (3.165) and using analogous estimation as in (3.174), we obtain

$$\begin{aligned}&\beta\int_Q \left|\nabla\theta_\ell^{\frac{r+1-\beta}{2}}\right|^2 + \int_Q |\mathbb{D}\boldsymbol{v}_\ell|^2 \\ &\quad\le C\int_Q \theta_\ell|\mathbb{B}_\ell||\mathbb{D}\boldsymbol{v}_\ell| + C\|\theta_\ell\|^{2-\beta}_{(2-\beta)q';Q}\left(\|\mathbb{D}\boldsymbol{v}_\ell\|^{\frac{2}{q}}_{2;Q} + 1\right) + C \\ &\quad\le C\left(\|\mathbb{D}\boldsymbol{v}_\ell\|^{1+\frac{1}{q}}_{2;Q} + \|\mathbb{D}\boldsymbol{v}_\ell\|_{2;Q}\right)\left(\left\|\nabla\theta_\ell^{\frac{r+1-\beta}{2}}\right\|^{\frac{2}{r_d+1-\beta}}_{2;Q} + 1\right) \\ &\qquad + C\left(\|\mathbb{D}\boldsymbol{v}_\ell\|^{\frac{2}{q}}_{2;Q} + 1\right)\left(\left\|\nabla\theta_\ell^{\frac{r+1-\beta}{2}}\right\|^{\frac{2(2-\beta)}{r_d+1-\beta}}_{2;Q} + 1\right) + C.\end{aligned}$$

If we choose β as in (3.173) and use Young's inequality, noticing that

$$\frac{2(2-\beta)}{r_d+1-\beta}q' < 2\frac{2q'-\beta q'}{2q'-\beta} = 2\left(1 - \frac{(q'-1)\beta}{2q'-\beta}\right) < 2,$$

we again conclude that (3.175) holds.

To summarize the estimates up to this point, we proved (3.154)–(3.163),

$$\left\|\nabla\theta_\ell^{\frac{R}{2}}\right\|_{L^2L^2} + \|\theta_\ell\|_{L^{R_d}L^{R_d}} \le C \tag{3.176}$$

and

$$\|\mathbb{D}\boldsymbol{v}_\ell\|_{L^pL^p} \le C. \tag{3.177}$$

Moreover, we deduce from (3.163), (3.177) and

$$\frac{q+\sigma}{q} = (q+\sigma)\left(1-\frac{2}{2q'}\right) < (q+\sigma)\left(1-\frac{2}{r_d+1}\right) = p \tag{3.178}$$

that

$$\|\mathbb{B}_\ell\|_{L^\infty L^\sigma} + \|\mathbb{B}_\ell\|_{L^{q+\sigma}L^{q+\sigma}} \le C. \tag{3.179}$$

Next, the combination of (4.33) and (4.35) yields

$$|\nabla \mathbb{B}_\ell| \le |\mathbb{B}_\ell^{\frac{1}{2}}||\mathbb{B}_\ell^{-\frac{1}{2}}\nabla\mathbb{B}_\ell\mathbb{B}_\ell^{-\frac{1}{2}}||\mathbb{B}_\ell^{\frac{1}{2}}| = (\operatorname{tr}\mathbb{B}_\ell)|\mathbb{B}_\ell^{-\frac{1}{2}}\nabla\mathbb{B}_\ell\mathbb{B}_\ell^{-\frac{1}{2}}|.$$

Then, since $q+\sigma > 2$, we deduce by appealing to the Hölder inequality and (3.90), (3.179) that

$$\|\nabla\mathbb{B}_\ell\|_{L^{s_0}L^{s_0}} \le C, \tag{3.180}$$

where

$$s_0 = \frac{1}{\frac{1}{q+\sigma}+\frac{1}{2}} = \frac{2(q+\sigma)}{q+\sigma+2}.$$

Next, we derive the uniform estimates for the time derivatives. To this end, we need to determine integrability of the non-linear terms in (3.103), (3.104) and (3.143).

It follows from an interpolation inequality, Korn's inequality, (3.146) and (3.177) that

$$\|\boldsymbol{v}_\ell\|_{L^{p(1+\frac{2}{d})}L^{p(1+\frac{2}{d})}} \le C\|\boldsymbol{v}_\ell\|_{L^\infty L^2}^{\frac{2}{d+2}}\|\mathbb{D}\boldsymbol{v}_\ell\|_{L^pL^p}^{\frac{d}{d+2}} \le C. \tag{3.181}$$

We remark that (3.40) and (A2) imply

$$p\left(1+\frac{2}{d}\right) > 2. \tag{3.182}$$

Indeed, this is obvious if $r_d \ge r_1$ (and $p=2$ or $p=2$)), while if $r_d < r_1$, we use (3.38) to estimate

$$p\left(1+\frac{2}{d}\right) = \frac{d+2}{d}\frac{q+\sigma}{2r_d'-1} > \frac{d+2}{d}\frac{q+\sigma}{\frac{d+2}{2d}(q+\sigma)} = 2.$$

Furthermore, the Hölder inequality, (3.176) and (3.163) yield

$$\|\theta_\ell\mathbb{B}_\ell\|_{L^{\varrho_1}L^{\varrho_1}} \le C,$$

where

$$\varrho_1 = \frac{1}{\frac{1}{R_d}+\frac{1}{q+\sigma}}.$$

Recalling (3.41), we note that

$$\varrho_1 = \frac{1}{\frac{2}{(r_d-1)(q+\sigma)}+\frac{1}{q+\sigma}} = \frac{q+\sigma}{\frac{r_d+1}{r_d-1}} = p \quad \text{if } r_d < r_1$$

and

$$\varrho_1 = \frac{1}{\frac{1}{r_d+1)}+\frac{1}{q+\sigma}} = \frac{1}{\frac{q+\sigma-2}{2(q+\sigma)}+\frac{1}{q+\sigma}}\Bigg) = 2) = p \quad \text{if } r_d = r_1,$$

and then obviously also $\varrho_1 > 2 = p$ if $r_d > r_1$. Hence, we read from (3.103) that

$$\|\partial_t \boldsymbol{v}_\ell\|_{L^{p_1} W^{-1,p_1}_{0,\mathrm{div}}} \leq C, \tag{3.183}$$

where p_1 is defined in (3.42).

Next, we focus on the non-linear terms in (3.104). There, we integrate by parts in the convective term with the help of (3.69). Then, using Hölder's inequality and (3.163), (3.181), we observe that

$$\|\mathbb{B}_\ell \otimes \boldsymbol{v}_\ell\|_{L^{s_1} L^{s_1}} \leq C, \tag{3.184}$$

with

$$s_1 = \frac{1}{\frac{1}{q+\sigma} + \frac{1}{p(1+\frac{2}{d})}} > \frac{1}{\frac{1}{q+\sigma} + \frac{1}{2}} = s_0, \tag{3.185}$$

where we used (3.182). Moreover, Young's inequality, (3.36) and (3.34), give

$$|\delta(\theta_\ell)\gamma(|\mathbb{B}_\ell - \mathbb{I}|)(\mathbb{B}_\ell - \mathbb{I})| \leq C(|\mathbb{B}_\ell|^{q+1} + 1),$$

hence, making use of (3.179), we obtain

$$\|\delta(\theta_\ell)\gamma(|\mathbb{B}_\ell - \mathbb{I}|)(\mathbb{B}_\ell - \mathbb{I})\|_{L^{s_2} L^{s_2}} \leq C, \tag{3.186}$$

where

$$s_2 = \frac{q+\sigma}{q+1}.$$

Furthermore, using (3.179), (3.177) and Hölder's inequality, we get

$$\|(a\mathbb{D}\boldsymbol{v}_\ell + \mathbb{W}\boldsymbol{v}_\ell)\mathbb{B}_\ell\|_{L^{s_3} L^{s_3}} \leq C, \tag{3.187}$$

where

$$s_3 = \frac{1}{\frac{1}{q+\sigma} + \frac{1}{p}} = \frac{q+\sigma}{1 + \frac{r_d+1}{r_d-1}} > \frac{q+\sigma}{q+1} = s_2 \tag{3.188}$$

(using $r_d > r_0$). Thus, we read from (3.104) using (3.180), (3.184), (3.185) and (3.186), (3.187), (3.188) that[4]

$$\|\partial_t \mathbb{B}_\ell\|_{(L^{s_0'} W^{1,s_0'} \cap L^{s_2'} L^{s_2'})^*} \leq C. \tag{3.189}$$

This property is important for obtaining a point-wise convergence of $\mathbb{B}_\ell$ using the Aubin-Lions lemma. To avoid any confusion regarding this argument and the space appearing in (3.189), we can use $s_2 \leq s_0$ (which follows from $\sigma \leq q$) and obvious embeddings to replace (3.189) by a weaker information

$$\|\partial_t \mathbb{B}_\ell\|_{L^{s_2} W^{-1,s_2}} \leq C, \tag{3.190}$$

which is still sufficient for the use of Aubin-Lions lemma (and for weak compactness of $\partial_t \mathbb{B}_\ell$ as $s_2 > 1$).

Finally, we examine the non-linearities related to (3.143). There, since ξ_ℓ is under control thanks to (3.154), the problematic terms could be only on the left

[4] For the definition and properties of intersection (and sums) of two normed spaces, we refer to Section 4.2 below.

hand side. To get an appropriate uniform control over the convective term, we need to show that η_ℓ is bounded in $L^2(Q)$ (or only slightly worse, recall (3.181) and (3.182)). To this end, we use inequalities (3.152), (3.153) once again to get

$$\eta_\ell \le \eta_\ell + \mu\psi_2(\mathbb{B}_\ell) = c_v \ln\theta_\ell \le c_v(\theta_\ell - 1).$$

This, together with (3.146) and (3.155), yields

$$\|\ln\theta_\ell\|_{L^\infty L^1} \le C. \tag{3.191}$$

Then, since (3.156) and (3.33) give

$$\|\nabla \ln\theta_\ell\|_{L^2L^2} \le C, \tag{3.192}$$

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

$$\|\ln\theta_\ell\|_{L^{2+\frac{2}{d}}L^{2+\frac{2}{d}}} \le C\|\ln\theta_\ell\|_{L^\infty L^1}^{\frac{1}{d+1}} \|\ln\theta_\ell\|_{L^2W^{1,2}}^{\frac{d}{d+1}} \le C. \tag{3.193}$$

Now we observe that a similar reasoning applies also for the quantity $\ln\det\mathbb{B}_\ell$. Indeed, using (3.155), (3.191), (3.163) and (3.133) 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

$$\|\ln\det\mathbb{B}_\ell\|_{L^\infty L^1} \le C. \tag{3.194}$$

Further, the estimate of its derivative follows immediately from (4.40) and (3.156) as

$$\|\nabla \ln\det\mathbb{B}_\ell\|_{L^2L^2} = \left\|\operatorname{tr}(\mathbb{B}_\ell^{-\frac{1}{2}}\nabla\mathbb{B}_\ell\mathbb{B}_\ell^{-\frac{1}{2}})\right\|_{L^2L^2} \le C. \tag{3.195}$$

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

$$\|\ln\det\mathbb{B}_\ell\|_{L^{2+\frac{2}{d}}L^{2+\frac{2}{d}}} \le C. \tag{3.196}$$

From (3.193), (3.196), (3.163) and (3.133), we finally deduce

$$\|\eta_\ell\|_{L^{s_4}L^{s_4}}, \quad \text{where} \quad s_4 = \min\left\{2 + \tfrac{2}{d}, q + \sigma\right\} > 2, \tag{3.197}$$

and thus

$$\|\boldsymbol{v}_\ell\eta_\ell\|_{L^{s_5}L^{s_5}} \le C, \quad \text{where} \quad s_5 = \left(\frac{d}{p(d+2)} + \frac{1}{s_4}\right)^{-1} > 1. \tag{3.198}$$

due to (3.182) and (3.197). 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 (3.192), (3.180), (3.156) and Poincaré's inequality that

$$\|\eta_\ell\|_{L^{s_0}W^{1,s_0}} \le C. \tag{3.199}$$

Looking at (3.143), we still need to verify that the flux terms are under control.

As for the term $\kappa(\theta_\ell)\nabla\ln\theta_\ell$, we use Hölder's inequality, (3.33), (3.176) and (3.156) to get

$$\|\kappa(\theta_\ell)\nabla\ln\theta_\ell\|_{\frac{2R_d}{r+R_d};Q} \leq \left\|\sqrt{\kappa(\theta_\ell)}\right\|_{\frac{2R_d}{r};Q}\left\|\sqrt{\kappa(\theta_\ell)}\nabla\ln\theta_\ell\right\|_{2;Q} \leq C, \tag{3.200}$$

where note, using (A1), that

$$\begin{aligned}R_d &= \frac{(r_d-1)(q+\sigma)}{2} = r_d + \frac{r_d(q+\sigma-2)-q-\sigma}{2}\\ &> r_d + \frac{q^2+q\sigma-2q+q+\sigma-2 - q^2-q\sigma+q+\sigma}{2(q-1)} = r_d + \frac{\sigma-1}{q-1} > r_d.\end{aligned} \tag{3.201}$$

Finally, let us derive an estimate on $\omega|\nabla\theta_\ell|^r\nabla\theta_\ell$, from which it follows that this term vanishes as $\omega\to 0_+$. The number s_6 defined by

$$s_6 = \frac{R_d(r+2)}{R_d(r+1)+r} = 1 + \frac{R_d-r}{R_d(r+1)+r}$$

is greater that one due to (3.201). Next, we remark that (3.156) yields

$$\omega\int_Q \frac{|\nabla\theta_\ell|^{r+2}}{\theta_\ell^2} \leq C. \tag{3.202}$$

Using this together with (3.176) and Hölder's inequality leads to

$$\begin{aligned}\|\omega|\nabla\theta_\ell|^r\nabla\ln\theta_\ell\|_{s_6;Q} &= \omega\left(\int_Q \frac{|\nabla\theta_\ell|^{\frac{R_d(r+2)(r+1)}{R_d(r+1)+r}}}{\theta_\ell^{\frac{2R_d(r+1)}{R_d(r+1)+r}}}\theta_\ell^{\frac{R_d r}{R_d(r+1)+r}}\right)^{\frac{1}{s_6}}\\ &\leq \omega\left\|\frac{|\nabla\theta_\ell|^{\frac{R_d(r+2)(r+1)}{R_d(r+1)+r}}}{\theta_\ell^{\frac{2R_d(r+1)}{R_d(r+1)+r}}}\right\|_{\frac{R_d(r+1)+r}{R_d(r+1)};Q}^{\frac{R_d(r+1)+r}{R_d(r+2)}}\left\|\theta_\ell^{\frac{R_d r}{R_d(r+1)+r}}\right\|_{\frac{R_d(r+1)+r}{r};Q}^{\frac{R_d(r+1)+r}{R_d(r+2)}}\\ &= \omega^{\frac{1}{r+2}}\int_Q\left(\omega\frac{|\nabla\theta_\ell|^{r+2}}{\theta_\ell^2}\right)^{\frac{r+1}{r+2}}\|\theta_\ell\|_{R_d;Q}^{\frac{r}{r+2}} \leq C\omega^{\frac{1}{r+2}}.\end{aligned} \tag{3.203}$$

From this and from (3.200), (3.195), (3.156), (3.143), we see, using the definition of a weak time derivative, that

$$\|\partial_t\eta_\ell\|_{L^1W^{-M,2}} \leq C, \tag{3.204}$$

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

Taking the final limit

Let us note that the estimates above are independent not only of ℓ, but also of ω (except for (3.202), which is used only to infer (3.203)). Indeed, since $0\leq g_\omega<1$, the presence of g_ω in some terms does not affect the corresponding estimates. Actually, the only information which is ω-dependent are the lower estimates on $\Lambda(\mathbb{B}_\ell)$ and θ_ℓ, but we used this information only quantitatively to verify that $\mathbb{B}_\ell^{-1}$

and θ_ℓ^{-1} are admissible test functions in (3.104) and (3.124), respectively. Hence, to spare us some work, we set

$$\omega = \frac{1}{\ell},$$

and hereby, it remains to take the limit $\ell \to \infty$ only.

By collecting the estimates (3.146), (3.177), (3.183), (3.179), (3.180), (3.190), (3.197), (3.199), (3.204) and using Lemma 4.1, we get the following convergence results:

$$\begin{aligned}
\boldsymbol{v}_\ell &\rightharpoonup \boldsymbol{v} && \text{weakly in } L^p(0,T;W^{1,p}_{0,\mathrm{div}}), && (3.205)\\
\boldsymbol{v}_\ell &\to \boldsymbol{v} && \text{strongly in } L^{p+\frac{2p}{d})}(Q) \text{ and a.e. in } Q, && (3.206)\\
\partial_t\boldsymbol{v}_\ell &\rightharpoonup \partial_t\boldsymbol{v} && \text{weakly in } L^{p_1}(0,T;W^{-1,p_1}_{0,\mathrm{div}}), &&\\
\mathbb{B}_\ell &\rightharpoonup \mathbb{B} && \text{weakly in } L^{s_0}(0,T;W^{1,s_0}(\Omega)), && (3.207)\\
\mathbb{B}_\ell &\to \mathbb{B} && \text{strongly in } L^{q+\sigma)}(Q) \text{ and a.e. in } Q, && (3.208)\\
\partial_t\mathbb{B}_\ell &\rightharpoonup \partial_t\mathbb{B} && \text{weakly in } L^{s_2}(0,T;W^{-1,s_2}(\Omega)), &&\\
\eta_\ell &\rightharpoonup \eta && \text{weakly in } L^{s_0}(0,T;W^{1,s_0}(\Omega)), &&\\
\eta_\ell &\to \eta && \text{strongly in } L^{s_4)}(Q) \text{ and a.e. in } Q, && (3.209)\\
\theta_\ell &\rightharpoonup \theta && \text{weakly in } L^{R_d}(Q). && (3.210)
\end{aligned}$$

Now we explain how to take the limit in equations (3.103), (3.104), (3.143), (3.145) and then, we also identify the corresponding initial conditions. First, we focus on taking the limit in the function $g_{\frac{1}{\ell}}$. From (3.126), (3.128) and (3.205), (3.207) (or (3.206), (3.208)), we obtain

$$\mathbb{B}\boldsymbol{x}\cdot\boldsymbol{x} \geq 0 \quad \text{for all } \boldsymbol{x}\in\mathbb{R}^d \quad \text{and} \quad \theta \geq 0 \quad \text{a.e. in } Q, \tag{3.211}$$

however, we need these properties with strict inequalities. To this end, we use the Fatou lemma, (3.208) and (3.194) 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

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

which, together with (3.211) implies

$$\mathbb{B}\boldsymbol{x}\cdot\boldsymbol{x} > 0 \quad \text{for all } \boldsymbol{x}\in\mathbb{R}^d \quad \text{a.e. in } Q. \tag{3.212}$$

Then, note that, by (3.133), we have

$$c_v\ln\theta_\ell = \eta_\ell + \mu\psi_2(\mathbb{B}_\ell) \tag{3.213}$$

and, by (3.209), (3.208), the right hand side of (3.213) converges a.e. in Q. Therefore, we also have

$$c_v\ln\theta_\ell \to \eta + \mu\psi_2(\mathbb{B}) \quad \text{a.e. in } Q$$

with the limit being finite a.e. in Q thanks to (3.212). This we can rewrite as

$$\theta_\ell \to \exp\left(\frac{1}{c_v}\Big(\eta + \mu\psi_2(\mathbb{B})\Big)\right) \quad \text{a.e. in } Q,$$

where the limit is positive and finite a.e. in Q. But looking at (3.210), this yields

$$\theta = \exp\left(\frac{1}{c_v}\Big(\eta + \mu\psi_2(\mathbb{B})\Big)\right) > 0 \quad \text{a.e. in } Q, \tag{3.214}$$

which is (3.56), and

$$\theta_\ell \to \theta \quad \text{a.e. in } Q. \tag{3.215}$$

By combination of (3.212), (3.214) and the point-wise convergence (3.207), (3.215) 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 point-wise a.e. in Q to 1. Hence, the Vitali theorem and $0 \le g_{\frac{1}{\ell}} < 1$, imply

$$g_{\frac{1}{\ell}}(\mathbb{B}_\ell, \theta_\ell) \to 1 \quad \text{strongly in } L^{\infty)}(Q). \tag{3.216}$$

Therefore, regarding the first two equations (3.103) and (3.104), we can take the limit in the same way as we did in the limit $n \to \infty$ (using the scheme (3.100)–(3.102)). Indeed, the integrability of the resulting non-linear limits was already verified when estimating $\partial_t \boldsymbol{v}_\ell$ and $\partial_t \mathbb{B}_\ell$ ((3.181)–(3.187)). This way, taking (3.216) into account, using the density of $\{\boldsymbol{w}_i\}_{i=1}^\infty$ in $W^{1,p_1'}_{0,\mathrm{div}}$, integrating by parts in the convective term of (3.104) and extending the functional $\partial_t \mathbb{B}$ to the space stated in (3.50) using (3.189) and

$$s_0' = \frac{2(q+\sigma)}{q+\sigma-2} \quad \text{and} \quad s_2' = \frac{q+\sigma}{\sigma-1},$$

we obtain precisely (3.57) and (3.58).

Next, we show how to take the limit in (3.143). In particular, we need to prove that

$$\begin{aligned} \eta_0^{\frac{1}{\ell}} &= c_v \ln \theta_0^{\frac{1}{\ell}} - \mu(\operatorname{tr} \mathbb{B}_0^{\frac{1}{\ell}} - d - \ln\det \mathbb{B}_0^{\frac{1}{\ell}}) \\ &\to c_v \ln \theta_0 - \mu(\operatorname{tr}\mathbb{B}_0 - d - \ln\det\mathbb{B}_0) = \eta_0, \quad \ell \to \infty, \end{aligned} \tag{3.217}$$

weakly in $L^1(\Omega)$, at least. Using (3.73) and (3.72), 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 (3.45) and (3.46). Moreover, the function $\eta_0^{1/\ell}$ converges point-wise a.e. in Ω due to (3.74) and (3.75).

Thus, the limit (3.217) indeed holds (even strongly in $L^1(\Omega)$) by the dominated convergence theorem. In order to take the limit in the convective term, we use (3.206), (3.209) and (3.198). Next, in order to identify the objects $\nabla \ln \theta$ and $\nabla \ln \det \mathbb{B}$, note first that (3.193), (3.196) with (3.215), (3.208) yield

$$\begin{aligned} \ln \theta_\ell &\rightharpoonup \ln \theta && \text{weakly in } L^{2+\frac{2}{d}}(Q), \\ \ln \det \mathbb{B}_\ell &\rightharpoonup \ln \det \mathbb{B} && \text{weakly in } L^{2+\frac{2}{d}}(Q). \end{aligned}$$

This, together with (3.195) and (3.192), implies

$$\nabla \ln \theta_\ell \rightharpoonup \nabla \ln \theta \qquad \text{weakly in } L^2(Q), \tag{3.218}$$

$$\nabla \ln \det \mathbb{B}_\ell \rightharpoonup \nabla \ln \det \mathbb{B} \qquad \text{weakly in } L^2(Q). \tag{3.219}$$

Then, for the term $\kappa(\theta_\ell)\nabla \ln \theta_\ell$, we use (3.30), (3.33), (3.176) and Vitali's theorem to find that

$$\sqrt{\kappa(\theta_\ell)} \rightharpoonup \sqrt{\kappa(\theta)} \quad \text{strongly in } L^{\frac{2R_d}{r})}(Q), \tag{3.220}$$

where we recall that $R_d > r_d$. As an immediate consequence of this and (3.218), we get

$$\sqrt{\kappa(\theta_\ell)}\nabla \ln \theta_\ell \rightharpoonup \sqrt{\kappa(\theta)}\nabla \ln \theta \quad \text{weakly in } L^1(Q). \tag{3.221}$$

However, this weak convergence is true (up to a subsequence) also in $L^2(Q)$ due to (3.156). Therefore, using again (3.220), we obtain

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

Next, the term containing $\omega|\nabla\theta_\ell|^r \nabla \ln \theta_\ell$ tends to zero by (3.203). Furthermore, in the term $\mu\lambda(\theta_\ell)\nabla \operatorname{tr} \mathbb{B}_\ell$, we use (3.30), (3.35), (3.215), Vitali's theorem and (3.207). Analogously, we take the limit in the term $\mu\lambda(\theta_\ell)\nabla \ln \det \mathbb{B}_\ell$, only we use (3.219) instead of (3.207).

Now we take the limit in the terms on the right hand side of (3.143). From (3.156), we deduce that there exists $K \in L^2(Q)$ such that

$$\sqrt{\frac{2\nu(\theta_\ell)}{\theta_\ell}}\mathbb{D}\boldsymbol{v}_\ell \rightharpoonup K \quad \text{weakly in } L^2(Q). \tag{3.222}$$

For $\varepsilon \in (0,1)$, let $h_\varepsilon : (0,\infty) \to [0,1]$ be a smooth function satisfying

$$h_\varepsilon(s) = \begin{cases} 1, & s > \varepsilon; \\ 0, & s < \frac{\varepsilon}{2}, \end{cases}$$

and define

$$f_{\varepsilon,\ell} = h_\varepsilon(\theta_\ell)\sqrt{\frac{2\nu(\theta_\ell)}{\theta_\ell}}\mathbb{D}\boldsymbol{v}_\ell, \qquad f_\varepsilon = h_\varepsilon(\theta)\sqrt{\frac{2\nu(\theta)}{\theta}}\mathbb{D}\boldsymbol{v}$$

For fixed $\varepsilon > 0$, the function

$$h_\varepsilon(\theta_\ell)\sqrt{\frac{2\nu(\theta_\ell)}{\theta_\ell}}$$

is bounded independently of ℓ and converges point-wise due to (3.215) and (3.30). Thus, using the Vitali theorem and (3.205), we find

$$f_{\varepsilon,\ell} \underset{\ell\to\infty}{\rightharpoonup} f_\varepsilon \quad \text{weakly in } L^1(Q). \tag{3.223}$$

Next note that, by Hölder's, Chebyshev's inequalities and (3.191) we have

$$\begin{aligned}\int_Q \left| f_{\varepsilon,\ell} - \sqrt{\frac{2\nu(\theta_\ell)}{\theta_\ell}}\mathbb{D}\boldsymbol{v}_\ell \right| &\le \int_{\{\theta_\ell<\varepsilon\}} \sqrt{\frac{2\nu(\theta_\ell)}{\theta_\ell}}|\mathbb{D}\boldsymbol{v}_\ell| \le C|\{\theta_\ell<\varepsilon\}|^{\frac12}\\ &\le C\,|\{-\ln\theta_\ell > -\ln\varepsilon\}|^{\frac12} \le \frac{C}{\sqrt{-\ln\varepsilon}}\left(\int_Q |\ln\theta_\ell|\right)^{\frac12} \le \frac{C}{\sqrt{-\ln\varepsilon}}\end{aligned}$$

for all $\ell\in\mathbb{N}$. Hence, using (3.222), (3.223) and weak lower semi-continuity, we get

$$\int_Q |f_\varepsilon - K| \le \liminf_{\ell\to\infty}\int_Q \left| f_{\varepsilon,\ell} - \sqrt{\frac{2\nu(\theta_\ell)}{\theta_\ell}}\mathbb{D}\boldsymbol{v}_\ell \right| \le \frac{C}{\sqrt{-\ln\varepsilon}},$$

and thus, for $\varepsilon\to 0_+$, we obtain

$$f_\varepsilon \to K \quad \text{strongly in } L^1(Q). \tag{3.224}$$

On the other hand, since $\theta>0$ a.e. in Q by (3.214), it is clear that

$$f_\varepsilon \to \sqrt{\frac{2\nu(\theta)}{\theta}}\mathbb{D}\boldsymbol{v} \quad \text{a.e. in } Q. \tag{3.225}$$

Therefore, from (3.224) and (3.225), we conclude

$$K = \sqrt{\frac{2\nu(\theta)}{\theta}}\mathbb{D}\boldsymbol{v},$$

which, using (3.222) and weak lower semi-continuity of $\|\cdot\|_{L^2L^2}$, finally gives

$$\liminf_{\ell\to\infty}\int_Q \frac{2\nu(\theta_\ell)}{\theta_\ell}|\mathbb{D}\boldsymbol{v}_\ell|^2 \ge \int_Q \frac{2\nu(\theta)}{\theta}|\mathbb{D}\boldsymbol{v}|^2.$$

In the next term $\kappa(\theta_\ell)|\nabla\ln\theta_\ell|^2$, we can use the weak lower semi-continuity directly since we already proved that (3.221) is valid in $L^2(Q)$. Moreover, the auxiliary term $\omega|\nabla\theta_\ell|^r|\nabla\ln\theta_\ell|^2$ is simply estimated from below by zero.

To take the limit in the term $\delta(\theta_\ell)\gamma(|\mathbb{B}_\ell-\mathbb{I}|)|\mathbb{B}_\ell^{\frac12} - \mathbb{B}_\ell^{-\frac12}|^2$, we use (3.215), (3.208) and apply Fatou's lemma.

To handle the limit in the last term of (3.143), we use again the function h_ε, but this time, we define

$$F_{\varepsilon,\ell} = h_\varepsilon(\det\mathbb{B}_\ell)\sqrt{\lambda(\theta_\ell)}\mathbb{B}_\ell^{-\frac12}\nabla\mathbb{B}_\ell\mathbb{B}_\ell^{-\frac12}, \qquad F_\varepsilon = h_\varepsilon(\det\mathbb{B})\sqrt{\lambda(\theta)}\mathbb{B}^{-\frac12}\nabla\mathbb{B}\mathbb{B}^{-\frac12}.$$

Let $G\in L^2(Q)$ such that

$$\sqrt{\lambda(\theta_\ell)}\mathbb{B}_\ell^{-\frac12}\nabla\mathbb{B}_\ell\mathbb{B}_\ell^{-\frac12} \rightharpoonup G \quad \text{weakly in } L^2(Q). \tag{3.226}$$

For $\varepsilon>0$ fixed, the function

$$\mathbb{H}_{\varepsilon,\ell} = \sqrt{h_\varepsilon(\det\mathbb{B}_\ell)\sqrt{\lambda(\theta_\ell)}}\mathbb{B}_\ell^{-\frac12}$$

is bounded and converges point-wise a.e. in Q due to (3.208) and (3.215). Thus, using Vitali's theorem and (3.207), we get

$$F_{\varepsilon,\ell} = \mathbb{H}_{\varepsilon,\ell}\nabla\mathbb{B}_\ell\mathbb{H}_{\varepsilon,\ell} \underset{\ell\to\infty}{\rightharpoonup} F_\varepsilon \quad \text{weakly in } L^1(Q). \tag{3.227}$$

Moreover, using Hölder's and Chebyshev's inequalities and (3.194), we find

$$\begin{aligned}\int_Q \Big|F_{\varepsilon,\ell} - \sqrt{\lambda(\theta_\ell)}\mathbb{B}_\ell^{-\frac12}\nabla\mathbb{B}_\ell\mathbb{B}_\ell^{-\frac12}\Big| &\le \int_{\{\det\mathbb{B}_\ell<\varepsilon\}} \sqrt{\lambda(\theta_\ell)}\Big|\mathbb{B}_\ell^{-\frac12}\nabla\mathbb{B}_\ell\mathbb{B}_\ell^{-\frac12}\Big| \\ &\le C|\{\det\mathbb{B}_\ell<\varepsilon\}|^{\frac12} \le C\,|\{-\ln\det\mathbb{B}_\ell > -\ln\varepsilon\}|^{\frac12} \\ &\le \frac{C}{\sqrt{-\ln\varepsilon}}\Big(\int_Q |\ln\det\mathbb{B}_\ell|\Big)^{\frac12} \le \frac{C}{\sqrt{-\ln\varepsilon}}\end{aligned}$$

for every $\ell\in\mathbb{N}$. Therefore, from (3.226), (3.227) and weak lower semi-continuity, we deduce

$$\int_Q |F_\varepsilon - G| \le \liminf_{\ell\to\infty}\int_Q \Big|F_{\varepsilon,\ell} - \sqrt{\lambda(\theta_\ell)}\mathbb{B}_\ell^{-\frac12}\nabla\mathbb{B}_\ell\mathbb{B}_\ell^{-\frac12}\Big| \le \frac{C}{\sqrt{-\ln\varepsilon}},$$

hence

$$F_\varepsilon \to G \quad \text{strongly in } L^1(Q).$$

Since $\det\mathbb{B} > 0$ a.e. by (3.212), we also have that

$$F_\varepsilon \to \sqrt{\lambda(\theta)}\mathbb{B}^{-\frac12}\nabla\mathbb{B}\mathbb{B}^{-\frac12} \quad \text{a.e. in } Q.$$

Thus, we identified that

$$G = \sqrt{\lambda(\theta)}\mathbb{B}^{-\frac12}\nabla\mathbb{B}\mathbb{B}^{-\frac12}$$

and, by (3.226) and weak lower semi-continuity, there holds

$$\liminf_{\ell\to\infty}\int_Q \lambda(\theta_\ell)\left|\mathbb{B}_\ell^{-\frac12}\nabla\mathbb{B}_\ell\mathbb{B}_\ell^{-\frac12}\right|^2 \ge \int_Q \lambda(\theta)\left|\mathbb{B}^{-\frac12}\nabla\mathbb{B}\mathbb{B}^{-\frac12}\right|^2.$$

Using the argumentation above to take the limit $\ell\to\infty$ in (3.143), we obtain (3.59).

Finally, to take the limit in (3.145), we first integrate it over $(0,t)$, $t\in(0,T)$, and use (3.121), (3.109) to get

$$\int_\Omega E_\ell(t) - \int_\Omega \big(\tfrac12|P_\ell\boldsymbol{v}_0|^2 + c_v\theta_0^{\frac1\ell}\big) = \int_0^t (\boldsymbol{f},\boldsymbol{v}_\ell). \tag{3.228}$$

Then, note that (3.206) and (3.215) yield

$$E_\ell(t) \to E(t) \quad \text{a.e. in } \Omega \tag{3.229}$$

for almost every $t\in(0,T)$ and also that

$$\boldsymbol{v}_\ell \to \boldsymbol{v} \quad \text{strongly in } L^2(0,t;L^2(\Omega)) \tag{3.230}$$

(cf. (3.182)). Thus, in order to take the limit $\ell\to\infty$ in (3.228), we use (3.229), Fatou's lemma and (3.75), $\|P_\ell\boldsymbol{v}_0 - \boldsymbol{v}_0\|_2 \to 0$ on the left hand side and (3.230) on the right hand side. This leads precisely to (3.60).

3.5.7 Attainment of initial conditions

To finish the proof, it remains to identify the initial conditions stated in the Theorem 3.2. Note in particular, that we search for the initial condition for the temperature while now we only have the entropy inequality at our disposal. Let us start by an observation that $\boldsymbol{v}$ and $\mathbb{B}$ are weakly continuous in time. Indeed, first of all, let us recall that

$$\boldsymbol{v} \in L^\infty(0,T;L^2(\Omega)), \quad \partial_t \boldsymbol{v} \in L^{p_1}(0,T;W^{-1,p_1}_{0,\mathrm{div}}(\Omega)), \tag{3.231}$$

where $p_1 > 1$ is given in Definition 3.1. Concerning $\mathbb{B}$, we have

$$\mathbb{B} \in L^\infty(0,T;L^\sigma(\Omega)), \quad \partial_t \mathbb{B} \in L^{\frac{q+\sigma}{q+1}}(0,T;W^{-1,\frac{q+\sigma}{q+1}}(\Omega)),$$

cf. (3.190) and (3.179). From this and (3.231) we obtain, by a standard argument known from the theory of Navier-Stokes equations (see e.g. [54, Sect. 3.8.]), that

$$\boldsymbol{v} \in \mathcal{C}_w([0,T];L^2(\Omega)) \quad \text{and} \quad \mathbb{B} \in \mathcal{C}_w([0,T];L^\sigma(\Omega)). \tag{3.232}$$

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 (3.74). This way, we obtain

$$\lim_{t\to 0_+} \int_\Omega \boldsymbol{v}(t) \cdot \boldsymbol{w} = \int_\Omega \boldsymbol{v}_0 \cdot \boldsymbol{w} \quad \text{for all } \boldsymbol{w} \in L^2(\Omega) \tag{3.233}$$

and

$$\lim_{t\to 0_+} \int_\Omega \mathbb{B}(t) \cdot \mathbb{W} = \lim_{\ell\to\infty} \int_\Omega \mathbb{B}_0^{\frac{1}{\ell}} \cdot \mathbb{W} = \int_\Omega \mathbb{B}_0 \cdot \mathbb{W} \quad \text{for all } \mathbb{W} \in L^{\sigma'}(\Omega). \tag{3.234}$$

Before we improve this information, it is worthy to realize that, unlike in the theory of Navier-Stokes(-Fourier) systems, we can not draw any information about $\limsup_{t\to 0_+} \|\boldsymbol{v}(t)\|_2^2$ from the (kinetic) energy estimate because of the presence of $\theta\mathbb{B}$ in (3.57). However, we can use roughly this idea for $\mathbb{B}$. Indeed, let us return to (3.162), where we apply the Young inequality to deduce that

$$\int_\Omega (\operatorname{tr} \mathbb{B}_\ell(t))^\sigma \le \int_\Omega (\operatorname{tr} \mathbb{B}_0^\omega)^\sigma + C \int_0^t \int_\Omega |\mathbb{D}\boldsymbol{v}_\ell|^{\frac{q+\sigma}{q}} + Ct$$

for all $t \in [0,T]$ (appealing also to (3.232)). Recalling (3.178) and (3.177), we see that application of Hölder's inequality yields

$$\int_\Omega (\operatorname{tr} \mathbb{B}_\ell(t))^\sigma \le \int_\Omega (\operatorname{tr} \mathbb{B}_0^\omega)^\sigma + Ct^{\varepsilon_1} + Ct$$

for certain $\varepsilon_1 > 0$.[5] Thus, using (3.208) and Fatou's lemma on the left hand side and (3.74) on the right hand side, we get, for all $t \in [0,T]$, that

$$\int_\Omega (\operatorname{tr} \mathbb{B}(t))^\sigma \le \liminf_{\ell\to\infty} \int_\Omega (\operatorname{tr} \mathbb{B}_0^{\frac{1}{\ell}})^\sigma + Ct^{\varepsilon_1} + Ct = \int_\Omega (\operatorname{tr} \mathbb{B}_0)^\sigma + Ct^{\varepsilon_1} + Ct.$$

Thus, by taking the limes superior, we arrive at

$$\limsup_{t\to 0_+} \int_\Omega (\operatorname{tr} \mathbb{B}(t))^\sigma \le \int_\Omega (\operatorname{tr} \mathbb{B}_0)^\sigma. \tag{3.235}$$

[5] This is possible whenever $\sigma < q$. In the special case $\sigma = q$, $p = 2$, one has to derive a version of (3.162) with the number σ replaced by $\sigma_1 < \sigma$.

It is evident that the functional

$$\mathbb{A} \mapsto \Big(\int_\Omega (\operatorname{tr}\mathbb{A})^\sigma\Big)^{\frac{1}{\sigma}} \quad \text{is a norm in} \quad \{\mathbb{A} \in L^\sigma(\Omega);\ \mathbb{A} \in \mathbb{R}^{d\times d}_{>0}\}. \tag{3.236}$$

Then, using (3.234), weak lower semi-continuity of that norm and (3.235), we obtain

$$\lim_{t\to 0_+} \int_\Omega (\operatorname{tr}\mathbb{B}(t))^\sigma = \int_\Omega (\operatorname{tr}\mathbb{B}_0)^\sigma. \tag{3.237}$$

Since the norm defined in (3.236) is, by (4.37), one of the equivalent norms in the space $L^\sigma(\Omega)$, which is uniformly convex as $\sigma > 1$, we conclude using the Radon-Riesz property, (3.234) and (3.237) that (3.62) holds.

Next we focus on obtaining the initial condition for θ. From (3.60), (3.45), (3.233) and weak lower semi-continuity, we get

$$\begin{aligned}
\int_\Omega (\tfrac{1}{2}|\boldsymbol{v}_0|^2 + c_v\theta_0) &\geq \operatorname*{ess\,lim\,sup}_{t\to 0_+} \int_\Omega (\tfrac{1}{2}|\boldsymbol{v}(t)|^2 + c_v\theta(t)) - \lim_{t\to 0_+}\int_0^t \|\boldsymbol{f}\|_2\|\boldsymbol{v}\|_2 \\
&\geq \liminf_{t\to 0_+} \int_\Omega \tfrac{1}{2}|\boldsymbol{v}(t)|^2 + \operatorname*{ess\,lim\,sup}_{t\to 0_+} \int_\Omega c_v\theta(t) \\
&\geq \int_\Omega \tfrac{1}{2}|\boldsymbol{v}_0|^2 + \operatorname*{ess\,lim\,sup}_{t\to 0_+} \int_\Omega c_v\theta(t),
\end{aligned}$$

hence

$$\operatorname*{ess\,lim\,sup}_{t\to 0_+} \int_\Omega \theta(t) \leq \int_\Omega \theta_0. \tag{3.238}$$

To obtain also the corresponding lower estimate, we need to extract the available information from the entropy inequality (3.59). To this end, we proceed analogously as we did between (3.146) and (3.150), with two small nuances, however. First, now we are working with an inequality and thus ϕ has to be non-negative. Second, instead of (3.147), we can only use that

$$u = \int_\Omega \eta\phi \in L^\infty(0,T).$$

Consequently, the fixed time $t \in (0,T)$ has to be chosen as a Lebesgue point of u and to conclude (3.149), we use the Lebesgue differentiation theorem (for which we can refer to [29, Corollary III.12.7]). This way, we obtain

$$\int_\Omega \eta(t)\phi + \int_0^t\int_\Omega \boldsymbol{j}\cdot\nabla\phi \geq \int_\Omega \eta_0\phi + \int_0^t\int_\Omega \xi\phi \quad \text{for all } \phi \in W^{1,\infty}(\Omega),\ \phi \geq 0, \tag{3.239}$$

a.e. in $(0,T)$ (in every Lebesgue point of $\int_\Omega \eta\phi$), where

$$\begin{aligned}
\boldsymbol{j} &= -\boldsymbol{v}\eta + \kappa(\theta)\nabla\ln\theta - \mu\lambda(\theta)\nabla(\operatorname{tr}\mathbb{B} - d - \ln\det\mathbb{B}), \\
\xi &= \Big(\frac{2\nu(\theta)}{\theta}|\mathbb{D}\boldsymbol{v}|^2 + \kappa(\theta)|\nabla\ln\theta|^2 \\
&\qquad + \mu\delta(\theta)\gamma(|\mathbb{B}-\mathbb{I}|)|\mathbb{B}^{\frac{1}{2}} - \mathbb{B}^{-\frac{1}{2}}|^2 + \mu\lambda(\theta)|\mathbb{B}^{-\frac{1}{2}}\nabla\mathbb{B}\mathbb{B}^{-\frac{1}{2}}|^2, \phi\Big)
\end{aligned}$$

are both integrable functions. Hence, by taking $\operatorname{ess\,lim\,inf}_{t\to 0_+}$ of (3.239), we deduce

$$\operatorname*{ess\,lim\,inf}_{t\to 0_+} \int_\Omega \eta(t)\phi \geq \int_\Omega \eta_0\phi,$$

which is (3.64).[6] Let us now fix $\varphi \in \mathcal{C}^1(\Omega)$ such that $\varphi \geq 0$ in Ω and $\int_\Omega \varphi = 1$. Since ψ_2 is convex (cf. (2.27)), we get from (3.64) and (3.234) (or (3.62)) that

$$\begin{aligned}\int_\Omega c_v \ln \theta_0 \varphi = \int_\Omega \eta_0 \varphi + \int_\Omega \mu \psi_2(\mathbb{B})\varphi &\leq \operatorname{ess}\liminf_{t\to 0_+} \int_\Omega \eta(t)\varphi + \liminf_{t\to 0_+} \int_\Omega \psi_2(\mathbb{B}(t))\varphi \\ &\leq \operatorname{ess}\liminf_{t\to 0_+} \int_\Omega c_v \ln \theta(t)\varphi.\end{aligned}$$

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

$$\begin{aligned}\exp\left(\frac{1}{2}\int_\Omega \ln \theta_0 \varphi\right) &\leq \exp\left(\frac{1}{2} \operatorname{ess}\liminf_{t\to 0_+} \int_\Omega \ln \theta(t)\varphi\right) \\ &= \operatorname{ess}\liminf_{t\to 0_+} \exp\left(\int_\Omega \ln \sqrt{\theta(t)}\varphi\right) \\ &\leq \operatorname{ess}\liminf_{t\to 0_+} \int_\Omega \sqrt{\theta(t)}\varphi.\end{aligned} \tag{3.240}$$

By (3.60), we have

$$\operatorname{ess}\sup_{(0,T)} \left\|\sqrt{\theta}\right\|_2 < \infty,$$

which means that

$$\left\|\sqrt{\theta(t)}\right\|_2 \leq C \quad \text{in } (0,T) \setminus N_0, \tag{3.241}$$

where C is independent of t and N_0 is a set of zero Lebesgue measure. Let us denote

$$L = \operatorname{ess}\limsup_{t\to 0_+} \left\|\sqrt{\theta(t)} - \sqrt{\theta_0}\right\|_2. \tag{3.242}$$

Then, since $|N_0| = 0$ (and thus the values at points $t \in N_0$ can not affect L), we can find a set $N \subset (0,T)$, such that $N \supset N_0$, $|N| = 0$, and there exists a sequence $\{t_n\}_{n=1}^\infty \subset (0,T) \setminus N$ for which

$$\lim_{n\to\infty} \left\|\sqrt{\theta(t_n)} - \sqrt{\theta_0}\right\|_2 = L. \tag{3.243}$$

Using (3.241) and reflexivity of $L^2(\Omega)$, we get a subsequence $\{s_n\}_{n=1}^\infty \subset \{t_n\}_{n=1}^\infty$ and a function $h \in L^2(\Omega)$, such that

$$\sqrt{\theta(s_n)} \rightharpoonup h \quad \text{weakly in } L^2(\Omega). \tag{3.244}$$

Therefore, we can write

$$\operatorname{ess}\liminf_{t\to 0_+} \int_\Omega \sqrt{\theta(t)}\varphi \leq \lim_{n\to\infty} \int_\Omega \sqrt{\theta(s_n)}\varphi = \int_\Omega h\varphi,$$

which, together with (3.240), gives

$$\exp\left(\frac{1}{2}\int_\Omega \ln \theta_0 \varphi\right) \leq \int_\Omega h\varphi. \tag{3.245}$$

[6] Using a similar technique, we could prove that η is essentially lower semi-continuous on $[0,T]$ in the weak topology of measures.

Let x_0 be a Lebesgue point of both $\ln\theta_0$ and h. Furthermore, let $B(\varepsilon)$ be the ball centered in x_0 with radius ε so small that $B(\varepsilon)\subset\Omega$ and choose $\varphi_\varepsilon\in\mathcal{C}^1(\mathbb{R}^d)$, $\varepsilon>0$, with the properties:

$$0\le\varphi_\varepsilon\le\frac{2}{|B(\varepsilon)|}\quad\text{in }\Omega,\quad \varphi=0\quad\text{in }\mathbb{R}^d\setminus B(\varepsilon)\quad\text{and}\quad\int_{\mathbb{R}^d}\varphi_\varepsilon=1.$$

Then, by the Lebesgue differentiation theorem, we obtain

$$\begin{aligned}\left|\int_\Omega\ln\theta_0\varphi_\varepsilon-(\ln\theta_0)(x_0)\right|&\le\int_\Omega|\ln\theta_0-(\ln\theta_0)(x_0)|\,\varphi_\varepsilon\\&\le\frac{2}{|B(\varepsilon)|}\int_{B(\varepsilon)}|\ln\theta_0-(\ln\theta_0)(x_0)|\to0\quad\text{as }\varepsilon\to0_+\end{aligned}\tag{3.246}$$

and similarly also

$$\left|\int_\Omega h\varphi_\varepsilon-h(x_0)\right|\to0\quad\text{as }\varepsilon\to0_+.\tag{3.247}$$

On choosing $\varphi=\varphi_\varepsilon$ in (3.245), taking $\varepsilon\to0_+$ and using (3.246), (3.247), we get

$$\sqrt{\theta_0(x_0)}=\exp(\tfrac12\ln\theta(x_0))\le h(x_0),$$

and, consequently, also

$$\sqrt{\theta_0}\le h\quad\text{a.e. in }\Omega.$$

From this, (3.243), (3.244) and (3.238), we deduce that

$$\begin{aligned}L^2=\lim_{n\to\infty}\left\|\sqrt{\theta(s_n)}-\sqrt{\theta_0}\right\|_2^2&\le\operatorname*{ess\,lim\,sup}_{t\to0_+}\int_\Omega\theta(t)-2\lim_{n\to\infty}\int_\Omega\sqrt{\theta(s_n)}\sqrt{\theta_0}+\int_\Omega\theta_0\\&\le2\int_\Omega\theta_0-2\int_\Omega h\sqrt{\theta_0}\le0,\end{aligned}$$

which, looking at the (3.242), yields

$$\operatorname*{ess\,lim}_{t\to0_+}\left\|\sqrt{\theta(t)}-\sqrt{\theta_0}\right\|_2=0.$$

Hence, by Hölder's inequality, we get

$$\begin{aligned}\operatorname*{ess\,lim\,sup}_{t\to0_+}\|\theta(t)-\theta_0\|_1&=\operatorname*{ess\,lim\,sup}_{t\to0_+}\int_\Omega\left|\left(\sqrt{\theta(t)}+\sqrt{\theta_0}\right)\left(\sqrt{\theta(t)}-\sqrt{\theta_0}\right)\right|\\&\le C\operatorname*{ess\,lim\,sup}_{t\to0_+}\left\|\sqrt{\theta(t)}-\sqrt{\theta_0}\right\|_2=0,\end{aligned}$$

which is (3.63).

Using information above, we can finally improve the initial condition for $\boldsymbol{v}$ as well. Indeed, from (3.60), (3.63) and (3.45), we obtain

$$\begin{aligned}\limsup_{t\to0_+}\int_\Omega\tfrac12|\boldsymbol{v}(t)|^2&\le\operatorname*{ess\,lim\,sup}_{t\to0_+}\int_\Omega(\tfrac12|\boldsymbol{v}(t)|^2+c_v\theta(t))-\operatorname*{ess\,lim\,inf}_{t\to0_+}\int_\Omega c_v\theta(t)\\&\le\int_\Omega(\tfrac12|\boldsymbol{v}_0|^2+c_v\theta_0)+\lim_{t\to0_+}\int_0^t(\boldsymbol{f},\boldsymbol{v})-\int_\Omega c_v\theta_0=\int_\Omega\tfrac12|\boldsymbol{v}_0|^2.\end{aligned}$$

Thus, using also (3.233), we conclude that

$$\limsup_{t\to0_+}\|\boldsymbol{v}(t)-\boldsymbol{v}_0\|_2^2=\limsup_{t\to0_+}\int_\Omega|\boldsymbol{v}(t)|^2+\int_\Omega|\boldsymbol{v}_0|^2-2\lim_{t\to0_+}\int_\Omega\boldsymbol{v}(t)\cdot\boldsymbol{v}_0\le0,$$

which implies (3.61).

The proof of Theorem 3.2 is complete. □

We remark that the differentiability of $\mathbb{B}$ could be slightly improved to

$$\nabla\mathbb{B} \in L^{\frac{\sigma}{2}}(Q)$$

if we used $\mathbb{B}^{\sigma-1}$ instead of $(\operatorname{tr}\mathbb{B})^{\sigma-1}\mathbb{I}$ as a test function in (3.58) (or rather its truncation, recall (3.157) and (3.159)). However, as long as $d \geq 2$, this additional information does not improve the integrability of $\mathbb{B}$ (via the Sobolev embeddings and interpolation), and thus conditions (A0)–(A2) can not be relaxed based on this information. Since the corresponding matrix computation (in order to show $\mathbb{B}^{\sigma-1} \in W^{1,1}(\Omega)$ and $\nabla\mathbb{B} \cdot \nabla\mathbb{B}^{\sigma-1} \geq C|\nabla\mathbb{B}^{\frac{\sigma}{2}}|^2$) seems quite involved (due to the non-commutativity of $\mathbb{B}$ and $\nabla\mathbb{B}$), we decided to omit this improvement.

It remains unclear whether a suitable weak solution constructed above fulfils the local balance of total/internal energy. It seems that the conditions (A0)–(A2) are insufficient to give any sense to these balances. However, it turns out that if the parameters r, q and ϱ are large enough (in certain sense), then the local balance of total/internal energy holds. In the next two sections we show that the sufficient conditions for this to happen are (B1), (B2) and (C), respectively.

3.6 Local balance of total energy

If we insisted on fulfilment of

$$\partial_t E + \boldsymbol{v} \cdot \nabla E - \operatorname{div}(\kappa(\theta)\nabla\theta) = \operatorname{div}(-\mathrm{p}\boldsymbol{v} + 2\nu(\theta)(\mathbb{D}\boldsymbol{v})\boldsymbol{v} + 2a\mu\theta\mathbb{B}\boldsymbol{v}) + \boldsymbol{f} \cdot \boldsymbol{v}$$

locally, we need to ensure that every its term can be defined in a weak (i.e., the distributional sense). Thus, in addition to the requirements imposed by Definition 3.1, it is necessary that the terms $|\boldsymbol{v}|^3$, $\kappa(\theta)\nabla\theta$, $\mathrm{p}\boldsymbol{v}$, $(\mathbb{D}\boldsymbol{v})\boldsymbol{v}$ and $\theta\mathbb{B}\boldsymbol{v}$ are integrable (or actually slightly better for the compactness arguments to work). In particular, we need to construct the pressure p which appears in (3.9) explicitly. To this end, we need to consider different boundary conditions than we did so far (there are indications that with Dirichlet boundary conditions for $\boldsymbol{v}$, the pressure is only a time-distribution, see [8]). For example, if we consider a Navier's slip $\boldsymbol{v} \cdot \boldsymbol{n} = 0$, $\boldsymbol{n} \times (\mathbb{T}\boldsymbol{n} + \boldsymbol{v}) \times \boldsymbol{n} = 0$ on $(0,T) \times \partial\Omega$ as in Chapter 5 or in [12], we can formally estimate the pressure using regularity of solutions to the following Neumann problem:

$$\Delta\mathrm{p} = \operatorname{div}\operatorname{div}(\boldsymbol{v} \otimes \boldsymbol{v} - 2\nu(\theta)\mathbb{D}\boldsymbol{v} - 2a\mu\theta\mathbb{B}) + \operatorname{div}\boldsymbol{f}.$$

If the domain Ω is sufficiently smooth, then it follows from this equation that the pressure p has the same integrability as the terms inside the double divergence operator on the right hand side (see e.g. [37]). Thus in fact, we only need to verify that the terms $|\boldsymbol{v}|^3$, $(\mathbb{D}\boldsymbol{v})\boldsymbol{v}$, $\theta\mathbb{B}\boldsymbol{v}$ and $\kappa(\theta)\nabla\theta$ are integrable (with some exponent greater than one).

The condition (3.39) (which is an equivalent version of (B2)) yields

$$p = \frac{q+\sigma}{2r_d' - 1} > \frac{q+\sigma}{\frac{d+2}{3d}(q+\sigma)} = \frac{3d}{d+2} \quad \text{if } r_d < r_1$$

while if $r_d > r_1$, we use $d < 4$ to write $2 = \frac{2d+4}{d+2} > \frac{3d}{d+2}$. Hence, property (3.65) ensures that $\boldsymbol{v} \in L^{3+\varepsilon}(Q)$ for some $\varepsilon > 0$ sufficiently small. Next, we rewrite the condition (B1) equivalently as

$$r_d' < 1 + \frac{d}{4d-2}(q+\sigma-2)$$

and then we estimate, using also $d \geq 2$, that

$$p = \frac{q+\sigma}{2r_d'-1} > \frac{q+\sigma}{q+\sigma-\frac{1}{d}}\frac{2d-1}{d} > 2 - \frac{2}{d+d} \geq 2 - \frac{2}{d+2} = 2\frac{d+1}{d+2}$$

(this inequality is obviously true also if $r_d \geq r_1$). Hence, we get

$$\frac{1}{p} + \frac{1}{p(1+\frac{2}{d})} = \frac{1+\frac{d}{d+2}}{p} < 1$$

and thus, with the help of Hölder's inequality, (3.47) and (3.65), we see that $(\mathbb{D}\boldsymbol{v})\boldsymbol{v} \in L^{1+\varepsilon}(Q)$. Furthermore, we remark that $\theta\mathbb{B} \in L^p(Q)$ (i.e., the same integrability as that of $\mathbb{D}\boldsymbol{v}$) if $r_d < r_1$. Indeed, this follows from (3.41) and (3.40) as

$$\frac{1}{R_d} + \frac{1}{q+\sigma} = \frac{\frac{2}{r_d-1}+1}{q+\sigma} = \frac{2r_d'-1}{q+\sigma} = \frac{1}{p}$$

and thus we deduce using (3.53) and (3.49) that also $\theta\mathbb{B}\boldsymbol{v} \in L^{1+\varepsilon}(Q)$. If $r_d \geq r_1$, then we write

$$\frac{1}{r_d+1} + \frac{1}{q+\sigma} + \frac{1}{2(1+\frac{2}{d})} < \frac{q+\sigma-2}{2(q+\sigma)} + \frac{1}{q+\sigma} + \frac{d}{2d+4} = \frac{2d+2}{2d+4} < 1$$

and we see by Hölder's inequality that also in this case we have $\theta\mathbb{B}\boldsymbol{v} \in L^{1+\varepsilon}(Q)$. Finally, to verify that the term $\kappa(\theta)\nabla\theta$ is integrable, we first use (3.33) to write

$$|\kappa(\theta)\nabla\theta| \leq C|\nabla\theta| + C|\theta^r\nabla\theta| \leq C|\nabla\theta| + C\theta^{r+1-\frac{R}{2}}|\nabla\theta^{\frac{R}{2}}|,$$

where note that $r+1-\frac{R}{2} > 0$ since

$$\frac{R}{2} = \frac{R_d}{2} - \frac{1}{d} = \frac{p(r_d-1)(2r_d'-1)}{4} - \frac{1}{d} < \frac{r_d+1}{2} - \frac{1}{d} = \frac{r+1}{2}$$

if $r_d < r_1$ and the other case is obvious. As (B1) gives

$$\begin{aligned} 2\left(r+1-\frac{R}{2}\right) &= 2r+2-\frac{(r_d-1)(q+\sigma)}{2}+\frac{2}{d} \\ &= R_d+4-\frac{2}{d}-(r_d-1)(q+\sigma-2) < R_d, \end{aligned}$$

the term $\theta^{r+1-\frac{R}{2}}|\nabla\theta^{\frac{R}{2}}|$ is integrable in view of (3.52), (3.53) and Hölder's inequality. To see that also $\nabla\theta \in L^{1+\varepsilon}(Q)$, let us distinguish two cases. If $R \geq 2$, we write

$$|\nabla\theta| = \theta|\nabla\ln\theta|,$$

use Hölder's inequality, (3.54), (3.53) and the fact that $R_d > 2$. On the other hand, if $R < 2$, we can write

$$|\nabla\theta| = \frac{2}{R}\theta^{1-\frac{R}{2}}|\nabla\theta^{\frac{R}{2}}|. \tag{3.248}$$

Then, note that (3.41) and (B1) imply

$$R_d > \frac{(r_d-1)(q+\sigma-2)}{2} > 2-\frac{1}{d},$$

hence

$$2-R = R_d + 2 + \frac{2}{d} - 2R_d < R_d + \frac{4}{d} - 2 \le R_d$$

due to $d \ge 2$. Therefore, if we apply (3.248), the Hölder inequality, (3.53) and (3.52), we also obtain $\nabla\theta \in L^{1+\varepsilon}(Q)$.

Based on these observations, we postulate the following theorem, which however, is not proved here. Due to different boundary conditions, we denote the trace operator by $\mathcal{T}$ and we replace the space $W^{1,p}_{0,\operatorname{div}}$ by

$$\begin{aligned} W^{1,p}_{\boldsymbol{n}} &= \overline{\{\boldsymbol{w}\in\mathcal{C}^\infty(\Omega),\ \boldsymbol{w}\cdot\boldsymbol{n}=0 \ \text{ on } \ \partial\Omega\}}^{\|\cdot\|_{1,p}},\\ W^{1,p}_{\boldsymbol{n},\operatorname{div}} &= \overline{\{\boldsymbol{w}\in\mathcal{C}^\infty(\Omega),\ \boldsymbol{w}\cdot\boldsymbol{n}=0 \ \text{ on } \ \partial\Omega, \quad \operatorname{div}\boldsymbol{v}=0 \ \text{ in } \ \Omega\}}^{\|\cdot\|_{1,p}}. \end{aligned}$$

Theorem 3.3. *Let the assumptions of Theorem 3.2 be satisfied and suppose that functions* $(\boldsymbol{v}, p, \mathbb{B}, \eta)$ *are a suitable weak solution in the sense of Definition 3.1 with properties* (3.47), (3.48) *replaced by*

$$\begin{aligned} \boldsymbol{v} &\in L^p(0,T;W^{1,p}_{\boldsymbol{n},\operatorname{div}})\cap\mathcal{C}_w([0,T];L^2(\Omega))\\ \partial_t\boldsymbol{v} &\in L^{p_1}(0,T;W^{-1,p_1}_{\boldsymbol{n},\operatorname{div}})\\ \mathcal{T}\boldsymbol{v} &\in L^2(0,T;L^2(\partial\Omega)) \end{aligned}$$

and with equation (3.57) *replaced by*

$$\begin{aligned} \int_0^T \Big(\langle\partial_t\boldsymbol{v},\boldsymbol{\varphi}\rangle - (\boldsymbol{v}\otimes\boldsymbol{v},\nabla\boldsymbol{\varphi}) + (2\nu(\theta)\mathbb{D}\boldsymbol{v},\nabla\boldsymbol{\varphi}) + \int_{\partial\Omega}\mathcal{T}\boldsymbol{v}\cdot\mathcal{T}\boldsymbol{\varphi}\Big)&\\ = \int_0^T (\mathrm{p}\mathbb{I} - 2a\mu\theta\mathbb{B},\nabla\boldsymbol{\varphi}) + \int_0^T(\boldsymbol{f},\boldsymbol{\varphi})&\\ \textit{for all} \quad \boldsymbol{\varphi}\in L^{p_1'}(0,T;W^{1,p_1'}_{\boldsymbol{n}}).& \end{aligned}$$

Moreover, let $d=2$, *or* $d=3$, *suppose that* Ω *is a domain of class* $\mathcal{C}^{1,1}$ *and assume that* (B1), (B2) *hold.*

Then, the functions E *and* E_0 *defined by*

$$E = \frac{1}{2}|\boldsymbol{v}|^2 + c_v\theta \quad \textit{and} \quad E_0 = \frac{1}{2}|\boldsymbol{v}_0|^2 + c_v\theta_0,$$

respectively, satisfy

$$E\in L^\infty(0,T;L^1(\Omega))\cap L^{p(\frac{1}{2}+\frac{1}{d})}(Q)$$

and

$$\begin{aligned} -(E_0,\phi)\varphi(0) - \int_0^T (E,\phi)\partial_t\varphi - \int_0^T(E\boldsymbol{v},\nabla\phi)\varphi&\\ + \int_0^T\int_{\partial\Omega}|\mathcal{T}\boldsymbol{v}|^2\phi\varphi + \int_0^T(\kappa(\theta)\nabla\theta,\nabla\phi)\varphi&\\ = \int_0^T(\mathrm{p}\boldsymbol{v} - 2\nu(\theta)(\mathbb{D}\boldsymbol{v})\boldsymbol{v} - 2a\mu\theta\mathbb{B}\boldsymbol{v},\nabla\phi)\varphi& \end{aligned} \tag{3.249}$$

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

To prove this theorem, one has to introduce an additional layer of approximation in (3.67)–(3.68), in which the pressure p is constructed. This can be achieved by relaxing the divergence free condition on $\boldsymbol{v}$ in the Galerkin approximation and by adding the constraint

$$\varepsilon \Delta \mathrm{p} = \operatorname{div} \boldsymbol{v}$$

instead. Then, as $\varepsilon \to 0_+$, one recovers $\operatorname{div} \boldsymbol{v} = 0$ and moreover, appealing to the regularity of solution to the Neumann-Poisson problem, one can deduce the a priori estimates for the pressure. See [12] for details. Besides that, one can proceed analogously as in the proof of Theorem 3.2. The compactness of all the terms appearing in (3.249) follows from (B1), (B2) as was verified above. A detailed proof of Theorem 3.3 falls outside the scope and aims of this thesis, and therefore is omitted.

3.7 Local balance of internal energy

Finally, we investigate when the local balance of internal energy

$$c_v \partial_t \theta + c_v \boldsymbol{v} \cdot \nabla \theta - \operatorname{div}(\kappa(\theta)\nabla\theta) \geq 2\nu(\theta)|\mathbb{D}\boldsymbol{v}|^2 + 2a\mu\theta\mathbb{B} \cdot \mathbb{D}\boldsymbol{v} \tag{3.250}$$

holds. To make the right hand side of this equation integrable, we need to impose conditions that are even stricter than those from the previous section. Let us now verify that the condition (C) is optimal for this purpose. First of all, recall that (C) is equivalent to $r_d > r_1$ and that we have $p = 2$ and $R_d = r_d + 1) > 2$ in this case. Hence, we easily deduce that the terms $\theta\boldsymbol{v}$ and $\kappa(\theta)\nabla\theta$ are well defined, using the ideas from the last section. Next, the viscous dissipation term is obviously integrable as (3.31) is assumed and $p = 2$. However, we can not improve this information, and therefore (3.250) is stated only as an inequality. Finally, the term $\theta\mathbb{B} \cdot \mathbb{D}\boldsymbol{v}$ is integrable since (C) gives

$$\begin{aligned}\frac{1}{R_d} + \frac{1}{q+\sigma} + \frac{1}{p} = \frac{1}{r_d + 1)} + \frac{1}{q+\sigma} + \frac{1}{p} &< \frac{1}{\frac{q+\sigma+2}{q+\sigma-2} + 1} + \frac{1}{q+\sigma} + \frac{1}{2} \\ &= \frac{q+\sigma-2}{2(q+\sigma)} + \frac{q+\sigma+2}{2(q+\sigma)} = 1.\end{aligned}$$

Now let us state the precise result.

Theorem 3.4. *Let the assumptions of Theorem 3.2 be satisfied and let $(\boldsymbol{v}, \mathbb{B}, \theta, \eta)$ be a corresponding suitable weak solution.*

If the condition (C) *holds, then*

$$\begin{aligned}-(c_v\theta_0, \phi)\varphi(0) - \int_0^T (\theta, \phi)\partial_t\varphi - \int_0^T (c_v\theta\boldsymbol{v}, \nabla\phi)\varphi + \int_0^T (\kappa(\theta)\nabla\theta, \nabla\phi)\varphi \\ \geq \int_0^T (2\nu(\theta)|\mathbb{D}\boldsymbol{v}|^2 + 2a\mu\theta\mathbb{B}\cdot\mathbb{D}\boldsymbol{v}, \phi)\varphi\end{aligned} \tag{3.251}$$

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

Proof. In the proof of Theorem 3.2 it was shown that the suitable weak solution $(\boldsymbol{v}, \mathbb{B}, \theta, \eta)$ can be constructed as weak limit of a sequence $(\boldsymbol{v}_\ell, \mathbb{B}_\ell, \theta_\ell, \eta_\ell)$ satisfying, among other things, the equation

$$
\begin{aligned}
-(c_v\theta_0^{\frac{1}{\ell}},\phi)\varphi(0) - \int_0^T(\theta_\ell,\phi)\partial_t\varphi - \int_0^T(c_v\theta_\ell\boldsymbol{v}_\ell,\nabla\phi)\varphi + \int_0^T(\kappa(\theta_\ell)\nabla\theta_\ell,\nabla\phi)\varphi \\
\geq \int_0^T(2\nu(\theta_\ell)|\mathbb{D}\boldsymbol{v}_\ell|^2 + 2a\mu g_{\frac{1}{\ell}}(\mathbb{B}_\ell,\theta_\ell)\theta_\ell\mathbb{B}_\ell\cdot\mathbb{D}\boldsymbol{v}_\ell,\phi)\varphi
\end{aligned}
\tag{3.252}
$$

for all $\varphi \in W^{1,\infty}(0,T)$ and every $\phi \in W^{1,\infty}(\Omega)$ (recall (3.124), take $\tau = \varphi\phi$ and integrate by parts in the first two terms). Moreover, as $p = 2$ and $R_d = r_d{+}1)$, we have the following information about the convergence of $(\boldsymbol{v}_\ell, \mathbb{B}_\ell, \theta_\ell)$ to $(\boldsymbol{v}, \mathbb{B}, \theta)$:

$$
\begin{aligned}
\boldsymbol{v}_\ell &\rightharpoonup \boldsymbol{v} && \text{weakly in } L^2(0,T;W^{1,2}_{0,\mathrm{div}}), && (3.253)\\
\boldsymbol{v}_\ell &\to \boldsymbol{v} && \text{strongly in } L^{2+\frac{4}{d})}(Q) \text{ and a.e. in } Q, && (3.254)\\
\mathbb{B}_\ell &\to \mathbb{B} && \text{strongly in } L^{q+\sigma)}(Q) \text{ and a.e. in } Q, && (3.255)\\
\theta_\ell &\to \theta && \text{strongly in } L^{R_d}(Q) \text{ and a.e. in } Q, && (3.256)\\
\nabla\theta_\ell^{\frac{R}{2}} &\rightharpoonup \nabla\theta^{\frac{R}{2}} && \text{weakly in } L^2(Q), && (3.257)\\
g_{\frac{1}{\ell}}(\mathbb{B}_\ell,\theta_\ell) &\to 1 && \text{strongly in } L^{\infty)}(Q). && (3.258)
\end{aligned}
$$

(recall (3.205)–(3.210), (3.215), (3.176) and (3.216)). It remains to take the limit $\ell \to \infty$ in (3.252).

Taking the limit in the first two terms of (3.252) is easy, we use (3.75) and (3.256), respectively. In the convective term we apply (3.256), (3.254) and the fact that $R_d = r_d + 1) > 2$. Next, we observe that

$$
0 < r + 1 - \frac{R}{2} = r + 1 - \frac{r+1)}{2} < r + 1 - \frac{r + 1 - \frac{1}{d}}{2} = \frac{r_d + 1 - \frac{1}{d}}{2} < \frac{R_d}{2}
$$

and thus

$$
\kappa(\theta_\ell)\theta_\ell^{1-\frac{R}{2}} \to \kappa(\theta)\theta^{1-\frac{R}{2}} \quad \text{strongly in } L^2(Q)
$$

by (3.33), (3.30), (3.256) and Vitali's theorem. Hence, using also (3.257) (and $\theta > 0$ a.e. in Q), we arrive at

$$
\kappa(\theta_\ell)\nabla\theta_\ell = \frac{2}{R}\theta_\ell^{1-\frac{R}{2}}\kappa(\theta_\ell)\nabla\theta_\ell^{\frac{R}{2}} \rightharpoonup \frac{2}{R}\theta^{1-\frac{R}{2}}\kappa(\theta)\nabla\theta^{\frac{R}{2}} = \kappa(\theta)\nabla\theta
$$

weakly in $L^1(Q)$.

To see that the term $\frac{1}{\ell}\kappa(\theta_\ell)\nabla\theta_\ell$ vanishes in the limit $\ell \to \infty$, we can use roughly the same argumentation as for the analogous term in the entropy inequality (recall (3.203)). First we need to realize that in the estimates that follow after (3.163), we can keep the term

$$
\omega\int_Q |\nabla\theta_\ell|^r\nabla\theta_\ell\cdot\nabla\tau_\beta = \beta\omega\int_Q\frac{|\nabla\theta|^{r+2}}{\theta^{\beta+1}}
$$

on the left hand side (this term was previously omitted on the first occasion since we were interested only in ω-uniform estimates). This way, we observe that the estimate (3.175) can be replaced by

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

(recall that we are now in the case $r_d > r_1$). Then, using an analogous estimation as in (3.203), we arrive at

$$\begin{aligned}\|\omega|\nabla\theta_\ell|^r\nabla\theta_\ell\|_{\alpha;Q} &= \omega\left(\int_Q \frac{|\nabla\theta_\ell|^{(r+1)\alpha}}{\theta_\ell^{\frac{\alpha(1+\beta)(r+1)}{r+2}}}\theta_\ell^{\frac{\alpha(1+\beta)(r+1)}{r+2}}\right)^{\frac{1}{\alpha}} \\ &\le \omega^{\frac{1}{r+2}}\left(\omega\int_Q \frac{|\nabla\theta_\ell|^{r+2}}{\theta_\ell^{\beta+1}}\right)^{\frac{r+1}{r+2}} \|\theta_\ell\|^{\frac{(1+\beta)(r+1)}{r+2}}_{(\frac{r+2}{\alpha(r+1)})'\frac{\alpha(1+\beta)(r+1)}{r+2}} \\ &\le C(\beta)\omega^{\frac{1}{r+2}}\|\theta_\ell\|^{\frac{(1+\beta)(r+1)}{r+2}}_{R_d;Q}\end{aligned}$$

where α solves

$$\left(\frac{r+2}{\alpha(r+1)}\right)'\frac{\alpha(1+\beta)(r+1)}{r+2} = R_d.$$

Since this is equivalent to

$$\alpha = \frac{\frac{r+2}{r+1}}{1+\frac{1+\beta}{R_d}}$$

and we have $R_d = r_d + 1) > r+1$ we see that $\beta > 0$ can be chosen so small that $\alpha > 1$, and hereby we get

$$\|\omega|\nabla\theta_\ell|^r\nabla\theta_\ell\|_{\alpha;Q} \le \frac{C}{\ell^{\frac{1}{r+2}}} \to 0 \quad \text{as} \quad \ell\to\infty.$$

Next, using (3.31), (3.256) and (3.253), it is easy to see that

$$\sqrt{2\nu(\theta_\ell)}\mathbb{D}\boldsymbol{v}_\ell \rightharpoonup \sqrt{2\nu(\theta)}\mathbb{D}\boldsymbol{v} \quad \text{weakly in } L^2(Q).$$

Then, by the weak lower semi-continuity, we get

$$\liminf_{\ell\to\infty}\int_Q 2\nu(\theta_\ell)|\mathbb{D}\boldsymbol{v}_\ell|^2 \ge \int_Q 2\nu(\theta)|\mathbb{D}\boldsymbol{v}|^2.$$

Finally, since

$$\frac{1}{R_d}+\frac{1}{q+\sigma} = \frac{1}{r_d+1)}+\frac{1}{q+\sigma} < \frac{1}{\frac{q+\sigma+2}{q+\sigma-2}+1}+\frac{1}{q+\sigma} = \frac{1}{2}$$

due to $r_d > r_1$ (which is equivalent to (C)), we deduce from (3.256) and (3.255) that

$$\theta_\ell\mathbb{B}_\ell \to \theta\mathbb{B} \quad \text{strongly in } L^{2+\varepsilon}(Q)$$

for some sufficiently small $\varepsilon > 0$. Hence, by (3.258), it is also true that

$$g(\mathbb{B}_\ell,\theta_\ell)\theta_\ell\mathbb{B}_\ell \to \theta\mathbb{B} \quad \text{strongly in } L^2(Q).$$

But this together with (3.253) implies

$$g(\mathbb{B}_\ell,\theta_\ell)\theta_\ell\mathbb{B}_\ell\cdot\mathbb{D}\boldsymbol{v}_\ell \rightharpoonup \theta\mathbb{B}\cdot\mathbb{D}\boldsymbol{v} \quad \text{weakly in } L^1(Q).$$

Therefore, we can indeed take the limit $\ell\to\infty$ in every term of (3.252), leading to (3.251), and the proof is finished. □

4. Auxiliary analytic tools

In this chapter, we prove all the auxiliary assertions that were used in the thesis. Based on similar results in simpler settings, it should be intuitively clear that these assertions hold true, however it seems difficult (or impossible in some cases) to find a precise reference.

For the purposes of this chapter, 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 $\mathbb{R}^d$, $d \in \mathbb{N}$.

4.1 Sequential (weak) compactness

To deduce that some subsequences of solutions constructed in Chapter 3 converge in an appropriate sense, we need to combine several classical results of functional analysis. To make the corresponding argumentation clear, let us formulate precisely the following auxiliary lemma, which is used many times in the proof of Theorem 3.2.

Lemma 4.1. *Let $1 < p_1, p_2, q_2 < \infty$ be such that*

$$\frac{dq_2}{d + q_2} \leq p_1 \leq p_2$$

and suppose also $1 < v < \infty$ and $1 \leq w \leq \infty$. Let the sequence $\{u_k\}_{k=1}^{\infty} \subset L^1(Q)$ be bounded in the sense that

$$\|\nabla u_k\|_{L^{p_1}L^{p_1}} + \|u_k\|_{L^{p_2}L^{q_2}} + \|\partial_t u_k\|_{L^w W^{-m,v}} \leq C \tag{4.1}$$

for some $m \in \mathbb{N}$.

Then, there exists a subsequence, which we do not relabel, and its limit u, such that

$$\begin{aligned}
u_k &\rightharpoonup u && \text{weakly in } L^{p_1}(I; W^{1,p_1}(\Omega)), && (4.2)\\
u_k &\rightharpoonup u && \text{weakly in } L^{p_2}(I; L^{q_2}(\Omega)), && (4.3)\\
u_k &\to u && \text{almost everywhere in } Q, && (4.4)\\
u_k &\to u && \text{strongly in } L^s(I; L^s(\Omega)) \quad \text{for all } 1 \leq s < S, && (4.5)
\end{aligned}$$

where

$$S = \frac{\frac{1}{d} + \frac{1}{q_2} - \frac{1}{p_2}}{\frac{1}{p_2 d} + \frac{1}{p_1 q_2} - \frac{1}{p_1 p_2}}. \tag{4.6}$$

Moreover, if $w > 1$, then also

$$\partial_t u_k \rightharpoonup \partial_t u \quad \text{weakly in } L^w(I; W^{-m,v}(\Omega)). \tag{4.7}$$

Furthermore, if $p_1 = \infty$ or $p_2 q_2 = \infty$ then the conclusion of the lemma still holds if the weak convergence in (4.2) or (4.3), respectively, is replaced by the weak star convergence.

Proof. Let us first consider the case $p_1 p_2 q_2 < \infty$. Note that, since $p_1 \le p_2$, we have $\|u_k\|_{L^{p_1}L^1} \le C\|u_k\|_{L^{p_2}L^{q_2}} \le C$, and thus we also get

$$\|u_k\|_{L^{p_1}W^{1,p_1}} \le C$$

from (4.1) and Poincaré's inequality (see, e.g., [31, Sec. 5.8., Theorem 1]). Since the spaces $L^{p_1}(I; W^{1,p_1}(\Omega))$ and $L^{p_2}(I; L^{q_2}(\Omega))$ are reflexive Banach spaces, there exists a (not relabelled) subsequence of $\{u_\ell\}$ which is weakly converging to u in these spaces by the Kakutani theorem (see [9, Theorem 3.18] or [29, V.4, Corollary 8]). This proves (4.2) and (4.3). Further, the compact embedding

$$W^{1,p_1}(\Omega) \hookrightarrow\hookrightarrow L^{p_1}(\Omega)$$

holds by the Rellich-Kondrachov theorem (cf. [1, Theorem 6.3 part I]). Thus, since we also control $\partial_t u_k$, we can apply the Aubin-Lions lemma (see, e.g., [66, Lemma 7.7]) to select another subsequence such that

$$u_k \to u \quad \text{strongly in } L^{p_1}(I; L^{p_1}(\Omega)).$$

Thus, by the properties of (Bochner-)Lebesgue spaces, there exists another subsequence which fulfils (4.4), hence it remains to show (4.5). To this end, let $p_1 < d$ and start with the Sobolev embedding

$$W^{1,p_1}(\Omega) \hookrightarrow L^{p_1^*}(\Omega), \quad \text{where } p_1^* = \frac{dp_1}{d - p_1} \tag{4.8}$$

(for a reference see, e.g., [1, Theorem 4.12]). Then, for the $\theta \in [0,1]$ and $S \in [1,\infty]$ solving

$$\frac{1-\theta}{p_1} + \frac{\theta}{p_2} = \frac{1-\theta}{p_1^*} + \frac{\theta}{q_2} = \frac{1}{S},$$

we observe that S satisfies (4.6) and

$$\|u_k\|_{L^S L^S} \le \|u_k\|^{1-\theta}_{L^{p_1}L^{p_1^*}} \|u_k\|^{\theta}_{L^{p_2}L^{q_2}}$$

by the Hölder (interpolation) inequality. Thus the sequence $\{u_k\}_{k=1}^\infty$ is bounded in $L^S(I; L^S(\Omega))$. Therefore, by Hölder's inequality, the sequence defined as $a_k = |u_k - u|^s$, $k \in \mathbb{N}$ satisfies

$$\int_E a_k \le \|a_k\|_{\frac{S}{s}} |E|^{1/(\frac{S}{s})'} \le C|E|^{\frac{S-s}{S}},$$

for any measurable $E \subset Q$, and thus a_k is uniformly integrable whenever $1 \le s < S$. Moreover, by (4.4), we get $a_k \to 0$ a.e. in Q. Hence, by Vitali's theorem (see [29, III.6, Theorem 15]), there exists yet another subsequence of $\{u_k\}_{k=1}^\infty$, for which $a_k \to 0$ strongly in $L^1(Q)$, which is (4.5). If $p_1 > d$, then (4.8) holds with $p_1^* = \infty$ and there is no other change in the proof[1]. Finally, if $p_1^* = d$, then (4.8) can be replaced by

$$W^{1,p_1}(\Omega) \hookrightarrow L^{p_3}(\Omega) \quad \text{for any } p_3 < \infty,$$

[1]Here we are obviously losing some information. To deduce a corresponding precise result in the case $p_1 > d$, one can use the interpolation between Lebesgue and Sobolev-Slobodeckij spaces. This is omitted since in Chapter 3 we actually only need the case $p_1 < d$.

which is again sufficient to prove (4.5).

The property (4.7) holds since the Banach space $L^w(I; W^{-m,v}(\Omega))$ is reflexive if and only if $1 < w, v < \infty$.

If $p_1 p_2 q_2 = \infty$, we think of the sequence $\{u_k\}_{k=1}^{\infty}$ as of a bounded subset of an appropriate dual space. Then, relying on separability of L^1 spaces and using the Banach theorem (cf. [4, Ch. 8, Theorem 3] or [9, Corollary 3.30]), we obtain weakly star converging subsequence(s). Also, we can use the Aubin-Lions lemma on a slightly worse space (to ensure its separability and reflexivity) if $p_1 = \infty$. We can argue similarly in the case $w = \infty$. Otherwise, there is no change in the proof. $\square$

4.2 Intersections of Sobolev-Bochner spaces

Suppose that $X \hookrightarrow H \hookrightarrow X^*$ is a Gelfand triple and consider the space $W^{1,p}(I; X)$ (defined in (3.2)) with $1 \le p \le \infty$. Then, it is known that the space $\mathcal{C}^1(I; X)$ is dense in $W^{1,p}(I; X)$. Also, we have the embedding $W^{1,p}(I; X) \hookrightarrow \mathcal{C}(I; H)$ and the corresponding "integration by parts" formula holds true. These classical results can be found in [74, Problem 23.10], [66, Ch. 7] or [35, Ch. IV]). In this section, we derive an analogous result with $W^{1,p}(I; X)$ replaced by the space

$$\mathcal{W} = \{u \in L^p(I; X) \cap L^q(I; Y);\ \partial_t u \in (L^p(I; X) \cap L^q(I; Y))^*\}, \quad 1 < p, q < \infty,$$

equipped with the norm

$$\|u\|_{\mathcal{W}} = \|u\|_{L^p X \cap L^q Y} + \|\partial_t u\|_{(L^p X \cap L^q Y)^*}.$$

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). Thus, we may assume that both X and Y admit the Gelfand triplet structure with a common Hilbert space H:

$$X \overset{\mathrm{d}}{\hookrightarrow} H \overset{\mathrm{d}}{\hookrightarrow} X^* \quad \text{and} \quad Y \overset{\mathrm{d}}{\hookrightarrow} H \overset{\mathrm{d}}{\hookrightarrow} Y^*$$

(even though this assumption could be relaxed if we needed similar results as presented below in a more general setting). This implies

$$X + Y \hookrightarrow H \hookrightarrow (X + Y)^* \tag{4.9}$$

(the space H is identified with H^* using the canonical isomorphism). Before we continue, let us first review some properties of sums and intersections of normed vector spaces.

Let A, B be normed vector spaces. The intersection $A \cap B$ is defined simply as the set intersection of A and B, equipped with the norm

$$\|u\|_{A\cap B} = \|u\|_A + \|u\|_B.$$

Furthermore the space $A + B$ is defined as

$$A + B = \{a + b;\ a \in A,\ b \in B\}$$

with the norm

$$\|u\|_{A+B} = \inf\{\|a\|_A + \|b\|_B;\ a \in A,\ b \in B,\ a+b=u\}.$$

We recall that the norm in the continuous dual space A^* is given by

$$\|u\|_{A^*} = \sup\{|\langle u, v\rangle|;\ v \in A,\ \|v\|_A \le 1\}.$$

Then, one has

$$(A \cap B)^* = A^* + B^* \tag{4.10}$$

and

$$(A + B)^* = A^* \cap B^*, \tag{4.11}$$

where the equalities must be understood as isomorphisms. Since we actually do not need (4.11), let us verify (4.10) only. If $u \in (A \cap B)^*$, then we can extend this functional to $u_E \in A^*$ (using the Hahn-Banach theorem) and thus, we can identify u_E with an element of $A^* + B^*$ (adding the zero element of B^*). Moreover, the extension u_E is chosen as to satisfy $\|u_E\|_{A^*} = \|u\|_{(A\cap B)^*}$. Then, we can estimate

$$\|u_E\|_{A^*+B^*} \le \|u_E\|_{A^*} + \|0\|_{B^*} = \|u\|_{(A\cap B)^*}$$

and the embedding

$$(A \cap B)^* \hookrightarrow A^* + B^*$$

follows. Vice versa, let $u \in A^* + B^*$, and choose $a \in A^*$ and $b \in B^*$ such that $a + b = u$. We denote the restrictions of the functionals a and b to $A \cap B$ by a_R and b_R, respectively. It is obvious that $a_R, b_R \in (A \cap B)^*$, and thus $u_R = a_R + b_R$ satisfies $u_R \in (A \cap B)^*$ as well. Moreover, note that

$$\langle u_R, v\rangle = \langle a_R, v\rangle + \langle b_R, v\rangle = \langle a, v\rangle + \langle b, v\rangle = \langle u, v\rangle \quad \text{for all } v \in A \cap B.$$

Hence, the definition of u_R is independent of the choice of a, b and we may write

$$\begin{aligned}
\|u_R\|_{(A\cap B)^*} &= \sup\{|\langle u_R, v\rangle|;\ v \in A\cap B,\ \|v\|_{A\cap B} \le 1\} \\
&\le \sup\{|\langle a, v\rangle|;\ v \in A\cap B,\ \|v\|_{A\cap B} \le 1\} \\
&\qquad + \sup\{|\langle b, v\rangle|;\ v \in A\cap B,\ \|v\|_{A\cap B} \le 1\} \\
&\le \sup\{|\langle a, v\rangle|;\ v \in A,\ \|v\|_A \le 1\} + \sup\{|\langle b, v\rangle|;\ v \in B,\ \|v\|_B \le 1\} \\
&= \|a\|_{A^*} + \|b\|_{B^*}.
\end{aligned}$$

Since $a \in A^*$ and $b \in B^*$ are arbitrary functionals satisfying $a + b = u$, we get

$$\|u_R\|_{(A\cap B)^*} \le \|u\|_{A^*+B^*}$$

and the embedding

$$A^* + B^* \hookrightarrow (A \cap B)^*$$

follows. Thus, identity (4.10) is verified.

Note that (4.9) implies

$$X + Y \hookrightarrow (X + Y)^* \hookrightarrow (X \cap Y)^*.$$

Thus, denoting $P = \max\{p, q\}$, it is to see that

$$\begin{aligned} L^p(I;X) \cap L^q(I;Y) &\hookrightarrow L^p(I;X+Y) \cap L^q(I;X+Y) \\ &= L^P(I;X+Y) \hookrightarrow L^P(I;(X\cap Y)^*) \end{aligned} \tag{4.12}$$

and also that

$$\begin{aligned} (L^p(I;X) \cap L^q(I;Y))^* &\hookrightarrow (L^p(I;X\cap Y) \cap L^q(I;X\cap Y))^* \\ &= L^P(I;X\cap Y)^* = L^{P'}(I;(X\cap Y)^*). \end{aligned}$$

Hence, the time derivative appearing in the definition of $\mathcal{W}$ is correctly defined by

$$\int_I \partial_t u\varphi = -\int_I u\partial_t\varphi \quad \text{for all } \varphi \in \mathcal{D}(I),$$

which is an equality in the space $(X \cap Y)^*$ (cf. (3.1)). Furthermore, from (4.10) and (3.3), we see that

$$(L^p(I;X) \cap L^q(I;Y))^* = L^{p'}(I;X^*) + L^{q'}(I;Y^*) \tag{4.13}$$

(up to an isomorphism) and therefore, the norm $\|\cdot\|_{\mathcal{W}}$ is equivalent with the norm

$$\|u\|'_{\mathcal{W}} = \|u\|_{L^pX\cap L^qY} + \|\partial_t u\|_{L^{p'}X^*+L^{q'}Y^*}.$$

Theorem 4.2. *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*

$$X \overset{\mathrm{d}}{\hookrightarrow} H \overset{\mathrm{d}}{\hookrightarrow} X^* \quad \textit{and} \quad Y \overset{\mathrm{d}}{\hookrightarrow} H \overset{\mathrm{d}}{\hookrightarrow} Y^*. \tag{4.14}$$

Then, we have the embeddings

$$\mathcal{C}^1(I;X\cap Y) \overset{\mathrm{d}}{\hookrightarrow} \mathcal{W} \hookrightarrow \mathcal{C}(I;H). \tag{4.15}$$

Moreover, the integration by parts formula

$$(u(t_2), v(t_2))_H - (u(t_1), v(t_1))_H = \int_{t_1}^{t_2} \langle \partial_t u, v\rangle + \int_{t_1}^{t_2} \langle \partial_t v, u\rangle \tag{4.16}$$

holds for any $u, v \in \mathcal{W}$ and any $t_1, t_2 \in I$.

Proof. Let us prove the first embedding in (4.15). If $u \in \mathcal{C}^1(I;X\cap Y)$, then

$$\begin{aligned} u \in \mathcal{C}^1(I;X\cap Y) &= \mathcal{C}^1(I;X\cap Y) \cap \mathcal{C}^1(I;X\cap Y) \\ &\hookrightarrow \mathcal{C}^1(I;X) \cap \mathcal{C}^1(I;Y) \hookrightarrow L^p(I;X) \cap L^q(I;Y) \end{aligned}$$

and, using (3.3), (4.12), we also get

$$\partial_t u \in \mathcal{C}(I;X\cap Y) \hookrightarrow L^{P'}(I;X\cap Y) = (L^P(I;X\cap Y)^*)^* \hookrightarrow (L^p(I;X)\cap L^q(I;Y))^*.$$

Hence, we obtain $\mathcal{C}^1(I;X\cap Y) \hookrightarrow \mathcal{W}$. To show that this embedding is dense, we take $u \in \mathcal{W}$ and extend outside I. Let $a < b \in \mathbb{R}$ be the left and right endpoints of I, respectively. We define the extension of u on $(2a-b, 2b-a)$ by

$$u(t) = \begin{cases} u(2a-t), & t \in (2a-b, a); \\ u(2b-t), & t \in (b, 2b-a). \end{cases}$$

Then, for $\varepsilon \in (0, \frac{b-a}{2})$, let $\varrho_\varepsilon \in \mathcal{C}^1(\mathbb{R})$ be a non-negative function supported in $(-\varepsilon, \varepsilon)$ satisfying

$$\varrho_\varepsilon(t) \le \frac{1}{\varepsilon} \quad \text{for all} \quad t \in \mathbb{R} \quad \text{and} \quad \int_{\mathbb{R}} \varrho_\varepsilon = 1.$$

Next, we define u_ε by a convolution as

$$u_\varepsilon(t) = u * \varrho_\varepsilon(t) = \int_{\mathbb{R}} u(s)\varrho_\varepsilon(t-s)\,\mathrm{d}s, \quad t \in I.$$

Note that $\varrho_\varepsilon(t-s)$ is non-zero only if $s \in (t-\varepsilon, t+\varepsilon) \subset (a - \frac{b-a}{2}, b + \frac{b-a}{2})$, which is an interval, where u has been defined. Since $\varrho_\varepsilon \in \mathcal{C}^1(\mathbb{R})$ and

$$\partial_t u_\varepsilon(t) = \int_{\mathbb{R}} u(s)\varrho_\varepsilon'(t-s)\,\mathrm{d}s,$$

it is evident that $u_\varepsilon \in \mathcal{C}^1(X \cap Y)$. Then, using the properties of ϱ_ε, we can write

$$\begin{aligned}
\int_I \|u_\varepsilon(t) - u(t)\|_X^p \,\mathrm{d}t &= \int_I \left\| \int_{\mathbb{R}} (u(s) - u(t))\varrho_\varepsilon(t-s)\,\mathrm{d}s \right\|_X^p \mathrm{d}t \\
&= \int_I \left\| \int_{\mathbb{R}} (u(t+s) - u(t))\varrho_\varepsilon(s)\,\mathrm{d}s \right\|_X^p \mathrm{d}t \\
&\le \int_I \left(\int_{-\varepsilon}^{\varepsilon} \|u(t+s) - u(t)\|_X \varrho_\varepsilon(s)\,\mathrm{d}s \right)^p \mathrm{d}t \\
&\le \int_I \left(\int_{-\varepsilon}^{\varepsilon} \varrho_\varepsilon(s)^{p'}\,\mathrm{d}s \right)^{p-1} \int_{-\varepsilon}^{\varepsilon} \|u(t+s) - u(t)\|_X^p \,\mathrm{d}s\,\mathrm{d}t \\
&\le \frac{2^p}{2\varepsilon} \int_{-\varepsilon}^{\varepsilon} \int_I \|u(t+s) - u(t)\|_X^p \,\mathrm{d}t\,\mathrm{d}s \\
&\le 2^p \sup_{|s|<\varepsilon} \int_I \|u(t+s) - u(t)\|_X^p,
\end{aligned}$$

where the right hand side tends to zero as $\varepsilon \to 0_+$ by the classical property of the translation operator in Bochner spaces (see [35, Ch. IV, Lemma 1.5.]). Analogously, we obtain that

$$\int_I \|u_\varepsilon - u\|_Y^q \to 0,$$

and thus

$$\|u_\varepsilon - u\|_{L^p X \cap L^q Y} \to 0$$

as $\varepsilon \to 0_+$. Next, note that $\partial_t u$ exists also in the interval $(a - \frac{b-a}{2}, b + \frac{b-a}{2})$ (thanks to the even extension of u outside I) and satisfies $\partial_t u \in L^{p'}(I; X^*) + L^{q'}(I; Y^*)$ (cf. (4.13)). Therefore, there exist $w_1 \in L^{p'}(I; X^*)$ and $w_2 \in L^{q'}(I; Y^*)$ with $w_1 + w_2 = \partial_t u$. Then, using the same estimate as above, we can show that

$$\int_I \|w_1 * \varrho_\varepsilon - w_1\|_{X^*}^{p'} + \int_I \|w_2 * \varrho_\varepsilon - w_2\|_{Y^*}^{q'} \to 0$$

as $\varepsilon \to 0_+$. This implies

$$\begin{aligned}
\|\partial_t u_\varepsilon - \partial_t u\|_{L^{p'} X^* + L^{q'} Y^*} &= \|(w_1 * \varrho_\varepsilon - w_1) + (w_2 * \varrho_\varepsilon - w_2)\|_{L^{p'} X^* + L^{q'} Y^*} \\
&\le \|w_1 * \varrho_\varepsilon - w_1\|_{L^{p'} X^*} + \|w_2 * \varrho_\varepsilon - w_2\|_{L^{q'} Y^*} \to 0
\end{aligned}$$

as $\varepsilon \to 0_+$, which finishes the proof of $\mathcal{C}^1(I; X \cap Y) \overset{\mathrm{d}}{\hookrightarrow} \mathcal{W}$.

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 (4.14), the duality in (4.16) 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 (4.16) is obvious in that case. Now we proceed as in [66, Lemma 7.3.]. Note that, for any $a, b \geq 0$, the inequality

$$x - y \leq \sqrt{|x^2 - y^2|} \tag{4.17}$$

holds. Indeed, it is obvious if $x \leq y$ and in the other case $x > y$, we can rewrite it as

$$(x - y)^2 \leq (x - y)(x + y),$$

which is again evident. Let $t \in I$ and in (4.16) we take $v = u \in \mathcal{C}^1(I; X \cap Y)$, $t_2 = t$ and $t_1 \in I$ such that

$$\|u(t_1)\|_H = \frac{1}{|I|} \int_I \|u\|_H \tag{4.18}$$

(using the mean value theorem). This way, we obtain

$$\|u(t)\|_H^2 - \|u(t_1)\|_H^2 = 2 \int_0^t \langle \partial_t u, u \rangle$$

and by taking $\sqrt{|\cdot|}$ of both sides and using (4.17), we get

$$\|u(t)\|_H - \|u(t_1)\|_H \leq \left(2 \left| \int_0^t \langle \partial_t u, u \rangle \right| \right)^{\frac{1}{2}} \leq \sqrt{2} \left(\int_I |\langle \partial_t u, u \rangle| \right)^{\frac{1}{2}}.$$

If we rearrange this, use (4.18), estimate the duality pairing and use Young's inequality, we obtain

$$\begin{aligned} \|u(t)\|_H &\leq \frac{1}{T} \int_I \|u\|_H + \sqrt{2} \|\partial_t u\|_{L^{p'} X^* + L^{q'} Y^*}^{\frac{1}{2}} \|u\|_{L^p X \cap L^q Y}^{\frac{1}{2}} \\ &\leq C(\|u\|_{L^1 H} + \|u\|_{\mathcal{W}}). \end{aligned} \tag{4.19}$$

Moreover, by (4.14), we have

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

and thus (4.19) yields

$$\|u\|_{\mathcal{C}H} \leq C \|u\|_{\mathcal{W}}. \tag{4.20}$$

Since $\mathcal{C}^1(I; X \cap Y)$ is dense in $\mathcal{W}$, the estimate (4.20) and identity (4.16) remain valid for all $u \in \mathcal{W}$. Moreover, if $u \in \mathcal{W}$, then we can take $v = u$ and $t_2 \to t_1$ in (4.16) to deduce that $u \in \mathcal{C}(I; H)$. Thus, the embedding $\mathcal{W} \hookrightarrow \mathcal{C}(I; H)$ holds and the proof is finished. $\square$

By taking $X = Y$ and $p = q$ in the previous theorem, we obtain the classical result as an obvious corollary.

Corollary 4.3. *Let $1 < p < \infty$. Suppose that X is a separable reflexive Banach space and that H is a separable Hilbert space such that $X \overset{\mathrm{d}}{\hookrightarrow} H \overset{\mathrm{d}}{\hookrightarrow} X^*$ is a Gelfand triple.*

Then, we have

$$\mathcal{C}^1(I; X) \overset{\mathrm{d}}{\hookrightarrow} W^{1,p}(I; X) \hookrightarrow \mathcal{C}(I; H).$$

Moreover, the integration by parts formula

$$(u(t_2), v(t_2))_H - (u(t_1), v(t_1))_H = \int_{t_1}^{t_2} \langle \partial_t u, v \rangle + \int_{t_1}^{t_2} \langle \partial_t v, u \rangle \tag{4.21}$$

holds for any $u, v \in W^{1,p}(I; X)$ and any $t_1, t_2 \in I$.

Let $H = L^2(\Omega)$. The formula (4.21) can be used to identify that

$$\langle \partial_t u, u \rangle = \frac{1}{2} \frac{\mathrm{d}}{\mathrm{d}t} \int_\Omega u^2 \tag{4.22}$$

a.e. in I. This is an useful identity when deriving an a priori estimate by multiplying a parabolic equation including the term $\partial_t u$ by its solution. However, in certain situations we need to consider also test functions of the form $\psi(u)$ and then, we would like to generalize (4.22) 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 situation where such an identification is possible.

Lemma 4.4. *Let $1 < p, q < \infty$. Suppose that $\psi : \mathbb{R} \to \mathbb{R}$ is a Lipschitz function.*[2] *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 $u \in W^{1,p}(I; W^{1,q}(\Omega))$, there holds

$$\Psi(u) \in \mathcal{C}(I; L^1(\Omega)) \tag{4.23}$$

and

$$\int_{t_1}^{t_2} \langle \partial_t u, \psi(u) \rangle = \int_\Omega \Psi(u(t_2)) - \int_\Omega \Psi(u(t_1)) \quad \textit{for all } t_1, t_2 \in I. \tag{4.24}$$

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, e.g. by [75, Theorem 2.1.11.], and thus the duality in (4.24) is well defined. Next, we apply Theorem 4.2 to find $u_\varepsilon \in \mathcal{C}^1(I; W^{1,q}(\Omega))$ satisfying

$$\|u_\varepsilon - u\|_{L^p W^{1,q}} + \|\partial_t u_\varepsilon - \partial_t u\|_{L^{p'} W^{-1,q'}} \to 0 \quad \text{as } \varepsilon \to 0_+. \tag{4.25}$$

[2]This means that there exists $L \geq 0$ such that $|\psi(x) - \psi(y)| \leq L|x - y|$ for all $x, y \in \mathbb{R}$.

Then, using the classical calculus (chain rule and changing the order of integration and differentiation), it is easy to see that the identity

$$\begin{aligned}\int_{t_1}^{t_2}\langle\partial_t u_\varepsilon,\psi(u_\varepsilon)\rangle &= \int_{t_1}^{t_2}\int_\Omega \psi(u_\varepsilon)\partial_t u_\varepsilon \\ &= \int_{t_1}^{t_2}\int_\Omega \partial_t\Psi(u_\varepsilon) = \int_\Omega \Psi(u_\varepsilon(t_2)) - \int_\Omega \Psi(u_\varepsilon(t_1))\end{aligned} \tag{4.26}$$

holds true for any $t_1, t_2 \in I$. Since ψ is Lipschitz (with some Lipschitz constant $L \geq 0$), we can 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 reflexive space, and thus, there exist a subsequence and its limit $\overline{\psi(u)} \in L^p(I; W^{1,q}(\Omega))$ such that

$$\psi(u_\varepsilon) \rightharpoonup \overline{\psi(u)} \quad \text{weakly in } L^p(I; W^{1,q}(\Omega)). \tag{4.27}$$

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 (4.25) and (4.27), we obtain

$$\begin{aligned}\int_{t_1}^{t_2}\langle\partial_t u_\varepsilon,\psi(u_\varepsilon)\rangle &= \int_{t_1}^{t_2}\langle\partial_t u_\varepsilon - \partial_t u,\psi(u_\varepsilon)\rangle + \int_{t_1}^{t_2}\langle\partial_t u,\psi(u_\varepsilon)\rangle \\ &\to \int_{t_1}^{t_2}\langle\partial_t u,\psi(u)\rangle\end{aligned} \tag{4.28}$$

as $\varepsilon \to 0_+$. Next, using the embedding

$$W^{1,p}(I; W^{1,q}(\Omega)) \hookrightarrow \mathcal{C}(I; L^2(\Omega))$$

from Corollary 4.3 and (4.25), we get, for any $t_0 \in I$, that

$$\|u(t) - u(t_0)\|_2 \to 0 \quad \text{as} \quad t \to t_0 \tag{4.29}$$

and

$$\|u_\varepsilon(t_0) - u(t_0)\|_2 \to 0 \quad \text{as} \quad \varepsilon \to 0_+. \tag{4.30}$$

Then, the Lipschitz continuity of ψ, Hölder's inequality and (4.29) yield

$$\begin{aligned}\int_\Omega |\Psi(u(t)) - \Psi(u(t_0))| &= \int_\Omega \left|\int_{u(t_0)}^{u(t)}\psi(s)\,\mathrm{d}s\right| \leq \int_\Omega\int_{u(t_0)}^{u(t)}(|\psi(0)| + L|s|) \\ &\leq \int_\Omega\int_{u(t_0)}^{u(t)} C(1 + |u(t_0)| + |u(t)|) \leq C\int_\Omega (1 + |u(t_0)| + |u(t)|)|u(t) - u(t_0)| \\ &\leq C\|1 + |u(t_0)| + |u(t)|\|_2\|u(t) - u(t_0)\|_2 \leq C\|u(t) - u(t_0)\|_2 \to 0\end{aligned} \tag{4.31}$$

as $\varepsilon \to 0_+$, which proves (4.23) (and thus, the values $\Phi(u(t))$, $t \in I$, are well defined). By an analogous estimate, using (4.30) instead of (4.29), we can prove that

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

for any $t \in I$. This and (4.28) used in (4.26) to take the limit $\varepsilon \to 0_+$ proves (4.24).

If ψ is bounded, we replace (4.31) 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. $\square$

We also consider a small modification of the above lemma. In (4.24), we can replace $\psi(\cdot)$ by $\psi(\cdot)\phi$, where $\phi \in W^{1,\infty}(\Omega)$, leading to the identity

$$\int_0^t \langle \partial_t u, \psi(u)\phi \rangle = \int_\Omega \int_w^{u(t)} \psi(s) \, \mathrm{d}s \, \phi - \int_\Omega \int_w^{u(0)} \psi(s) \, \mathrm{d}s \, \phi \quad \text{for all } t \in I. \tag{4.32}$$

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

4.3 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 $\mathbb{A}$. Thus, we also have that

$$|\mathbb{A}| = \sqrt{\operatorname{tr}(\mathbb{A}^T \mathbb{A})} = \sqrt{\operatorname{tr}(\mathbb{A}\mathbb{A}^T)}, \quad \text{where } \operatorname{tr} \mathbb{A} = \sum_{i=1}^d (\mathbb{A})_{ii}$$

and

$$|\mathbb{A}^T \mathbb{A}| \leq |\mathbb{A}|^2. \tag{4.33}$$

As a consequence of the Schur decomposition, every symmetric (and thus normal) matrix admits a spectral decomposition of the form

$$\mathbb{A} = QDQ^T,$$

where D is a diagonal matrix containing the eigenvalues of $\mathbb{A}$ and Q is a unitary matrix satisfying $Q^{-1} = Q^T$, $|Q| = 1$, see [39, Theorem 2.5.4] or [51, Theorem 1.13]. Moreover, since $\mathbb{A}$ is positive definite, the eigenvalues of $\mathbb{A}$ are strictly positive. Thus, we may consistently define any real power α of $\mathbb{A}$ by

$$\mathbb{A}^\alpha = QD^\alpha Q^T.$$

Then, the matrix $\mathbb{A}^\alpha$ is again symmetric and positive definite.

In the next lemma, we collect some basic algebraic facts.

Lemma 4.5. *Let $\mathbb{A}_i \in \mathbb{R}^{d\times d}$, $i = 1,2,3,4$, $d \in \mathbb{N}$ be some matrices and let $\mathbb{A} = \mathbb{A}^T \in \mathbb{R}^{d\times d}_{\mathrm{sym}}$ be a symmetric matrix. Then, the identities*

$$\text{(i)} \qquad \mathbb{A}_1\mathbb{A}_2\mathbb{A}_3 \cdot \mathbb{A}_4 = \mathbb{A}_2 \cdot \mathbb{A}_1^T\mathbb{A}_4\mathbb{A}_3^T; \tag{4.34}$$

$$\text{(ii)} \qquad |\mathbb{A}|^2 = \operatorname{tr}\mathbb{A}^2 \tag{4.35}$$

hold. Moreover, if $\mathbb{A} \in \mathbb{R}^{d\times d}_{>0}$ is a positive definite matrix, then the estimates

$$\text{(iii)} \qquad 0 \le \operatorname{tr}\mathbb{A} - d - \ln\det\mathbb{A}; \tag{4.36}$$

$$\text{(iv)} \qquad |\mathbb{A}| \le \operatorname{tr}\mathbb{A} \le \sqrt{d}|\mathbb{A}|; \tag{4.37}$$

$$\text{(v)} \qquad \min\Big\{1, d^{\frac{1-\alpha}{2}}\Big\}|\mathbb{A}|^\alpha \le |\mathbb{A}^\alpha| \le \max\Big\{1, d^{\frac{1-\alpha}{2}}\Big\}|\mathbb{A}|^\alpha \quad \textit{for any } \alpha \in [0,\infty).$$

hold.

Proof. The identity (i) is a consequence of the well known identities for the trace operator. Indeed, we can write

$$\begin{aligned}\mathbb{A}_1\mathbb{A}_2\mathbb{A}_3 \cdot \mathbb{A}_4 &= \operatorname{tr}(\mathbb{A}_1\mathbb{A}_2\mathbb{A}_3\mathbb{A}_4^T) = \operatorname{tr}(\mathbb{A}_2\mathbb{A}_3\mathbb{A}_4^T\mathbb{A}_1)\\ &= \mathbb{A}_2 \cdot (\mathbb{A}_3\mathbb{A}_4^T\mathbb{A}_1)^T = \mathbb{A}_2 \cdot \mathbb{A}_1^T\mathbb{A}_4\mathbb{A}_3^T.\end{aligned}$$

The property (ii) then follows from (i) and symmetry of $\mathbb{A}$ since

$$|\mathbb{A}|^2 = \mathbb{A}\cdot\mathbb{A} = \mathbb{A}^2 \cdot \mathbb{I} = \operatorname{tr}\mathbb{A}^2.$$

Let us denote the eigenvalues of $\mathbb{A}$ by $\{\lambda_i\}_{i=1}^d$. Then, it is well known and easy to see using the spectral decomposition that invariants of $\mathbb{A}$ satisfy

$$\operatorname{tr}\mathbb{A} = \sum_{i=1}^d \lambda_i, \qquad |\mathbb{A}| = \sqrt{\sum_{i=1}^d \lambda_i^2}, \qquad \det\mathbb{A} = \prod_{i=1}^d \lambda_i.$$

Thus, we get (iii) from the fact that $x \mapsto x - 1 - \ln x \ge 0$ for all $x > 0$ (this function attains its minimum at $x = 1$). Furthermore, by (4.33), (4.35) and Young's inequality, we get

$$\begin{aligned}|\mathbb{A}| &= |(\mathbb{A}^{\frac12})^T\mathbb{A}^{\frac12}| \le |\mathbb{A}^{\frac12}|^2 = \operatorname{tr}\mathbb{A}\\ &= \sum_{i=1}^d \lambda_i \le \sqrt{d}\sqrt{\sum_{i=1}^d \lambda_i^2} = \sqrt{d}|D| = \sqrt{d}|QDQ^T| = \sqrt{d}|\mathbb{A}|,\end{aligned}$$

which proves (iv). Next, for $\alpha \in [0,\infty)$, we denote $\sigma(\alpha) = \sum_{i=1}^d \lambda_i^{2\alpha}$. If we use concavity of the power function $x \mapsto x^\alpha$ for $\alpha \in [0,1]$ twice (first time in the form $\varepsilon x^\alpha \le (\varepsilon x)^\alpha$, $\varepsilon \in (0,1)$), we get the inequality

$$\sigma(1)^\alpha = \sum_{i=1}^d \frac{\lambda_i^2}{\sigma(1)}\sigma(1)^\alpha \le \sum_{i=1}^d \lambda_i^{2\alpha} = \sigma(\alpha) = d\sum_{i=1}^d \frac{\lambda_i^{2\alpha}}{d} \le d\left(\sum_{i=1}^d \frac{\lambda_i^2}{d}\right)^\alpha = d^{1-\alpha}\sigma(1)^\alpha.$$

Thus, since

$$\sigma(\alpha)^{\frac12} = |D^\alpha| = \sqrt{QD^\alpha Q^T \cdot QD^\alpha Q^T} = \sqrt{\mathbb{A}^\alpha \cdot \mathbb{A}^\alpha} = |\mathbb{A}^\alpha|,$$

we obtain

$$|\mathbb{A}|^\alpha = \sigma(1)^{\frac{\alpha}{2}} \le \sigma(\alpha)^{\frac{1}{2}} = |\mathbb{A}^\alpha| = \sigma(\alpha)^{\frac{1}{2}} \le d^{\frac{1-\alpha}{2}}(\sigma(1))^{\frac{\alpha}{2}} = d^{\frac{1-\alpha}{2}}|\mathbb{A}|^\alpha.$$

Analogously, for $\alpha \in [1,\infty)$, using the convexity of $x \mapsto x^\alpha$, we obtain

$$d^{\frac{1-\alpha}{2}}|\mathbb{A}|^\alpha \le |\mathbb{A}^\alpha| \le |\mathbb{A}|^\alpha$$

and the proof of (v) is finished. □

Next, we assume that

$$\mathbb{A} : \Omega \to \mathbb{R}^{d\times d}_{>0}, \quad d \in \mathbb{N},$$

is a Sobolev mapping defined in a domain $\Omega \subset \mathbb{R}^d$, $d \in \mathbb{N}$, with a Lipschitz boundary. Thus, the derivatives of $\mathbb{A}$ exist in the distributional sense and they are integrable in Ω. The result below is stated only for the domain Ω just for simplicity. However, it is easy to see that an analogous assertion will hold also in the case where the domain Ω and the symbol $\nabla\mathbb{A}$ are replaced with $(0,T)$ (i.e., only an one-dimensional domain) and $\partial_t\mathbb{A}$, respectively. This fact is occasionally used in Chapter 3.

Lemma 4.6. *Let $\mathbb{A}, \mathbb{A}^{-1} \in W^{1,2}(\Omega)$ be positive definite a.e. in Ω. Then the following identities hold:*

$$\text{(i)} \qquad \nabla\mathbb{A}^{-1} = -\mathbb{A}^{-1}\nabla\mathbb{A}\mathbb{A}^{-1}; \tag{4.38}$$

$$\text{(ii)} \qquad -\nabla\mathbb{A}\cdot\nabla\mathbb{A}^{-1} = |\mathbb{A}^{-\frac{1}{2}}\nabla\mathbb{A}\mathbb{A}^{-\frac{1}{2}}|^2; \tag{4.39}$$

$$\text{(iii)} \qquad \nabla \ln\det\mathbb{A} = \mathbb{A}^{-1}\cdot\nabla\mathbb{A} = \operatorname{tr}(\mathbb{A}^{-\frac{1}{2}}\nabla\mathbb{A}\mathbb{A}^{-\frac{1}{2}}) \tag{4.40}$$

a.e. in Ω.

Proof. Let $\varphi \in \mathcal{D}(\Omega)$. Using integration by parts twice, we obtain

$$\begin{aligned}\int_\Omega \mathbb{A}\nabla\mathbb{A}^{-1}\varphi + \int_\Omega \nabla\mathbb{A}\mathbb{A}^{-1}\varphi &= -\int_\Omega \nabla(\mathbb{A}\varphi)\mathbb{A}^{-1} + \int_\Omega \nabla\mathbb{A}\mathbb{A}^{-1}\varphi \\ &= -\int_\Omega \nabla\varphi = -\int_{\partial\Omega} \varphi\boldsymbol{n} = 0,\end{aligned}$$

hence, we find

$$\mathbb{A}\nabla\mathbb{A}^{-1} = -\nabla\mathbb{A}\mathbb{A}^{-1} \quad \text{a.e. in } \Omega \tag{4.41}$$

by the fundamental theorem of variational calculus. The identity (i) then follows upon multiplying (4.41) by $\mathbb{A}^{-1}$ from the left.

The identity (ii) is an easy consequence of (i) and (4.34) since

$$\begin{aligned}-\nabla\mathbb{A}\cdot\nabla\mathbb{A}^{-1} &= \nabla\mathbb{A}\cdot\mathbb{A}^{-1}\nabla\mathbb{A}\mathbb{A}^{-1} = \nabla\mathbb{A}\cdot\mathbb{A}^{-\frac{1}{2}}\mathbb{A}^{-\frac{1}{2}}\nabla\mathbb{A}\mathbb{A}^{-\frac{1}{2}}\mathbb{A}^{-\frac{1}{2}} \\ &= \mathbb{A}^{-\frac{1}{2}}\nabla\mathbb{A}\mathbb{A}^{-\frac{1}{2}}\cdot\mathbb{A}^{-\frac{1}{2}}\nabla\mathbb{A}\mathbb{A}^{-\frac{1}{2}} = |\mathbb{A}^{-\frac{1}{2}}\nabla\mathbb{A}\mathbb{A}^{-\frac{1}{2}}|^2.\end{aligned}$$

To prove (iii) (which is a version of well known Jacobi's formula) we consider a suitable approximation of $\mathbb{A}$. For $\lambda, \varepsilon > 0$, we define

$$\mathbb{A}_{\lambda,\varepsilon} = \lambda\mathbb{I} + \mathbb{A}_\varepsilon,$$

where $\mathbb{A}_\varepsilon \in \mathcal{C}^1(\Omega)$ satisfies

$$\mathbb{A}_\varepsilon \to \mathbb{A} \quad \text{a.e. in } \Omega \quad \text{and} \quad \|\mathbb{A} - \mathbb{A}_\varepsilon\|_{1,2} \to 0 \quad \text{as } \varepsilon \to 0_+. \tag{4.42}$$

Since $\mathbb{A}_\varepsilon$ can be constructed using a convolution and $\mathbb{A}\boldsymbol{x} \cdot \boldsymbol{x} > 0$ for all $\boldsymbol{x} \in \mathbb{R}^d$, we may suppose that

$$\mathbb{A}_\varepsilon \boldsymbol{x} \cdot \boldsymbol{x} > 0 \quad \text{for all } \boldsymbol{x} \in \mathbb{R}^d. \tag{4.43}$$

Consequently, we also get

$$\mathbb{A}_{\lambda,\varepsilon} \boldsymbol{x} \cdot \boldsymbol{x} = \mathbb{A}_\varepsilon \boldsymbol{x} \cdot \boldsymbol{x} + \lambda|\boldsymbol{x}|^2 > \lambda|\boldsymbol{x}|^2 \quad \text{for all } \boldsymbol{x} \in \mathbb{R}^d$$

and $\mathbb{A}_{\lambda,\varepsilon}, \mathbb{A}_{\lambda,\varepsilon}^{-1} \in \mathcal{C}^1(\Omega)$. Then, we recall the Jacobi formula

$$\nabla \det \mathbb{A}_{\lambda,\varepsilon} = \det \mathbb{A}_{\lambda,\varepsilon} \, \mathbb{A}_{\lambda,\varepsilon}^{-T} \cdot \nabla \mathbb{A}_\varepsilon$$

which we divide by $\det \mathbb{A}_{\lambda,\varepsilon}$ and rewrite as

$$\int_\Omega \nabla \ln \det \mathbb{A}_{\lambda,\varepsilon} \varphi = \int_\Omega \mathbb{A}_{\lambda,\varepsilon}^{-1} \cdot \nabla \mathbb{A}_{\lambda,\varepsilon} \varphi \quad \text{for all } \varphi \in \mathcal{D}(\Omega) \tag{4.44}$$

(see [51, Theorems 8.1, 8.2]). Next, denoting the eigenvalues of $\mathbb{A}_\varepsilon$ by $\{\lambda_i\}_{i=1}^d$ (which are positive due to (4.43)) and using (4.37), we observe that

$$|\mathbb{A}_{\lambda,\varepsilon}^{-1}| \le \operatorname{tr}\left((\lambda\mathbb{I} + \mathbb{A}_\varepsilon)^{-1}\right) = \sum_{i=1}^d \frac{1}{\lambda + \lambda_i} \le \frac{d}{\lambda}.$$

Therefore, we have

$$\left\|\mathbb{A}_{\lambda,\varepsilon}^{-1}\right\|_{L^\infty L^\infty} \le C(\lambda)$$

and moreover it follows from (4.44) and the properties of $\mathbb{A}_\varepsilon$ that

$$\|\nabla \ln \det \mathbb{A}_{\lambda,\varepsilon}\|_2 \le C(\lambda).$$

Thus, by selecting appropriate subsequences we get

$$\mathbb{A}_{\lambda,\varepsilon}^{-1} \rightharpoonup K_1 \quad \text{weakly in } L^{\infty)}(\Omega), \tag{4.45}$$

$$\nabla \ln \det \mathbb{A}_{\lambda,\varepsilon} \rightharpoonup K_2 \quad \text{weakly in } L^2(\Omega) \tag{4.46}$$

as $\varepsilon \to 0_+$. Due to (4.42), we have $K_1 = (\lambda\mathbb{I} + \mathbb{A})^{-1}$. Furthermore, note that

$$\begin{aligned} |\ln \det \mathbb{A}_{\lambda,\varepsilon}| = d|\ln(\det \mathbb{A}_{\lambda,\varepsilon})^{\frac{1}{d}}| &\le d|\det \mathbb{A}_{\lambda,\varepsilon}|^{\frac{1}{d}} + d|\det \mathbb{A}_{\lambda,\varepsilon}^{-1}|^{\frac{1}{d}} \\ &\le C(|\mathbb{A}_{\lambda,\varepsilon}| + |\mathbb{A}_{\lambda,\varepsilon}^{-1}|), \end{aligned}$$

which implies, using also (4.42), that

$$\ln \det \mathbb{A}_{\lambda,\varepsilon} \rightharpoonup \ln \det(\lambda\mathbb{I} + \mathbb{A}) \quad \text{weakly in } L^2(\Omega).$$

Hence, we identify that $K_2 = \nabla \ln \det(\lambda\mathbb{I} + \mathbb{A})$. If we use $\nabla \mathbb{A}_{\lambda,\varepsilon} = \nabla \mathbb{A}_\varepsilon$, (4.42), (4.45) and (4.46) to take the limit in (4.44), we obtain

$$\int_\Omega \nabla \ln \det(\lambda\mathbb{I} + \mathbb{A})\varphi = \int_\Omega (\lambda\mathbb{I} + \mathbb{A})^{-1} \cdot \nabla \mathbb{A} \, \varphi \quad \text{for all } \varphi \in \mathcal{D}(\Omega). \tag{4.47}$$

Denoting the eigenvalues of $\mathbb{A}$ by $\{\mu_i\}_{i=1}^d$ and using (4.37), we find that

$$|(\lambda\mathbb{I}+\mathbb{A})^{-1}| \le \operatorname{tr}(\lambda\mathbb{I}+\mathbb{A})^{-1} = \sum_{i=1}^{d} \frac{1}{\lambda+\mu_i} \le \operatorname{tr}\mathbb{A}^{-1} \le \sqrt{d}|\mathbb{A}^{-1}|.$$

Together with the Sobolev embedding $W^{1,2}(\Omega) \hookrightarrow L^{\frac{2d}{d-2}}(\Omega)$, this yields

$$\left\|(\lambda\mathbb{I}+\mathbb{A})^{-1}\right\|_{\frac{2d}{d-2}} \le C.$$

Thus, from Hölder's inequality and (4.47), we see that

$$\|\nabla \ln \det(\lambda\mathbb{I}+\mathbb{A})\|_{\frac{d}{d-1}} \le C.$$

Then, using analogous procedure as above, we eventually find that

$$\begin{aligned} (\lambda\mathbb{I}+\mathbb{A})^{-1} &\rightharpoonup \mathbb{A}^{-1} \quad \text{weakly in } L^{\frac{2d}{d-2}}(\Omega), \\ \nabla \ln\det(\lambda\mathbb{I}+\mathbb{A}) &\rightharpoonup \nabla \ln\det\mathbb{A} \quad \text{weakly in } L^{\frac{d}{d-1}}(\Omega), \end{aligned}$$

which is sufficient to take the limit $\lambda \to 0_+$ in (4.47) and obtain

$$\int_\Omega \nabla \ln\det\mathbb{A}\varphi = \int_\Omega \mathbb{A}^{-1}\cdot\nabla\mathbb{A}\,\varphi \quad \text{for all } \varphi \in \mathcal{D}(\Omega).$$

Consequently, the first part of (4.40) is proved. The second equality in (4.40) follows directly from (4.34). $\square$

5. Analysis of viscoelastic fluids in the isothermal setting

This stand-alone chapter contains the work [6] titled *Large data existence theory for three-dimensional unsteady flows of rate-type viscoelastic fluids with stress diffusion.*

5.1 Introduction

We aim to establish a global-in-time and large-data existence theory, within the context of weak solutions, to a class of homogeneous incompressible rate-type viscoelastic fluids flowing in a closed three-dimensional container. The studied class of models can be seen as the Navier-Stokes system coupled with a viscoelastic rate-type fluid model that shares the properties of both Oldroyd-B and Giesekus models and is completed with a diffusion term. Such models are frequently encountered in the theory of non-Newtonian fluid mechanics, see [30, 27] and further references cited in [27].

In order to precisely formulate the problems investigated in this chapter, we start by introducing the necessary notation. For a bounded domain $\Omega \subset \mathbb{R}^3$ with the Lipschitz boundary $\partial\Omega$ and a time interval of the length $T > 0$, we set $Q = (0,T) \times \Omega$ for a time-space cylinder and $\Sigma = (0,T) \times \partial\Omega$ for a part of its boundary. The symbol $\boldsymbol{n}$ denotes the outward unit normal vector on $\partial\Omega$ and, for any vector $\boldsymbol{z}$, the vector $\boldsymbol{z}_\tau$ denotes the projection of the vector to a tangent plane on $\partial\Omega$, i.e., $\boldsymbol{z}_\tau = \boldsymbol{z} - (\boldsymbol{z} \cdot \boldsymbol{n})\boldsymbol{n}$. Then, for a given density of the external body forces $\boldsymbol{f} : Q \to \mathbb{R}^3$, a given initial velocity $\boldsymbol{v}_0 : \Omega \to \mathbb{R}^3$ and a given initial extra stress tensor $\mathbb{B}_0 : \Omega \to \mathbb{R}^{d\times d}_{>0}$ (here $\mathbb{R}^{d\times d}_{>0}$ denotes the set of symmetric positive definite (3×3)-matrices), we look for a vector field $\boldsymbol{v} : Q \to \mathbb{R}^3$, a scalar field $p : Q \to \mathbb{R}$ and a positive definite matrix field $\mathbb{B} : Q \to \mathbb{R}^{d\times d}_{>0}$ solving the following system in Q:

$$\operatorname{div} \boldsymbol{v} = 0, \tag{5.1}$$

$$\partial_t \boldsymbol{v} + \boldsymbol{v} \cdot \nabla \boldsymbol{v} - \nu \Delta \boldsymbol{v} + \nabla p = 2\mu a \operatorname{div}((1-\gamma)(\mathbb{B} - \mathbb{I}) + \gamma(\mathbb{B}^2 - \mathbb{B})) + \boldsymbol{f}, \tag{5.2}$$

$$\begin{aligned}\partial_t \mathbb{B} + \boldsymbol{v} \cdot \nabla \mathbb{B} + \delta_1(\mathbb{B} - \mathbb{I}) + \delta_2(\mathbb{B}^2 - \mathbb{B}) - \lambda \Delta \mathbb{B} \\ = \tfrac{a+1}{2}(\nabla \boldsymbol{v} \mathbb{B} + (\nabla \boldsymbol{v} \mathbb{B})^T) + \tfrac{a-1}{2}(\mathbb{B} \nabla \boldsymbol{v} + (\mathbb{B} \nabla \boldsymbol{v})^T),\end{aligned} \tag{5.3}$$

and being completed by the following boundary conditions on Σ:

$$\begin{aligned} \boldsymbol{v} \cdot \boldsymbol{n} &= 0, \\ -\sigma \boldsymbol{v}_\tau &= \Big(\big(\nu \nabla \boldsymbol{v} + \nu(\nabla \boldsymbol{v})^T + 2\mu a(1-\gamma)(\mathbb{B} - \mathbb{I}) + 2\mu a\gamma(\mathbb{B}^2 - \mathbb{B})\big)\, \boldsymbol{n}\Big)_\tau, \\ \boldsymbol{n} \cdot \nabla \mathbb{B} &= \mathbb{O}, \qquad \text{(here } \mathbb{O} \text{ stands for zero } 3 \times 3\text{-matrix)} \end{aligned} \tag{5.4}$$

and by the initial conditions in Ω:

$$\boldsymbol{v}(0, \cdot) = \boldsymbol{v}_0, \tag{5.5}$$

$$\mathbb{B}(0, \cdot) = \mathbb{B}_0. \tag{5.6}$$

The parameters $\gamma \in (0,1)$, $\nu, \lambda, \sigma > 0$, $\delta_1, \delta_2 \geq 0$ and $a \in \mathbb{R}$ are given numbers.

The main result of this chapter can be stated as:

Let $\boldsymbol{v}_0$ and $\mathbb{B}_0$ be such that the initial total energy is bounded. Then, for sufficiently regular $\boldsymbol{f}$, there exists a global-in-time weak solution to (5.1)–(5.6).

Although the above result is stated vaguely, we would like to emphasize that we are going to establish the **long-time** existence of a weak solution for **large data** and for **three-dimensional** flows. A more precise and rigorous version of the above result including the correct function spaces and the properly defined weak formulation is stated in Theorem 5.2 below.

We complete the introductory part by providing the physical background relevant to the studied problem and by recalling earlier results relevant to the problem (5.1)–(5.6) analyzed here.

5.2 Mathematical and physical background

The system (5.1)–(5.4) can be rewritten into a more concise form once one recognizes some physical quantities. First of all, let

$$\mathbb{D}\boldsymbol{v} = \tfrac{1}{2}(\nabla\boldsymbol{v} + (\nabla\boldsymbol{v})^T) \quad \text{and} \quad \mathbb{W}\boldsymbol{v} = \tfrac{1}{2}(\nabla\boldsymbol{v} - (\nabla\boldsymbol{v})^T)$$

denote the symmetric and antisymmetric parts of the velocity gradient $\nabla\boldsymbol{v}$, respectively. Then, looking at the equation (5.2), we see that (5.2) is obtained from a general form of the balance of linear momentum, namely

$$\varrho\dot{\boldsymbol{v}} = \operatorname{div}\mathbb{T} + \varrho\boldsymbol{f}, \tag{5.7}$$

once we set the density $\varrho = 1$ and require that the Cauchy stress tensor $\mathbb{T}$ has the form

$$\mathbb{T} = -p\mathbb{I} + 2\nu\mathbb{D}\boldsymbol{v} + 2a\mu((1-\gamma)(\mathbb{B}-\mathbb{I}) + \gamma(\mathbb{B}^2 - \mathbb{B})). \tag{5.8}$$

In (5.7), $\dot{\boldsymbol{v}}$ stands for the material time derivative of $\boldsymbol{v}$, i.e., $\dot{\boldsymbol{v}} = \partial_t\boldsymbol{v} + \boldsymbol{v}\cdot\nabla\boldsymbol{v}$. Defining similarly the material time derivative of a tensor $\mathbb{B}$ as

$$\dot{\mathbb{B}} = \partial_t\mathbb{B} + \boldsymbol{v}\cdot\nabla\mathbb{B},$$

we can recognize the presence of a general objective derivative in (5.3). Namely, defining

$$\overset{\circ}{\mathbb{B}} = \dot{\mathbb{B}} - a(\mathbb{D}\boldsymbol{v}\mathbb{B} + \mathbb{B}\mathbb{D}\boldsymbol{v}) - (\mathbb{W}\boldsymbol{v}\mathbb{B} - \mathbb{B}\mathbb{W}\boldsymbol{v}),$$

we can rewrite the system (5.1)–(5.3) into a more familiar form as

$$\operatorname{div}\boldsymbol{v} = 0, \tag{5.9}$$

$$\dot{\boldsymbol{v}} = \operatorname{div}\mathbb{T} + \boldsymbol{f}, \tag{5.10}$$

$$\overset{\circ}{\mathbb{B}} + \delta_1(\mathbb{B}-\mathbb{I}) + \delta_2(\mathbb{B}^2 - \mathbb{B}) = \lambda\Delta\mathbb{B}, \tag{5.11}$$

which is supposed to hold true in Q and which is completed by the initial conditions (5.5), (5.6) fulfilled in Ω and by the boundary conditions (5.4) on Σ that take the form:

$$\boldsymbol{v}\cdot\boldsymbol{n} = 0, \tag{5.12}$$

$$(\mathbb{T}\boldsymbol{n})_\tau = -\sigma\boldsymbol{v}_\tau, \tag{5.13}$$

$$\boldsymbol{n}\cdot\nabla\mathbb{B} = \mathbb{O}. \tag{5.14}$$

We provide several comments regarding (5.8)–(5.11) as well as the boundary conditions (5.12)–(5.14). The Navier slip boundary condition (5.13) is considered here just for simplicity; note that for smooth domains, namely if $\Omega \in \mathcal{C}^{1,1}$, we can introduce the pressure p as an integrable function, e.g., by using an additional layer of approximation as in [12], see also [16, 15] or [8] which discuss the treatment of the pressure in evolutionary models subject to the Navier boundary condition. Nevertheless, since we always deal with formulation without the pressure (see Definition 5.1), we can also treat the Dirichlet boundary condition, as well as very general implicitly specified boundary conditions see, e.g., [58, 13, 14] or [8]. The Neumann boundary condition for $\mathbb{B}$ is considered here only for simplicity and without any specific physical meaning.

A further aspect, which makes the above system more complicated than the Navier-Stokes equation is the form of the Cauchy stress tensor $\mathbb{T}$ as in (5.8). The term $-p\mathbb{I} + 2\nu\mathbb{D}\boldsymbol{v}$ corresponds to the standard Newtonian fluid flow model with a constant kinematic viscosity ν. The next part of the Cauchy stress, which depends linearly on $\mathbb{B}$, appears in all the viscoelastic rate-type fluid models - see, e.g., [53, (7.20b), (8.20e)], [40, (6.43e)] or [27, (43a)]. On the other hand, the addition of the term $2a\mu\gamma(\mathbb{B}^2 - \mathbb{B})$ is, to our best knowledge, considered here for the first time. The fact that we require that γ is positive (and strictly less than 1) plays a key role in the analysis of the problem, as is shown below. Note that the linearization of $\mathbb{T}$ with respect to $\mathbb{B}$ when $\mathbb{B}$ is close to the identity $\mathbb{I}$ yields

$$\mathbb{T} = -p\mathbb{I} + 2\nu\mathbb{D}\boldsymbol{v} + 2a\mu(\mathbb{B} - \mathbb{I})$$

and we recover the standard form of $\mathbb{T}$ (after possible redefinition of the pressure).

The quantity $\mathbb{B}$ takes into account the elastic responses of the fluid and the equation (5.11) describes its evolution in the current configuration (Eulerian coordinates), just as the velocity $\boldsymbol{v}$. It is frequent to call the tensor $\mu(\mathbb{B} - \mathbb{I})$ the extra stress or conformation tensor and to denote it by $\boldsymbol{\tau}$. More importantly, since the material derivative of $\mathbb{B}$ is not objective, it must be "corrected" and this is the reason, why in (5.11) the derivative $\overset{\circ}{\mathbb{B}}$ appears. The parameter a in the definition of $\overset{\circ}{\mathbb{B}}$ determines the type of the objective derivative. The case $a = 1$ leads to the upper convected Oldroyd derivative, that has favourable physical properties and that leads to a clear interpretation of $\mathbb{B}$ within the thermodynamical framework developed in [63], see also [64, 55, 56, 57]. Next, the case $a = 0$ leads to the corrotational Jaumann-Zaremba derivative and this is the only case for which the analysis is much simpler than in other cases. Furthermore, if $a \in [-1, 1]$, one obtains the entire class of Gordon-Schowalter derivatives. However, it turns out that the physical properties of these derivatives are irrelevant for the analysis presented below (except the case $a = 0$), therefore we may take any $a \in \mathbb{R}$. For $a = 1$ and $\lambda = 0$ we distinguish two cases: if $\delta_1 > 0$ and $\delta_2 = 0$ we obtain the classical Oldroyd-B model while if $\delta_1 = 0$ and $\delta_2 > 0$ we get the Giesekus model. Next, by considering $a \in [-1, 1]$, we obtain the class of Johnson-Segalman models. If we further let $\lambda > 0$, we are introducing diffusive variants of the previous models. It has been observed that including the diffusion term in (5.11) is physically reasonable, see, e.g., [30] or [27] and references therein. However, up to now, it has been unknown what precise form should the diffusion term take and also whether it actually helps in the analysis of the model. Our main result provides a partial answer to this question, namely: for $\gamma \in (0, 1)$ and with the diffusion term being of the form $\Delta\mathbb{B}$ (or more generally, a linear second order operator), the global existence of a weak solution is available.

Compared to the equations describing flows of standard Oldroyd-B viscoleastic rate-type fluids, there are two deviations in the set of equations (5.9)–(5.11) studied hereafter. We provide a few comments on these differences.

The first deviation concerns the incorporation of the stress diffusion term, i.e. the term $-\Delta\mathbb{B}$, into the equations. Following the pioneering work of [30] it is clear that a quantity related to $|\nabla\mathbb{B}|^2$ has to be added into the list of underlying dissipation mechanisms. On the other hand, the precise form in which stress diffusion should appear depends on the choice of a thermodynamical approach and specific assumptions. In fact, using the thermodynamical concepts as in [53] or [27], one can derive models, where the stress diffusion term takes the form $-\mathbb{B}\Delta\mathbb{B} - \Delta\mathbb{B}\mathbb{B}$, $-\mathbb{B}^{\frac{1}{2}}\Delta\mathbb{B}\mathbb{B}^{\frac{1}{2}}$ etc., however, we would prefer $-\Delta\mathbb{B}$ simply because it coincides with the form proposed by [30], and, from the perspective of PDE analysis and numerical approximation, one prefers to deal with stress diffusion that leads to a linear operator.

The second deviation from usual viscoelastic models consists in the presence of the term $(\mathbb{B}^2 - \mathbb{B})$ in the Cauchy stress tensor, see (5.8). This term arises if we slightly modify energy storage mechanism and apply the thermodynamic approach as developed in [53]. In what follows, we give a clear interpretation and a thermodynamic derivation of our model.

5.3 Thermodynamical derivation of the model

Viscoelastic models with (nonlinear) stress diffusion, but without the term $\mathbb{B}^2$ in the stress tensor are derived, e.g., in [53] and [27] even in the temperature-dependent case. Here, we briefly explain the approach in a simplified isothermal setting (sufficient for the purpose of this study), referring to the cited works for the derivation in a complete thermal setting and for more details.

First, we postulate the constitutive equation for the Helmholtz free energy in the form

$$\psi(\mathbb{B}) = \mu((1-\gamma)(\operatorname{tr}\mathbb{B} - 3 - \ln\det\mathbb{B}) + \tfrac{1}{2}\gamma|\mathbb{B} - \mathbb{I}|^2), \tag{5.15}$$

where $\mu > 0$ and $\gamma \in [0,1]$ is a parameter interpolating between two forms of the energy. The choice $\gamma = 0$ would lead to a standard Oldroyd-B diffusive model. To our best knowledge, the case $\gamma > 0$ was not considered before in literature. The term $\frac{1}{2}\gamma|\mathbb{B} - \mathbb{I}|^2$, which is newly included in ψ is obviously convex with the minimum at $\mathbb{B} = \mathbb{I}$ and depends only on $\operatorname{tr}\mathbb{B}$ and on $\operatorname{tr}(\mathbb{B}\mathbb{B})$, i.e., on invariants of $\mathbb{B}$, therefore it does not violate any of the basic principles of continuum physics. Moreover, such an addition does not affect the first three terms in the asymptotic expansion of ψ near $\mathbb{I}$, on the logarithmic scale. To see this, let $\mathbb{H}$ denote the Hencky logarithmic tensor satisfying $e^{\mathbb{H}} = \mathbb{B}$ (which exists due to the positive definiteness of $\mathbb{B}$). Using Jacobi's identity, we compute that

$$\operatorname{tr}\mathbb{B} - 3 - \ln\det\mathbb{B} = \operatorname{tr}(e^{\mathbb{H}} - \mathbb{I} - \mathbb{H}) = \operatorname{tr}(\tfrac{1}{2}\mathbb{H}^2 + O(\mathbb{H}^3)).$$

On the other hand, we easily get

$$\tfrac{1}{2}|\mathbb{B} - \mathbb{I}|^2 = \tfrac{1}{2}\operatorname{tr}(e^{2\mathbb{H}} - 2e^{\mathbb{H}} + \mathbb{I}) = \operatorname{tr}(\tfrac{1}{2}\mathbb{H}^2 + O(\mathbb{H}^3)),$$

hence we also have

$$(1-\gamma)(\operatorname{tr}\mathbb{B}-3-\ln\det\mathbb{B})+\tfrac{1}{2}\gamma|\mathbb{B}-\mathbb{I}|^2=\operatorname{tr}(\tfrac{1}{2}\mathbb{H}^2+O(\mathbb{H}^3))$$

and we see that for $\mathbb{B}$ being close to identity, the form of ψ is almost independent of the choice of parameter γ and the second part of ψ in (5.15) can be just understood as a correction for large values of $\mathbb{B}$.

Next, we show how the constitutive equation for $\mathbb{T}$ (see (5.8)) appears naturally if we start with the choice of the Helmholtz free energy (5.15) and require that the form of the equation for $\mathbb{B}$ is given by (5.11). For the derivation, we followed the approach developed in [53] that stems from the balance equations and requires the knowledge of how the material stores the energy, but we simplify the derivation presented there by assuming that the density is constant (in fact we set for simplicity $\varrho = 1$ and hence $\operatorname{div}\boldsymbol{v} = 0$) and the flow is isothermal, i.e., the temperature θ is constant as well. Under these assumptions the balance equations of continuum physics (for linear and angular momenta, energy and for formulation of the second law of thermodynamics) take the form

$$\begin{aligned}
&\overset{\bullet}{\boldsymbol{v}} = \operatorname{div}\mathbb{T}, \quad \mathbb{T}=\mathbb{T}^T,\\
&\overset{\bullet}{e} = \mathbb{T}\cdot\mathbb{D}\boldsymbol{v} - \operatorname{div}\boldsymbol{j}_e,\\
&\overset{\bullet}{\eta} = \xi - \operatorname{div}\boldsymbol{j}_\eta \qquad \text{with } \xi \geq 0,
\end{aligned}$$

where e is the (specific) internal energy, η is the entropy, ξ is the rate of entropy production, $\mathbb{T}$ is the Cauchy stress tensor and the quantities $\boldsymbol{j}_e$, $\boldsymbol{j}_\eta$ represent the internal and the entropy fluxes, respectively. Since the quantities ψ, e, θ and η are related through the thermodynamical identity

$$e = \psi + \theta\eta,$$

we can easily deduce from above identities that

$$\theta\xi = \theta\overset{\bullet}{\eta} + \operatorname{div}(\theta\boldsymbol{j}_\eta) = \mathbb{T}\cdot\mathbb{D}\boldsymbol{v} - \operatorname{div}(\boldsymbol{j}_e - \theta\boldsymbol{j}_\eta) - \overset{\bullet}{\psi}. \tag{5.16}$$

To evaluate the last term, we rewrite (5.11) as

$$-\overset{\bullet}{\mathbb{B}} = -\lambda\Delta\mathbb{B} - a(\mathbb{D}\boldsymbol{v}\mathbb{B}+\mathbb{B}\mathbb{D}\boldsymbol{v}) - (\mathbb{W}\boldsymbol{v}\mathbb{B}-\mathbb{B}\mathbb{W}\boldsymbol{v}) + \delta_1(\mathbb{B}-\mathbb{I}) + \delta_2(\mathbb{B}^2-\mathbb{B}). \tag{5.17}$$

Next, it follows from (5.15) that

$$\frac{\partial\psi(\mathbb{B})}{\partial\mathbb{B}} = \mathbb{J},$$

where $\mathbb{J}$ is defined by

$$\mathbb{J} = \mu(1-\gamma)(\mathbb{I}-\mathbb{B}^{-1}) + \mu\gamma(\mathbb{B}-\mathbb{I}).$$

Consequently, taking the inner product of (5.17) with $\mathbb{J}$ we observe that (since $\mathbb{B}\mathbb{J} = \mathbb{J}\mathbb{B}$, the term with $\mathbb{W}\boldsymbol{v}$ vanishes)

$$\begin{aligned}
-\overset{\bullet}{\psi} &= -\lambda\Delta\mathbb{B}\cdot\mathbb{J} - a(\mathbb{D}\boldsymbol{v}\mathbb{B}+\mathbb{B}\mathbb{D}\boldsymbol{v})\cdot\mathbb{J} - (\mathbb{W}\boldsymbol{v}\mathbb{B}-\mathbb{B}\mathbb{W}\boldsymbol{v})\cdot\mathbb{J}\\
&\quad + \delta_1(\mathbb{B}-\mathbb{I})\cdot\mathbb{J} + \delta_2(\mathbb{B}^2-\mathbb{B})\cdot\mathbb{J}\\
&= -\lambda\operatorname{div}(\nabla\psi(\mathbb{B})) - a(\mathbb{D}\boldsymbol{v}\mathbb{B}+\mathbb{B}\mathbb{D}\boldsymbol{v})\cdot\mathbb{J}\\
&\quad + \delta_1(\mathbb{B}-\mathbb{I})\cdot\mathbb{J} + \delta_2(\mathbb{B}^2-\mathbb{B})\cdot\mathbb{J} + \lambda\nabla\mathbb{B}\cdot\nabla\mathbb{J}.
\end{aligned} \tag{5.18}$$

To evaluate the terms on the last line, we use the symmetry and the positive definiteness of the matrix $\mathbb{B}$ to obtain

$$\begin{aligned}
(\mathbb{B}-\mathbb{I})\cdot\mathbb{J} &= \mu(1-\gamma)|\mathbb{B}^{\frac12}-\mathbb{B}^{-\frac12}|^2+\mu\gamma|\mathbb{B}-\mathbb{I}|^2,\\
(\mathbb{B}^2-\mathbb{B})\cdot\mathbb{J} &= \mu(1-\gamma)|\mathbb{B}-\mathbb{I}|^2+\mu\gamma|\mathbb{B}^{\frac32}-\mathbb{B}^{\frac12}|^2,\\
\nabla\mathbb{B}\cdot\nabla\mathbb{J} &= \mu\gamma|\nabla\mathbb{B}|^2-\mu(1-\gamma)\nabla\mathbb{B}\cdot\nabla\mathbb{B}^{-1}\\
&= \mu\gamma|\nabla\mathbb{B}|^2+\mu(1-\gamma)\nabla\mathbb{B}\cdot\mathbb{B}^{-1}\nabla\mathbb{B}\mathbb{B}^{-1}\\
&= \mu\gamma|\nabla\mathbb{B}|^2+\mu(1-\gamma)|\mathbb{B}^{-\frac12}\nabla\mathbb{B}\mathbb{B}^{-\frac12}|^2.
\end{aligned}\tag{5.19}$$

Similarly, we obtain

$$a(\mathbb{B}\mathbb{D}\boldsymbol{v}+\mathbb{D}\boldsymbol{v}\mathbb{B})\cdot\mathbb{J}=\Big[2\mu a((1-\gamma)(\mathbb{B}-\mathbb{I})+\gamma(\mathbb{B}^2-\mathbb{B}))\Big]\cdot\mathbb{D}\boldsymbol{v}.\tag{5.20}$$

Thus, using (5.18)–(5.20) in (5.16), we conclude that

$$\begin{aligned}
\theta\xi = &-\operatorname{div}(\lambda\nabla\psi(\mathbb{B})+\boldsymbol{j}_e-\theta\boldsymbol{j}_\eta)\\
&+\Big[\mathbb{T}-2a\mu((1-\gamma)(\mathbb{B}-\mathbb{I})+\gamma(\mathbb{B}^2-\mathbb{B}))\Big]\cdot\mathbb{D}\boldsymbol{v}\\
&+\mu\lambda(\gamma|\nabla\mathbb{B}|^2+(1-\gamma)|\mathbb{B}^{-\frac12}\nabla\mathbb{B}\mathbb{B}^{-\frac12}|^2)\\
&+\mu\left((1-\gamma)\delta_1|\mathbb{B}^{\frac12}-\mathbb{B}^{-\frac12}|^2+\gamma\delta_2|\mathbb{B}^{\frac32}-\mathbb{B}^{\frac12}|^2\right)\\
&\quad+\mu\left(((1-\gamma)\delta_2+\gamma\delta_1)|\mathbb{B}-\mathbb{I}|^2\right).
\end{aligned}\tag{5.21}$$

Hence, assuming that the fluxes fulfil

$$\lambda\nabla\psi(\mathbb{B})+\boldsymbol{j}_e-\theta\boldsymbol{j}_\eta=0,\tag{5.22}$$

and setting (compare with (5.8))

$$\mathbb{T}=-p\mathbb{I}+2\nu\mathbb{D}\boldsymbol{v}+2a\mu((1-\gamma)(\mathbb{B}-\mathbb{I})+\gamma(\mathbb{B}^2-\mathbb{B})),$$

the identity (5.21) reduces to (noticing that $-p\mathbb{I}\cdot\mathbb{D}\boldsymbol{v}=-p\operatorname{div}\boldsymbol{v}=0$)

$$\begin{aligned}
\theta\xi = &\ \mu\lambda(\gamma|\nabla\mathbb{B}|^2+(1-\gamma)|\mathbb{B}^{-\frac12}\nabla\mathbb{B}\mathbb{B}^{-\frac12}|^2)+2\nu|\mathbb{D}\boldsymbol{v}|^2\\
&+\mu\left((1-\gamma)\delta_1|\mathbb{B}^{\frac12}-\mathbb{B}^{-\frac12}|^2+\gamma\delta_2|\mathbb{B}^{\frac32}-\mathbb{B}^{\frac12}|^2\right)\\
&+\mu\left(((1-\gamma)\delta_2+\gamma\delta_1)|\mathbb{B}-\mathbb{I}|^2\right),
\end{aligned}\tag{5.23}$$

which gives the nonnegative rate of the entropy production. Moreover, we have seen how the form of the Cauchy stress tensor $\mathbb{T}$ in (5.8) is dictated by the second line in (5.21). Furthermore, we can also see in (5.23) (and also in the last line of (5.19)) how the choice of the free energy (5.15) affects the entropy production due to the presence of the diffusive term $\Delta\mathbb{B}$ in (5.3).

5.4 Concept of weak solution

In order to introduce the proper concept of weak solution, we first derive the basic energy estimates based on the observations from the previous section. First, taking the scalar product of (5.10) and $\boldsymbol{v}$, we deduce the kinetic energy identity

$$\frac12\partial_t|\boldsymbol{v}|^2+\frac12\operatorname{div}(|\boldsymbol{v}|^2\boldsymbol{v})-\operatorname{div}(\mathbb{T}\boldsymbol{v})+\mathbb{T}\cdot\mathbb{D}\boldsymbol{v}=\boldsymbol{f}\cdot\boldsymbol{v}$$

and replacing the term $\mathbb{T}\cdot\mathbb{D}\boldsymbol{v}$ from the equation (5.16), and using then also (5.22) and (5.23), we finally obtain

$$\begin{aligned}
&\partial_t(\psi + \tfrac{1}{2}|\boldsymbol{v}|^2) + \operatorname{div}((\psi + \tfrac{1}{2}|\boldsymbol{v}|^2)\boldsymbol{v}) - \operatorname{div}(\mathbb{T}\boldsymbol{v} + \lambda\nabla\psi(\mathbb{B})) + 2\nu|\mathbb{D}\boldsymbol{v}|^2 \\
&\quad + \mu\lambda\left(\gamma|\nabla\mathbb{B}|^2 + (1-\gamma)|\mathbb{B}^{-\frac{1}{2}}\nabla\mathbb{B}\mathbb{B}^{-\frac{1}{2}}|^2\right) \\
&\quad + \mu\left((1-\gamma)\delta_1|\mathbb{B}^{\frac{1}{2}} - \mathbb{B}^{-\frac{1}{2}}|^2 + \gamma\delta_2|\mathbb{B}^{\frac{3}{2}} - \mathbb{B}^{\frac{1}{2}}|^2 + ((1-\gamma)\delta_2 + \gamma\delta_1)|\mathbb{B} - \mathbb{I}|^2\right) \\
&\qquad = \boldsymbol{f}\cdot\boldsymbol{v}.
\end{aligned}$$

Integrating the above identity over Ω, using integration by parts and the boundary conditions (5.12)–(5.14), we obtain

$$\begin{aligned}
&\frac{\mathrm{d}}{\mathrm{d}t}\int_\Omega \left(\tfrac{1}{2}|\boldsymbol{v}|^2 + \psi(\mathbb{B})\right) + 2\nu\int_\Omega |\mathbb{D}\boldsymbol{v}|^2 + \sigma\int_{\partial\Omega}|\boldsymbol{v}|^2 \\
&\quad + \mu\lambda\int_\Omega \left(\gamma|\nabla\mathbb{B}|^2 + (1-\gamma)|\mathbb{B}^{-\frac{1}{2}}\nabla\mathbb{B}\mathbb{B}^{-\frac{1}{2}}|^2\right) \\
&\qquad + \mu\int_\Omega \Big((1-\gamma)\delta_1|\mathbb{B}^{\frac{1}{2}} - \mathbb{B}^{-\frac{1}{2}}|^2 + \gamma\delta_2|\mathbb{B}^{\frac{3}{2}} - \mathbb{B}^{\frac{1}{2}}|^2 \\
&\qquad\qquad + ((1-\gamma)\delta_2 + \gamma\delta_1)|\mathbb{B} - \mathbb{I}|^2\Big) = \int_\Omega \boldsymbol{f}\cdot\boldsymbol{v}.
\end{aligned} \tag{5.24}$$

The identity (5.24) indicates the proper choice of the function spaces for the solution $(\boldsymbol{v}, \mathbb{B})$ and the form of the (weak) formulation of the solution to (5.1)–(5.6).

Definition 5.1. *Let $T > 0$ and assume that $\Omega \subset \mathbb{R}^3$ is a Lipschitz domain. Let $\gamma \in (0,1)$, $\nu, \sigma, \lambda > 0$, $\delta_1, \delta_2 \geq 0$, $a \in \mathbb{R}$, and $\boldsymbol{f} \in L^2(0,T; W^{-1,2}_{\boldsymbol{n},\mathrm{div}})$, $\boldsymbol{v}_0 \in L^2_{\boldsymbol{n},\mathrm{div}}(\Omega)$. Furthermore, let $\mathbb{B}_0 \in L^2(\Omega)$ be such that*

$$-\int_\Omega \ln\det\mathbb{B}_0 < \infty.$$

Then, we say that a couple $(\boldsymbol{v}, \mathbb{B}) : Q \to \mathbb{R}^3\times\mathbb{R}^{d\times d}_{>0}$ is a weak solution to (5.1)–(5.6) if the following hold:

$$\begin{aligned}
&\boldsymbol{v} \in L^2(0,T; W^{1,2}_{\boldsymbol{n},\mathrm{div}}) \cap L^\infty(0,T; L^2(\Omega)), \quad \partial_t\boldsymbol{v} \in L^{\frac{4}{3}}(0,T; W^{-1,2}_{\boldsymbol{n},\mathrm{div}}), \\
&\mathbb{B} \in L^2(0,T; W^{1,2}(\Omega)) \cap L^\infty(0,T; L^2(\Omega)), \quad \partial_t\mathbb{B} \in L^{\frac{4}{3}}(0,T; W^{-1,2}(\Omega));
\end{aligned}$$

For all $\boldsymbol{\varphi} \in L^4(0,T; W^{1,2}_{\boldsymbol{n},\mathrm{div}})$ we have

$$\begin{aligned}
&\int_0^T \langle\partial_t\boldsymbol{v}, \boldsymbol{\varphi}\rangle + \int_Q \boldsymbol{v}\cdot\nabla\boldsymbol{v}\cdot\boldsymbol{\varphi} + \sigma\int_0^T\int_{\partial\Omega} \mathcal{T}\boldsymbol{v}\cdot\mathcal{T}\boldsymbol{\varphi} \\
&\quad = -\int_Q (2\nu\mathbb{D}\boldsymbol{v} + 2a\mu((1-\gamma)(\mathbb{B} - \mathbb{I}) + \gamma(\mathbb{B}^2 - \mathbb{B})))\cdot\nabla\boldsymbol{\varphi} + \int_0^T\langle\boldsymbol{f}, \boldsymbol{\varphi}\rangle;
\end{aligned} \tag{5.25}$$

For all $\mathbb{A} \in L^4(0,T; W^{1,2}(\Omega))$, $\mathbb{A} = \mathbb{A}^T$, we have

$$\begin{aligned}
&\int_0^T \langle\partial_t\mathbb{B}, \mathbb{A}\rangle + \int_Q (\boldsymbol{v}\cdot\nabla\mathbb{B} + 2\mathbb{B}\mathbb{W}\boldsymbol{v} - 2a\,\mathbb{B}\mathbb{D}\boldsymbol{v})\cdot\mathbb{A} \\
&\qquad + \int_Q (\delta_1(\mathbb{B} - \mathbb{I}) + \delta_2(\mathbb{B}^2 - \mathbb{B}))\cdot\mathbb{A} + \lambda\int_Q \nabla\mathbb{B}\cdot\nabla\mathbb{A} = 0;
\end{aligned} \tag{5.26}$$

The initial conditions are satisfied in the following sense

$$\lim_{t\to 0_+} (\|\boldsymbol{v}(t) - \boldsymbol{v}_0\|_2 + \|\mathbb{B}(t) - \mathbb{B}_0\|_2) = 0. \tag{5.27}$$

Moreover, we say that the solution satisfies the energy inequality if, for all $t \in (0,T)$*:*

$$\begin{aligned}\int_\Omega \left(\frac{|\boldsymbol{v}(t)|^2}{2} + \psi(\mathbb{B}(t))\right) + \int_0^t \left(2\nu\|\mathbb{D}\boldsymbol{v}\|_2^2 + \sigma\|\mathcal{T}\boldsymbol{v}\|_{2,\partial\Omega}^2\right) \\ + \mu\lambda\int_0^t \left((1-\gamma)\left\|\mathbb{B}^{-\frac{1}{2}}\nabla\mathbb{B}\mathbb{B}^{-\frac{1}{2}}\right\|_2^2 + \gamma\|\nabla\mathbb{B}\|_2^2\right) \\ + \mu\int_0^t \Big((1-\gamma)\delta_1\left\|\mathbb{B}^{\frac{1}{2}} - \mathbb{B}^{-\frac{1}{2}}\right\|_2^2 + \gamma\delta_2\left\|\mathbb{B}^{\frac{3}{2}} - \mathbb{B}^{\frac{1}{2}}\right\|_2^2 \\ + (\gamma\delta_1 + (1-\gamma)\delta_2)\|\mathbb{B} - \mathbb{I}\|_2^2\Big) \\ \le \int_\Omega \left(\frac{|\boldsymbol{v}_0|^2}{2} + \psi(\mathbb{B}_0)\right) + \int_0^t \langle \boldsymbol{f}, \boldsymbol{v}\rangle.\end{aligned} \tag{5.28}$$

In the above definition we used the following notation. By $L^p(\Omega)$ and $W^{n,p}(\Omega)$, $1 \le p \le \infty$, $n \in \mathbb{N}$, we denote the usual Lebesgue and Sobolev space, with their usual norms denoted as $\|\cdot\|_p$ and $\|\cdot\|_{n,p}$, respectively. The trace operator that maps $W^{1,p}(\Omega)$ into $L^q(\partial\Omega)$, for certain $q \ge 1$, is denoted by $\mathcal{T}$. Further, we set $W^{-1,p'}(\Omega) = (W^{1,p}(\Omega))^*$, where $p' = p/(p-1)$. We use the same notation for the function spaces of scalar-, vector-, or tensor-valued functions, but we distinguish the functions themselves using different fonts such as a for scalars, $\boldsymbol{a}$ for vectors and $\mathbb{A}$ for tensors. Also, we do not specify the meaning of the duality pairing $\langle\cdot,\cdot\rangle$, assuming that it is clear from the context. Moreover, for certain subspaces of vector valued functions, we use the following notation:

$$\begin{aligned} C^\infty_{\boldsymbol{n}} &= \{\boldsymbol{w} : \Omega \to \mathbb{R}^3 : \boldsymbol{w} \text{ infinitely differentiable, } \boldsymbol{w}\cdot\boldsymbol{n} = 0 \text{ on } \partial\Omega\}, \\ C^\infty_{\boldsymbol{n},\mathrm{div}} &= \{\boldsymbol{w} \in C^\infty_{\boldsymbol{n}} : \operatorname{div}\boldsymbol{w} = 0 \text{ in } \Omega\}, \\ L^2_{\boldsymbol{n},\mathrm{div}} &= \overline{C^\infty_{\boldsymbol{n},\mathrm{div}}}^{\|\cdot\|_2}, \quad W^{1,2}_{\boldsymbol{n},\mathrm{div}} = \overline{C^\infty_{\boldsymbol{n},\mathrm{div}}}^{\|\cdot\|_{1,2}}, \quad W^{3,2}_{\boldsymbol{n},\mathrm{div}} = \overline{C^\infty_{\boldsymbol{n},\mathrm{div}}}^{\|\cdot\|_{3,2}}, \\ W^{-1,2}_{\boldsymbol{n},\mathrm{div}} &= (W^{1,2}_{\boldsymbol{n},\mathrm{div}})^*, \quad W^{-3,2}_{\boldsymbol{n},\mathrm{div}} = (W^{3,2}_{\boldsymbol{n},\mathrm{div}})^*. \end{aligned}$$

Occasionally, we denote the standard inner products in $L^2(\Omega)$ and $L^2(\partial\Omega)$ as $(\cdot,\cdot)$ and $(\cdot,\cdot)_{\partial\Omega}$, respectively. The Bochner spaces of mappings from $(0,T)$ to a Banach space X is denoted as $L^p(0,T;X)$ with the norm $\|\cdot\|_{L^p(0,T;X)} = (\int_0^T \|\cdot\|_X^p)^{\frac{1}{p}}$. If $X = L^q(\Omega)$, or $X = W^{k,q}(\Omega)$, we write just $\|\cdot\|_{L^pL^q}$, or $\|\cdot\|_{L^pW^{k,q}}$, respectively. The space $\mathcal{C}_{\mathrm{weak}}(0,T;X) \subset L^\infty(0,T;X)$ denotes a space of weakly continuous functions, i.e., for every $f \in \mathcal{C}_{\mathrm{weak}}(0,T;X)$ and every $g \in X^*$ there holds

$$\lim_{t\to t_0} \langle f(t), g\rangle = \langle f(t_0), g\rangle.$$

The symbol $\mathbb{R}^{3\times 3}_{\mathrm{sym}}$ denotes the set of symmetric 3×3 real matrices. Furthermore, by $\mathbb{R}^{d\times d}_{>0}$ we denote the subset of $\mathbb{R}^{3\times 3}_{\mathrm{sym}}$ which consists of positive definite matrices, i.e., those which satisfy

$$\mathbb{A}\boldsymbol{z}\cdot\boldsymbol{z} > 0 \quad \text{for all } \boldsymbol{z} \in \mathbb{R}^3 \setminus \{0\}.$$

5.5 Existence of a weak solution

The key result of this chapter is the following

Theorem 5.2. *Let $T > 0$ and assume that $\Omega \subset \mathbb{R}^3$ is a Lipschitz domain. Suppose that $\gamma \in (0,1)$, $\nu, \sigma, \lambda > 0$, $\delta_1, \delta_2 \geq 0$, $a \in \mathbb{R}$, and $\boldsymbol{f} \in L^2(0,T; W^{-1,2}_{\boldsymbol{n},\mathrm{div}})$, $\boldsymbol{v}_0 \in L^2_{\boldsymbol{n},\mathrm{div}}(\Omega)$. Furthermore, let $\mathbb{B}_0 \in L^2(\Omega)$ be such that*

$$-\int_\Omega \ln \det \mathbb{B}_0 < \infty.$$

Then there exists a weak solution to (5.1)–(5.6) *satisfying the energy inequality.*

Let us briefly explain the main difficulties connected with the analysis of the system (5.9)–(5.13) and our ideas how to solve them. In the standard models where $\gamma = 0$, to get an a priori estimate for $\mathbb{B}$, the appropriate test function to take in (5.11) is $\mathbb{I} - \mathbb{B}^{-1}$. Then, using (5.9) and (5.10) tested by $\boldsymbol{v}$, one can eliminate the problematic terms, such as $\mathbb{B} \cdot \mathbb{D}\boldsymbol{v}$ coming from the objective derivative. However, the non-negative quantity to be controlled, which comes from the diffusion term, turns out to be just $|\mathbb{B}^{-\frac{1}{2}}\nabla\mathbb{B}\mathbb{B}^{-\frac{1}{2}}|^2$ and this provides little to no information. In particular, the terms $\nabla\boldsymbol{v}\mathbb{B}$ appearing in (5.11) are going to be just integrable and it is unclear if one can show strong convergence of $\mathbb{B}$. Instead, one would like to test also by $\mathbb{B}$ to achieve control over $|\nabla\mathbb{B}|^2$. But this is not possible, since the resulting term $\nabla\boldsymbol{v}\mathbb{B} \cdot \mathbb{B}$ cannot be estimated without some serious simplifications (such as boundedness of $\nabla\boldsymbol{v}$, two or one dimensional setting or small data). Quite remarkably, this problem is solved simply by adding $\frac{1}{2}\gamma|\mathbb{B} - \mathbb{I}|^2$ into the constitutive form for ψ. More precisely, considering $\gamma \in (0,1)$, we observe that the appropriate test function in (5.11) is in fact $(1-\gamma)(\mathbb{I}-\mathbb{B}^{-1})+\gamma(\mathbb{B}-\mathbb{I})$. Indeed, the terms from the objective derivative cancel again due to the presence of $\gamma(\mathbb{B}^2 - \mathbb{B})$ in $\mathbb{T}$. But now, we also get $\gamma|\nabla\mathbb{B}|^2$ under control, which is much better information than in the case $\gamma = 0$ and it yields compactness of all the terms appearing in (5.10) and (5.11). We have seen above that such a modification of ψ, and consequently of $\mathbb{T}$, is not ad-hoc and that it rests on solid physical grounds.

The second and also the last major difficulty which we encounter is how one can justify testing of (5.11) by $\mathbb{B}^{-1}$ on the approximate (discrete level), where $\mathbb{B}^{-1}$ might not even exist. This we overcome by designing a delicate approximation scheme, which takes into account the smallest eigenvalue of $\mathbb{B}$, and also by noting that testing (5.11) only by $\mathbb{B}$ yields sufficiently strong a priori estimates for the initial limit passage (in the Galerkin approximation of $\mathbb{B}$).

Up to now, there have been no results on global existence of weak solutions to Oldroyd-B models in three dimensions, including either the standard, or diffusive variants. The closest result so far is probably [59, Theorem 4.1], however there it is assumed that $\delta_2 > 0$ and $\lambda = 0$ (Giesekus model), whereas we treat also the case $\delta_2 = 0$, but with $\lambda > 0$ (diffusive Oldroyd-B or Giesekus model). Moreover, in [59], only the weak sequential stability of a hypothetical approximation is proved. We, on the other hand, provide the complete existence proof, including the construction of approximate solutions (which, in viscoelasticity, is generally a non-trivial task). In the article [49], Lions and Masmoudi prove the global existence in three dimensions, but only for $a = 0$ (corrotational case), which is known to be much easier. The local in time existence of regular solutions for the non-diffusive variants of the models above ($\lambda = 0$) is proved in the pioneering work [38, Theorem 2.4.]. There, also the global existence for small data is shown. In two dimensions, the problem is solved in [25] in the case $\lambda > 0$, $\delta_1 > 0$, $\delta_2 = 0$ (diffusive Oldroyd-B model). There are also global large data existence results in three dimensions for slightly different classes of diffusive rate-type viscoelastic models, but under some simplifying assumptions. For example, in [19] and [11], the authors consider the case where $\mathbb{B} = b\mathbb{I}$. This assumption, however, turns (5.11) into a much simpler scalar equation. Moreover, note that if $\mathbb{B} = b\mathbb{I}$, then the equations (5.10) and (5.11) decouple (which is not the case in [19] and [11] since there the considered constitutive relation for $\mathbb{T}$ is more complicated than here). Furthermore, in [50], the authors consider yet another class of Peterlin viscoelastic models with stress diffusion and prove existence of a global two- or three-dimensional solution. However, the free energy associated with these models depends only on the trace of the extra stress tensor. This is a significant simplification, which can even be seen as unphysical. See also [23] for various modifications of Oldroyd-B viscoelastic models, for which an existence theory is available. Finally, in [5] (see also [42]), the global existence of a weak solution is shown for a certain regularized Oldroyd-B model (including a cut-off or nonlinear $p-$Laplace operator in the diffusive term in $\mathbb{B}$). Thus, one might argue that since the case $\gamma > 0$ could be also seen as a regularization of the original model, we are just proving an existence of a solution to another regularization. However, this argument is not, in our opinion, correct for several reasons. First of all, the "regularization" $\gamma > 0$ does not touch the equation (5.11) at all. Second, it is not obvious why the nonlinear term $\gamma(\mathbb{B} - \mathbb{I})^2$ should have any regularization effect. And, perhaps most importantly, we already showed in Section 5.3 that the model with $\gamma > 0$ is physically well founded and worthy of studying in its own right.

Since the topic is quite new and unexplored, we decided, for brevity and clarity of presentation to consider only the isothermal case. However, we believe that the framework and ideas presented here are robust enough to provide an existence analysis also for the full thermodynamical model if the evolution of the internal energy is described correctly. This is the subject of our forthcoming study.

Remark. Finally, we close this section with several concluding remarks on possible extensions, but we do not provide their proofs in this paper.

(i) Theorem 5.2 holds also in arbitrary dimensions $d > 3$ (in $d \leq 2$, it is known), however with worse function spaces for the time derivatives and better for the test functions. Indeed, the only dimension-specific argument in the proof below is in the derivation of interpolation inequalities, which are then used to estimate $\partial_t \boldsymbol{v}$ and $\partial_t \mathbb{B}$. Moreover, all of the non-linear terms in (5.25), (5.26) are integrable for arbitrary d if the test functions are smooth. In addition, if $d = 2$, then we can prove the existence of a weak solution satisfying even the energy equality, i.e., (5.28) holds with the equality sign.

(ii) When Ω has $C^{1,1}$ boundary, then, in addition, there exists a pressure $p \in L^{\frac{5}{3}}(Q)$, which appears in (5.2). Then, the test functions in (5.25) need not be divergence-free if we include the term $\int_\Omega p \operatorname{div} \boldsymbol{\varphi}$ in (5.25). This follows in a standard way, using the Helmholtz decomposition of $\boldsymbol{v}$ (see e.g. [8] for details).

(iii) It is possible to replace (5.12), (5.13) by the no-slip boundary condition $\boldsymbol{v} = \mathbf{0}$ on $\partial\Omega$. Then, we only need to change the space $W^{1,2}_{\boldsymbol{n}}$ to $W^{1,2}_0$, and so on. However, then it seems that the pressure p can be only obtained as a distribution (see [8]).

Proof. Throughout the proof, we simplify notation by assuming

$$\lambda = \mu = \nu = \sigma = 1$$

and refer to Section 5.3 for a detailed computation for general parameters. To shorten all formulae, we also denote

$$\begin{aligned} \mathbb{S}(\mathbb{A}) &= (1-\gamma)(\mathbb{A}-\mathbb{I}) + \gamma(\mathbb{A}^2-\mathbb{A}) \qquad && \text{for } \mathbb{A} \in \mathbb{R}^{3\times 3}, \\ \mathbb{R}(\mathbb{A}) &= \delta_1(\mathbb{A}-\mathbb{I}) + \delta_2(\mathbb{A}^2-\mathbb{A}) && \text{for } \mathbb{A} \in \mathbb{R}^{3\times 3}. \end{aligned}$$

The general scheme of the proof is the following: In order to invert the matrix $\mathbb{B}$ and to avoid problems with low integrability in the objective derivative, we introduce the special cut-off function

$$\rho_\varepsilon(\mathbb{A}) = \frac{\max\{0, \Lambda(\mathbb{A}) - \varepsilon\}}{\Lambda(\mathbb{A})(1+\varepsilon|\mathbb{A}|^3)} \quad \text{for } \mathbb{A} \in \mathbb{R}^{3\times 3}_{\mathrm{sym}},$$

where $\Lambda(\mathbb{A})$ denotes a minimal eigenvalue of $\mathbb{A}$ (whose spectrum is real due to its symmetry)[1]. Since eigenvalues of a matrix depend continuously on its entries, the function ρ_ε is continuous. Moreover, for any positive definite matrix $\mathbb{A}$ there holds $\rho_\varepsilon(\mathbb{A}) \to 1$ as $\varepsilon \to 0_+$. We construct a solution by an approximation scheme with parameters k, l and ε, where $k, l \in \mathbb{N}$ correspond to the Galerkin approximation for $\boldsymbol{v}$ and $\mathbb{B}$, respectively, and ε corresponds to the presence of the cut-off function ρ_ε in certain terms. The first limit we take is $l \to \infty$, which corresponds to the limit in the equation for $\mathbb{B}$. This way, the limiting object $\mathbb{B}$ is infinite-dimensional and, using the properties of ρ_ε, we prove that $\mathbb{B}^{-1}$ exists. With the help of this information, we derive the energy estimates that are uniform with respect to all the parameters. Next, we let $\varepsilon \to 0_+$ in order to remove the truncation function and finally we take $k \to \infty$, which corresponds to the limiting procedure in the equation for the velocity $\boldsymbol{v}$.

Galerkin approximation

Following e.g., [52, Appendix A.4], we know that there exists a basis $\{\boldsymbol{w}_i\}_{i=1}^\infty$ of $W^{3,2}_{\boldsymbol{n},\mathrm{div}}$, which is orthonormal in $L^2(\Omega)$ and orthogonal in $W^{3,2}_{\boldsymbol{n},\mathrm{div}}$. Moreover, the projection $P_k : L^2(\Omega) \to \mathrm{span}\{\boldsymbol{w}_i\}_{i=1}^k$, defined as

$$P_k\boldsymbol{\varphi} = \sum_{i=1}^k (\boldsymbol{\varphi}, \boldsymbol{w}_i)\boldsymbol{w}_i, \quad \boldsymbol{\varphi} \in L^2(\Omega),$$

is continuous in $L^2(\Omega)$ and also in $W^{3,2}_{\boldsymbol{n},\mathrm{div}}$ independently of k, i.e.,

$$\|P_k\boldsymbol{\varphi}\|_2 \le C\|\boldsymbol{\varphi}\|_2 \qquad \|P_k\boldsymbol{\varphi}\|_{W^{3,2}_{\boldsymbol{n},\mathrm{div}}} \le C\|\boldsymbol{\varphi}\|_{W^{3,2}_{\boldsymbol{n},\mathrm{div}}}$$

[1] We set $\rho_\varepsilon(\mathbb{A}) = 0$ if $\Lambda(\mathbb{A}) = 0$.

for all $\boldsymbol{\varphi} \in W^{3,2}_{\boldsymbol{n},\mathrm{div}}$, where the constant C is independent of k. Furthermore, by the standard embedding, we also have that $W^{3,2}_{\boldsymbol{n},\mathrm{div}} \hookrightarrow W^{2,6}(\Omega) \hookrightarrow W^{1,\infty}(\Omega)$. Similarly, we construct the basis $\{\mathbb{W}_j\}_{j=1}^{\infty}$ of $W^{1,2}(\Omega)$, which is L^2-orthonormal, $W^{1,2}$-orthogonal and the projection

$$Q_l\mathbb{A} = \sum_{j=1}^{l} (\mathbb{A}, \mathbb{W}_j)\mathbb{W}_j, \quad \mathbb{A} \in L^2(\Omega),$$

is continuous in $L^2(\Omega)$ and in $W^{1,2}(\Omega)$ independently of l.

Then for fixed $k, l \in \mathbb{N}$ and $\varepsilon \in (0,1)$, we look for the functions $\boldsymbol{v}^{k,l}_{\varepsilon}, \mathbb{B}^{k,l}_{\varepsilon}$ of the form

$$\boldsymbol{v}^{k,l}_{\varepsilon}(t,x) = \sum_{i=1}^{k} c^{k,l,\varepsilon}_i(t)\boldsymbol{w}_i(x) \quad \text{and} \quad \mathbb{B}^{k,l}_{\varepsilon}(t,x) = \sum_{j=1}^{l} d^{k,l,\varepsilon}_j(t)\mathbb{W}_j(x),$$

where $c^{k,l,\varepsilon}_i, d^{k,l,\varepsilon}_j$, $i = 1, \ldots, k$, $j = 1, \ldots, l$, are unknown functions of time, and we require that $\boldsymbol{v}^{k,l}_{\varepsilon}, \mathbb{B}^{k,l}_{\varepsilon}$ (and consequently the functions $c^{k,l,\varepsilon}_i(t)$ and $d^{k,l,\varepsilon}_j(t)$) satisfy the following system of $(k + l)$ ordinary differential equations in time interval $(0, T)$:

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}t}(\boldsymbol{v}^{k,l}_{\varepsilon}, \boldsymbol{w}_i) + ((\boldsymbol{v}^{k,l}_{\varepsilon} \cdot \nabla)\boldsymbol{v}^{k,l}_{\varepsilon}, \boldsymbol{w}_i) + 2(\mathbb{D}\boldsymbol{v}^{k,l}_{\varepsilon}, \nabla \boldsymbol{w}_i) + (\mathcal{T}\boldsymbol{v}^{k,l}_{\varepsilon}, \mathcal{T}\boldsymbol{w}_i)_{\partial\Omega} \\ = -2a(\rho_\varepsilon(\mathbb{B}^{k,l}_{\varepsilon})\mathbb{S}(\mathbb{B}^{k,l}_{\varepsilon}), \nabla \boldsymbol{w}_i) + \langle \boldsymbol{f}, \boldsymbol{w}_i\rangle \qquad \text{for } i = 1, \ldots, k,\end{aligned} \tag{5.29}$$

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}t}(\mathbb{B}^{k,l}_{\varepsilon}, \mathbb{W}_j) + ((\boldsymbol{v}^{k,l}_{\varepsilon} \cdot \nabla)\mathbb{B}^{k,l}_{\varepsilon}, \mathbb{W}_j) + (\rho_\varepsilon(\mathbb{B}^{k,l}_{\varepsilon})\mathbb{R}(\mathbb{B}^{k,l}_{\varepsilon}), \mathbb{W}_j) + (\nabla\mathbb{B}^{k,l}_{\varepsilon}, \nabla\mathbb{W}_j) \\ = 2(\rho_\varepsilon(\mathbb{B}^{k,l}_{\varepsilon})\mathbb{B}^{k,l}_{\varepsilon}(a\mathbb{D}\boldsymbol{v}^{k,l}_{\varepsilon} - \mathbb{W}\boldsymbol{v}^{k,l}_{\varepsilon}), \mathbb{W}_j) \qquad \text{for } j = 1, \ldots, l.\end{aligned} \tag{5.30}$$

Due to the L^2-orthonormality of the bases $\{\boldsymbol{w}_i\}_{i=1}^{\infty}$ and $\{\mathbb{W}_j\}_{j=1}^{\infty}$, the system (5.29)–(5.30) can be rewritten as a nonlinear system of ordinary differential equations for $c^{k,l,\varepsilon}_i$ and $d^{k,l,\varepsilon}_j$, where $i = 1, \ldots, k$ and $j = 1, \ldots, l$, and we equip this system with the initial conditions

$$c^{k,l,\varepsilon}_i(0) = (\boldsymbol{v}_0, \boldsymbol{w}_i) \qquad \text{and} \qquad d^{k,l,\varepsilon}_j(0) = (\mathbb{B}^{\varepsilon}_0, \mathbb{W}_j). \tag{5.31}$$

Here, $\mathbb{B}^{\varepsilon}_0$ is defined by

$$\mathbb{B}^{\varepsilon}_0(x) = \begin{cases} \mathbb{B}_0(x) & \text{if } \Lambda(\mathbb{B}_0(x)) > \varepsilon, \\ \mathbb{I} & \text{elsewhere.} \end{cases}$$

Since $\mathbb{B}_0(x) \in \mathbb{R}^{d\times d}_{>0}$ for almost every $x \in \Omega$, we have that $\Lambda(\mathbb{B}_0(x)) > 0$ for almost all $x \in \Omega$. Consequently, using the fact $\mathbb{B}_0 \in L^2(\Omega)$, we obtain, as $\varepsilon \to 0_+$, that

$$\|\mathbb{B}^{\varepsilon}_0 - \mathbb{B}_0\|_2^2 = \int_{\Lambda(\mathbb{B}_0)\leq\varepsilon} |\mathbb{I} - \mathbb{B}_0|^2 \to 0$$

Note also that the initial conditions (5.31) can be rewritten as $\boldsymbol{v}^{k,l}_{\varepsilon}(0) = P_k\boldsymbol{v}_0$ and $\mathbb{B}^{k,l}_{\varepsilon}(0) = Q_l\mathbb{B}^{\varepsilon}_0$.

For the system (5.29)–(5.31), Carathéodory's theorem can be applied and therefore there exists $T^* > 0$ and absolutely continuous functions $c^{k,l,\varepsilon}_i$, $d^{k,l,\varepsilon}_j$ satisfying (5.31) and (5.29)–(5.30) almost everywhere in $(0, T^*)$. If T^* is the maximal time, for which the solution exists, and $T^* < T$, then at least one of the functions $c^{k,l,\varepsilon}_i$, $d^{k,l,\varepsilon}_j$ must blow up as $t \to T^*_-$. But using the estimate presented below (see (5.36) valid for all $t \in (0, T^*)$), this never happens. Thus, we can set $T^* = T$.

5.5.1 Limit $l \to \infty$

In this part, we simplify the notation and denote the approximating solution, constructed in the previous section, by setting $(\boldsymbol{v}_l, \mathbb{B}_l) = (\boldsymbol{v}_\varepsilon^{k,l}, \mathbb{B}_\varepsilon^{k,l})$. We start by proving estimates independent of l. Since $\mathbb{B}_l(t)$ and $\boldsymbol{v}_l(t)$ belong for almost all t to the linear hull of $\{\mathbb{W}_j\}_{j=1}^l$ and $\{\boldsymbol{w}_i\}_{i=1}^k$, respectively, we can use $\boldsymbol{v}_l$ instead of $\boldsymbol{w}_i$ in (5.29) and $\mathbb{B}_l$ instead of $\mathbb{W}_j$ in (5.30) to deduce

$$\frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}t}\|\mathbb{B}_l\|_2^2 + \|\nabla\mathbb{B}_l\|_2^2 = 2a(\rho_\varepsilon(\mathbb{B}_l)\mathbb{B}_l\mathbb{D}\boldsymbol{v}_l, \mathbb{B}_l) - (\rho_\varepsilon(\mathbb{B}_l)\mathbb{R}(\mathbb{B}_l), \mathbb{B}_l), \tag{5.32}$$

$$\frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}t}\|\boldsymbol{v}_l\|_2^2 + 2\|\mathbb{D}\boldsymbol{v}_l\|_2^2 + \|\mathcal{T}\boldsymbol{v}_l\|_{2,\partial\Omega}^2 = -2a(\rho_\varepsilon(\mathbb{B}_l)\mathbb{S}(\mathbb{B}_l), \mathbb{D}\boldsymbol{v}_l) + \langle \boldsymbol{f}, \boldsymbol{v}_l\rangle, \tag{5.33}$$

where we used the integration by parts formula and the facts that $\operatorname{div}\boldsymbol{v}_l = 0$ and $\mathcal{T}\boldsymbol{v}\cdot\boldsymbol{n} = 0$. Next, it follows from the definition of ρ_ε, $\mathbb{R}$ and $\mathbb{S}$ that

$$\rho_\varepsilon(\mathbb{B}_l)\left(|\mathbb{S}(\mathbb{B}_l)| + |\mathbb{R}(\mathbb{B}_l)||\mathbb{B}_l| + |\mathbb{B}_l|^2\right) \le C\frac{1+|\mathbb{B}_l|^3}{1+\varepsilon|\mathbb{B}_l|^3} \le C(\varepsilon). \tag{5.34}$$

Here, the notation $C(\varepsilon)$ emphasizes that the constant C depends on ε; we keep this notation in what follows. Summing (5.32) and (5.33) and using the estimate (5.34) to bound the term on the right-hand side, we obtain with the help of Hölder's, Young's and Korn's inequalities that

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(\|\boldsymbol{v}_l\|_2^2 + \|\mathbb{B}_l\|_2^2\right) + \|\mathbb{D}\boldsymbol{v}_l\|_2^2 + \|\mathcal{T}\boldsymbol{v}_l\|_{2,\partial\Omega}^2 + \|\nabla\mathbb{B}_l\|_2^2 \le C(\varepsilon) + C\|\boldsymbol{f}\|_{W^{-1,2}_{\boldsymbol{n},\mathrm{div}}}^2.$$

After integrating over $(0,T)$ with respect to time, we obtain the following bound:

$$\begin{aligned}\sup_{t\in(0,T)}&\left(\|\boldsymbol{v}_l\|_2^2 + \|\mathbb{B}_l\|_2^2\right) + \int_0^T \left(\|\mathbb{D}\boldsymbol{v}_l\|_2^2 + \|\mathcal{T}\boldsymbol{v}_l\|_{2,\partial\Omega}^2 + \|\nabla\mathbb{B}_l\|_2^2\right)\\ &\le C(\varepsilon) + \|P_k\boldsymbol{v}_0\|_2^2 + \|Q_l\mathbb{B}_0^\varepsilon\|_2^2 + C\int_0^T \|\boldsymbol{f}\|_{W^{-1,2}_{\boldsymbol{n},\mathrm{div}}}^2 \le C(\varepsilon),\end{aligned} \tag{5.35}$$

where the last inequality follows from the continuity of the projections P_k and Q_l and from the assumptions on data, namely that

$$\|\boldsymbol{v}_0\|_2^2 + \|\mathbb{B}_0\|_2^2 + \|\ln\det\mathbb{B}_0\|_1 + C\int_0^T \|\boldsymbol{f}\|_{W^{-1,2}_{\boldsymbol{n},\mathrm{div}}}^2 < \infty.$$

Next, we focus on the estimate for time derivatives. First, it follows from L^2-orthonormality of the bases and the estimate (5.35) that

$$\sum_{i=1}^k c_i(t)^2 + \sum_{j=1}^l d_j(t)^2 \le C(\varepsilon). \tag{5.36}$$

Then, since $\boldsymbol{v}_l$ is a linear combination of $\{\boldsymbol{w}_i\}_{i=1}^k \subset W^{1,\infty}(\Omega)$, we can estimate

$$\|\boldsymbol{v}_l\|_{L^\infty W^{1,\infty}} \le \operatorname*{ess\,sup}_{t\in(0,T)}\sum_{i=1}^k |c_i(t)|\|\boldsymbol{w}_i\|_{1,\infty} \le C(\varepsilon,k), \tag{5.37}$$

and we can deduce from (5.29) that

$$\|\partial_t\boldsymbol{v}_l\|_{L^\infty W^{1,\infty}} \le C(\varepsilon,k). \tag{5.38}$$

Finally, it follows from (5.30) and (5.35) that

$$\|\partial_t \mathbb{B}_l\|_{L^2W^{-1,2}} \le C(\varepsilon, k). \tag{5.39}$$

Using (5.35), (5.37)–(5.39) and Banach-Alaoglu's theorem, we can find subsequences (which we do not relabel) and corresponding weak limits (denoted with the subscript ε), such that, for $l \to \infty$, we get

$$\begin{aligned}
\boldsymbol{v}_l &\rightharpoonup \boldsymbol{v}_\varepsilon && \text{weakly in } L^2(0,T; W^{1,2}_{\boldsymbol{n},\mathrm{div}}), && (5.40)\\
\boldsymbol{v}_l &\overset{*}{\rightharpoonup} \boldsymbol{v}_\varepsilon && \text{weakly}^* \text{ in } L^\infty(0,T; W^{1,\infty}(\Omega)), && (5.41)\\
\partial_t \boldsymbol{v}_l &\overset{*}{\rightharpoonup} \partial_t \boldsymbol{v}_\varepsilon && \text{weakly}^* \text{ in } L^\infty(0,T; W^{1,\infty}(\Omega)), && (5.42)\\
\mathcal{T}\boldsymbol{v}_l &\rightharpoonup \mathcal{T}\boldsymbol{v}_\varepsilon && \text{weakly in } L^2(0,T; L^2(\partial\Omega)), && (5.43)\\
\mathbb{B}_l &\rightharpoonup \mathbb{B}_\varepsilon && \text{weakly in } L^2(0,T; W^{1,2}(\Omega)), && (5.44)\\
\partial_t \mathbb{B}_l &\rightharpoonup \partial_t \mathbb{B}_\varepsilon && \text{weakly in } L^2(0,T; W^{-1,2}(\Omega)). && (5.45)
\end{aligned}$$

Moreover, it follows from (5.40), (5.42), (5.44), (5.45) and from the Aubin-Lions lemma that for some further subsequences, we have

$$\begin{aligned}
\boldsymbol{v}_l &\to \boldsymbol{v}_\varepsilon && \text{strongly in } L^2(Q), && (5.46)\\
\mathbb{B}_l &\to \mathbb{B}_\varepsilon && \text{strongly in } L^2(Q) \text{ and a.e. in } Q, && (5.47)\\
\rho_\varepsilon(\mathbb{B}_l) &\to \rho_\varepsilon(\mathbb{B}_\varepsilon) && \text{a.e. in } Q. && (5.48)
\end{aligned}$$

Using the convergence results (5.40)–(5.48), it is rather standard to let $l \to \infty$ in (5.29)–(5.30). This way, for almost all $t \in (0,T)$, we obtain

$$\begin{aligned}
(\partial_t \boldsymbol{v}_\varepsilon, \boldsymbol{w}_i) + (\boldsymbol{v}_\varepsilon \cdot \nabla \boldsymbol{v}_\varepsilon, \boldsymbol{w}_i) + 2(\mathbb{D}\boldsymbol{v}_\varepsilon, \nabla \boldsymbol{w}_i) + (\mathcal{T}\boldsymbol{v}_\varepsilon, \mathcal{T}\boldsymbol{w}_i)_{\partial\Omega}&\\
= -2a(\rho_\varepsilon(\mathbb{B}_\varepsilon)\mathbb{S}(\mathbb{B}_\varepsilon), \nabla \boldsymbol{w}_i) + \langle \boldsymbol{f}, \boldsymbol{w}_i\rangle&
\end{aligned} \tag{5.49}$$

for $i = 1, \ldots, k$, and

$$\begin{aligned}
\langle \partial_t \mathbb{B}_\varepsilon, \mathbb{A}\rangle + (\boldsymbol{v}_\varepsilon \cdot \nabla \mathbb{B}_\varepsilon, \mathbb{A}) + (\nabla \mathbb{B}_\varepsilon, \nabla \mathbb{A})&\\
= 2(\rho_\varepsilon(\mathbb{B}_\varepsilon)\mathbb{B}_\varepsilon(a\mathbb{D}\boldsymbol{v}_\varepsilon - \mathbb{W}\boldsymbol{v}_\varepsilon), \mathbb{A}) - (\rho_\varepsilon(\mathbb{B}_\varepsilon)\mathbb{R}(\mathbb{B}_\varepsilon), \mathbb{A})&
\end{aligned} \tag{5.50}$$

for all $\mathbb{A} \in W^{1,2}(\Omega)$. Also, from (5.44) and (5.45), we get $\mathbb{B}_\varepsilon \in \mathcal{C}(0,T; L^2(\Omega))$ and it is standard to show that $\mathbb{B}_\varepsilon(0,\cdot) = \mathbb{B}_0^\varepsilon$ and $\boldsymbol{v}_\varepsilon(0,\cdot) = P_k \boldsymbol{v}_0$.

5.5.2 Limit $\varepsilon \to 0$

In this part we consider the solutions $(\boldsymbol{v}_\varepsilon, \mathbb{B}_\varepsilon)$ constructed in the preceding section for $\varepsilon \in (0,1)$ and we study their behaviour as $\varepsilon \to 0_+$. To do so, we first have to derive estimates that are uniform with respect to ε. Following the ideas used before in the derivation of the model, we wish to test (5.50) by the function

$$\mathbb{J}_\varepsilon = (1-\gamma)(\mathbb{I} - \mathbb{B}_\varepsilon^{-1}) + \gamma(\mathbb{B}_\varepsilon - \mathbb{I}). \tag{5.51}$$

This test function, however, contains $\mathbb{B}_\varepsilon^{-1}$ and we need to justify that it exists (for any $\varepsilon \in (0,1)$).

Estimates for the inverse matrix, still ε-dependent

First, we prove that $\Lambda(\mathbb{B}_\varepsilon) \geq \varepsilon$. For this purpose, let $\boldsymbol{z} \in \mathbb{R}^3$ be arbitrary and consider[2]

$$\mathbb{A} = (\mathbb{B}_\varepsilon \boldsymbol{z} \cdot \boldsymbol{z} - \varepsilon|\boldsymbol{z}|^2)_- \, (\boldsymbol{z} \otimes \boldsymbol{z}), \quad \text{where } (\boldsymbol{z} \otimes \boldsymbol{z})_{ij} = z_i z_j \tag{5.52}$$

in (5.50). Due to the properties of $\mathbb{B}_\varepsilon$ (see (5.44)), we know that $\mathbb{A}$ belongs to $L^2(0,T;W^{1,2}(\Omega))$ and we can use it as a test function in (5.50). Upon inserting $\mathbb{A}$ into (5.50), we integrate the result over $(0,\tau)$ with some fixed $\tau \in (0,T)$. We evaluate all terms in (5.50) separately. For the time derivative, we have

$$\begin{aligned}
\int_0^\tau \langle \partial_t \mathbb{B}_\varepsilon, \mathbb{A}\rangle &= \int_0^\tau \Big\langle \partial_t(\mathbb{B}_\varepsilon \boldsymbol{z} \cdot \boldsymbol{z} - \varepsilon|\boldsymbol{z}|^2), (\mathbb{B}_\varepsilon \boldsymbol{z} \cdot \boldsymbol{z} - \varepsilon|\boldsymbol{z}|^2)_- \Big\rangle \\
&= \frac{1}{2}\Big\|(\mathbb{B}_\varepsilon(\tau)\boldsymbol{z} \cdot \boldsymbol{z} - \varepsilon|\boldsymbol{z}|^2)_-\Big\|_2^2 - \frac{1}{2}\Big\|(\mathbb{B}_0^\varepsilon \boldsymbol{z} \cdot \boldsymbol{z} - \varepsilon|\boldsymbol{z}|^2)_-\Big\|_2^2 \\
&= \frac{1}{2}\Big\|(\mathbb{B}_\varepsilon(\tau)\boldsymbol{z} \cdot \boldsymbol{z} - \varepsilon|\boldsymbol{z}|^2)_-\Big\|_2^2,
\end{aligned} \tag{5.53}$$

where, for the last equality, the definition of $\mathbb{B}_0^\varepsilon$ was used. Furthermore, we obtain

$$\begin{aligned}
\int_Q \nabla \mathbb{B}_\varepsilon \cdot \nabla \mathbb{A} &= \int_0^\tau \int_\Omega \nabla(\mathbb{B}_\varepsilon - \varepsilon \mathbb{I}) \cdot \nabla((\mathbb{B}_\varepsilon \boldsymbol{z} \cdot \boldsymbol{z} - \varepsilon|\boldsymbol{z}|^2)_- \, (\boldsymbol{z} \otimes \boldsymbol{z})) \\
&= \int_0^\tau \Big\|\nabla(\mathbb{B}_\varepsilon \boldsymbol{z} \cdot \boldsymbol{z} - \varepsilon|\boldsymbol{z}|^2)_-\Big\|_2^2
\end{aligned} \tag{5.54}$$

and

$$\begin{aligned}
\int_Q \boldsymbol{v}_\varepsilon \cdot \nabla \mathbb{B}_\varepsilon \cdot \mathbb{A} &= \int_0^\tau \int_\Omega \boldsymbol{v}_\varepsilon \cdot \nabla(\mathbb{B}_\varepsilon \boldsymbol{z} \cdot \boldsymbol{z} - \varepsilon|\boldsymbol{z}|^2)(\mathbb{B}_\varepsilon \boldsymbol{z} \cdot \boldsymbol{z} - \varepsilon|\boldsymbol{z}|^2)_- \\
&= \frac{1}{2}\int_0^\tau \int_\Omega \boldsymbol{v}_\varepsilon \cdot \nabla((\mathbb{B}_\varepsilon \boldsymbol{z} \cdot \boldsymbol{z} - \varepsilon|\boldsymbol{z}|^2)_-^2 \\
&= -\frac{1}{2}\int_0^\tau \int_\Omega ((\mathbb{B}_\varepsilon \boldsymbol{z} \cdot \boldsymbol{z} - \varepsilon|\boldsymbol{z}|^2)_-^2 \operatorname{div} \boldsymbol{v}_\varepsilon = 0,
\end{aligned} \tag{5.55}$$

integrating by parts and using the fact that $\operatorname{div} \boldsymbol{v}_\varepsilon = 0$ and $\mathcal{T}\boldsymbol{v}_\varepsilon = 0$. Since

$$\mathbb{B}_\varepsilon \boldsymbol{z} \cdot \boldsymbol{z} \geq \Lambda(\mathbb{B}_\varepsilon)|\boldsymbol{z}|^2 \qquad \text{a.e. in } Q,$$

we also observe, that

$$0 \geq (\Lambda(\mathbb{B}_\varepsilon) - \varepsilon)_+(\mathbb{B}_\varepsilon \boldsymbol{z} \cdot \boldsymbol{z} - \varepsilon|\boldsymbol{z}|^2)_- \geq (\Lambda(\mathbb{B}_\varepsilon) - \varepsilon)_+(\Lambda(\mathbb{B}_\varepsilon) - \varepsilon)_- \, |\boldsymbol{z}|^2 = 0.$$

Hence, we get

$$\rho_\varepsilon(\mathbb{B}_\varepsilon)\mathbb{A} = 0 \qquad \text{a.e. in } Q. \tag{5.56}$$

Consequently, inserting $\mathbb{A}$ of the form (5.52) into (5.50), we see that the right-hand side is identically zero. Therefore, relations (5.53), (5.54), (5.55) and (5.56) yield

$$\begin{aligned}
&\Big\|(\mathbb{B}_\varepsilon \boldsymbol{z} \cdot \boldsymbol{z} - \varepsilon|\boldsymbol{z}|^2)_-\Big\|_2^2(\tau) \\
&\quad \leq \Big\|(\mathbb{B}_\varepsilon \boldsymbol{z} \cdot \boldsymbol{z} - \varepsilon|\boldsymbol{z}|^2)_-\Big\|_2^2(\tau) + 2\int_0^\tau \Big\|\nabla(\mathbb{B}_\varepsilon \boldsymbol{z} \cdot \boldsymbol{z} - \varepsilon|\boldsymbol{z}|^2)_-\Big\|_2^2 = 0,
\end{aligned}$$

[2] In this subsection, we use the notation $(f)_+ = \max\{0, f\}$ and $(f)_- = \min\{0, f\}$.

which implies

$$\mathbb{B}_\varepsilon \boldsymbol{z} \cdot \boldsymbol{z} \geq \varepsilon |\boldsymbol{z}|^2 \quad \text{for every } \boldsymbol{z} \in \mathbb{R}^3 \text{ and a.e. in } Q. \tag{5.57}$$

Thus, we have the following estimate for the minimal eigenvalue of $\mathbb{B}_\varepsilon$:

$$\Lambda(\mathbb{B}_\varepsilon) \geq \inf_{0 \neq \boldsymbol{z} \in \mathbb{R}^3} \frac{\mathbb{B}_\varepsilon \boldsymbol{z} \cdot \boldsymbol{z}}{|\boldsymbol{z}|^2} \geq \varepsilon.$$

Therefore, the inverse matrix $\mathbb{B}_\varepsilon^{-1}$ is well defined and satisfies

$$|\mathbb{B}_\varepsilon^{-1}| \leq \frac{C}{\varepsilon} \quad \text{a.e. in Q.} \tag{5.58}$$

Furthermore, since

$$\nabla \mathbb{B}_\varepsilon^{-1} = \mathbb{B}_\varepsilon^{-1} \mathbb{B}_\varepsilon \nabla \mathbb{B}_\varepsilon^{-1} = \mathbb{B}_\varepsilon^{-1} \nabla(\mathbb{B}_\varepsilon \mathbb{B}_\varepsilon^{-1}) - \mathbb{B}_\varepsilon^{-1} (\nabla \mathbb{B}_\varepsilon) \mathbb{B}_\varepsilon^{-1} = -\mathbb{B}_\varepsilon^{-1} (\nabla \mathbb{B}_\varepsilon) \mathbb{B}_\varepsilon^{-1},$$

we conclude from (5.35) and (5.58), that

$$\int_Q |\nabla \mathbb{B}_\varepsilon^{-1}|^2 \leq \int_Q |\mathbb{B}_\varepsilon^{-1}|^4 |\nabla \mathbb{B}_\varepsilon|^2 \leq C(\varepsilon).$$

Hence, the inverse of $\mathbb{B}_\varepsilon$ exists and $\mathbb{B}_\varepsilon^{-1} \in L^2(0,T; W^{1,2}(\Omega))$.

Estimates independent of (ε, k)

At this point, we can test (5.50) with $\mathbb{J}_\varepsilon$ defined in (5.51). This way, we obtain

$$\begin{aligned}\langle \partial_t \mathbb{B}_\varepsilon, \mathbb{J}_\varepsilon \rangle &+ (\boldsymbol{v}_\varepsilon \cdot \nabla \mathbb{B}_\varepsilon, \mathbb{J}_\varepsilon) + (\nabla \mathbb{B}_\varepsilon, \nabla \mathbb{J}_\varepsilon) \\ &= 2(\rho_\varepsilon(\mathbb{B}_\varepsilon)\mathbb{B}_\varepsilon(a\mathbb{D}\boldsymbol{v}_\varepsilon - \mathbb{W}\boldsymbol{v}_\varepsilon), \mathbb{J}_\varepsilon) - (\rho_\varepsilon(\mathbb{B}_\varepsilon)\mathbb{R}(\mathbb{B}_\varepsilon), \mathbb{J}_\varepsilon).\end{aligned}$$

Next, we evaluate all terms. Here, we follow very closely the procedure developed in Section 5.3, see the derivation of (5.18) and consequent identities. Since

$$\mathbb{J}_\varepsilon = \frac{\partial \psi(\mathbb{B}_\varepsilon)}{\partial \mathbb{B}_\varepsilon},$$

where ψ is defined in (5.15), it is clear that

$$\begin{aligned}\langle \partial_t \mathbb{B}_\varepsilon, \mathbb{J}_\varepsilon \rangle &= \frac{\mathrm{d}}{\mathrm{d}t} \int_\Omega \psi(\mathbb{B}_\varepsilon), \\ (\boldsymbol{v}_\varepsilon \cdot \nabla \mathbb{B}_\varepsilon, \mathbb{J}_\varepsilon) &= \int_\Omega \boldsymbol{v}_\varepsilon \cdot \nabla \psi(\mathbb{B}_\varepsilon) = 0.\end{aligned}$$

Next, recalling (5.19), we get

$$\begin{aligned}(\rho_\varepsilon(\mathbb{B}_\varepsilon)\mathbb{R}(\mathbb{B}_\varepsilon), \mathbb{J}_\varepsilon) = \int_\Omega \rho_\varepsilon(\mathbb{B}_\varepsilon) \Big(&\delta_1(1-\gamma)|\mathbb{B}_\varepsilon^{\frac{1}{2}} - \mathbb{B}_\varepsilon^{-\frac{1}{2}}|^2 + (\delta_1\gamma + \delta_2(1-\gamma))|\mathbb{B}_\varepsilon - \mathbb{I}|^2 \\ &+\delta_2\gamma|\mathbb{B}_\varepsilon^{\frac{3}{2}} - \mathbb{B}_\varepsilon^{\frac{1}{2}}|^2\Big), \\ (\nabla \mathbb{B}_\varepsilon, \nabla \mathbb{J}_\varepsilon) = \gamma \|\nabla \mathbb{B}_\varepsilon\|_2^2 &+ (1-\gamma)\|\mathbb{B}_\varepsilon^{-\frac{1}{2}} \nabla \mathbb{B}_\varepsilon \mathbb{B}_\varepsilon^{-\frac{1}{2}}\|_2^2\end{aligned}$$

and due to the fact that $\mathbb{B}_\varepsilon\mathbb{J}_\varepsilon = \mathbb{J}_\varepsilon\mathbb{B}_\varepsilon$ we also have

$$
\begin{aligned}
(\rho_\varepsilon(\mathbb{B}_\varepsilon)(\mathbb{W}\boldsymbol{v}_\varepsilon\mathbb{B}_\varepsilon - \mathbb{B}_\varepsilon\mathbb{W}\boldsymbol{v}_\varepsilon), \mathbb{J}_\varepsilon) &= 0,\\
a(\rho_\varepsilon(\mathbb{B}_\varepsilon)(\mathbb{D}\boldsymbol{v}_\varepsilon\mathbb{B}_\varepsilon + \mathbb{B}_\varepsilon\mathbb{D}\boldsymbol{v}_\varepsilon), \mathbb{J}_\varepsilon) &= 2a(\rho_\varepsilon(\mathbb{B}_\varepsilon)\mathbb{D}\boldsymbol{v}_\varepsilon, \mathbb{B}_\varepsilon\mathbb{J}_\varepsilon)\\
&= 2a(\rho_\varepsilon(\mathbb{B}_\varepsilon)\mathbb{D}\boldsymbol{v}_\varepsilon, (1-\gamma)(\mathbb{B}_\varepsilon - \mathbb{I}) + \gamma(\mathbb{B}_\varepsilon^2 - \mathbb{B}_\varepsilon))\\
&= 2a(\rho_\varepsilon(\mathbb{B}_\varepsilon)\mathbb{S}(\mathbb{B}_\varepsilon), \mathbb{D}\boldsymbol{v}_\varepsilon),
\end{aligned}
$$

where we used the fact that the trace of $\mathbb{D}\boldsymbol{v}_\varepsilon$ is identically zero. Hence, using $\mathbb{A} = \mathbb{J}_\varepsilon$ (defined in (5.51)) in (5.50) and taking into account the above identities, we deduce that

$$
\begin{aligned}
\frac{\mathrm{d}}{\mathrm{d}t}\int_\Omega \psi(\mathbb{B}_\varepsilon) &+ (1-\gamma)\left\|\mathbb{B}_\varepsilon^{-\frac12}\nabla\mathbb{B}_\varepsilon\mathbb{B}_\varepsilon^{-\frac12}\right\|_2^2 + \gamma\|\nabla\mathbb{B}_\varepsilon\|_2^2\\
&+(\gamma\delta_1 + (1-\gamma)\delta_2)\left\|\sqrt{\rho_\varepsilon(\mathbb{B}_\varepsilon)}(\mathbb{B}_\varepsilon - \mathbb{I})\right\|_2^2\\
&\quad+(1-\gamma)\delta_1\left\|\sqrt{\rho_\varepsilon(\mathbb{B}_\varepsilon)}(\mathbb{B}_\varepsilon^{\frac12} - \mathbb{B}_\varepsilon^{-\frac12})\right\|_2^2\\
&\qquad+\gamma\delta_2\left\|\sqrt{\rho_\varepsilon(\mathbb{B}_\varepsilon)}(\mathbb{B}_\varepsilon^{\frac32} - \mathbb{B}_\varepsilon^{\frac12})\right\|_2^2 = 2a(\rho_\varepsilon(\mathbb{B}_\varepsilon)\mathbb{S}(\mathbb{B}_\varepsilon), \mathbb{D}\boldsymbol{v}_\varepsilon).
\end{aligned}
\tag{5.59}
$$

Similarly as in previous section, replacing $\boldsymbol{w}_i$ in (5.49) by $\boldsymbol{v}_\varepsilon$, we get

$$
\frac12\frac{\mathrm{d}}{\mathrm{d}t}\|\boldsymbol{v}_\varepsilon\|_2^2 + 2\|\mathbb{D}\boldsymbol{v}_\varepsilon\|_2^2 + \|\boldsymbol{\mathcal{T}}\boldsymbol{v}_\varepsilon\|_{2,\partial\Omega}^2 = \langle\boldsymbol{f}, \boldsymbol{v}_\varepsilon\rangle - 2a(\rho_\varepsilon(\mathbb{B}_\varepsilon)\mathbb{S}(\mathbb{B}_\varepsilon), \mathbb{D}\boldsymbol{v}_\varepsilon). \tag{5.60}
$$

Thus, summing (5.59) and (5.60) and integrating the result with respect to time $t \in (0,\tau)$, we deduce the identity

$$
\begin{aligned}
&\frac12\|\boldsymbol{v}_\varepsilon(\tau)\|_2^2 + \int_\Omega \psi(\mathbb{B}_\varepsilon(\tau))\\
&\quad+\int_0^\tau\Big(2\|\mathbb{D}\boldsymbol{v}_\varepsilon\|_2^2 + \|\boldsymbol{\mathcal{T}}\boldsymbol{v}_\varepsilon\|_{2,\partial\Omega}^2 + (1-\gamma)\left\|\mathbb{B}_\varepsilon^{-\frac12}\nabla\mathbb{B}_\varepsilon\mathbb{B}_\varepsilon^{-\frac12}\right\|_2^2 + \gamma\|\nabla\mathbb{B}_\varepsilon\|_2^2\\
&\quad+(\gamma\delta_1 + (1-\gamma)\delta_2)\left\|\sqrt{\rho_\varepsilon(\mathbb{B}_\varepsilon)}(\mathbb{B}_\varepsilon - \mathbb{I})\right\|_2^2\\
&\quad+(1-\gamma)\delta_1\left\|\sqrt{\rho_\varepsilon(\mathbb{B}_\varepsilon)}(\mathbb{B}_\varepsilon^{\frac12} - \mathbb{B}_\varepsilon^{-\frac12})\right\|_2^2 + \gamma\delta_2\left\|\sqrt{\rho_\varepsilon(\mathbb{B}_\varepsilon)}(\mathbb{B}_\varepsilon^{\frac32} - \mathbb{B}_\varepsilon^{\frac12})\right\|_2^2\Big)\\
&= \frac12\|P_k\boldsymbol{v}_0\|_2^2 + \int_\Omega\psi(\mathbb{B}_0^\varepsilon) + \int_0^\tau\langle\boldsymbol{f}, \boldsymbol{v}_\varepsilon\rangle \le \frac12\|\boldsymbol{v}_0\|_2^2 + \int_\Omega\psi(\mathbb{B}_0) + \int_0^\tau\langle\boldsymbol{f}, \boldsymbol{v}_\varepsilon\rangle,
\end{aligned}
\tag{5.61}
$$

where, for the last inequality we used the continuity of P_k, the definition of $\mathbb{B}_0^\varepsilon$ and the fact that $\psi(\mathbb{I}) = 0$.

From (5.61), we get, using Korn's, Sobolev's, Hölder's and Young's inequalities, that

$$
\|\boldsymbol{v}_\varepsilon\|_{L^\infty L^2} + \|\boldsymbol{v}_\varepsilon\|_{L^2L^6} + \|\boldsymbol{v}_\varepsilon\|_{L^2W^{1,2}} + \|\mathbb{B}_\varepsilon\|_{L^2W^{1,2}} + \|\mathbb{B}_\varepsilon\|_{L^2L^6} \le C, \tag{5.62}
$$

where the constant C depends only on Ω, $\boldsymbol{v}_0$, $\mathbb{B}_0$ and $\boldsymbol{f}$. Furthermore, the interpolation inequalities yield

$$
\|\boldsymbol{v}_\varepsilon\|_{L^{\frac{10}{3}}L^{\frac{10}{3}}} + \|\boldsymbol{v}_\varepsilon\|_{L^4L^3} + \|\mathbb{B}_\varepsilon\|_{L^{\frac{10}{3}}L^{\frac{10}{3}}} + \|\mathbb{B}_\varepsilon\|_{L^4L^3} + \|\mathbb{B}_\varepsilon\|_{L^{\frac83}L^4} \le C. \tag{5.63}
$$

Finally, we focus on the estimate for time derivatives. Let $\boldsymbol{\varphi} \in L^4(0,T;W^{3,2}_{\boldsymbol{n},\mathrm{div}})$ be such that $\|\boldsymbol{\varphi}\|_{L^4W^{3,2}} \le 1$. Then, since $\boldsymbol{v}_\varepsilon$ is a linear combination of $\{\boldsymbol{w}_i\}_{i=1}^k$, we obtain, using (5.49), Hölder's inequality, (5.61), (5.63) and $W^{3,2}$-continuity of P_k, that

$$\int_0^T \langle \partial_t \boldsymbol{v}_\varepsilon, \boldsymbol{\varphi}\rangle \le C,$$

hence

$$\|\partial_t \boldsymbol{v}_\varepsilon\|_{L^{\frac{4}{3}}W^{-3,2}_{\boldsymbol{n},\mathrm{div}}} \le C. \tag{5.64}$$

Similarly, by considering $\mathbb{A} \in L^4(0,T;W^{1,2}(\Omega))$ in (5.50), we get

$$\|\partial_t \mathbb{B}_\varepsilon\|_{L^{\frac{4}{3}}W^{-1,2}} \le C. \tag{5.65}$$

Taking the limit $\varepsilon \to 0$.

From (5.62), (5.64), (5.65), the Banach-Alaoglu theorem and the Aubin-Lions lemma, we obtain the existence of a couple $(\boldsymbol{v}_k, \mathbb{B}_k)$ satisfying the following convergence results[3]

$$\begin{aligned}
\boldsymbol{v}_\varepsilon &\rightharpoonup \boldsymbol{v}_k && \text{weakly in } L^2(0,T;W^{1,2}_{\boldsymbol{n},\mathrm{div}}),\\
\partial_t \boldsymbol{v}_\varepsilon &\rightharpoonup \partial_t \boldsymbol{v}_k && \text{weakly in } L^{\frac{4}{3}}(0,T;W^{-3,2}_{\boldsymbol{n},\mathrm{div}}),\\
\mathcal{T}\boldsymbol{v}_\varepsilon &\rightharpoonup \mathcal{T}\boldsymbol{v}_k && \text{weakly in } L^2(0,T;L^2(\partial\Omega)),\\
\mathbb{B}_\varepsilon &\rightharpoonup \mathbb{B}_k && \text{weakly in } L^2(0,T;W^{1,2}(\Omega)),\\
\partial_t \mathbb{B}_\varepsilon &\rightharpoonup \partial_t \mathbb{B}_k && \text{weakly in } L^{\frac{4}{3}}(0,T;W^{-1,2}(\Omega)),\\
\boldsymbol{v}_\varepsilon &\to \boldsymbol{v}_k && \text{strongly in } L^3(Q) \text{ and a.e. in } Q, && (5.66)\\
\mathbb{B}_\varepsilon &\to \mathbb{B}_k && \text{strongly in } L^3(Q) \text{ and a.e. in } Q. && (5.67)
\end{aligned}$$

Using (5.67) and letting $\varepsilon \to 0_+$ in (5.57), we obtain

$$\mathbb{B}_k \boldsymbol{z}\cdot\boldsymbol{z} \ge 0 \quad \text{a.e. in } Q \text{ and for all } \boldsymbol{z} \in \mathbb{R}^3.$$

Hence $\Lambda(\mathbb{B}_k) \ge 0$ and $\det \mathbb{B}_k \ge 0$ a.e. in Q. Therefore, using (5.67) again and the continuity of ψ, there exists (still possibly infinite) limit

$$\psi(\mathbb{B}_\varepsilon) \to \psi(\mathbb{B}_k) \quad \text{a.e. in } Q.$$

However, since $\psi \ge 0$, Fatou's lemma implies that, for almost every $t \in (0,T)$, we have

$$\int_\Omega \psi(\mathbb{B}_k)(t) \le \liminf_{\varepsilon\to 0_+} \int_\Omega \psi(\mathbb{B}_\varepsilon)(t) \le C.$$

Thus, we deduce that

$$\|\psi(\mathbb{B}_k)\|_{L^\infty L^1} \le C. \tag{5.68}$$

If there existed a set $E \subset Q$ of a positive measure, where $\Lambda(\mathbb{B}_k) = 0$, then also $-\ln\det\mathbb{B}_k = \infty$ on that set, which contradicts (5.68). Thus, we have

$$\Lambda(\mathbb{B}_k) > 0 \text{ a.e. in } Q. \tag{5.69}$$

[3] The convergence results (5.66), (5.67) are true in any space $L^p(Q)$, $1 \le p < \frac{10}{3}$, as can be seen from (5.63) and Vitali's theorem. The space $L^3(\Omega)$ is chosen for simplicity; in our proof, we need $p > 2$.

Then it directly follows from the continuity of Λ, that $\rho_\varepsilon(\mathbb{B}_\varepsilon) \to 1$ a.e. in Q. Then, since $\rho_\varepsilon(\mathbb{B}_\varepsilon) \le 1$, we further get, by Vitali's theorem, that

$$\rho_\varepsilon(\mathbb{B}_\varepsilon) \to 1 \quad \text{strongly in } L^p(Q) \text{ for all } p \in [1, \infty).$$

Using the established convergence results, it is easy to let $\varepsilon \to 0_+$ in (5.49) and (5.50) and obtain, for almost all $t \in (0, T)$, that

$$\begin{aligned}&\langle \partial_t \boldsymbol{v}_k, \boldsymbol{w}_i \rangle + (\boldsymbol{v}_k \cdot \nabla \boldsymbol{v}_k, \boldsymbol{w}_i) + 2(\mathbb{D}\boldsymbol{v}_k, \nabla \boldsymbol{w}_i)\\ &\qquad = -(\mathcal{T}\boldsymbol{v}_k, \mathcal{T}\boldsymbol{w}_i)_{\partial\Omega} - 2a(\mathbb{S}(\mathbb{B}_k), \nabla \boldsymbol{w}_i) + \langle \boldsymbol{f}, \boldsymbol{w}_i \rangle, \quad \text{for } i = 1, \ldots, k,\end{aligned}$$

and that

$$\begin{aligned}&\langle \partial_t \mathbb{B}_k, \mathbb{A} \rangle + (\boldsymbol{v}_k \cdot \nabla \mathbb{B}_k, \mathbb{A}) + (\nabla \mathbb{B}_k, \nabla \mathbb{A})\\ &\qquad = 2(\mathbb{B}_k(a\mathbb{D}\boldsymbol{v}_k - \mathbb{W}\boldsymbol{v}_k), \mathbb{A}) - (\mathbb{R}(\mathbb{B}_k), \mathbb{A}) \quad \text{for all } \mathbb{A} \in W^{1,2}(\Omega).\end{aligned}$$

Furthermore, we can take the limit in the estimates (5.61), (5.63), (5.64) and (5.65) using either the weak lower semi-continuity of norms or, in the terms which depend on $\mathbb{B}_\varepsilon$, e.g. $\int_Q \rho_\varepsilon(\mathbb{B}_\varepsilon)|\mathbb{B}_\varepsilon^{\frac{3}{2}} - \mathbb{B}_\varepsilon^{\frac{1}{2}}|^2$, we apply (5.69) to conclude the pointwise limit and then use Fatou's lemma. Thus, inequalities (5.61), (5.63), (5.64) and (5.65) continue to hold in the same form, but for $(\boldsymbol{v}_k, \mathbb{B}_k)$ instead of $(\boldsymbol{v}_\varepsilon, \mathbb{B}_\varepsilon)$ and with 1 instead of $\rho_\varepsilon(\mathbb{B}_\varepsilon)$. In particular, for almost all $t \in (0, T)$, we have

$$\begin{aligned}&\frac{1}{2}\|\boldsymbol{v}_k(\tau)\|_2^2 + \int_\Omega \psi(\mathbb{B}_k(\tau))\\ &\quad + \int_0^\tau \bigg(2\|\mathbb{D}\boldsymbol{v}_k\|_2^2 + \|\mathcal{T}\boldsymbol{v}_k\|_{2,\partial\Omega}^2 + (1-\gamma)\left\|\mathbb{B}_k^{-\frac{1}{2}} \nabla \mathbb{B}_k \mathbb{B}_k^{-\frac{1}{2}}\right\|_2^2 + \gamma\|\nabla \mathbb{B}_k\|_2^2\\ &\qquad + (\gamma\delta_1 + (1-\gamma)\delta_2)\|\mathbb{B}_k - \mathbb{I}\|_2^2\\ &\qquad + (1-\gamma)\delta_1\left\|\mathbb{B}_k^{\frac{1}{2}} - \mathbb{B}_k^{-\frac{1}{2}}\right\|_2^2 + \gamma\delta_2\left\|\mathbb{B}_k^{\frac{3}{2}} - \mathbb{B}_k^{\frac{1}{2}}\right\|_2^2\bigg)\\ &\le \frac{1}{2}\|\boldsymbol{v}_0\|_2^2 + \int_\Omega \psi(\mathbb{B}_0) + \int_0^\tau \langle \boldsymbol{f}, \boldsymbol{v}_k \rangle.\end{aligned}$$

The attainment of initial conditions is standard (see the last section for details in a more complicated case).

5.5.3 Limit $k \to \infty$

Since we start from the same a priori estimates as in the previous section, we follow, step by step, the procedure developed when taking the limit $\varepsilon \to 0_+$. The only difference is that the term $\rho_\varepsilon(\mathbb{B}_\varepsilon)$ is not present. Thus, using the density of $\{\boldsymbol{w}_i\}_{i=1}^\infty$ in $W^{3,2}_{\boldsymbol{n},\mathrm{div}}$, we obtain, after letting $k \to \infty$, for almost all $t \in (0, T)$, that

$$\begin{aligned}&\langle \partial_t \boldsymbol{v}, \boldsymbol{\varphi} \rangle + (\boldsymbol{v} \cdot \nabla \boldsymbol{v}, \boldsymbol{\varphi}) + 2(\mathbb{D}\boldsymbol{v}, \nabla \boldsymbol{\varphi})\\ &\qquad = -(\mathcal{T}\boldsymbol{v}, \mathcal{T}\boldsymbol{\varphi})_{\partial\Omega} - 2a(\mathbb{S}(\mathbb{B}), \nabla \boldsymbol{\varphi}) + \langle \boldsymbol{f}, \boldsymbol{\varphi} \rangle \quad \text{for all } \boldsymbol{\varphi} \in W^{3,2}_{\boldsymbol{n},\mathrm{div}}\end{aligned} \tag{5.70}$$

and that

$$\begin{aligned}&\langle \partial_t \mathbb{B}, \mathbb{A} \rangle + (\boldsymbol{v} \cdot \nabla \mathbb{B}, \mathbb{A}) + (\nabla \mathbb{B}, \nabla \mathbb{A})\\ &\qquad = 2(\mathbb{B}(a\mathbb{D}\boldsymbol{v} - \mathbb{W}\boldsymbol{v}), \mathbb{A}) - (\mathbb{R}(\mathbb{B}), \mathbb{A}) \quad \text{for all } \mathbb{A} \in W^{1,2}(\Omega).\end{aligned}$$

Moreover, from the weak lower semi-continuity of norms, we obtain the energy inequality (5.28) for almost all $t \in (0,T)$. Furthermore, the same argument as above implies that $\mathbb{B}$ is positive definite a.e. in Q. Now observe that, by Hölder's inequality and (5.63), all the terms in (5.70) except the first one, are integrable for every $\boldsymbol{\varphi} \in L^4(0,T;W^{1,2}_{\boldsymbol{n},\mathrm{div}}) \hookrightarrow L^4(0,T;L^6(\Omega))$. Indeed, for example for the non-linear terms, we get

$$\int_Q |\boldsymbol{v}\cdot\nabla\boldsymbol{v}\cdot\boldsymbol{\varphi}| \le \|\boldsymbol{v}\|_{L^4L^3}\|\nabla\boldsymbol{v}\|_{L^2L^2}\|\boldsymbol{\varphi}\|_{L^4L^6}$$

and

$$\int_Q |\mathbb{S}(\mathbb{B})\cdot\nabla\boldsymbol{\varphi}| \le C\|\mathbb{B}\|^2_{L^{\frac{8}{3}}L^4}\|\nabla\boldsymbol{\varphi}\|_{L^4L^2}.$$

Hence, the functional $\partial_t\boldsymbol{v}$ can be uniquely extended to $\partial_t\boldsymbol{v} \in L^{\frac{4}{3}}(0,T;W^{-1,2}_{\boldsymbol{n},\mathrm{div}})$ and we can use the density argument to conclude (5.25). Analogously, we obtain (5.26). Hence, it remains to show that (5.28) holds for all $t \in (0,T)$ and that the initial data fulfil (5.27).

Energy inequality for all $t \in (0,T)$

First, we observe, that due to (5.62), (5.64) and (5.65), we have that

$$\boldsymbol{v} \in \mathcal{C}_{\mathrm{weak}}(0,T;L^2(\Omega)) \quad \text{and} \quad \mathbb{B} \in \mathcal{C}_{\mathrm{weak}}(0,T;L^2(\Omega)). \tag{5.71}$$

Next, we notice that the function ψ is convex on the convex set $\mathbb{R}^{d\times d}_{>0}$. Indeed, evaluating the second Fréchet derivative of ψ, we get

$$\frac{\partial^2\psi(\mathbb{A})}{\mathbb{A}^2} = (1-\gamma)\mathbb{A}^{-1}\otimes\mathbb{A}^{-1} + \gamma\mathbb{I}\otimes\mathbb{I} \quad \text{for all } \mathbb{A} \in \mathbb{R}^{d\times d}_{>0},$$

which is obviously a positive definite operator for any $\gamma \in [0,1]$ and consequently, ψ must be convex on $\mathbb{R}^{d\times d}_{>0}$.

Further, we integrate (5.28) over $(t_1, t_1+\delta)$, where $t_1 \in (0,T)$, and divide the result by δ. Using also an elementary inequality

$$\int_0^{t_1} g \le \frac{1}{\delta}\int_{t_1}^{t_1+\delta}\left(\int_0^t g\right)\,\mathrm{d}t$$

valid for every integrable non-negative g, we get

$$\begin{aligned}
&\frac{1}{2\delta}\int_{t_1}^{t_1+\delta}\|\boldsymbol{v}(t)\|_2^2 + \frac{1}{\delta}\int_{t_1}^{t_1+\delta}\int_\Omega \psi(\mathbb{B}(t)) \\
&\quad + \int_0^{t_1}\Big(2\|\mathbb{D}\boldsymbol{v}\|_2^2 + \|\mathcal{T}\boldsymbol{v}\|^2_{2,\partial\Omega} + (1-\gamma)\Big\|\mathbb{B}^{-\frac{1}{2}}\nabla\mathbb{B}\mathbb{B}^{-\frac{1}{2}}\Big\|_2^2 + \gamma\|\nabla\mathbb{B}\|_2^2 \\
&\qquad + (\gamma\delta_1 + (1-\gamma)\delta_2)\|\mathbb{B}-\mathbb{I}\|_2^2 \\
&\qquad\quad + (1-\gamma)\delta_1\Big\|\mathbb{B}^{\frac{1}{2}} - \mathbb{B}^{-\frac{1}{2}}\Big\|_2^2 + \gamma\delta_2\Big\|\mathbb{B}^{\frac{3}{2}} - \mathbb{B}^{\frac{1}{2}}\Big\|_2^2\Big) \\
&\qquad\qquad \le \frac{1}{2}\|\boldsymbol{v}_0\|_2^2 + \int_\Omega \psi(\mathbb{B}_0) + \frac{1}{\delta}\int_{t_1}^{t_1+\delta}\int_0^\tau \langle \boldsymbol{f}, \boldsymbol{v}\rangle.
\end{aligned}$$

Finally, we let $\delta \to 0_+$. The limit on the right hand side is standard and consequently, if we show that

$$\frac{1}{2}\|\boldsymbol{v}(t_1)\|_2^2 + \int_\Omega \psi(\mathbb{B}(t_1)) \le \liminf_{\delta\to 0_+}\frac{1}{\delta}\int_{t_1}^{t_1+\delta}\left(\frac{\|\boldsymbol{v}(t)\|_2^2}{2} + \int_\Omega \psi(\mathbb{B}(t))\right), \tag{5.72}$$

then (5.28) holds for all $t \in (0,T)$. To show (5.72), we notice that due to (5.71)

$$\begin{aligned} \boldsymbol{v}(t) &\rightharpoonup \boldsymbol{v}(t_1) \quad \text{weakly in } L^2(\Omega) \text{ as } t \to t_1, \\ \mathbb{B}(t) &\rightharpoonup \mathbb{B}(t_1) \quad \text{weakly in } L^2(\Omega) \text{ as } t \to t_1, \end{aligned} \tag{5.73}$$

Consequently, due to the weak lower semicontinuity and the convexity of ψ we also have for all $t \in (0,T)$

$$\int_\Omega |\boldsymbol{v}(t)|^2 + \psi(\mathbb{B}(t)) \le C.$$

Hence denoting by $\Omega_M \subset \Omega$ the set where $|\boldsymbol{v}(t_1,\cdot)| + |\mathbb{B}(t_1,\cdot)| + |\mathbb{B}^{-1}(t_1,\cdot)| \le M$, it follows from the previous estimate that $|\Omega \setminus \Omega_M| \to 0$ as $M \to \infty$. Hence, since ψ is nonnegative and convex, we have for all $t \in (t_1, t_1+\delta)$ that

$$\begin{aligned} &\int_\Omega \frac{|\boldsymbol{v}(t)|^2}{2} + \psi(\mathbb{B}(t)) \ge \int_{\Omega_M} \frac{|\boldsymbol{v}(t)|^2}{2} + \psi(\mathbb{B}(t)) \\ &\ge \int_{\Omega_M} \frac{|\boldsymbol{v}(t_1)|^2}{2} + \psi(\mathbb{B}(t_1)) + \int_{\Omega_M} \boldsymbol{v}(t_1)\cdot(\boldsymbol{v}(t)-\boldsymbol{v}(t_1)) + \frac{\partial \psi(\mathbb{B}(t_1))}{\partial \mathbb{B}} \cdot (\mathbb{B}(t) - \mathbb{B}(t_1)). \end{aligned}$$

Since, $\boldsymbol{v}(t_1)$ and $\partial_{\mathbb{B}}\psi(\mathbb{B}(t_1))$ are bounded on Ω_M, we can integrate the above estimate over $(t_1, t_1+\delta)$ and it follows from (5.73) that

$$\liminf_{\delta \to 0_+} \frac{1}{\delta} \int_{t_1}^{t_1+\delta} \int_\Omega \frac{|\boldsymbol{v}(t)|^2}{2} + \psi(\mathbb{B}(t)) \ge \int_{\Omega_M} \frac{|\boldsymbol{v}(t_1)|^2}{2} + \psi(\mathbb{B}(t_1)).$$

Hence, letting $M \to \infty$, we deduce (5.72) and the proof of (5.28) is complete.

Attainment of initial conditions

First, it is standard to show from the construction and from the weak continuity (5.73), that for arbitrary $\boldsymbol{\varphi}, \mathbb{A} \in L^2(\Omega)$ there holds

$$\lim_{t\to 0_+} (\boldsymbol{v}(t), \boldsymbol{\varphi}) = (\boldsymbol{v}_0, \boldsymbol{\varphi}) \quad \text{and} \quad \lim_{t\to 0_+} (\mathbb{B}(t), \mathbb{A}) = (\mathbb{B}_0, \mathbb{A}). \tag{5.74}$$

Next, using the convexity of ψ and (5.74) (and consequently weak lower semicontinuity of the corresponding integral) and letting $t \to 0_+$ in (5.28), we deduce that

$$\begin{aligned} \|\boldsymbol{v}_0\|_2^2 + 2\int_\Omega \psi(\mathbb{B}_0) &\le \liminf_{t\to 0_+} \left(\|\boldsymbol{v}(t)\|_2^2 + 2\int_\Omega \psi(\mathbb{B}(t)) \right) \\ &\le \limsup_{t\to 0_+} \left(\|\boldsymbol{v}(t)\|_2^2 + 2\int_\Omega \psi(\mathbb{B}(t)) \right) \le \|\boldsymbol{v}_0\|_2^2 + 2\int_\Omega \psi(\mathbb{B}_0). \end{aligned} \tag{5.75}$$

We claim that this implies that

$$\|\boldsymbol{v}_0\|_2^2 = \lim_{t\to 0_+} \|\boldsymbol{v}(t)\|_2^2 \quad \text{and} \quad \int_\Omega \psi(\mathbb{B}_0) = \lim_{t\to 0_+} \int_\Omega \psi(\mathbb{B}(t)). \tag{5.76}$$

Indeed, assume for a moment that

$$\|\boldsymbol{v}_0\|_2^2 < \liminf_{t\to 0_+} \|\boldsymbol{v}(t)\|_2^2.$$

But then it follows from (5.75) that

$$\int_\Omega \psi(\mathbb{B}_0) > \liminf_{t\to 0_+} \int_\Omega \psi(\mathbb{B}(t)),$$

which contradicts (5.74) and convexity of ψ. Consequently, (5.76) holds.

It directly follows from $(5.74)_1$ and $(5.76)_1$ that

$$\lim_{t\to 0_+} \|\boldsymbol{v}(t) - \boldsymbol{v}_0\|_2^2 = 0.$$

To claim the same result also for $\mathbb{B}$, we simply split ψ as follows

$$\psi(\mathbb{A}) = \frac{\gamma}{2}|\mathbb{A} - \mathbb{I}|^2 + (1-\gamma)(\operatorname{tr}\mathbb{A} - 3 - \ln\det\mathbb{A}) =: \gamma\psi_1(\mathbb{A}) + (1-\gamma)\psi_2(\mathbb{A}).$$

Similarly as above, it is easy to observe that ψ_1 as well as ψ_2 are convex on the set of positive definite matrices. Therefore, $(5.76)_2$ and $(5.74)_2$ imply

$$\begin{aligned} \int_\Omega |\mathbb{B}_0 - \mathbb{I}|^2 = 2\int_\Omega \psi_1(\mathbb{B}_0) = 2\lim_{t\to 0_+}\int_\Omega \psi_1(\mathbb{B}(t)) = \lim_{t\to 0_+}\int_\Omega |\mathbb{B}(t) - \mathbb{I}|^2, \\ \int_\Omega \psi_2(\mathbb{B}_0) = \lim_{t\to 0_+}\int_\Omega \psi_2(\mathbb{B}(t)). \end{aligned} \tag{5.77}$$

Finally, (5.74) and $(5.77)_1$ lead to

$$\begin{aligned} \lim_{t\to 0_+} \|\mathbb{B}(t) - \mathbb{B}_0\|_2^2 &= \lim_{t\to 0_+} \|(\mathbb{B}(t) - \mathbb{I}) + (\mathbb{I} - \mathbb{B}_0)\|_2^2 \\ &= \lim_{t\to 0_+} \left(\|\mathbb{B}(t) - \mathbb{I}\|_2^2 + \|\mathbb{B}_0 - \mathbb{I}\|_2^2 - 2\int_\Omega (\mathbb{B}(t) - \mathbb{I})\cdot(\mathbb{B}_0 - \mathbb{I}) \right) \\ &= 0, \end{aligned}$$

which finishes the proof of (5.27) and consequently also the proof of Theorem 5.2. $\square$

References

[1] R.A. Adams, J.J.F. Fournier, *Sobolev spaces*, second ed., Pure and Applied Mathematics (Amsterdam), vol. 140, Elsevier/Academic Press, Amsterdam, 2003.

[2] M. Anand, K. Rajagopal, *A shear-thinning viscoelastic fluid model for describing the flow of blood*, International Journal of Cardiovascular Medicine and Science **4** (2004), no. 2, 59–68.

[3] N. Arada, A. Sequeira, *Strong Steady Solutions for a Generalized Oldroyd-B Model with Shear-Dependent Viscosity in a Bounded Domain*, Math. Models Methods Appl. Sci. **13** (2003), no. 09, 1303–1323.

[4] S. Banach, *Theory of linear operations*, 1 ed., North-Holland Mathematical Library 38, North Holland, 1987.

[5] J.W. Barrett, S. Boyaval, *Existence and approximation of a (regularized) Oldroyd-B model*, Math. Models Methods Appl. Sci. **21** (2011), no. 09, 1783–1837.

[6] M. Bathory, M. Bulíček, J. Málek, *Large data existence theory for three-dimensional unsteady flows of rate-type viscoelastic fluids with stress diffusion*, Submitted to Adv. Nonlinear Anal., preprint at `https://arxiv.org/abs/2002.11224v2`, 2020.

[7] R.B. Bird, P.J. Carreau, *A nonlinear viscoelastic model for polymer solutions and melts—I*, Chemical Engineering Science **23** (1968), no. 5, 427–434.

[8] J. Blechta, J. Málek, 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), no. 2, 1232–1289.

[9] H. Brezis, *Functional analysis, Sobolev spaces and partial differential equations*, Universitext, Springer, New York, 2011.

[10] T. Buckmaster, V. Vicol, *Nonuniqueness of weak solutions to the Navier-Stokes equation*, Ann. of Math. (2) **189** (2019), no. 1, 101–144.

[11] M. Bulíček, E. Feireisl, J. Málek, *On a class of compressible viscoelastic rate-type fluids with stress-diffusion*, Nonlinearity **32** (2019), no. 12, 4665–4681.

[12] M. Bulíček, J. Málek, 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 J. Math. Anal. **41** (2009), no. 2, 665–707.

[13] M. Bulíček, J. Málek, *Internal flows of incompressible fluids subject to stick-slip boundary conditions*, Vietnam J. Math. **45** (2017), no. 1-2, 207–220.

[14] ———, *Large data analysis for Kolmogorov's two-equation model of turbulence*, Nonlinear Anal. Real World Appl. **50** (2019), 104–143.

[15] M. Bulíček, J. Málek, J. Žabenský, *On generalized Stokes' and Brinkman's equations with a pressure-and shear-dependent viscosity and drag coefficient*, Nonlinear Anal. Real World Appl. **26** (2015), 109–132.

[16] M. Bulíček, J. Žabenský, *Large data existence theory for unsteady flows of fluids with pressure- and shear-dependent viscosities*, Nonlinear Anal. **127** (2015), 94–127.

[17] M. Bulíček, E. Feireisl, J. Málek, *A Navier-Stokes-Fourier system for incompressible fluids with temperature dependent material coefficients*, Nonlinear Anal. Real World Appl. **10** (2009), no. 2, 992–1015.

[18] M. Bulíček, 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), no. 6, 589–619.

[19] M. Bulíček, J. Málek, V. Průša, E. Süli, *PDE analysis of a class of thermodynamically compatible viscoelastic rate-type fluids with stress-diffusion*, Mathematical analysis in fluid mechanics—selected recent results, Contemp. Math., vol. 710, Amer. Math. Soc., Providence, RI, 2018, pp. 25–51.

[20] ———, *On incompressible heat-conducting viscoelastic rate-type fluids with stress–diffusion and purely spherical elastic response*, (to appear), 2020.

[21] L. Caffarelli, R. Kohn, L. Nirenberg, *Partial regularity of suitable weak solutions of the Navier-Stokes equations*, Comm. Pure Appl. Math. **35** (1982), no. 6, 771–831.

[22] L. Campo-Deaño, R.P.A. Dullens, D.G.A.L. Aarts, F.T. Pinho, M.S.N. Oliveira, *Viscoelasticity of blood and viscoelastic blood analogues for use in polydymethylsiloxane in vitro models of the circulatory system*, Biomicrofluidics **7** (2013), no. 3, 034102.

[23] L. Chupin, *Global Strong Solutions for Some Differential Viscoelastic Models*, SIAM J. Appl. Math. **78** (2018), no. 6, 2919–2949.

[24] E.A. Coddington, N. Levinson, *Theory of ordinary differential equations*, McGraw-Hill Book Company, Inc., New York-Toronto-London, 1955.

[25] P. Constantin, M. Kliegl, *Note on Global Regularity for Two-Dimensional Oldroyd-B Fluids with Diffusive Stress*, Arch. Ration. Mech. Anal. **206** (2012), no. 3, 725–740.

[26] J. Diestel, *Geometry of Banach spaces—selected topics*, Lecture Notes in Mathematics, Vol. 485, Springer-Verlag, Berlin-New York, 1975.

[27] M. Dostalík, V. Průša, T. Skřivan, *On diffusive variants of some classical viscoelastic rate-type models*, AIP Conference Proceedings **2107** (2019).

[28] J. Duchon, R. Robert, *Inertial energy dissipation for weak solutions of incompressible Euler and Navier-Stokes equations*, Nonlinearity **13** (2000), no. 1, 249–255.

[29] N. Dunford, J.T. Schwartz, *Linear operators. Part I*, Wiley Classics Library, John Wiley & Sons, Inc., New York, 1988, General theory, With the assistance of William G. Bade and Robert G. Bartle, Reprint of the 1958 original, A Wiley-Interscience Publication.

[30] A.W. El-Kareh, L.G. Leal, *Existence of solutions for all Deborah numbers for a non-Newtonian model modified to include diffusion*, J. Non-Newton. Fluid Mech. **33** (1989), no. 3, 257–287.

[31] L.C. Evans, *Partial differential equations*, Graduate Studies in Mathematics, vol. 19, American Mathematical Society, Providence, RI, 1998.

[32] E. Feireisl, *Relative entropies in thermodynamics of complete fluid systems*, Discrete Contin. Dyn. Syst. **32** (2012), no. 9, 3059–3080.

[33] E. Feireisl, A. Novotný, Y. Sun, *On the motion of viscous, compressible, and heat-conducting liquids*, J. Math. Phys. **57** (2016), no. 8, 083101.

[34] E. Feireisl, A. Novotný, *Singular limits in thermodynamics of viscous fluids*, Advances in Mathematical Fluid Mechanics, Birkhäuser/Springer, Cham, 2017, second edition.

[35] H. Gajewski, K. Gröger, K. Zacharias, *Nichtlineare Operatorgleichungen und Operatordifferentialgleichungen*, Akademie-Verlag, Berlin, 1974, Mathematische Lehrbücher und Monographien, II. Abteilung, Mathematische Monographien, Band 38.

[36] R.J. Gordon, W.R. Schowalter, *Anisotropic fluid theory: A different approach to the dumbbell theory of dilute polymer solutions*, Trans. Soc. Rheol. **16** (1972), no. 1, 79–97.

[37] P. Grisvard, *Elliptic problems in nonsmooth domains*, Monographs and Studies in Mathematics, vol. 24, Pitman (Advanced Publishing Program), Boston, MA, 1985.

[38] C. Guillopé, J.C. Saut, *Existence results for the flow of viscoelastic fluids with a differential constitutive law*, Nonlinear Anal. **15** (1990), no. 9, 849–869.

[39] R.A. Horn, C.R. Johnson, *Matrix analysis*, Second edition, Cambridge University Press, Cambridge, 2013.

[40] J. Hron, V. Miloš, V. Průša, O. Souček, K. Tůma, *On thermodynamics of incompressible viscoelastic rate type fluids with temperature dependent material coefficients*, Internat. J. Non-Linear Mech. **95** (2017), 193–208.

[41] M. Johnson, D. Segalman, *A model for viscoelastic fluid behavior which allows non-affine deformation*, J. Non-Newton. Fluid Mech. **2** (1977), 255–270.

[42] O. Kreml, M. Pokorný, P. Šalom, *On the global existence for a regularized model of viscoelastic non-Newtonian fluid*, Colloq. Math. **139** (2015), no. 2, 149–163.

[43] O. Kreml, *Mathematical Analysis of Models for Viscoelastic Fluids*, Ph.D. thesis, Charles University, 2010.

[44] K. Kubo, H. Kanehisa, Y. Kawakami, T. Fukunaga, *Influence of static stretching on viscoelastic properties of human tendon structures in vivo*, Journal of Applied Physiology **90** (2001), no. 2, 520–527, PMID: 11160050.

[45] A. Kufner, O. John, S. Fučík, *Function spaces*, Noordhoff International Publishing, Leyden; Academia, Prague, 1977, Monographs and Textbooks on Mechanics of Solids and Fluids; Mechanics: Analysis.

[46] S. Kuznetsov, M. Pankow, K. Peters, H.Y.S. Huang, *Strain state dependent anisotropic viscoelasticity of tendon-to-bone insertion*, Math. Biosci. **308** (2019), 1–7.

[47] A. Leonov, *Nonequilibrium thermodynamics and rheology of viscoelastic polymer media*, Rheol. Acta **15** (1976), 85–98.

[48] J. Leray, *Sur le mouvement d'un liquide visqueux emplissant l'espace*, Acta Math. **63** (1934), no. 1, 193–248.

[49] P.L. Lions, N. Masmoudi, *Global solutions for some oldroyd models of non-newtonian flows*, Chin. Ann. Math. Ser. B **21** (2000), no. 2, 131–146.

[50] M. Lukáčová-Medviďová, H. Mizerová, Š. Nečasová, M. Renardy, *Global existence result for the generalized Peterlin viscoelastic model*, SIAM J. Math. Anal. **49** (2017), no. 4, 2950–2964.

[51] J.R. Magnus, H. Neudecker, *Matrix differential calculus with applications in statistics and econometrics*, Wiley Series in Probability and Statistics, John Wiley & Sons, Ltd., Chichester, 1999, Revised reprint of the 1988 original.

[52] J. Málek, J. Nečas, M. Rokyta, M. Růžička, *Weak and Measure-valued Solutions to Evolutionary PDEs*, Chapman & Hall, 1996.

[53] J. Málek, V. Průša, T. Skřivan, E. Süli, *Thermodynamics of viscoelastic rate-type fluids with stress diffusion*, Physics of Fluids **30** (2018).

[54] J. Málek, K.R. Rajagopal, *Mathematical issues concerning the Navier-Stokes equations and some of its generalizations*, Evolutionary equations. Vol. II, Handb. Differ. Equ., Elsevier/North-Holland, Amsterdam, 2005, pp. 371–459.

[55] J. Málek, K.R. Rajagopal, K. Tůma, *On a variant of the Maxwell and Oldroyd-B models within the context of a thermodynamic basis*, Internat. J. Non-Linear Mech. **76** (2015), 42 – 47.

[56] ______, *Derivation of the variants of the Burgers model using a thermodynamic approach and appealing to the concept of evolving natural configurations*, Fluids **3** (2018), no. 4.

[57] J. Málek, V. Průša, *Derivation of equations for continuum mechanics and thermodynamics of fluids*, Handbook of mathematical analysis in mechanics of viscous fluids, Springer, Cham, 2018, pp. 3–72.

[58] E. Maringová, J. Žabenský, *On a Navier-Stokes-Fourier-like system capturing transitions between viscous and inviscid fluid regimes and between no-slip and perfect-slip boundary conditions*, Nonlinear Anal. Real World Appl. **41** (2018), 152–178.

[59] N. Masmoudi, *Global existence of weak solutions to macroscopic models of polymeric flows*, J. Math. Pures Appl. (9) **96** (2011), no. 5, 502–520.

[60] 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.

[61] J.G. Oldroyd, *On the formulation of rheological equations of state*, Proc. Roy. Soc. London Ser. A **200** (1950), 523–541.

[62] K.R. Rajagopal, *Remarks on the notion of "pressure"*, Internat. J. Non-Linear Mech. **71** (2015), 165 – 172.

[63] K.R. Rajagopal, A.R. Srinivasa, *A thermodynamic frame work for rate type fluid models*, J. Non-Newton. Fluid Mech. **88** (2000), no. 3, 207–227.

[64] ______, *On thermomechanical restrictions of continua*, Proc. R. Soc. Lond., Ser. A, Math. Phys. Eng. Sci. **460** (2004), no. 2042, 631–651.

[65] I.J. Rao, K.R. Rajagopal, *A thermodynamic framework for the study of crystallization in polymers*, Z. Angew. Math. Phys. **53** (2002), no. 3, 365–406.

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

[67] W. Rudin, *Real and complex analysis*, third ed., McGraw-Hill Book Co., New York, 1987.

[68] M. Řehoř, V. Průša, K. Tůma, *On the response of nonlinear viscoelastic materials in creep and stress relaxation experiments in the lubricated squeeze flow setting*, Physics of Fluids **28** (2016), no. 10, 103102.

[69] C. Truesdell, W. Noll, *The non-linear field theories of mechanics*, third ed., Springer-Verlag, Berlin, 2004, Edited and with a preface by Stuart S. Antman.

[70] C. Truesdell, K.R. Rajagopal, *An introduction to the mechanics of fluids*, Modeling and Simulation in Science, Engineering and Technology, Birkhäuser Boston, Inc., Boston, MA, 2000.

[71] P. Wapperom, M. Hulsen, *Thermodynamics of viscoelastic fluids: the temperature equation*, J. Rheol. **42** (1998), no. 5, 999–1019 (English).

[72] V.I. Yudovich, *Eleven great problems of mathematical hydrodynamics*, Mosc. Math. J. **3** (2003), no. 2, 711–737, 746, Dedicated to Vladimir I. Arnold on the occasion of his 65th birthday.

[73] S. Zaremba, *Sur une forme perfectionee de la theorie de la relaxation*, Bull. Int. Acad. Sci. Cracovie (1903), 594–614.

[74] 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.

[75] W.P. Ziemer, *Weakly differentiable functions*, Graduate Texts in Mathematics, vol. 120, Springer-Verlag, New York, 1989, Sobolev spaces and functions of bounded variation.

List of author's publications

M. Bathory, *Joint weak type interpolation on Lorentz-Karamata spaces*, Math. Inequal. Appl. **21** (2018), no. 2, 385–419.

M. Bathory, *Outflow Boundary Condition Leading to Minimal Energy Dissipation for an Incompressible Flow*, WDS'17 Proceedings of Contributed Papers — Physics (Prague) (J. Šafránková, J. Pavlů, eds.), Prague, Matfyzpress, 2017, pp. 7–12.

M. Bathory, M. Bulíček, *Optimal outflow boundary condition for a stationary flow of an incompressible fluid*, Preprint NCMM/2018/11, 2018.

M. Bathory, M. Bulíček, O. Souček, *Existence and qualitative theory for nonlinear elliptic systems with a nonlinear interface condition used in electrochemistry*, Z. Angew. Math. Phys **71** (2020), 74.

M. Bathory, M. Bulíček, J. Málek, *Large data existence theory for three-dimensional unsteady flows of rate-type viscoelastic fluids with stress diffusion*, Submitted, preprint at `https://arxiv.org/abs/2002.11224v2`, 2020.