ENERGY CONSISTENT DISCONTINUOUS GALERKIN METHODS FOR THE NAVIER–STOKES–KORTEWEG SYSTEM

JAN GIESSELMANN, CHARALAMBOS MAKRIDAKIS, AND TRISTAN PRYER

ABSTRACT. We design consistent discontinuous Galerkin finite element schemes for the approximation of the Euler–Korteweg and the Navier–Stokes–Korteweg systems. We show that the scheme for the Euler–Korteweg system is energy and mass conservative and that the scheme for the Navier–Stokes–Korteweg system is mass conservative and monotonically energy dissipative. In this case the dissipation is isolated to viscous effects, that is, there is no numerical dissipation. In this sense the methods are consistent with the energy dissipation of the continuous PDE systems.

1. INTRODUCTION

In this work we propose a new class of finite element methods for the Navier–Stokes–Korteweg system which are *by design consistent with the energy dissipation structure of the problem.* The methods are of arbitrary high order of accuracy and provide physically relevant approximations free of numerical artifacts. It seems that these are the first methods in the literature to enjoy these properties.

Liquid vapour flow occurs in many technical applications and natural phenomena. A particularly interesting and challenging case is when the fluid undergoes phase transition, i.e., there is mass transfer between the phases, which is driven by thermodynamics. The applications of these phenomena are extremely varied, for example, it is applicable to modelling the fuel injection system in modern car engines and also to the study of cloud formation. The modelling of these phenomena can be traced back to [vdW, Kor01], however, there remain open questions, for example, what is the correct model for the given application at hand.

The compressible flow of a single substance containing both a liquid and vapour phase undergoing a phase transition can be modelled by different techniques. One widely used approach for the treatment of these problems, which emerged in the last few decades (see [AMW98] and references therein) is the so-called *diffuse interface approach.* In this philosophy the phases are separated by a (thin) interfacial layer across which the fields vary smoothly. The benefit of this approach is that there is only one set of PDEs solved on the whole domain whose solution already includes the position of the interfacial layer. However, these models must include

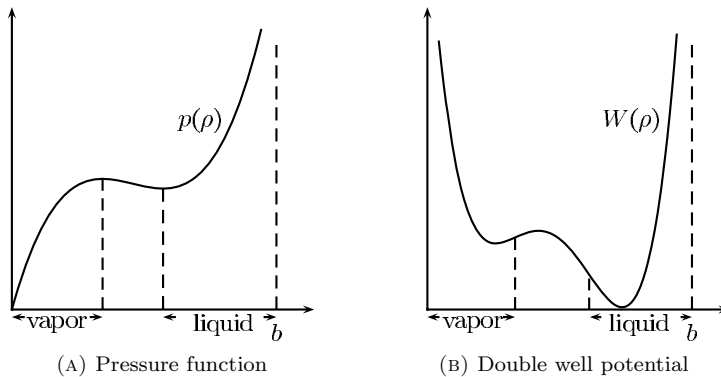(A) Pressure function   (B) Double well potential

FIGURE 1. The relation between the pressure function and the double well potential.

a parameter distinguishing when we are in one phase or another. In most diffuse interface models this is a more or less arbitrary indicator function based on the mass or volume fraction of one of the constituents.

In this contribution we will consider the isothermal Navier–Stokes–Korteweg system which is a diffuse interface model, but here the mass density serves as a phase indicator, it originates in the work of Korteweg [Kor01] and van der Walls [vdW] and was derived in modern terminology in [DS85, TN92, JLCD01]. This model includes surface tension effects by a third order term in the momentum balance which corresponds to a non-local (gradient) contribution in the energy functional. Another feature of compressible diffuse interface models is a non-monotone constitutive relation for the pressure. This corresponds to a non-convex local part of the energy; see equation (2.7). As can be seen from Figure 1 the phases of the problem (liquid/vapour) are the corresponding regions where the pressure function is monotonically increasing.

The Korteweg type third order term together with the non-monotonicity of the pressure function cause several issues in the numerical treatment of this problem. In previous numerical studies [JTB02, Die07, BP] it has been observed that "classical" explicit-in-time finite volume (FV) and discontinuous Galerkin (DG) schemes which use standard fluxes used in the computational conservation laws introduce several numerical artifacts.

The first artifact is non-monotonicity of the energy. The Euler–Korteweg model is energy conservative over time whereas the Navier–Stokes–Korteweg model is monotonically energy dissipative. In the Navier–Stokes–Korteweg model all the dissipation is due to viscous effects (see Lemma 2.3). The classical FV and DG methods applied to the Navier–Stokes–Korteweg system lead to a non-monotone behaviour of the energy. This is mainly due to the fact that these "classical" schemes introduce standard diffusion in the mass conservation equation as a stabilising mechanism. While for convex energies standard diffusion, in fact, leads to energy dissipation, it may lead to an increase in energy for multiphase flow [Die07, DGR]. In fact, standard diffusion is also present in the finite element method proposed in [BP].

The second artifact is so-called *parasitic currents*, i.e., the schemes are not well balanced, as they do not preserve the correct equilibria. Parasitic currents occur when equilibrium is approached and the numerical velocity field does not vanish uniformly, but in the interfacial layer large velocities whose magnitude is dependent on the gridsize and inversely dependent on the width of interfacial layer appear [Die07, §5]. As the interfacial layer is extremely thin this effect cannot be neglected in practical computations.

Both the non-monotone behaviour of the Navier–Stokes–Korteweg energy and the parasitic currents are due to numerical regularisation terms which are not adapted to the variational structure of the problem; see [DGR] for a study on regularisation terms taking into account the underlying variational structure of the problem. For previous works on scalar dispersive equations by discontinuous Galerkin methods we refer to [CS08, BCKX11, XS11].

The key target in the work at hand is to consider a high-order DG discretisation of the problem which aims at preserving the energy dissipation inequality satisfied by the original problem and avoiding the introduction of any artificial diffusion terms. By achieving this goal we can treat the case where the system preserves the energy exactly. In addition, we can address the case where the system has natural dissipation and the energy is diminishing. Our schemes are therefore *energy consistent* in the sense that they are consistent with the energy dissipation structure of the Navier–Stokes–Korteweg system. The resulting schemes are free from the above mentioned artifacts of other approximating methods in the literature and are successful in computing the physically relevant solution. It is to be noted that our approach does not hinge on an adaptation of "entropy conservative schemes" developed for conservation laws, [Tad03]. The non-monotone pressure function makes a direct application of this approach unfeasible in our case. Conservative DG schemes for the scalar generalized KdV equation were suggested recently in [BCKX11]. To achieve our goals we follow a constructive step-by-step approach. Motivated by the proof of energy conservation at the continuous level we introduce a new mixed formulation for the Navier–Stokes–Korteweg system. This mixed formulation will be the basis of our discrete schemes. Discretisation in space is achieved by employing a DG approach with generic discrete fluxes. Then we specifically identify the properties and thus the fluxes which yield *energy consistent* schemes.

We then consider a Crank–Nicolson type time discretisation and identify a precise time discrete method which is energy consistent. By combining the ideas of space and time discretization we obtain the fully discrete schemes with the desired properties.

We also note that higher order temporal discretisations are possible by appropriately modifying Gauss–Legendre Runge–Kutta methods, for example. Although the framework presented herein can be used in deriving such methods, we shall not insist on this issue.

As a further application we consider the case when the Euler–Korteweg system has discontinuous initial conditions. In this case an energy consistent method may be inappropriate due to the propagation of oscillations. We give an example of a choice of fluxes which will naturally dissipate a small amount of energy.

In this work we will not present an error analysis for the scheme we derive. To the authors knowledge there is no error analysis for discontinuous Galerkin schemes for such dissipative-dispersive systems in multidimensions; for one-dimensional, scalar

results see [CS08, BCKX11]. The derivation of such an error analysis is, thus, an open problem. We do, however, numerically benchmark the approximation against a known solution of the Euler–Korteweg system, showing optimal convergence in $L_\infty (0, T; L_2(\Omega))$ for the approximations of both the density and velocity.

The structure of the paper is as follows: In §2 we introduce the Navier–Stokes–Korteweg model problem as well as some of its conservative properties. We give the mixed formulation and necessary notation which will be used throughout the paper. In addition we describe the mixed formulation in the *broken Sobolev* framework necessary for the construction of the DG scheme. In §3 we detail the construction of the energy consistent DG scheme initially in the spatially semidiscrete case. We then move on to the temporal semidiscrete case in §4 and combine the results to obtain an energy consistent fully discrete scheme in §5. In §6 we perform various numerical experiments to test the convergence, conservativity and computational properties of the scheme.

2. Model problem, mixed formulation and discretisation

In this section we formulate the model problem, fix notation and give some basic assumptions. Let $\Omega \subset \mathbb{R}^d$, with $d = 1, 2, 3$ be a bounded domain. We then begin by introducing the Sobolev spaces [Cia78, Eva98]

$$(2.1) \qquad H^k(\Omega) := \left\{ \phi \in L_2(\Omega) : \ D^{\boldsymbol{\alpha}} \phi \in L_2(\Omega), \ \text{for } |\boldsymbol{\alpha}| \le k \right\},$$

which are equipped with norms and semi-norms

$$(2.2) \qquad \|u\|_k^2 := \|u\|_{H^k(\Omega)}^2 = \sum_{|\boldsymbol{\alpha}| \le k} \|D^{\boldsymbol{\alpha}} u\|_{L_2(\Omega)}^2 ,$$

$$(2.3) \qquad |u|_k^2 := |u|_{H^k(\Omega)}^2 = \sum_{|\boldsymbol{\alpha}| = k} \|D^{\boldsymbol{\alpha}} u\|_{L_2(\Omega)}^2 ,$$

respectively, where $\boldsymbol{\alpha} = \{\alpha_1, ..., \alpha_d\}$ is a multi-index, $|\boldsymbol{\alpha}| = \sum_{i=1}^d \alpha_i$ and derivatives $D^{\boldsymbol{\alpha}}$ are understood in a weak sense. In addition, let
$$(2.4)$$
$$H_0^1 := \left\{ \phi \in H^1(\Omega) : \ \phi|_{\partial\Omega} = 0 \right\} \ \text{and} \ H_{\boldsymbol{n}}^1(\Omega) := \left\{ \boldsymbol{\phi} \in \left[H^1(\Omega) \right]^d : \ (\boldsymbol{\phi}|_{\partial\Omega})^{\mathsf{T}} \boldsymbol{n} = 0 \right\},$$

where $\boldsymbol{n}$ denotes the outward pointing normal to $\partial\Omega$.

We use the convention that for a multivariate function, u, the quantity ∇u is a column vector consisting of first-order partial derivatives with respect to the spatial coordinates. The divergence operator, div, acts on a vector-valued multivariate function and $\Delta u := \operatorname{div}(\nabla u)$ is the generalised Laplacian operator. We also note that when the Laplacian acts on a vector-valued multivariate function, it is meant componentwise. Moreover, for a vector field $\boldsymbol{v}$, we denote its Jacobian by $D\boldsymbol{v}$. We also make use of the following notation for time dependant Sobolev spaces:

$$(2.5) \qquad L_2(0, T; H^k(\Omega)) := \left\{ u : [0, T] \to H^k(\Omega) : \ \int_0^T \|u(t)\|_k^2 \ \mathrm{d}t < \infty \right\} .$$

2.1. **Model problem.** Consider a fluid in the domain Ω with density ρ and velocity $\boldsymbol{v}$. The Navier–Stokes–Korteweg system is made up of the balances of mass and momentum of said fluid, that is,

$$(2.6) \qquad \begin{aligned} \partial_t \rho + \operatorname{div}(\rho \boldsymbol{v}) &= 0, \\ \partial_t(\rho \boldsymbol{v}) + \operatorname{div}(\rho \boldsymbol{v} \otimes \boldsymbol{v}) + \nabla p(\rho) &= \mu \Delta \boldsymbol{v} + \gamma \rho \nabla \Delta \rho, \end{aligned} \qquad \text{in} \quad \Omega \times (0, T),$$

where p is a non–monotone pressure function, μ is a viscosity coefficient and γ a capillarity coefficient. The pressure function p is linked to a double well-potential $W = W(\rho)$ via the relation (Figure 1)

$$(2.7) \qquad p(\rho) = \rho W'(\rho) - W(\rho).$$

Let $\boldsymbol{n}$ be the outward pointing normal to $\partial\Omega$, suppose the system (2.6) is given with boundary conditions

$$(2.8) \qquad \boldsymbol{v} = \boldsymbol{0} \text{ and } (\nabla\rho)^{\mathsf{T}}\boldsymbol{n} = 0 \text{ on } \partial\Omega \times (0, T)$$

and initial conditions

$$(2.9) \qquad \rho(\cdot, 0) = \rho^0, \quad \boldsymbol{v}(\cdot, 0) = \boldsymbol{v}^0 \text{ in } \Omega$$

for given functions $\rho^0 \in \mathrm{H}^1(\Omega)$ and $\boldsymbol{v}^0 \in (\mathrm{H}^1(\Omega))^d$ such that $W(\rho^0) \in \mathrm{L}_1(\Omega)$. The system (2.6) conserves mass as well as satisfying a momentum balance together with an energy dissipation equality, i.e.,

$$(2.10) \qquad \mathrm{d}_t\left(\int_\Omega \rho \, \mathrm{d}\boldsymbol{x}\right) = 0,$$

$$(2.11) \qquad \mathrm{d}_t\left(\int_\Omega \rho\boldsymbol{v} \, \mathrm{d}\boldsymbol{x}\right) = \int_{\partial\Omega} \left(-p(\rho) + \gamma\rho\Delta\rho + \mu(\mathrm{D}\boldsymbol{v})\right)\boldsymbol{n} \, \mathrm{d}s,$$

$$(2.12) \qquad \mathrm{d}_t\left(\int_\Omega W(\rho) + \frac{1}{2}\rho\,|\boldsymbol{v}|^2 + \frac{\gamma}{2}\,|\nabla\rho|^2 \, \mathrm{d}\boldsymbol{x}\right) = -\mu\int_\Omega |\mathrm{D}\boldsymbol{v}|^2 \, \mathrm{d}\boldsymbol{x},$$

respectively. The energy dissipation equality is only valid for smooth solutions. In case the system permits shocks, they would trigger additional energy dissipation and (2.12) would have to be replaced by an inequality. While the first two equalities follow by integrating the mass and momentum balance (2.6). The derivation of the energy dissipation equality is a little bit more involved. For completeness the result is formulated as Lemma 2.3. Moreover, the proof of Lemma 2.3 serves as a guideline for the construction of energy consistent discrete schemes.

2.2. **Assumption** (finite Helmholtz energy). *From here on we will assume that for a given ρ we have that $W(\rho) \in \mathrm{L}_1(0, T; \mathrm{L}_1(\Omega))$.*

2.3. **Lemma** (energy dissipation equality). *For every smooth solution $(\rho, \boldsymbol{v}) \in \mathrm{L}_2(0, T; \mathrm{H}^3(\Omega)) \times \mathrm{L}_2(0, T; \mathrm{H}^2(\Omega))^d$ of (2.6) such that $(\partial_t\rho, \partial_t\boldsymbol{v}) \in \mathrm{L}_2(0, T; \mathrm{L}_2(\Omega)) \times \mathrm{L}_2(0, T; \mathrm{L}_2(\Omega))^d$ which satisfies the boundary conditions (2.8) we have*

$$(2.13) \qquad \mathrm{d}_t\left(\int_\Omega W(\rho) + \frac{1}{2}\rho\,|\boldsymbol{v}|^2 + \frac{\gamma}{2}\,|\nabla\rho|^2 \, \mathrm{d}\boldsymbol{x}\right) = -\mu\int_\Omega |\mathrm{D}\boldsymbol{v}|^2 \, \mathrm{d}\boldsymbol{x}.$$

Proof. Let us first note that the second equation of (2.6) can be reformulated as

$$(2.14) \qquad \rho\partial_t\boldsymbol{v} + \mathrm{div}(\rho\boldsymbol{v} \otimes \boldsymbol{v}) - \mathrm{div}(\rho\boldsymbol{v})\boldsymbol{v} + \rho\nabla W'(\rho) - \mu\Delta\boldsymbol{v} - \gamma\rho\nabla\Delta\rho = 0.$$

Multiplying the first equation of (2.6) by $W'(\rho) + \frac{1}{2}\,|\boldsymbol{v}|^2 - \gamma\Delta\rho$ we see that

$$(2.15) \qquad \begin{aligned} 0 = {}& W'(\rho)\partial_t\rho + W'(\rho)\,\mathrm{div}\,(\rho\boldsymbol{v}) + \frac{1}{2}\,|\boldsymbol{v}|^2\,\partial_t\rho + \frac{1}{2}\,|\boldsymbol{v}|^2\,\mathrm{div}\,(\rho\boldsymbol{v}) \\ & - \gamma\Delta\rho\partial_t\rho - \gamma\Delta\rho\,\mathrm{div}\,(\rho\boldsymbol{v})\,. \end{aligned}$$

Then by multiplying (2.14) by $\boldsymbol{v}$ and summing together with (2.15) we obtain

$$
\begin{aligned}
(2.16) \quad 0 = {}& W'(\rho)\partial_t\rho + \frac{1}{2}\partial_t\rho\,|\boldsymbol{v}|^2 - \gamma\Delta\rho\partial_t\rho + W'(\rho)\operatorname{div}(\rho\boldsymbol{v}) - \frac{1}{2}\operatorname{div}(\rho\boldsymbol{v})\,|\boldsymbol{v}|^2 \\
& - \gamma\operatorname{div}(\rho\boldsymbol{v})\Delta\rho + \rho\boldsymbol{v}^\mathsf{T}\partial_t\boldsymbol{v} + \boldsymbol{v}^\mathsf{T}\operatorname{div}(\rho\boldsymbol{v}\otimes\boldsymbol{v}) + \rho\boldsymbol{v}^\mathsf{T}\nabla W'(\rho) \\
& - \mu\boldsymbol{v}^\mathsf{T}\Delta\boldsymbol{v} - \gamma\rho\boldsymbol{v}^\mathsf{T}\nabla\Delta\rho.
\end{aligned}
$$

We integrate (2.16) over Ω and by Green's formula we have

$$
\begin{aligned}
(2.17) \quad 0 = {}& \int_\Omega W'(\rho)\partial_t\rho + \frac{1}{2}\partial_t\rho\,|\boldsymbol{v}|^2 + \rho\boldsymbol{v}^\mathsf{T}\partial_t\boldsymbol{v} + \frac{\gamma}{2}\partial_t\,|\nabla\rho|^2 + \mu\,|\mathrm{D}\boldsymbol{v}|^2\,\mathrm{d}\boldsymbol{x} \\
& + \int_{\partial\Omega}\left(\left(-\gamma\partial_t\rho\nabla\rho + \frac{1}{2}\rho\boldsymbol{v}\,|\boldsymbol{v}|^2 - \gamma\rho\boldsymbol{v}\Delta\rho + \rho\boldsymbol{v}W'(\rho)\right)^\mathsf{T} - \mu\boldsymbol{v}^\mathsf{T}\mathrm{D}\boldsymbol{v}\right)\boldsymbol{n}\,\mathrm{d}s.
\end{aligned}
$$

The boundary integral in (2.17) vanishes because of the boundary conditions. $\qquad\square$

2.4. *Remark* (stable steady states). The energy dissipation equality gives rise to the fact that the (stable) steady states of (2.6) are minimizers of the energy functional

$$
(2.18) \qquad E[\rho, \boldsymbol{v}] := \int_\Omega W(\rho) + \frac{1}{2}\rho\,|\boldsymbol{v}|^2 + \frac{\gamma}{2}\,|\nabla\rho|^2\,\mathrm{d}\boldsymbol{x},
$$

under the constraint

$$
(2.19) \qquad \int_\Omega \rho\,\mathrm{d}\boldsymbol{x} = m,
$$

for some given $m > 0$ and therefore satisfy the Euler–Lagrange equations

$$
(2.20) \qquad \boldsymbol{v} = \boldsymbol{0},
$$

$$
(2.21) \qquad W'(\rho) - \gamma\Delta\rho = \lambda,
$$

where λ is the Lagrange multiplier associated with the mass conservation contraint (2.19).

Note that (2.21) is equivalent to

$$
\begin{aligned}
(2.22) \quad \boldsymbol{0} &= \nabla\left(W'(\rho) - \gamma\Delta\rho\right) \\
&= \nabla p(\rho) - \gamma\rho\nabla\Delta\rho
\end{aligned}
$$

using the relation

$$
(2.23) \qquad \nabla p(\rho) = \rho\nabla W'(\rho),
$$

which is readily derived from (2.7).

2.5. **Classical solvability of the problem.** The well–posedness of the Navier–Stokes–Korteweg system and similar systems was considered by several authors [BGDDJ07, BDL03, DD01, Fei02, HL96, Kot08]. For completeness we will state some results.

2.6. **Theorem** (existence of a solution to the Euler–Korteweg system [BGDDJ07])**.** *Let $s > \frac{d}{2} + 1$ and*

$$
(2.24) \qquad H_s := H^{s+1}(\mathbb{R}^d) \times H^s(\mathbb{R}^d, \mathbb{R}^d).
$$

Suppose the initial data $(\rho_0, \boldsymbol{v}_0) \in (\underline{\rho}(0), \underline{\boldsymbol{v}}(0)) + H_s$ where $\underline{\rho}, \underline{\boldsymbol{v}}$ is a special solution such that $\underline{\rho}$ is bounded away from zero and the Hessian of $\underline{\rho}$, $\mathrm{D}\nabla\underline{\rho}$, as well as the Jacobian of $\underline{\boldsymbol{v}}$, $\mathrm{D}\underline{\boldsymbol{v}}$ are both $C([0,T], H^{s+3}(\mathbb{R}^d, \mathbb{R}^{d\times d}))$ for some $T > 0$. Then the Euler–Korteweg system admits a unique solution $(\rho, \boldsymbol{v}) \in (\underline{\rho}, \underline{\boldsymbol{v}}) + C^1([0,T], H_{s-2}) \cap C([0,T], H_s)$ satisfying the initial data $(\rho_0, \boldsymbol{v}_0)$.

2.7. Theorem (existence of a solution to the Navier–Stokes–Korteweg system [DD01])**.** *Let $B^s = B^s_{2,1}(\mathbb{R}^d)$ denote the homogeneous Besov space. Let $\bar{\rho} > 0$ be a reference density such that $p'(\bar{\rho}) > 0$. Suppose also that the initial data $\rho_0, \boldsymbol{v}_0$ satisfies $\rho_0 - \bar{\rho} \in B^{d/2}$, $\rho_0 \geq c > 0$ and $\boldsymbol{v}_0 \in (B^{d/2-1})^d$.*

Then there exists a $T > 0$ such that the Navier–Stokes–Korteweg system has a unique solution $(\rho, \boldsymbol{v})$ with initial data $\rho_0, \boldsymbol{v}_0$ such that $\rho - \bar{\rho} \in C([0,T), B^{d/2}) \cap L^1([0,T), B^{d/2+2})$ and $\boldsymbol{v} \in C([0,T), (B^{d/2-1}))^d \cap L^1([0,T), (B^{d/2+1}))^d$.

2.8. Remark. Theorems 2.6 and 2.7 motivate us to construct numerical schemes which are adapted to the smooth situation. In particular, enforcing the energy dissipation equality proven in Lemma 2.3.

2.9. Mixed formulation. To mimic the proof of Lemma 2.3 at the discrete level it will be essential to have at our disposal a numerical formulation in which the velocity, $\boldsymbol{v}$, and the variation of the energy with respect to ρ, which depend non-linearly on the original variables ρ and $\rho \boldsymbol{v}$, are permitted as test functions. Indeed, this is our main motivation to reformulate the Navier–Stokes–Korteweg system (2.6) as a mixed system of PDEs by the introduction of two auxilliary variables, τ and $\boldsymbol{q}$, and using the relation of the pressure function and the double well potential (2.23).

The mixed formulation is then to seek $(\rho, \boldsymbol{v}, \tau, \boldsymbol{q})$ such that

$$(2.25) \qquad \partial_t \rho + \operatorname{div}(\rho \boldsymbol{v}) = 0,$$

$$(2.26) \qquad \rho \partial_t \boldsymbol{v} + \operatorname{div}(\rho \boldsymbol{v} \otimes \boldsymbol{v}) - \operatorname{div}(\rho \boldsymbol{v})\, \boldsymbol{v} + \rho \nabla \tau - \frac{1}{2}\rho \nabla |\boldsymbol{v}|^2 - \mu \Delta \boldsymbol{v} = 0,$$

$$(2.27) \qquad \tau - W'(\rho) + \gamma \operatorname{div}(\boldsymbol{q}) - \frac{1}{2}|\boldsymbol{v}|^2 = 0,$$

$$(2.28) \qquad \boldsymbol{q} - \nabla \rho = 0,$$

which is coupled with the boundary conditions

$$(2.29) \qquad \boldsymbol{v} = \boldsymbol{0} \text{ and } \boldsymbol{q}^\mathsf{T} \boldsymbol{n} = 0 \text{ on } \partial\Omega \times (0,T)$$

and the initial conditions (2.9).

2.10. Remark (non-conservative momentum formulation). Notice that the formulation (2.26) is not written in conservative form for the momentum. In view of this fact it becomes challenging to design discretisations which conserve the momentum, as well. This is discussed in detail in Proposition 3.11.

Our use of a non-conservative discretisation is due to the following: In the derivation of the energy dissipation Lemma 2.3 we tacitly used (2.23). This requires the chain rule, which is clearly not available to us at the discrete level unless we use a non-conservative discretisation of the pressure gradient. It is well known for hyperbolic conservation laws that non-conservative schemes may converge to an incorrect solution in the presence of shocks [HL94]. Due to the energy dissipation equality given in Lemma 2.3, given smooth initial data, solutions should have sufficient regularity, specifically $\rho \in L_\infty([0,T], H^1(\Omega))$, such that shocks do not form. Indeed, in the numerical experiments given in §6 we observe that initial discontinuities are immediately smoothed out in the case $\mu > 0$. The case $\mu = 0$ and non-smooth initial data will be discussed in the sequel (see Remark 3.9).

2.11. *Remark* (alternate notation). We note that the second and third term on the left-hand side of (2.26) can be rewritten as

$$\operatorname{div}(\rho\boldsymbol{v}\otimes\boldsymbol{v}) - \operatorname{div}(\rho\boldsymbol{v})\,\boldsymbol{v} = \rho\left(\boldsymbol{v}^{\mathsf{T}}\nabla\right)\boldsymbol{v}, \tag{2.30}$$

which is the standard notation in the incompressible scenario.

2.12. Discretisation. Let $\mathscr{T}$ be a conforming, shape regular triangulation of Ω, namely, $\mathscr{T}$ is a finite family of sets such that

  (1) $K \in \mathscr{T}$ implies K is an open simplex (segment for $d = 1$, triangle for $d = 2$, tetrahedron for $d = 3$),
  (2) for any $K, J \in \mathscr{T}$ we have that $\overline{K} \cap \overline{J}$ is a full subsimplex (i.e., it is either $\emptyset$, a vertex, an edge, a face, or the whole of $\overline{K}$ and $\overline{J}$) of both $\overline{K}$ and $\overline{J}$, and
  (3) $\bigcup_{K\in\mathscr{T}} \overline{K} = \overline{\Omega}$.

We use the convention where $h : \Omega \to \mathbb{R}$ denotes the *meshsize function* of $\mathscr{T}$, i.e.,

$$h(\boldsymbol{x}) := \max_{\overline{K}\ni\boldsymbol{x}} h_K, \tag{2.31}$$

where h_K is the diameter of an element K. We let $\mathscr{E}$ be the skeleton (set of common interfaces) of the triangulation $\mathscr{T}$ and say $e \in \mathscr{E}$ if e is on the interior of Ω and $e \in \partial\Omega$ if e lies on the boundary $\partial\Omega$.

2.13. Definition (broken Sobolev spaces, trace spaces). We introduce the broken Sobolev space

$$\mathrm{H}^k(\mathscr{T}) := \left\{ \phi : \ \phi|_K \in \mathrm{H}^k(K), \text{ for each } K \in \mathscr{T} \right\}, \tag{2.32}$$

similarly for $\mathrm{H}_0^1(\mathscr{T})$ and $\mathrm{H}_{\boldsymbol{n}}^1(\mathscr{T})$.

We also make use of functions defined in these broken spaces restricted to the skeleton of the triagulation. This requires an appropriate trace space:

$$\mathcal{T}(\mathscr{E}) := \prod_{K\in\mathscr{T}} \mathrm{L}_2(\partial K) \subset \prod_{K\in\mathscr{T}} \mathrm{H}^{\frac{1}{2}}(K). \tag{2.33}$$

Let $\mathbb{P}^p(\mathscr{T})$ denote the space of piecewise polynomials of degree p over the triangulation $\mathscr{T}$. We then introduce the *finite element spaces*,

$$\mathbb{V} := \mathbb{DG}(\mathscr{T}, p) = \mathbb{P}^p(\mathscr{T}), \tag{2.34}$$

$$\mathring{\mathbb{V}} := \mathbb{V} \cap \mathrm{H}_0^1(\mathscr{T}), \tag{2.35}$$

$$\overset{\boldsymbol{n}}{\mathbb{V}} := \mathbb{V}^d \cap \mathrm{H}_{\boldsymbol{n}}^1(\mathscr{T}), \tag{2.36}$$

to be the usual spaces of (discontinuous) piecewise polynomial functions. For simplicity we will assume that $\mathbb{V}$ is constant in time.

2.14. Definition (jumps and averages). We may define average and jump operators over $\mathcal{T}(\mathscr{E})$ for arbitrary scalar, $v \in \mathcal{T}(\mathscr{E})$, and vector-valued functions, $\boldsymbol{v} \in \mathcal{T}(\mathscr{E})^d$.

$$\begin{aligned} \{\!\!\{\,\cdot\,\}\!\!\} : \quad \mathcal{T}(\mathscr{E}) \ &\to\ \mathrm{L}_2(\mathscr{E}), \\ v \ &\mapsto\ \tfrac{1}{2}\left(v|_{K_1} + v|_{K_2}\right), \end{aligned} \tag{2.37}$$

$$\begin{aligned} \{\!\!\{\,\cdot\,\}\!\!\} : \quad (\mathcal{T}(\mathscr{E}))^d \ &\to\ (\mathrm{L}_2(\mathscr{E}))^d, \\ \boldsymbol{v} \ &\mapsto\ \tfrac{1}{2}\left(\boldsymbol{v}|_{K_1} + \boldsymbol{v}|_{K_2}\right), \end{aligned} \tag{2.38}$$

$$(2.39) \qquad \begin{aligned} \llbracket \cdot \rrbracket : \quad (\mathcal{T}(\mathscr{E})) \quad &\rightarrow \quad (\mathrm{L}_2(\mathscr{E}))^d, \\ v \quad &\mapsto \quad v|_{K_1} \boldsymbol{n}_{K_1} + v|_{K_2} \boldsymbol{n}_{K_2}, \end{aligned}$$

$$(2.40) \qquad \begin{aligned} \llbracket \cdot \rrbracket : \quad (\mathcal{T}(\mathscr{E}))^d \quad &\rightarrow \quad (\mathrm{L}_2(\mathscr{E})), \\ \boldsymbol{v} \quad &\mapsto \quad (\boldsymbol{v}|_{K_1})^{\mathsf{T}} \boldsymbol{n}_{K_1} + (\boldsymbol{v}|_{K_2})^{\mathsf{T}} \boldsymbol{n}_{K_2}, \end{aligned}$$

$$(2.41) \qquad \begin{aligned} \llbracket \cdot \rrbracket_{\otimes} : \quad (\mathcal{T}(\mathscr{E} \cup \partial \Omega))^d \quad &\rightarrow \quad (\mathrm{L}_2(\mathscr{E}))^{d \times d}, \\ \boldsymbol{v} \quad &\mapsto \quad \boldsymbol{v}|_{K_1} \otimes \boldsymbol{n}_{K_1} + \boldsymbol{v}|_{K_2} \otimes \boldsymbol{n}_{K_2}, \end{aligned}$$

where $\boldsymbol{n}_{K_i}$ denotes the outward pointing normal to K_i. Note that on the boundary of the domain $\partial \Omega$ the jump and average operators are defined as

$$(2.42) \qquad \llbracket v \rrbracket \Big|_{\partial \Omega} := v\boldsymbol{n} \qquad \llbracket \boldsymbol{v} \rrbracket \Big|_{\partial \Omega} := \boldsymbol{v}^{\mathsf{T}} \boldsymbol{n},$$

$$(2.43) \qquad \{\!\!\{ v \}\!\!\} \Big|_{\partial \Omega} := v \qquad \{\!\!\{ \boldsymbol{v} \}\!\!\} \Big|_{\partial \Omega} := \boldsymbol{v}.$$

2.15. Elementwise formulation and discrete fluxes.

As a next step towards the construction of a numerical scheme we give the elementwise variational formulation to the problem in mixed form (2.25)–(2.28). It is required to find $(\rho, \boldsymbol{v}) \in \mathrm{L}_2(0, T; \mathrm{H}^1(\mathscr{T})) \times \left(\mathrm{L}_2(0, T; \mathrm{H}_0^1(\mathscr{T}))\right)^d$ with $(\partial_t \rho, \partial_t \boldsymbol{v}) \in \mathrm{L}_2(0, T; \mathrm{L}_2(\mathscr{T})) \times (L_2(0, T; L_2(\mathscr{T}))^d$ and $(\tau, \boldsymbol{q}) \in \mathrm{L}_2(0, T; \mathrm{H}^1(\mathscr{T})) \times \mathrm{L}_2(0, T; \mathrm{H}_{\boldsymbol{n}}^1(\mathscr{T}))$ such that $W'(\rho) \in \mathrm{L}_2(0, T; \mathrm{L}_2(\mathscr{T}))$ and

$$(2.44) \qquad \begin{aligned} 0 &= \int_{\Omega} (\partial_t \rho + \mathrm{div}\,(\rho \boldsymbol{v}))\, \psi\, \mathrm{d}\boldsymbol{x} + \int_{\mathscr{E}} F_1(\rho, \boldsymbol{v}, \tau, \boldsymbol{q}, \psi)\, \mathrm{d}s \quad \forall \psi \in \mathrm{H}^1(\mathscr{T}), \\ 0 &= \int_{\Omega} \left(\rho \partial_t \boldsymbol{v} + \mathrm{div}\,(\rho \boldsymbol{v} \otimes \boldsymbol{v}) - \mathrm{div}\,(\rho \boldsymbol{v})\,\boldsymbol{v} + \rho \nabla \tau - \frac{1}{2}\rho \nabla |\boldsymbol{v}|^2\right)^{\mathsf{T}} \boldsymbol{\chi}\, \mathrm{d}\boldsymbol{x} \\ &\quad + \int_{\mathscr{E}} F_2(\rho, \boldsymbol{v}, \tau, \boldsymbol{q}, \boldsymbol{\chi})\, \mathrm{d}s + \mu B(\boldsymbol{v}, \boldsymbol{\chi}) \quad \forall \boldsymbol{\chi} \in \left(\mathrm{H}_0^1(\mathscr{T})\right)^d, \\ 0 &= \int_{\Omega} \left(\tau - W'(\rho) + \gamma \, \mathrm{div}\,(\boldsymbol{q}) - \frac{1}{2}|\boldsymbol{v}|^2\right) \xi\, \mathrm{d}\boldsymbol{x} \\ &\quad + \int_{\mathscr{E}} F_3(\rho, \boldsymbol{v}, \tau, \boldsymbol{q}, \xi)\, \mathrm{d}s \quad \forall \xi \in \mathrm{H}^1(\mathscr{T}), \\ 0 &= \int_{\Omega} (\boldsymbol{q} - \nabla \rho)^{\mathsf{T}} \boldsymbol{\zeta}\, \mathrm{d}\boldsymbol{x} + \int_{\mathscr{E}} F_4(\rho, \boldsymbol{v}, \tau, \boldsymbol{q}, \boldsymbol{\zeta})\, \mathrm{d}s \quad \forall \boldsymbol{\zeta} \in \mathrm{H}_{\boldsymbol{n}}^1(\mathscr{T}), \end{aligned}$$

where

$$(2.45) \qquad \begin{aligned} F_1, F_3 &: \mathrm{H}^1(\mathscr{T}) \times \mathrm{H}_0^1(\mathscr{T})^d \times \mathrm{H}^1(\mathscr{T}) \times \mathrm{H}_{\boldsymbol{n}}^1(\mathscr{T}) \times \mathrm{H}^1(\mathscr{T}) \rightarrow \mathrm{L}_2(\mathscr{E}), \\ F_2 &: \mathrm{H}^1(\mathscr{T}) \times \mathrm{H}_0^1(\mathscr{T})^d \times \mathrm{H}^1(\mathscr{T}) \times \mathrm{H}_{\boldsymbol{n}}^1(\mathscr{T}) \times \mathrm{H}_0^1(\mathscr{T})^d \rightarrow \mathrm{L}_2(\mathscr{E}), \\ F_4 &: \mathrm{H}^1(\mathscr{T}) \times \mathrm{H}_0^1(\mathscr{T})^d \times \mathrm{H}^1(\mathscr{T}) \times \mathrm{H}_{\boldsymbol{n}}^1(\mathscr{T}) \times \mathrm{H}_{\boldsymbol{n}}^1(\mathscr{T}) \rightarrow \mathrm{L}_2(\mathscr{E}), \end{aligned}$$

are appropriate choices of elementwise fluxes to be chosen in the sequel to suit our purposes; the operators div and ∇ are understood henceforth to be defined elementwise and $B : \left(\mathrm{H}_0^1(\mathscr{T})\right)^d \times \left(\mathrm{H}_0^1(\mathscr{T})\right)^d \rightarrow \mathbb{R}$ is a bilinear form, corresponding to a weak formulation of the Laplacian. We will also assume that the fluxes $F_1, \ldots, F_4$ only depend on the traces of their arguments and are linear in the test functions. We would like to mention that the spaces for the variational formulation are chosen such that all integrals in (2.44) are well defined. We do not claim that there is a well-posedness analysis for (2.44) with the given spaces; (2.44) serves only as a basis to define the spatial discrete DG scheme in the next section.

3. Development of energy consistent numerical methods: the spatially discrete case

In this section we will detail the methodology behind the construction of the energy consistent finite element scheme. We present our main results which show the conditions a generic scheme applied to the variational formulation (2.44) with no diffusion (i.e., $\mu = 0$) is mass and energy conservative. If a scheme conserves mass for $\mu = 0$ this does not change for $\mu \neq 0$, as the mass conservation equation is not affected by a reasonable discretization of the viscosity. Moreover, if a scheme conserves energy for $\mu = 0$, for $\mu \neq 0$ all energy dissipation is due to viscosity. For simplicity we first detail the calculations for the spatially discrete case, then construct a temporally discrete scheme. We also give a condition when a scheme falling under our framework can also conserve momentum.

3.1. **Spatially discrete scheme.** Throughout the calculations in this section we will regularly refer to the following proposition.

3.2. **Proposition** (elementwise integration). *Let*

$$(3.1) \qquad \mathrm{H}^{\mathrm{div}}(\mathscr{T}) := \left\{ \boldsymbol{p} \in (\mathrm{L}_2(\mathscr{T}))^d : \ \mathrm{div}\,\boldsymbol{p} \in \mathrm{L}_2(\mathscr{T}) \right\}.$$

Suppose $\boldsymbol{p} \in \mathrm{H}^{\mathrm{div}}(\mathscr{T})$ and $\phi \in \mathrm{H}^1(\mathscr{T})$, then

$$(3.2) \qquad \sum_{K \in \mathscr{T}} \int_K \mathrm{div}\,(\boldsymbol{p})\,\phi\,\mathrm{d}\boldsymbol{x} = \sum_{K \in \mathscr{T}} \left(-\int_K \boldsymbol{p}^{\mathsf{T}} \nabla \phi\,\mathrm{d}\boldsymbol{x} + \int_{\partial K} \phi \boldsymbol{p}^{\mathsf{T}} \boldsymbol{n}_K\,\mathrm{d}s \right).$$

In particular, we have $\boldsymbol{p} \in \mathcal{T}(\mathscr{E})^d$ and $\phi \in \mathcal{T}(\mathscr{E})$, and the following identity holds:
$$(3.3)$$
$$\sum_{K \in \mathscr{T}} \int_{\partial K} \phi \boldsymbol{p}^{\mathsf{T}} \boldsymbol{n}_K\,\mathrm{d}s = \int_{\mathscr{E}} [\![\boldsymbol{p}]\!]\,\{\!\!\{\,\phi\,\}\!\!\}\,\mathrm{d}s + \int_{\mathscr{E} \cup \partial\Omega} [\![\phi]\!]^{\mathsf{T}}\,\{\!\!\{\,\boldsymbol{p}\,\}\!\!\}\,\mathrm{d}s = \int_{\mathscr{E} \cup \partial\Omega} [\![\boldsymbol{p}\phi]\!]\,\mathrm{d}s.$$

3.3. **General numerical scheme.** A generic spatially discrete DG formulation to the problem in mixed form (2.25)–(2.28) is to find $\rho_h, \tau_h : [0,T] \to \mathbb{V}$ and $\boldsymbol{v}_h : [0,T] \to \mathring{\mathbb{V}}^d$ and $\boldsymbol{q}_h : [0,T] \to \overset{n}{\mathbb{V}}$ such that

$$(3.4)$$
$$0 = \int_\Omega \left(\partial_t \rho_h + \mathrm{div}\,(\rho_h \boldsymbol{v}_h) \right) \Psi\,\mathrm{d}\boldsymbol{x} + \int_{\mathscr{E}} F_1\,(\rho_h, \boldsymbol{v}_h, \tau_h, \boldsymbol{q}_h, \Psi)\,\mathrm{d}s \quad \forall\,\Psi \in \mathbb{V},$$

$$0 = \int_\Omega \left(\rho_h \partial_t \boldsymbol{v}_h + \mathrm{div}\,(\rho_h \boldsymbol{v}_h \otimes \boldsymbol{v}_h) - \mathrm{div}\,(\rho_h \boldsymbol{v}_h)\,\boldsymbol{v}_h + \rho_h \nabla \tau_h - \frac{1}{2} \rho_h \nabla \left| \boldsymbol{v}_h \right|^2 \right)^{\mathsf{T}} \mathbf{X}\,\mathrm{d}\boldsymbol{x}$$
$$+ \int_{\mathscr{E}} F_2\,(\rho_h, \boldsymbol{v}_h, \tau_h, \boldsymbol{q}_h, \mathbf{X})\,\mathrm{d}s + \mu B_h(\boldsymbol{v}_h, \mathbf{X}) \quad \forall\,\mathbf{X} \in \mathring{\mathbb{V}}^d,$$

$$0 = \int_\Omega \left(\tau_h - W'(\rho_h) + \gamma\,\mathrm{div}\,(\boldsymbol{q}_h) - \frac{1}{2} \left| \boldsymbol{v}_h \right|^2 \right) \Xi\,\mathrm{d}\boldsymbol{x}$$
$$+ \int_{\mathscr{E}} F_3\,(\rho_h, \boldsymbol{v}_h, \tau_h, \boldsymbol{q}_h, \Xi)\,\mathrm{d}s \quad \forall\,\Xi \in \mathbb{V},$$

$$0 = \int_\Omega (\boldsymbol{q}_h - \nabla \rho_h)^{\mathsf{T}} \mathbf{Z}\,\mathrm{d}\boldsymbol{x} + \int_{\mathscr{E}} F_4\,(\rho_h, \boldsymbol{v}_h, \tau_h, \boldsymbol{q}_h, \mathbf{Z})\,\mathrm{d}s \quad \forall\,\mathbf{Z} \in \overset{n}{\mathbb{V}},$$

where $B_h : \mathring{\mathbb{V}}^d \times \mathring{\mathbb{V}}^d \to \mathbb{R}$ is a discretization of B.

3.4. **Consistency and conservation.** Now we will give abstract properties of the fluxes which determine whether the scheme is consistent and conserves mass, momentum or energy.

3.5. **Definition** (consistency). A generic scheme having the form (3.4) is said to be consistent provided

$$(3.5) \qquad F_i(\rho, \boldsymbol{v}, \tau, \boldsymbol{q}, \cdot) \equiv 0 \text{ for } i = 1, \dots, 4$$

for all smooth functions $\rho, \tau \in \mathrm{L}_2(0, T; \mathrm{H}^1(\Omega))$ and $\boldsymbol{v}, \boldsymbol{q} \in \left[\mathrm{L}_2(0, T; \mathrm{H}^1(\Omega))\right]^d$.

3.6. **Theorem** (conservation). *For $\mu = 0$ a generic scheme of the form (3.4) conserves:*

(1) *Mass, that is,*

$$(3.6) \qquad \mathrm{d}_t \left(\int_\Omega \rho_h \, \mathrm{d}\boldsymbol{x} \right) = 0$$

if and only if

$$(3.7) \quad \int_{\mathscr{E}} F_1(\rho_h, \boldsymbol{v}_h, \tau_h, \boldsymbol{q}_h, 1) \, \mathrm{d}s = -\int_{\mathscr{E}} [\![\rho_h \boldsymbol{v}_h]\!] \, \mathrm{d}s \quad \forall \, \rho_h, \tau_h \in \mathbb{V}, \boldsymbol{v}_h \in \mathring{\mathbb{V}}^d, \boldsymbol{q}_h \in \overset{\boldsymbol{n}}{\mathbb{V}},$$

where 1 is the the constant element of $\mathrm{H}^1(\mathscr{T})$ which is 1 everywhere.

(2) *Energy, that is,*

$$(3.8) \qquad \mathrm{d}_t \left(\int_\Omega W(\rho_h) + \frac{1}{2}\rho_h \, |\boldsymbol{v}_h|^2 + \frac{\gamma}{2} \, |\boldsymbol{q}_h|^2 \, \mathrm{d}\boldsymbol{x} \right) = 0$$

if and only if

$$(3.9) \quad \begin{aligned} &\int_{\mathscr{E}} F_1(\rho_h, \boldsymbol{v}_h, \tau_h, \boldsymbol{q}_h, \tau_h) + F_2(\rho_h, \boldsymbol{v}_h, \tau_h, \boldsymbol{q}_h, \boldsymbol{v}_h) + [\![\rho_h \tau_h \boldsymbol{v}_h]\!] \, \mathrm{d}s = 0, \\ &\int_{\mathscr{E}} F_3(\rho_h, \boldsymbol{v}_h, \tau_h, \boldsymbol{q}_h, \partial_t \rho_h) - \gamma D_t F_4(\rho_h, \boldsymbol{v}_h, \tau_h, \boldsymbol{q}_h, \boldsymbol{q}_h) + \gamma [\![\partial_t \rho_h \boldsymbol{q}_h]\!] \, \mathrm{d}s = 0, \end{aligned}$$

for all $\rho_h, \tau_h : [0, T] \to \mathbb{V}, \boldsymbol{v}_h : [0, T] \to \mathring{\mathbb{V}}^d$ and $\boldsymbol{q}_h : [0, T] \to \overset{\boldsymbol{n}}{\mathbb{V}}$. Note that we use the notation $D_t F_4(\rho_h, \boldsymbol{v}_h, \tau_h, \boldsymbol{q}_h, \boldsymbol{Z})$ for the time derivative since $\rho_h, \boldsymbol{v}_h, \tau_h, \boldsymbol{q}_h$ are time dependent but $\boldsymbol{Z}$ is independent of time, as $\overset{\boldsymbol{n}}{\mathbb{V}}$ is independent of time.

3.7. **Corollary** (energy dissipation). *Let B_h be a coercive discretisation of B, then for $\mu > 0$ a generic scheme of the form (3.4) conserves mass, if and only if (3.7) is satisfied and it satisfies the energy dissipation equality*

$$(3.10) \qquad \mathrm{d}_t \left(\int_\Omega W(\rho_h) + \frac{1}{2}\rho_h \, |\boldsymbol{v}_h|^2 + \frac{\gamma}{2} \, |\boldsymbol{q}_h|^2 \, \mathrm{d}\boldsymbol{x} \right) = -\mu B_h(\boldsymbol{v}_h, \boldsymbol{v}_h) \leq 0$$

if and only if the following holds:

$$(3.11) \quad \begin{aligned} &\int_{\mathscr{E}} F_1(\rho_h, \boldsymbol{v}_h, \tau_h, \boldsymbol{q}_h, \tau_h) + F_2(\rho_h, \boldsymbol{v}_h, \tau_h, \boldsymbol{q}_h, \boldsymbol{v}_h) + [\![\rho_h \tau_h \boldsymbol{v}_h]\!] \, \mathrm{d}s = 0, \\ &\int_{\mathscr{E}} F_3(\rho_h, \boldsymbol{v}_h, \tau_h, \boldsymbol{q}_h, \partial_t \rho_h) - \gamma D_t F_4(\rho_h, \boldsymbol{v}_h, \tau_h, \boldsymbol{q}_h, \boldsymbol{q}_h) + \gamma [\![\partial_t \rho_h \boldsymbol{q}_h]\!] \, \mathrm{d}s = 0. \end{aligned}$$

We are now in a position to choose fluxes which give certain desired properties, in this case, energy consistency and mass conservation.

3.8. **Corollary** (consistency, conservation and dissipation). *For $\mu = 0$ the following spatially discrete scheme*

(3.12)

$$0 = \int_\Omega \left(\partial_t \rho_h + \mathrm{div}\,(\rho_h \boldsymbol{v}_h) \right) \Psi\, \mathrm{d}\boldsymbol{x} - \int_\mathscr{E} [\![\rho_h \boldsymbol{v}_h]\!]\, \{\!\!\{\, \Psi \,\}\!\!\}\ \mathrm{d}s \quad \forall\, \Psi \in \mathbb{V},$$

$$0 = \int_\Omega \left(\rho_h \partial_t \boldsymbol{v}_h + \mathrm{div}\,(\rho_h \boldsymbol{v}_h \otimes \boldsymbol{v}_h) - \mathrm{div}\,(\rho_h \boldsymbol{v}_h)\,\boldsymbol{v}_h + \rho_h \nabla \tau_h - \frac{1}{2} \rho_h \nabla\, |\boldsymbol{v}_h|^2 \right)^{\!\!\top} \mathbf{X}\, \mathrm{d}\boldsymbol{x}$$

$$- \int_\mathscr{E} [\![\tau_h]\!]^\top\, \{\!\!\{\, \rho_h \mathbf{X} \,\}\!\!\}\ \mathrm{d}s + \mu B_h(\boldsymbol{v}_h, \mathbf{X}) \quad \forall\, \mathbf{X} \in \mathring{\mathbb{V}}^d,$$

$$0 = \int_\Omega \left(\tau_h - W'(\rho_h) + \gamma\, \mathrm{div}\,(\boldsymbol{q}_h) - \frac{1}{2}\, |\boldsymbol{v}_h|^2 \right) \Xi\, \mathrm{d}\boldsymbol{x}$$

$$- \int_\mathscr{E} \gamma\, [\![\boldsymbol{q}_h]\!]\, \{\!\!\{\, \Xi \,\}\!\!\}\ \mathrm{d}s \quad \forall\, \Xi \in \mathbb{V},$$

$$0 = \int_\Omega (\boldsymbol{q}_h - \nabla \rho_h)^\top \mathbf{Z}\, \mathrm{d}\boldsymbol{x} + \int_\mathscr{E} [\![\rho_h]\!]^\top\, \{\!\!\{\, \mathbf{Z} \,\}\!\!\}\ \mathrm{d}s \quad \forall\, \mathbf{Z} \in \mathbb{V}^n,$$

is consistent and conserves mass (3.6) *and energy* (3.8).

Let $B_h(u, w)$ be the symmetric interior penalty method for the componentwise Laplacian given by

(3.13)

$$B_h(\boldsymbol{u}, \boldsymbol{w}) = \int_\Omega \mathrm{D}\boldsymbol{u} : \mathrm{D}\boldsymbol{w}\, \mathrm{d}\boldsymbol{x} - \int_{\mathscr{E} \cup \partial\Omega} \{\!\!\{\, \mathrm{D}\boldsymbol{w} \,\}\!\!\} : [\![\boldsymbol{u}]\!]_\otimes + \{\!\!\{\, \mathrm{D}\boldsymbol{u} \,\}\!\!\} : [\![\boldsymbol{w}]\!]_\otimes - \frac{\sigma}{h} [\![\boldsymbol{u}]\!]_\otimes : [\![\boldsymbol{w}]\!]_\otimes\, \mathrm{d}s,$$

where : *denotes the Frobenius inner product between two $d\times d$ matrices, i.e., $\boldsymbol{X} : \boldsymbol{Y} :=$ trace $(\boldsymbol{X}^\top \boldsymbol{Y})$ and the average of a matrix is defined analogously to the average of a vector, as in Definition 2.14. It is well known for large (enough) σ that this is a coercive discretisation of the componentwise Laplacian. Thus, for $\mu > 0$, the numerical scheme* (3.12) *with* (3.13) *is consistent, conserves mass* (3.6) *and dissipates energy* (3.10).

Proof of Theorem 3.6. Let us first consider the proof of conservation of mass. By using $\Psi \equiv 1$ as a test function in $(3.4)_1$ we want to show

(3.14)

$$0 = \mathrm{d}_t \left(\int_\Omega \rho_h\, \mathrm{d}\boldsymbol{x} \right) = \int_\Omega \partial_t \rho_h\, \mathrm{d}\boldsymbol{x} = - \int_\Omega \mathrm{div}(\rho_h \boldsymbol{v}_h)\, \mathrm{d}\boldsymbol{x} - \int_\mathscr{E} F_1(\rho_h, \boldsymbol{v}_h, \tau_h, \boldsymbol{q}_h, 1)\ \mathrm{d}s$$

$$= - \int_\mathscr{E} [\![\rho_h \boldsymbol{v}_h]\!]\, \{\!\!\{\, 1 \,\}\!\!\}\ \mathrm{d}\boldsymbol{x} - \int_\mathscr{E} F_1(\rho_h, \boldsymbol{v}_h, \tau_h, \boldsymbol{q}_h, 1)\ \mathrm{d}s,$$

by Proposition 3.2 with $\boldsymbol{p} = \rho_h \boldsymbol{v}_h$ and $\phi = 1$ and noting $\boldsymbol{v}_h = \boldsymbol{0}$ on $\partial\Omega$. Hence the scheme conserves mass if the condition (3.7) is true.

Let us now turn to the conservation of energy. Define

(3.15)

$$E(\rho_h, \boldsymbol{v}_h, \boldsymbol{q}_h) := \int_\Omega W(\rho_h) + \frac{1}{2} \rho_h\, |\boldsymbol{v}_h|^2 + \frac{\gamma}{2}\, |\boldsymbol{q}_h|^2\ \mathrm{d}\boldsymbol{x}.$$

Again we want to show that

(3.16)

$$0 = \mathrm{d}_t E(\rho_h, \boldsymbol{v}_h, \boldsymbol{q}_h).$$

Explicitly computing the time derivative

(3.17)
$$\mathrm{d}_t E(\rho_h, \boldsymbol{v}_h, \boldsymbol{q}_h) = \int_\Omega W'(\rho_h)\partial_t\rho_h + \frac{1}{2}\partial_t\rho_h\,|\boldsymbol{v}_h|^2 + \rho_h(\boldsymbol{v}_h)^\mathsf{T}\partial_t\boldsymbol{v}_h + \gamma(\boldsymbol{q}_h)^\mathsf{T}\partial_t\boldsymbol{q}_h\,\mathrm{d}\boldsymbol{x}.$$

In view of $(3.4)_4$ and Proposition 3.2 we see that

(3.18)
$$\mathrm{d}_t E(\rho_h, \boldsymbol{v}_h, \boldsymbol{q}_h) = \int_\Omega W'(\rho_h)\partial_t\rho_h + \frac{1}{2}\partial_t\rho_h\,|\boldsymbol{v}_h|^2 + \rho_h(\boldsymbol{v}_h)^\mathsf{T}\partial_t\boldsymbol{v}_h + \gamma(\nabla\partial_t\rho_h)^\mathsf{T}\boldsymbol{q}_h\,\mathrm{d}\boldsymbol{x},$$
$$-\gamma\int_\mathscr{E} \mathrm{D}_t F_4(\rho_h, \boldsymbol{v}_h, \tau_h, \boldsymbol{q}_h, \boldsymbol{q}_h)\,\mathrm{d}s,$$
$$= \int_\Omega W'(\rho_h)\partial_t\rho + \frac{1}{2}\partial_t\rho_h\,|\boldsymbol{v}_h|^2 + \rho_h(\boldsymbol{v}_h)^\mathsf{T}\partial_t\boldsymbol{v}_h - \gamma(\partial_t\rho_h)^\mathsf{T}\,\mathrm{div}\,\boldsymbol{q}_h\,\mathrm{d}\boldsymbol{x}$$
$$-\gamma\int_\mathscr{E} \mathrm{D}_t F_4(\rho_h, \boldsymbol{v}_h, \tau_h, \boldsymbol{q}_h, \boldsymbol{q}_h) - [\![(\partial_t\rho_h)\boldsymbol{q}_h]\!]\,\mathrm{d}s,$$

as $\boldsymbol{q}_h{}^\mathsf{T}\boldsymbol{n} = 0$ on $\partial\Omega$. Making use of $(3.4)_2$ and $(3.4)_3$ we see that

(3.19)
$$\mathrm{d}_t E(\rho_h, \boldsymbol{v}_h, \boldsymbol{q}_h) = \int_\Omega \tau_h\partial_t\rho_h - \boldsymbol{v}_h{}^\mathsf{T}\,\mathrm{div}(\rho_h\boldsymbol{v}_h\otimes\boldsymbol{v}_h) + \mathrm{div}(\rho_h\boldsymbol{v}_h)\,|\boldsymbol{v}_h|^2 - \rho_h\boldsymbol{v}_h{}^\mathsf{T}\nabla\tau_h$$
$$+\frac{1}{2}\rho_h\boldsymbol{v}_h{}^\mathsf{T}\nabla\left(|\boldsymbol{v}_h|^2\right)\,\mathrm{d}\boldsymbol{x}$$
$$-\int_\mathscr{E} \gamma\mathrm{D}_t F_4(\rho_h, \boldsymbol{v}_h, \tau_h, \boldsymbol{q}_h, \boldsymbol{q}_h) - \gamma\,[\![(\partial_t\rho_h)\boldsymbol{q}_h]\!]$$
$$- F_3(\rho_h, \boldsymbol{v}_h, \tau_h, \boldsymbol{q}_h, \partial_t\rho_h) + F_2(\rho_h, \boldsymbol{v}_h, \tau_h, \boldsymbol{q}_h, \boldsymbol{v}_h)\,\mathrm{d}s.$$

Now by $(3.4)_1$ and Proposition 3.2 we see, as $\boldsymbol{v}_h \in \mathring{\mathbb{V}}^d$, that

(3.20)
$$\mathrm{d}_t E(\rho_h, \boldsymbol{v}_h, \boldsymbol{q}_h) = \int_\Omega -\mathrm{div}(\rho_h\boldsymbol{v}_h)\tau_h - \boldsymbol{v}_h{}^\mathsf{T}\,\mathrm{div}(\rho_h\boldsymbol{v}_h\otimes\boldsymbol{v}_h) + \mathrm{div}(\rho_h\boldsymbol{v}_h)\,|\boldsymbol{v}_h|^2$$
$$-\rho_h\boldsymbol{v}_h\nabla\tau_h + \frac{1}{2}\rho_h\boldsymbol{v}_h\nabla\left(|\boldsymbol{v}_h|^2\right)\,\mathrm{d}\boldsymbol{x}$$
$$-\int_\mathscr{E} \gamma\mathrm{D}_t F_4(\rho_h, \boldsymbol{v}_h, \tau_h, \boldsymbol{q}_h, \boldsymbol{q}_h) - \gamma\,[\![(\partial_t\rho_h)\boldsymbol{q}_h]\!]$$
$$- F_3(\rho_h, \boldsymbol{v}_h, \tau_h, \boldsymbol{q}_h, \partial_t\rho_h)\,\mathrm{d}s$$
$$-\int_\mathscr{E} F_2(\rho_h, \boldsymbol{v}_h, \tau_h, \boldsymbol{q}_h, \boldsymbol{v}_h) + F_1(\rho_h, \boldsymbol{v}_h, \tau_h, \boldsymbol{q}_h, \tau_h)\,\mathrm{d}s$$
$$= -\int_\mathscr{E} \gamma\mathrm{D}_t F_4(\rho_h, \boldsymbol{v}_h, \tau_h, \boldsymbol{q}_h, \boldsymbol{q}_h) - \gamma\,[\![(\partial_t\rho_h)\boldsymbol{q}_h]\!]$$
$$- F_3(\rho_h, \boldsymbol{v}_h, \tau_h, \boldsymbol{q}_h, \partial_t\rho_h)\,\mathrm{d}s$$
$$-\int_\mathscr{E} F_2(\rho_h, \boldsymbol{v}_h, \tau_h, \boldsymbol{q}_h, \boldsymbol{v}_h) + F_1(\rho_h, \boldsymbol{v}_h, \tau_h, \boldsymbol{q}_h, \tau_h)$$
$$+ [\![\rho_h\tau_h\boldsymbol{v}_h]\!]\,\mathrm{d}s.$$

2084   JAN GIESSELMANN, CHARALAMBOS MAKRIDAKIS, AND TRISTAN PRYER

Thus, an energy conserving scheme has to satisfy

(3.21)
$$0 = \int_{\mathscr{E}} -\gamma D_t F_4(\rho_h, \boldsymbol{v}_h, \tau_h, \boldsymbol{q}_h, \boldsymbol{q}_h) + \gamma \left[\!\left[(\partial_t \rho_h) \boldsymbol{q}_h \right]\!\right] + F_3(\rho_h, \boldsymbol{v}_h, \tau_h, \boldsymbol{q}_h, \partial_t \rho_h) \, \mathrm{d}s$$
$$- \int_{\mathscr{E}} F_2(\rho_h, \boldsymbol{v}_h, \tau_h, \boldsymbol{q}_h, \boldsymbol{v}_h) + F_1(\rho_h, \boldsymbol{v}_h, \tau_h, \boldsymbol{q}_h, \tau_h) + \left[\!\left[\rho_h \tau_h \boldsymbol{v}_h \right]\!\right] \, \mathrm{d}s.$$

Note that when (3.21) holds, F_4 cannot depend on $\boldsymbol{v}_h, \tau_h, \boldsymbol{q}_h$. Furthermore, every summand in $D_t F_4$ and F_3 has to depend on $\partial_t \rho_h$. As the trace of $\partial_t \rho_h$ is independent of the traces of $\rho_h, \boldsymbol{v}_h, \tau_h, \boldsymbol{q}_h$ the quantities containing $\partial_t \rho_h$ must cancel each other, and the ones not containing $\partial_t \rho_h$ must cancel each other. $\qquad\square$

Proof of Corollary 3.7. The proof of conservation of mass is exactly the same as in the proof of Theorem 3.6 because $(3.4)_1$ does not depend on μ. For the dissipation of energy the only difference to the proof of Theorem 3.6 is that, when $(3.4)_2$ is tested with $\mathbf{X} = \boldsymbol{v}_h$ an additional summand $\mu B_h(\boldsymbol{v}_h, \boldsymbol{v}_h)$ is created and this term is not altered by the subsequent calculations. $\qquad\square$

3.9. *Remark* (choice of fluxes). Theorem 3.6 and Corollary 3.7 give conditions on the fluxes such that the numerical scheme will conserve (the case $\mu = 0$) and dissipate (the case $\mu > 0$) energy, respectively.

Even for $\mu = 0$ the energy estimate (Lemma 2.3) gives an a priori estimate for ρ in $L^\infty([0,T]; H^1(\Omega))$ provided the initial energy is finite. Thus, we do not expect discontinuities to form (in ρ), given that the initial data are smooth. Of course, for $\mu = 0$ we cannot exclude discontinuities in $\boldsymbol{v}$. However, in the case where the initial data is discontinuous, for example, shocks may form. We may use the generic framework developed in the Proof of Theorem 3.6 to design the fluxes such that in the case of irregular initial conditions we may dissipate a small amount of energy, i.e.,

(3.22)
$$\mathrm{d}_t \left(\int_\Omega W(\rho_h) + \frac{1}{2} \rho_h \left| \boldsymbol{v}_h \right|^2 + \frac{\gamma}{2} \left| \boldsymbol{q}_h \right|^2 \, \mathrm{d}\boldsymbol{x} \right) \leq 0.$$

The conditions on the fluxes necessary for (3.22) to hold are

(3.23)
$$\int_{\mathscr{E}} F_1(\rho_h, \boldsymbol{v}_h, \tau_h, \boldsymbol{q}_h, \tau_h) + F_2(\rho_h, \boldsymbol{v}_h, \tau_h, \boldsymbol{q}_h, \boldsymbol{v}_h) + \left[\!\left[\rho_h \tau_h \boldsymbol{v}_h \right]\!\right] \, \mathrm{d}s \geq 0,$$
$$\int_{\mathscr{E}} F_3(\rho_h, \boldsymbol{v}_h, \tau_h, \boldsymbol{q}_h, \partial_t \rho_h) - \gamma D_t F_4(\rho_h, \boldsymbol{v}_h, \tau_h, \boldsymbol{q}_h, \boldsymbol{q}_h) + \gamma \left[\!\left[\partial_t \rho_h \boldsymbol{q}_h \right]\!\right] \, \mathrm{d}s \leq 0.$$

This may also prove useful as a stabilising mechanism for some temporal discretisations, e.g., in the case where explicit time stepping is considered.

3.10. **Example.** For $\alpha, \beta > 0$ the choice of fluxes

(3.24)
$$F_1(\rho_h, \boldsymbol{v}_h, \tau_h, \boldsymbol{q}_h, \Psi) = - \left[\!\left[\rho_h \boldsymbol{v}_h \right]\!\right] \left\{\!\!\left\{ \Psi \right\}\!\!\right\} + \alpha \left[\!\left[\tau_h \right]\!\right]^\mathsf{T} \left[\!\left[\Psi \right]\!\right],$$
$$F_2(\rho_h, \boldsymbol{v}_h, \tau_h, \boldsymbol{q}_h, \mathbf{X}) = - \left[\!\left[\tau_h \right]\!\right]^\mathsf{T} \left\{\!\!\left\{ \rho_h \mathbf{X} \right\}\!\!\right\} + \beta \left[\!\left[\boldsymbol{v}_h \right]\!\right] \left[\!\left[\mathbf{X} \right]\!\right],$$
$$F_3(\rho_h, \boldsymbol{v}_h, \tau_h, \boldsymbol{q}_h, \Xi) = -\gamma \left[\!\left[\boldsymbol{q}_h \right]\!\right] \left\{\!\!\left\{ \Xi \right\}\!\!\right\},$$
$$F_4(\rho_h, \boldsymbol{v}_h, \tau_h, \boldsymbol{q}_h, \mathbf{Z}) = \left[\!\left[\rho_h \right]\!\right]^\mathsf{T} \left\{\!\!\left\{ \mathbf{Z} \right\}\!\!\right\},$$

in (3.4) yields a scheme which is consistent, conserves mass and dissipates energy.

License or copyright restrictions may apply to redistribution; see https://www.ams.org/journal-terms-of-use

Further to Remark 2.10 we will now show that it is very restrictive for schemes given by (3.4) to be consistent, conserve momentum and energy and, in fact, we will prove that the fluxes specified in Corollary 3.7 and Example 3.10 do not conserve momentum.

3.11. Proposition (momentum conservation). *Let $\boldsymbol{e}_i$ be the ith coordinate vector of $\mathbb{R}^d$. A generic scheme of the form* (3.4) *is momentum conservative, i.e., satisfies*

$$(3.25) \qquad \mathrm{d}_t \left(\int_\Omega \rho_h \boldsymbol{v}_h \, \mathrm{d}\boldsymbol{x} \right) = \boldsymbol{0}$$

if and only if

(3.26)

$$
\begin{aligned}
0 = &-\int_{\mathscr{E}} F_1(\rho_h, \boldsymbol{v}_h, \tau_h, \boldsymbol{q}_h, \boldsymbol{e}_i^\mathsf{T} \boldsymbol{v}_h) + F_2(\rho_h, \boldsymbol{v}_h, \tau_h, \boldsymbol{q}_h, \boldsymbol{e}_i) \\
&\qquad + F_3(\rho_h, \boldsymbol{v}_h, \tau_h, \boldsymbol{q}_h, \partial_{x_i} \rho_h) + F_4(\rho_h, \boldsymbol{v}_h, \tau_h, \boldsymbol{q}_h, \nabla \partial_{x_i} \rho_h) \, \mathrm{d}s \\
&+ \int_{\mathscr{E} \cup \partial\Omega} - [\![\rho_h \tau_h \boldsymbol{e}_i]\!] + \left[\!\!\left[\left(W(\rho_h) + \frac{1}{2} \rho_h \left| \boldsymbol{v}_h \right|^2 \right) \boldsymbol{e}_i \right]\!\!\right] - [\![\rho_h \boldsymbol{e}_i^\mathsf{T} (\boldsymbol{v}_h \otimes \boldsymbol{v}_h)]\!] \\
&\qquad - \gamma \, [\![\partial_{x_i} \rho_h \boldsymbol{q}_h]\!] + \gamma \, [\![\rho_h \nabla \partial_{x_i} \rho_h]\!] \\
&\qquad - \gamma \left[\!\!\left[\rho_h \Delta \rho_h \boldsymbol{e}_i + \frac{1}{2} \left| \nabla \rho_h \right|^2 \boldsymbol{e}_i - \partial_{x_i} \rho_h \nabla \rho_h \right]\!\!\right] \, \mathrm{d}s.
\end{aligned}
$$

The proof consists of calculating of the change of momentum in direction $\boldsymbol{e}_i$ for $i = 1, \ldots, d$ for the generic scheme (3.4).

3.12. *Remark* (restrictions for momentum conservation). Since the right-hand side of (3.26) depends non-linearly on ρ_h, it is not expected that the jump terms involving $\nabla \rho_h$ and $\boldsymbol{q}_h$ will necessarily cancel with the other terms appearing. Only in the case $F_4 \equiv 0$ will the $\nabla \rho_h$ and $\boldsymbol{q}_h$ terms cancel each other. In this case $(3.9)_2$ gives us the condition that

$$(3.27) \qquad F_3(\rho_h, \boldsymbol{v}_h, \tau_h, \boldsymbol{q}_h, \Xi) = -\gamma \, [\![\boldsymbol{q}_h \Xi]\!]$$

which excludes consistency. In the event that $F_4 \not\equiv 0$, the terms involving $\nabla \rho_h$ in (3.26) would be required to cancel independently, yielding a condition on F_3 which again excludes consistency.

4. Development of energy consistent numerical methods: the temporal discrete case

For the readers convenience, we will present an argument for designing the temporally discrete scheme in the spatially continuous setting. To obtain a fully discrete version the spatial and temporal discretisations have to be combined which is straightforward as presented in (5.2) and Theorem 5.1.

We subdivide the time interval $[0, T]$ into a partition of N consecutive adjacent subintervals whose endpoints are denoted $t_0 = 0 < t_1 < \cdots < t_N = T$. The n-th timestep is defined as $k_n := t_{n+1} - t_n$. We will consistently use the shorthand $F^n(\cdot) := F(\cdot, t_n)$ for a generic time function F. We also denote $F^{n+\frac{1}{2}} := \frac{1}{2} \left(F^n + F^{n+1} \right)$.

4.1. **Theorem** (temporally discrete energy consistent scheme). *Given initial conditions $\rho^0, \boldsymbol{v}^0, \tau^0$ and $\boldsymbol{q}^0$ the temporal semi-discrete scheme is: For $n \in \mathbb{N}$, find $\rho^{n+1}, \boldsymbol{v}^{n+1}, \tau^{n+1}$ and $\boldsymbol{q}^{n+1}$ such that $\boldsymbol{v}^{n+1} = \boldsymbol{0}$ and $(\boldsymbol{q}^{n+1})^{\mathsf{T}} \boldsymbol{n} = 0$ on $\partial\Omega$ and*

(4.1)

$$0 = \frac{\rho^{n+1} - \rho^n}{k_n} + \operatorname{div}\left(\rho^{n+\frac{1}{2}} \boldsymbol{v}^{n+\frac{1}{2}}\right),$$

$$0 = \rho^{n+\frac{1}{2}}\left(\frac{\boldsymbol{v}^{n+1} - \boldsymbol{v}^n}{k_n}\right) + \operatorname{div}\left(\rho^{n+\frac{1}{2}} \boldsymbol{v}^{n+\frac{1}{2}} \otimes \boldsymbol{v}^{n+\frac{1}{2}}\right) - \operatorname{div}\left(\rho^{n+\frac{1}{2}} \boldsymbol{v}^{n+\frac{1}{2}}\right)\boldsymbol{v}^{n+\frac{1}{2}}$$

$$+ \rho^{n+\frac{1}{2}} \nabla \tau^{n+\frac{1}{2}} - \frac{1}{2}\rho^{n+\frac{1}{2}} \nabla\left(\left|\boldsymbol{v}^{n+\frac{1}{2}}\right|^2\right) - \mu\Delta\boldsymbol{v}^{n+\frac{1}{2}},$$

$$0 = \tau^{n+\frac{1}{2}} - \frac{W(\rho^{n+1}) - W(\rho^n)}{\rho^{n+1} - \rho^n} + \gamma\operatorname{div}\left(\boldsymbol{q}^{n+\frac{1}{2}}\right) - \frac{1}{4}\left(\left|\boldsymbol{v}^{n+1}\right|^2 + \left|\boldsymbol{v}^n\right|^2\right),$$

$$0 = \boldsymbol{q}^{n+1} - \nabla\rho^{n+1}.$$

This scheme satisfies the following energy dissipation property for all $0 \le n \le N$:

$$\int_\Omega W(\rho^n) + \frac{1}{2}\rho^n \left|\boldsymbol{v}^n\right|^2 + \frac{\gamma}{2}\left|\boldsymbol{q}^n\right|^2 \,\mathrm{d}\boldsymbol{x} = \int_\Omega W(\rho^0) + \frac{1}{2}\rho^0 \left|\boldsymbol{v}^0\right|^2 + \frac{\gamma}{2}\left|\boldsymbol{q}^0\right|^2 \,\mathrm{d}\boldsymbol{x}$$

(4.2)
$$- \mu\sum_{j=0}^{n-1} k_j \int_\Omega \left|\mathrm{D}\boldsymbol{v}^{j+\frac{1}{2}}\right|^2 \,\mathrm{d}\boldsymbol{x}.$$

Proof. We proceed by multiplying $(4.1)_1$ by $\tau^{n+\frac{1}{2}}$ and $(4.1)_2$ by $\boldsymbol{v}^{n+\frac{1}{2}}$, we integrate over the domain Ω and take the sum. We obtain

(4.3)
$$0 = \int_\Omega \mathscr{I}_1 + \mathscr{I}_2 + \mathscr{I}_3 + \mathscr{I}_4 \,\mathrm{d}\boldsymbol{x}$$

with

(4.4)
$$\mathscr{I}_1 := \frac{\rho^{n+1} - \rho^n}{k_n}\left(\frac{W(\rho^{n+1}) - W(\rho^n)}{\rho^{n+1} - \rho^n} - \gamma\operatorname{div}\left(\boldsymbol{q}^{n+\frac{1}{2}}\right) + \frac{1}{4}\left(\left|\boldsymbol{v}^{n+1}\right|^2 + \left|\boldsymbol{v}^n\right|^2\right)\right)$$

$$+ \rho^{n+\frac{1}{2}}\left(\boldsymbol{v}^{n+\frac{1}{2}}\right)^{\mathsf{T}}\left(\frac{\boldsymbol{v}^{n+1} - \boldsymbol{v}^n}{k_n}\right),$$

(4.5)
$$\mathscr{I}_2 := \operatorname{div}\left(\rho^{n+\frac{1}{2}} \boldsymbol{v}^{n+\frac{1}{2}}\right)\tau^{n+\frac{1}{2}} + \rho^{n+\frac{1}{2}}\left(\boldsymbol{v}^{n+\frac{1}{2}}\right)^{\mathsf{T}} \nabla\tau^{n+\frac{1}{2}},$$

(4.6)
$$\mathscr{I}_3 := \left(\boldsymbol{v}^{n+\frac{1}{2}}\right)^{\mathsf{T}} \operatorname{div}\left(\rho^{n+\frac{1}{2}} \boldsymbol{v}^{n+\frac{1}{2}} \otimes \boldsymbol{v}^{n+\frac{1}{2}}\right) - \operatorname{div}\left(\rho^{n+\frac{1}{2}} \boldsymbol{v}^{n+\frac{1}{2}}\right)\left|\boldsymbol{v}^{n+\frac{1}{2}}\right|^2$$

$$- \frac{1}{2}\rho^{n+\frac{1}{2}}\left(\boldsymbol{v}^{n+\frac{1}{2}}\right)^{\mathsf{T}} \nabla\left(\left|\boldsymbol{v}^{n+\frac{1}{2}}\right|^2\right),$$

(4.7)
$$\mathscr{I}_4 := -\mu\int_\Omega\left(\boldsymbol{v}^{n+\frac{1}{2}}\right)^{\mathsf{T}}\left(\Delta\boldsymbol{v}^{n+\frac{1}{2}}\right)\,\mathrm{d}\boldsymbol{x}.$$

One may readily check that

$$k_n \int_\Omega \mathscr{I}_1 \, \mathrm{d}\boldsymbol{x} = \int_\Omega W(\rho^{n+1}) + \frac{1}{2}\rho^{n+1} \left| \boldsymbol{v}^{n+1} \right|^2 + \frac{\gamma}{2} \left| \boldsymbol{q}^{n+1} \right|^2 \, \mathrm{d}\boldsymbol{x}$$

$$- \int_\Omega W(\rho^n) + \frac{1}{2}\rho^n \left| \boldsymbol{v}^n \right|^2 + \frac{\gamma}{2} \left| \boldsymbol{q}^n \right|^2 \, \mathrm{d}\boldsymbol{x}$$

(4.8)

$$- \gamma \int_{\partial\Omega} (\rho^{n+1} - \rho^n)(\boldsymbol{q}^{n+1} + \boldsymbol{q}^n)^\mathsf{T} \boldsymbol{n} \, \mathrm{d}s$$

$$= \int_\Omega W(\rho^{n+1}) + \frac{1}{2}\rho^{n+1} \left| \boldsymbol{v}^{n+1} \right|^2 + \frac{\gamma}{2} \left| \boldsymbol{q}^{n+1} \right|^2 \, \mathrm{d}\boldsymbol{x}$$

$$- \int_\Omega W(\rho^n) + \frac{1}{2}\rho^n \left| \boldsymbol{v}^n \right|^2 + \frac{\gamma}{2} \left| \boldsymbol{q}^n \right|^2 \, \mathrm{d}\boldsymbol{x}.$$

Moreover,

$$(4.9) \quad \int_\Omega \mathscr{I}_2 \, \mathrm{d}\boldsymbol{x} = \mathrm{div}\left(\rho^{n+\frac{1}{2}} \boldsymbol{v}^{n+\frac{1}{2}} \tau^{n+\frac{1}{2}} \right) \mathrm{d}\boldsymbol{x} = \int_{\partial\Omega} \rho^{n+\frac{1}{2}} \tau^{n+\frac{1}{2}} \left(\boldsymbol{v}^{n+\frac{1}{2}} \right)^\mathsf{T} \boldsymbol{n} \, \mathrm{d}s = 0.$$

Furthermore, we see that $\mathscr{I}_3$ satisfies

$$\mathscr{I}_3 = \sum_{i,j=1}^d \partial_{x_i} \left(\rho^{n+\frac{1}{2}} v_i^{n+\frac{1}{2}} v_j^{n+\frac{1}{2}} \right) v_j^{n+\frac{1}{2}} - \partial_{x_i} \left(\rho^{n+\frac{1}{2}} v_i^{n+\frac{1}{2}} \right) \left(v_j^{n+\frac{1}{2}} \right)^2$$

(4.10)

$$- \frac{1}{2}\rho^{n+\frac{1}{2}} v_i^{n+\frac{1}{2}} \partial_{x_i} \left(\left(v_j^{n+\frac{1}{2}} \right)^2 \right) = 0.$$

Finally, for $\mathscr{I}_4$ we find that

$$\int_\Omega \mathscr{I}_4 = \mu \int_\Omega \left| \mathrm{D}\boldsymbol{v}^{n+\frac{1}{2}} \right|^2 \, \mathrm{d}\boldsymbol{x} - \mu \int_{\partial\Omega} \left(\boldsymbol{v}^{n+\frac{1}{2}} \right)^\mathsf{T} \left(\mathrm{D}\boldsymbol{v}^{n+\frac{1}{2}} \right) \boldsymbol{n} \, \mathrm{d}s$$

(4.11)

$$= \mu \int_\Omega \left| \mathrm{D}\boldsymbol{v}^{n+\frac{1}{2}} \right|^2 \, \mathrm{d}\boldsymbol{x}.$$

Inserting (4.8)–(4.10) into (4.3) yields

$$0 = \int_\Omega W(\rho^{n+1}) + \frac{1}{2}\rho^{n+1} \left| \boldsymbol{v}^{n+1} \right|^2 + \frac{\gamma}{2} \left| \boldsymbol{q}^{n+1} \right|^2 \, \mathrm{d}\boldsymbol{x}$$

(4.12)

$$- \int_\Omega W(\rho^n) + \frac{1}{2}\rho^n \left| \boldsymbol{v}^n \right|^2 + \frac{\gamma}{2} \left| \boldsymbol{q}^n \right|^2 \, \mathrm{d}\boldsymbol{x} + \mu k_n \int_\Omega \left| \mathrm{D}\boldsymbol{v}^{n+\frac{1}{2}} \right|^2 \, \mathrm{d}\boldsymbol{x},$$

concluding the proof. $\qquad\square$

5. Development of consistent numerical methods: the fully discrete case

In this section we combine our spatial and temporal discretisations to provide a fully discrete numerical method for the Euler–Korteweg and Navier–Stokes–Korteweg systems.

Let $\mathrm{P}_\mathbb{V} : \mathrm{H}^1(\mathscr{T}) \to \mathbb{V}, \mathrm{P}_{\overset{\circ}{\mathbb{V}}} : \mathrm{H}_0^1(\mathscr{T})^d \to \overset{\circ}{\mathbb{V}}{}^d$ and $\mathrm{P}_{\overset{\boldsymbol{n}}{\mathbb{V}}} : \mathrm{H}_{\boldsymbol{n}}^1(\mathscr{T}) \to \overset{\boldsymbol{n}}{\mathbb{V}}$ be the L_2 projection operators into $\mathbb{V}, \overset{\circ}{\mathbb{V}}{}^d$ and $\overset{\boldsymbol{n}}{\mathbb{V}}$ respectively. We combine the arguments given in §3 and §4 to obtain a fully discrete scheme which, given

$$(5.1) \qquad \rho_h^0 := \mathrm{P}_\mathbb{V} \rho^0, \quad \boldsymbol{v}_h^0 := \mathrm{P}_{\overset{\circ}{\mathbb{V}}} \boldsymbol{v}^0, \quad \tau_h^0 := \mathrm{P}_\mathbb{V} \tau^0 \text{ and } \quad \boldsymbol{q}_h^0 := \mathrm{P}_{\overset{\boldsymbol{n}}{\mathbb{V}}} \boldsymbol{q}^0,$$

requires us to find a sequence of functions $\rho_h^{n+1}, \tau_h^{n+1} \in \mathbb{V}, \boldsymbol{v}_h^{n+1} \in \overset{\circ}{\mathbb{V}}{}^d$ and $\boldsymbol{q}_h^{n+1} \in \overset{n}{\mathbb{V}}$ such that

(5.2)

$$0 = \int_\Omega \left(\frac{\rho_h^{n+1} - \rho_h^n}{k_n} + \operatorname{div}\left(\rho_h^{n+\frac{1}{2}} \boldsymbol{v}_h^{n+\frac{1}{2}}\right) \right) \Psi \, d\boldsymbol{x}$$
$$+ \int_{\mathcal{E}} F_1\left(\rho_h^{n+\frac{1}{2}}, \boldsymbol{v}_h^{n+\frac{1}{2}}, \tau_h^{n+\frac{1}{2}}, \boldsymbol{q}_h^{n+\frac{1}{2}}, \Psi\right) ds \quad \forall\, \Psi \in \mathbb{V},$$

$$0 = \int_\Omega \left(\left(\rho_h^{n+\frac{1}{2}} \frac{\boldsymbol{v}_h^{n+1} - \boldsymbol{v}_h^n}{k_n} \right) + \operatorname{div}\left(\rho_h^{n+\frac{1}{2}} \boldsymbol{v}_h^{n+\frac{1}{2}} \otimes \boldsymbol{v}_h^{n+\frac{1}{2}}\right) \right)^{\mathsf{T}} \mathbf{X} \, d\boldsymbol{x}$$
$$+ \int_\Omega \left(-\operatorname{div}\left(\rho_h^{n+\frac{1}{2}} \boldsymbol{v}_h^{n+\frac{1}{2}}\right) \boldsymbol{v}_h^{n+\frac{1}{2}} + \rho_h^{n+\frac{1}{2}} \nabla \tau_h^{n+\frac{1}{2}} - \frac{1}{2} \rho_h^{n+\frac{1}{2}} \nabla \left|\boldsymbol{v}_h^{n+\frac{1}{2}}\right|^2 \right)^{\mathsf{T}} \mathbf{X} \, d\boldsymbol{x}$$
$$+ \int_{\mathcal{E}} F_2\left(\rho_h^{n+\frac{1}{2}}, \boldsymbol{v}_h^{n+\frac{1}{2}}, \tau_h^{n+\frac{1}{2}}, \boldsymbol{q}_h^{n+\frac{1}{2}}, \mathbf{X}\right) ds + \mu B_h(\boldsymbol{v}_h^{n+\frac{1}{2}}, \mathbf{X}) \quad \forall\, \mathbf{X} \in \overset{\circ}{\mathbb{V}}{}^d,$$

$$0 = \int_\Omega \left(\tau_h^{n+\frac{1}{2}} - \frac{W(\rho_h^{n+1}) - W(\rho_h^n)}{\rho_h^{n+1} - \rho_h^n} + \gamma \operatorname{div}\left(\boldsymbol{q}_h^{n+\frac{1}{2}}\right) - \frac{1}{4}\left(\left|\boldsymbol{v}_h^{n+1}\right|^2 + \left|\boldsymbol{v}_h^n\right|^2\right) \right) \Xi \, d\boldsymbol{x}$$
$$+ \int_{\mathcal{E}} F_3\left(\rho_h^{n+\frac{1}{2}}, \boldsymbol{v}_h^{n+\frac{1}{2}}, \tau_h^{n+\frac{1}{2}}, \boldsymbol{q}_h^{n+\frac{1}{2}}, \Xi\right) ds \quad \forall\, \Xi \in \mathbb{V},$$

$$0 = \int_\Omega \left(\boldsymbol{q}_h^{n+1} - \nabla \rho_h^{n+1}\right)^{\mathsf{T}} \mathbf{Z} \, d\boldsymbol{x} + \int_{\mathcal{E}} F_4\left(\rho_h^{n+1}, \boldsymbol{v}_h^{n+1}, \tau_h^{n+1}, \boldsymbol{q}_h^{n+1}, \mathbf{Z}\right) ds \quad \forall\, \mathbf{Z} \in \overset{n}{\mathbb{V}}.$$

5.1. **Theorem** (fully discrete energy consistency). *Under the assumptions on the fluxes* (3.7)*,* (3.9) *given in Theorem* 3.6*, the solution of the scheme* (5.2) *conserves mass, i.e.,*

$$\int_\Omega \rho_h^n \, d\boldsymbol{x} = \int_\Omega \rho_h^0 \, d\boldsymbol{x} \qquad \text{for } 0 \le n \le N$$

and satisfies the energy dissipation equality

(5.3) $$\int_\Omega W(\rho_h^{n+1}) + \frac{1}{2}\rho_h^{n+1}\left|\boldsymbol{v}_h^{n+1}\right|^2 + \frac{\gamma}{2}\left|\boldsymbol{q}_h^{n+1}\right|^2 \, d\boldsymbol{x}$$
$$- \int_\Omega W(\rho_h^n) + \frac{1}{2}\rho_h^n\left|\boldsymbol{v}_h^n\right|^2 + \frac{\gamma}{2}\left|\boldsymbol{q}_h^n\right|^2 \, d\boldsymbol{x} = -\mu k_n B_h(\boldsymbol{v}_h^{n+\frac{1}{2}}, \boldsymbol{v}_h^{n+\frac{1}{2}}).$$

Proof. The proof is merely combining the results of Theorem 3.6, Corollary 3.7 and Theorem 4.1. $\qquad\square$

6. Numerical experiments

In this section we conduct a series of numerical experiments aimed at testing the robustness of the method. There are four experiments which investigate the behaviour of the discrete energy for the Euler–Korteweg (§6.4) and the Navier–Stokes–Korteweg systems (§6.5), benchmarking the algorithm against a travelling wave solution of the Euler–Korteweg system (§6.6), observing that there are no parasitic currents in long time simulations. Moreover, we conduct some simulations for $d = 2$ (§6.7).

In each of these experiments, except the second part of Test 1, we consider the fully discrete scheme (5.2) with the numerical fluxes given in Corollary 3.8.

6.1. Implementation issues.

6.1. **Implementation issues.** The numerical experiments were conducted using the DOLFIN interface for FEniCS [LW10]. The graphics were generated using Gnuplot and ParaView .

In each of the numerical experiments we fix W to be the following quartic double well potential

$$(6.1) \qquad W(\rho) = \frac{1}{4} (\rho - 1)^2 (\rho - 2)^2$$

with minima at $\rho = 1$ and $\rho = 2$.

6.2. *Remark* (the quotient of the double well). In the computational implementation we did not use the difference quotient $\frac{W(\rho^{n+1}) - W(\rho^n)}{\rho^{n+1} - \rho^n}$ appearing in (4.1) as it is ill-defined for $\rho^{n+1} = \rho^n$ and badly conditioned when $|\rho^{n+1} - \rho^n|$ is small. Instead we use a sufficiently high order approximation of this term. For (6.1) we use the following Taylor expansion representation

$$(6.2) \qquad \frac{W(\rho^{n+1}) - W(\rho^n)}{\rho^{n+1} - \rho^n} = W'(\rho^{n+\frac{1}{2}}) + \frac{1}{24} W'''(\rho^{n+\frac{1}{2}}) \left(\rho^{n+1} - \rho^n\right)^2,$$

which is exact. We note that when W is not polynomial a sufficiently high-order truncation of the Taylor expansion can be achieved such that the change in energy is of high order with respect to the timestep. This allows the construction of a method with arbitrarily small deviations of the energy with respect to the timestep.

6.3. *Remark* (numerical conservation). In each of the subsequent numerical experiments we assemble the discrete system (5.2) as a non-linear system of equations. The solution to the non-linear system was approximated by a Newton solver and on each Newton step the linear system of equations was approximated using a stabilised conjugate gradient solver with an incomplete LU preconditioner. Both solvers were set to a tolerance of 10^{-10}. Since the discretisations presented are implicit methods, the notion of conservation is only true up to a certain tolerance. In this case, the quantity will be conserved up to the tolerance of the solvers, 10^{-10}.

6.4. **Test 1: conservativity and dissipation for the Euler-Korteweg system.** In this case we take $\mu = 0$. We are then studying the conservativity property of the numerical method proposed for the Euler–Korteweg system in Corollary 3.8. We take $\Omega = [0, 1]$ and consider an initial condition given by a step function

$$(6.3) \qquad \rho_0(x) = \begin{cases} 1.1 & \text{if } x \leq 0.5, \\ 1.9 & \text{otherwise}, \end{cases} \qquad v_0 \equiv 0.$$

We take $\gamma = 10^{-4}$, $h = 10^{-4}$ and $k_n = k = 10^{-3}$ for each n. We consider the case when $\mathbb{V}$ consists of piecewise linear discontinuous functions, i.e., $p = 1$. Figure 2 shows the energy and mass conservativity of the simulation. In addition, we study the numerical scheme with the fluxes given in Example 3.10, chosen to dissipate a small amount of energy due to the non-smoothness of the initial conditions. Figure 3 shows the energy dissipation and mass conservation.

(A) Initial condition, $t = 0$

(B) $t = 0.1$

(C) $t = 0.05$

(D) $t = 0.1$

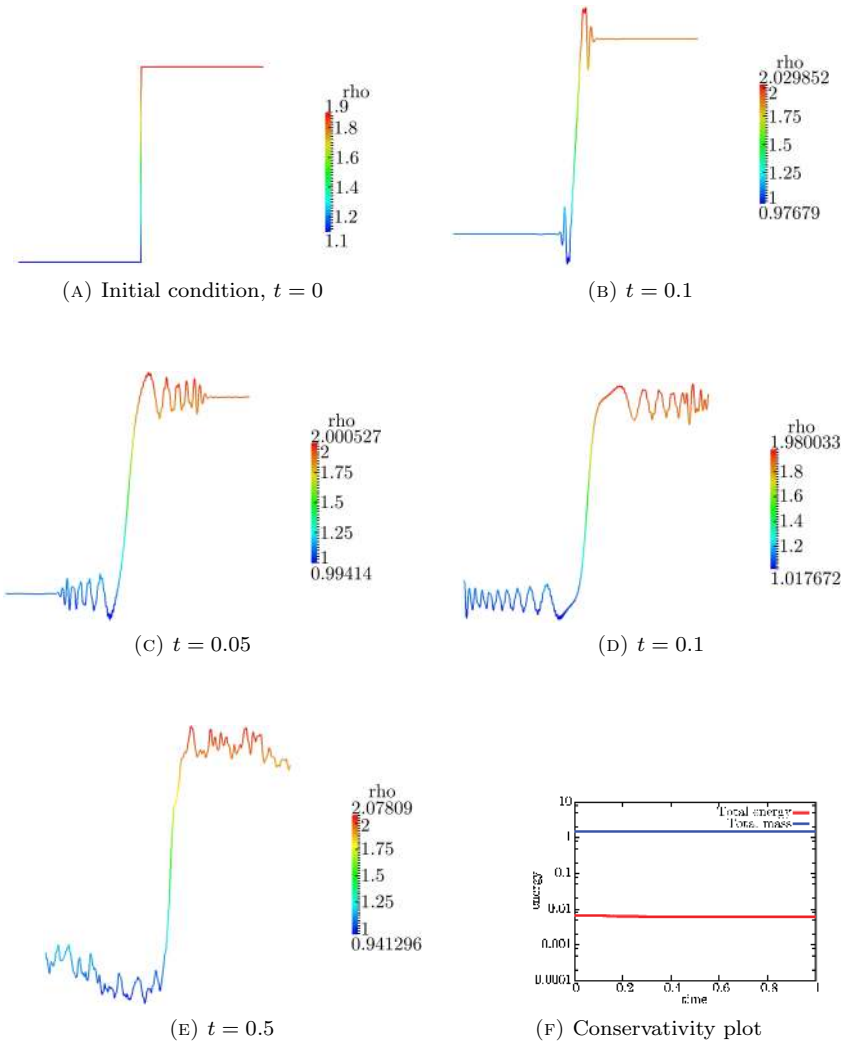(E) $t = 0.5$

(F) Conservativity plot

FIGURE 2. 6.4. Test 1: Numerical experiment showing the conservation of mass and energy for the numerical method proposed in Corollary 3.8 for the Euler–Korteweg system (i.e., $\mu = 0$). Due to the energy conservativity the Euler–Korteweg simulation will never achieve a steady state, the oscillations will continue to propagate.

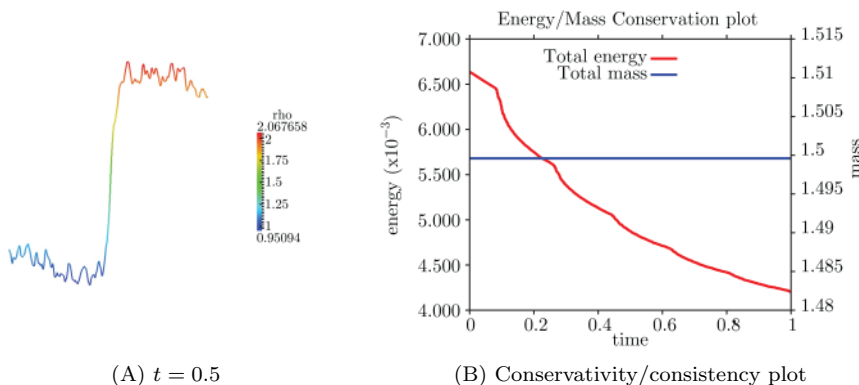(A) $t = 0.5$ (B) Conservativity/consistency plot

FIGURE 3. 6.4. Test 1: Numerical experiment showing the conservation of mass and dissipation of energy for the numerical method with fluxes given in Example 3.10 for the Euler–Korteweg system (i.e., $\mu = 0$). Due to the non-smoothness of the initial conditions we add a small amount of artificial viscocity, we take $\alpha = \beta = 10^{-5}$, as such, the oscillations will no longer propagate. We only show the solution plot at $t = 0.5$ here as they are very similar to those of Figure 2.

6.5. **Test 2: monotone energy dissipation for the Navier–Stokes–Korteweg system.** In this case we take $\mu > 0$ and study the dissipation property for the full Navier–Stokes–Korteweg system given by (5.2). We take $\Omega = [0, 1]$ and consider the initial conditions (6.3). We fix $\gamma = 10^{-4}, h = 10^{-4}$ and $k_n = k = 10^{-3}$. We consider the case when $p = 1$.

We test the effect of the ratio of viscocity to capillarity, i.e., μ/γ, on the dynamics of the simulation. To that end we run the simulation for $\mu = 10^{-7}$ (Figure 4), $\mu = 10^{-6}$ (Figure 5) and $\mu = 10^{-5}$ (Figure 6).

6.6. **Test 3: benchmarking.** In this test we look to benchmark the numerical algorithm against a steady state solution of the Euler–Korteweg system on the domain $\Omega = [-1, 1]$.

For the double well given by (6.1) a steady state solution to the Euler–Korteweg system is given by

$$(6.4) \qquad \rho(x, t) = \frac{3}{2} - \frac{1}{2} \tanh\left(\frac{x}{2\sqrt{2\gamma}}\right), \qquad v(x, t) \equiv 0 \quad \forall\, t$$

with appropriate initial data. Note that on the boundary $\nabla \rho$ is not zero but of negligable value (for small values of γ). Tables 1–3 detail three experiments aimed at testing the convergence properties for the scheme for $\gamma = 10^{-4}$ (Table 1), $\gamma = 10^{-5}$ (Table 2) and $\gamma = 10^{-6}$ (Table 3), using piecewise linear discontinuous elements ($p = 1$). In Table 4 we study convergence of the method for piecewise constant elements ($p = 0$) and in Table 5 we look at piecewise quadratic ($p = 2$).

(A) Initial condition, $t = 0$

(B) $t = 0.01$

(C) $t = 0.05$

(D) $t = 0.1$

(E) $t = 0.5$
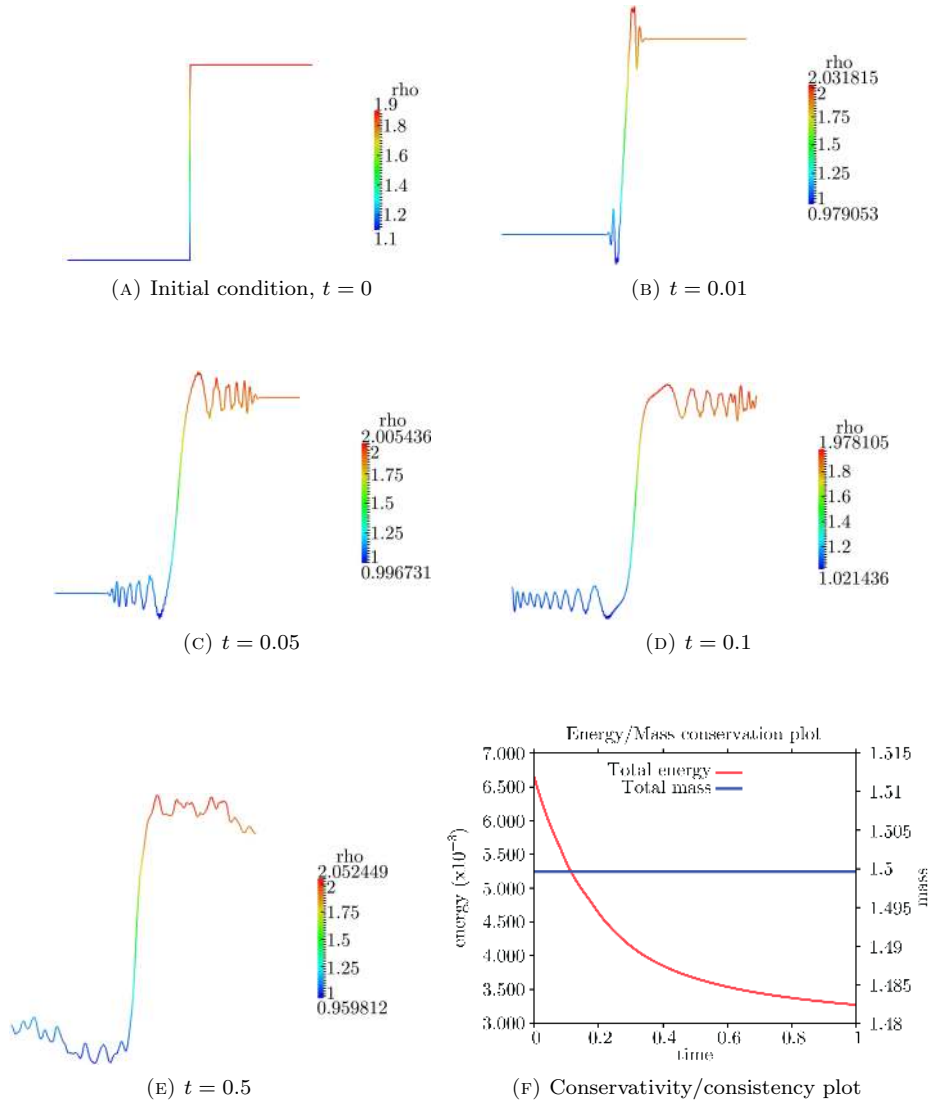
(F) Conservativity/consistency plot

FIGURE 4. 6.5. Test 2: Numerical experiment showing the conservation of mass and dissipation of energy for the numerical method proposed in Corollary 3.8 for the Navier–Stokes–Korteweg system. In this test we take $\mu = 10^{-7}$. Notice that the energy dissipation allows the Navier–Stokes–Korteweg simulation to achieve a steady state. At $t = 50$ the maximal value of the velocity is of magnitude 10^{-5}. Notice also that μ is chosen sufficiently small such that the dynamics are comparible with that of Figure 2 albeit with smeared out oscillations.

(A) Initial condition, $t = 0$

(B) $t = 0.01$

(C) $t = 0.05$

(D) $t = 0.1$
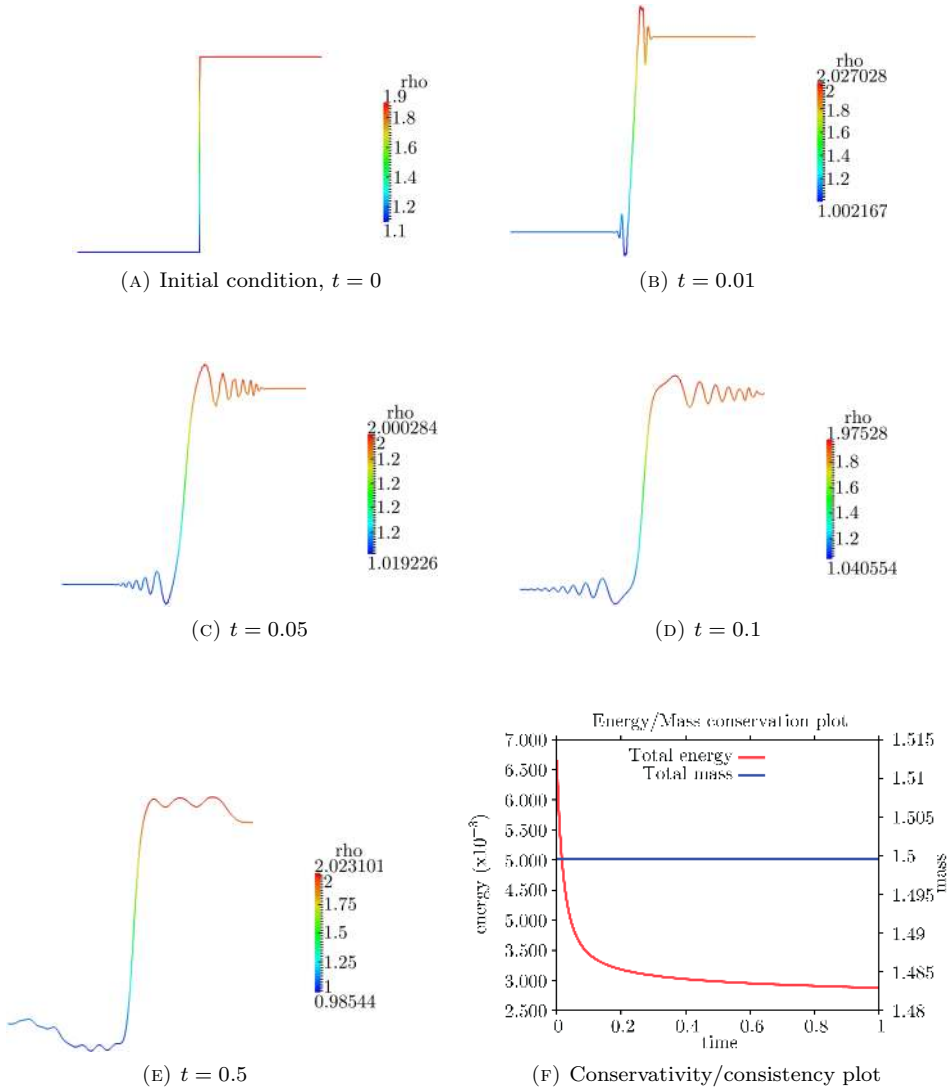
(E) $t = 0.5$

(F) Conservativity/consistency plot

FIGURE 5. 6.5. Test 2: Numerical experiment showing the effect of the ratio of viscocity to capillarity on the dynamics of the simulation. The simulation is the same as in Figure 4 with the exception that $\mu = 10^{-6}$. Notice the oscillations have become smeared out. The maximal value of velocity is of magnitude 10^{-5} at $t = 14$.

(A) Initial condition, $t = 0$

(B) $t = 0.01$

(C) $t = 0.05$

(D) $t = 0.1$

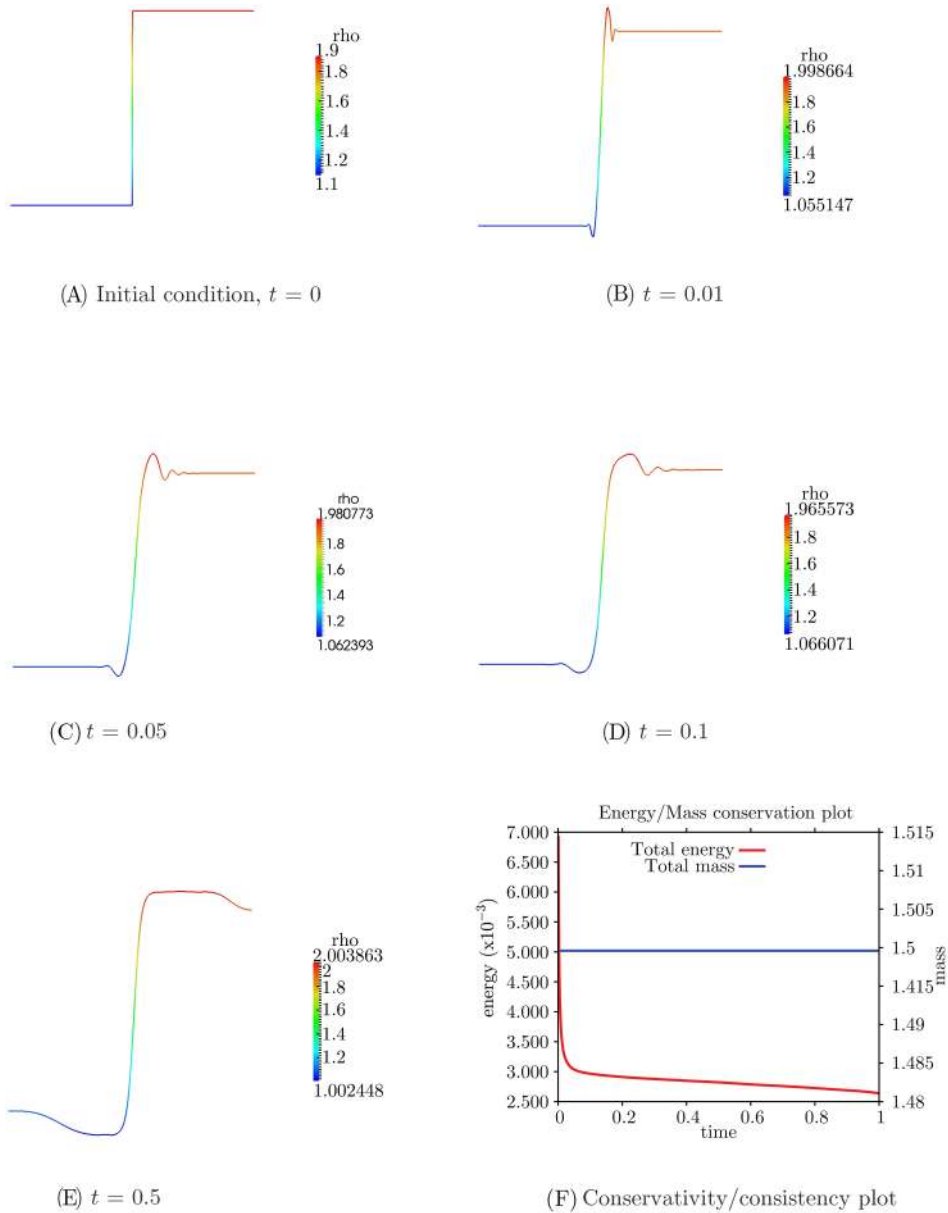(E) $t = 0.5$

(F) Conservativity/consistency plot

FIGURE 6. 6.5. Test 2: Numerical experiment showing the effect of the ratio of viscocity to capillarity on the dynamics of the simulation. The simulation is the same as in Figure 4 with the exception that $\mu = 10^{-5}$. Notice the oscillations have become heavily reduced in very short time due to the massive dissipation in energy initially. The maximal value of velocity is of magnitude 10^{-5} at $t = 2$.

TABLE 1. In this test we benchmark a stationary solution of the Euler–Korteweg system using the discretisation (5.2) with piecewise linear elements ($p = 1$), choosing $k = 1/N$. This is done by formulating (5.2) as a system of non-linear equations, the solution to this is then approximated by a Newton method with tolerance set at 10^{-10}. At each Newton step the solution to the linear system of equations is approximated using a stabilised conjugate gradient iterative solver with an successively overrelaxed preconditioner, also set at a tolerance of 10^{-10}. We look at the $L_\infty(0, T; L_2(\Omega))$ errors of the discrete variables ρ_h and v_h, and use $e_\rho := \rho - \rho_h$ and $e_v := v - v_h$. In this test we choose $\gamma = 10^{-4}$.

N	$\|e_\rho\|_{L_\infty(L_2)}$	EOC	$\|e_v\|_{L_\infty(L_2)}$	EOC
32	6.258e-3	0.000	6.194e-4	0.000
64	3.028e-4	4.369	4.631e-5	3.742
128	4.565e-5	2.730	1.105e-5	2.067
256	1.155e-5	1.983	3.691e-6	1.582
512	2.945e-6	1.972	9.916e-7	1.896
1024	7.368e-7	2.000	2.528e-7	1.972
2048	1.842e-7	2.000	6.324e-8	1.999
4096	4.605e-8	2.000	1.580e-8	2.009

TABLE 2. The test is the same as in Table 1 with the exception that we take $\gamma = 10^{-5}$.

N	$\|e_\rho\|_{L_\infty(L_2)}$	EOC	$\|e_v\|_{L_\infty(L_2)}$	EOC
32	7.017e-3	0.000	1.315e-3	0.000
64	2.469e-3	1.506	5.819e-4	1.176
128	4.411e-4	2.485	7.672e-5	2.923
256	2.885e-5	3.935	5.693e-6	3.752
512	6.5970-6	2.129	1.295e-6	2.136
1024	1.668e-6	1.984	3.228e-7	2.004
2048	4.161e-7	2.003	8.017e-8	2.010
4096	1.040e-7	2.001	2.001e-8	2.003

TABLE 3. The test is the same as in Table 1 with the exception that we take $\gamma = 10^{-6}$.

N	$\|e_\rho\|_{L_\infty(L_2)}$	EOC	$\|e_v\|_{L_\infty(L_2)}$	EOC
32	1.883e-2	0.000	1.488e-3	0.000
64	9.071e-3	1.054	8.134e-4	0.871
128	3.807e-3	1.253	3.820e-4	1.090
256	1.005e-3	1.922	9.110e-5	2.051
512	6.486e-5	3.954	5.118e-6	4.171
1024	4.907e-6	3.724	6.809e-7	2.910
2048	1.016e-6	2.272	1.445e-7	2.236
4096	2.439e-7	2.059	3.446e-8	2.068

TABLE 4. This test is the same as in Table 3 with the exception that we take $p = 0$.

N	$\|e_\rho\|_{L_\infty(L_2)}$	EOC	$\|e_v\|_{L_\infty(L_2)}$	EOC
32	4.167e-3	0.000	2.134e-4	0.000
64	2.083e-3	1.000	1.188e-4	0.845
128	1.039e-3	1.002	5.820e-5	1.029
256	6.464e-4	0.689	2.164e-5	1.427
512	5.610e-4	0.202	1.444e-5	0.584
1024	2.415e-4	1.199	9.165e-6	0.656
2048	1.116e-4	1.114	4.445e-6	1.044
4096	5.239e-5	1.091	2.246e-6	0.985

TABLE 5. This test is the same as in Table 3 with the exception that we take $p = 2$ and we take $k = 1/N^2$.

N	$\|e_\rho\|_{L_\infty(L_2)}$	EOC	$\|e_v\|_{L_\infty(L_2)}$	EOC
32	4.045e-3	0.000	5.316e-4	0.000
64	1.548e-3	1.386	2.334e-4	1.188
128	3.238e-4	2.257	6.219e-5	1.908
256	1.198e-5	4.031	8.116e-6	2.938
512	2.711e-6	2.144	8.643e-7	3.231
1024	3.710e-7	2.869	1.117e-7	2.884
2048	4.611e-8	3.008	1.414e-8	2.981
4096	5.639e-9	3.031	1.716e-9	3.043

6.7. Test 4: simulations for $d = 2$ and parasitic currents. In this test we consider the case $d = 2$. We take $\Omega = [0,1]^2$ and look at the following initial condition

$$(6.5) \qquad \rho_0(\boldsymbol{x}) = \begin{cases} 2 \text{ if } (x,y) \in [0.3, 0.7]^2, \\ 1 \text{ otherwise,} \end{cases} \qquad v_0 \equiv 0,$$

and examine its evolution.

We expect due to the non-local part of the energy that interfacial layers of size $\sim \sqrt{\gamma}$ form; see [Ste88, ORS90] for an energy argument. This process smoothes the profile. Moreover, the length of the interface is reduced such that the quadratic "droplet" becomes circular.

We take $\gamma = \mu = 0.0005$, $h \approx 0.02$ and $k_n = k = 0.001$ for all n. In addition, we take $p = 0$, that is, $\mathbb{V}$ consists of piecewise constant functions. Figure 7 shows the behaviour of the energy and mass of the numerical solution together with the solution plot of ρ_h at various times. The solution is overlaid with the velocity $\boldsymbol{v}_h$ as a glyph plot.

(A) $t = 0.001$

(B) $t = 0.1$

(C) $t = 0.25$

(D) $t = 0.5$

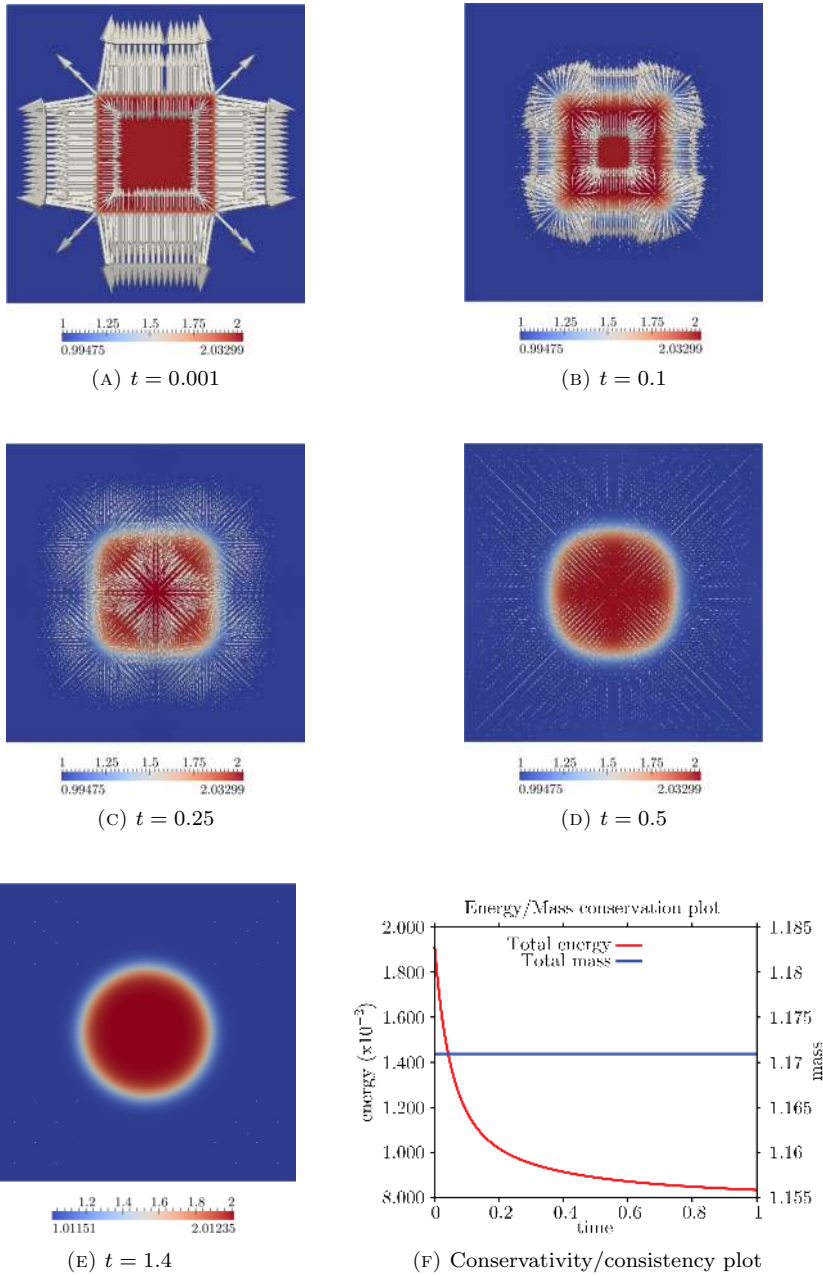(E) $t = 1.4$

(F) Conservativity/consistency plot

FIGURE 7. 6.7. Test 4: The solution, ρ_h to the Navier–Stokes–Korteweg system with initial conditions (6.5) at various values of t, overlayed with the velocity $\boldsymbol{v}_h$. Notice that there are no parasitic currents appearing in the interfacial layer. The velocity tends to zero over the entire domain as time increases. The energy-mass plot of the simulation is also given.

References

[AMW98]   D. M. Anderson, G. B. McFadden, and A. A. Wheeler, *Diffuse-interface methods in fluid mechanics*, Annu. Rev. Fluid Mech., vol. 30, Annual Reviews, Palo Alto, CA, 1998, pp. 139–165, DOI 10.1146/annurev.fluid.30.1.139. MR1609626 (99a:76048)

[BCKX11]   J.L. Bona, H. Chen, O.A. Karakashian, and Y. Xing. *Conservative, discontinuous-galerkin methods for the generalized Korteweg-de Vries equation*. Math. Comp., **82** (2013), no. 283, 1401-1432. MR3042569

[BDL03]   Didier Bresch, Benoît Desjardins, and Chi-Kun Lin, *On some compressible fluid models: Korteweg, lubrication, and shallow water systems*, Comm. Partial Differential Equations **28** (2003), no. 3-4, 843–868, DOI 10.1081/PDE-120020499. MR1978317 (2005b:76109)

[BGDDJ07] Sylvie Benzoni-Gavage, Raphaël Danchin, Stéphane Descombes, and Didier Jamet, *Stability issues in the Euler-Korteweg model*, Control methods in PDE-dynamical systems, Contemp. Math., vol. 426, Amer. Math. Soc., Providence, RI, 2007, pp. 103–127, DOI 10.1090/conm/426/08186. MR2311523 (2008k:35368)

[BP]   Malte Braack and Andreas Prohl, *Stable discretization of a diffuse interface model for liquid-vapor flows with surface tension*, ESAIM Math. Model. Num. Anal. **47** (2013), no. 2, 401-420. MR3021692. doi:10.1051/m2an/2012032.

[Cia78]   Philippe G. Ciarlet, *The finite element method for elliptic problems,* Studies in Mathematics and its Applications, Vol. 4,, North-Holland Publishing Co., Amsterdam, 1978. MR0520174 (58 #25001)

[CS08]   Yingda Cheng and Chi-Wang Shu, *A discontinuous Galerkin finite element method for time dependent partial differential equations with higher order derivatives*, Math. Comp. **77** (2008), no. 262, 699–730, DOI 10.1090/S0025-5718-07-02045-5. MR2373176 (2008m:65252)

[DD01]   Raphaël Danchin and Benoît Desjardins, *Existence of solutions for compressible fluid models of Korteweg type*, Ann. Inst. H. Poincaré Anal. Non Linéaire **18** (2001), no. 1, 97–133, DOI 10.1016/S0294-1449(00)00056-1 (English, with English and French summaries). MR1810272 (2002c:76081)

[DGR]   Katharina Degel, Jan Giesselmann, and Christian Rohde. Parasitic currents in compressible liquid vapour flow. in preparation.

[Die07]   D. Diehl. *Higher order schemes for simulation of compressible liquid–vapor flows with phase change.* PhD thesis, Universität Freiburg, 2007. http://www.freidok.uni-freiburg.de/volltexte/3762/.

[DS85]   J. E. Dunn and J. Serrin, *On the thermomechanics of interstitial working*, Arch. Rational Mech. Anal. **88** (1985), no. 2, 95–133, DOI 10.1007/BF00250907. MR775366 (86f:73051)

[Eva98]   Lawrence C. Evans, *Partial differential equations*, Graduate Studies in Mathematics, vol. 19, American Mathematical Society, Providence, RI, 1998. MR1625845 (99e:35001)

[Fei02]   Eduard Feireisl, *Compressible Navier-Stokes equations with a non-monotone pressure law*, J. Differential Equations **184** (2002), no. 1, 97–108, DOI 10.1006/jdeq.2001.4137. MR1929148 (2003h:35209)

[HL94]   Thomas Y. Hou and Philippe G. LeFloch, *Why nonconservative schemes converge to wrong solutions: error analysis*, Math. Comp. **62** (1994), no. 206, 497–530, DOI 10.2307/2153520. MR1201068 (94g:65093)

[HL96]   Harumi Hattori and Dening Li, *The existence of global solutions to a fluid dynamic model for materials for Korteweg type*, J. Partial Differential Equations **9** (1996), no. 4, 323–342. MR1426082 (98b:35145)

[JLCD01]   D. Jamet, O. Lebaigue, N. Coutris, and J. M. Delhaye, *The second gradient method for the direct numerical simulation of liquid-vapor flows with phase change*, J. Comput. Phys. **169** (2001), no. 2, 624–651, DOI 10.1006/jcph.2000.6692. MR1836527 (2002b:76078)

[JTB02]   D. Jamet, D. Torres, and J.U. Brackbill, *On the theory and computation of surface tension: The elimination of parasitic currents through energy conservation in the second-gradient method*, J. Comp. Phys, 182:262–276, 2002.

[Kor01] Diederick Johannes Korteweg, *Sur la forme que prennent les équations du mouvement des fluids si lón tient compte des forces capillaires causés par les variations de densité*, Arch. Néerl, 2:1–24, 1901.

[Kot08] Matthias Kotschote, *Strong solutions for a compressible fluid model of Korteweg type*, Ann. Inst. H. Poincaré Anal. Non Linéaire **25** (2008), no. 4, 679–696, DOI 10.1016/j.anihpc.2007.03.005 (English, with English and French summaries). MR2436788 (2009f:76143)

[LW10] Anders Logg and Garth N. Wells, *DOLFIN: automated finite element computing*, ACM Trans. Math. Software **37** (2010), no. 2, Art. 20, 28, DOI 10.1145/1731022.1731030. MR2738227 (2011i:65219)

[ORS90] N. C. Owen, J. Rubinstein, and P. Sternberg, *Minimizers and gradient flows for singularly perturbed bi-stable potentials with a Dirichlet condition*, Proc. Roy. Soc. London Ser. A **429** (1990), no. 1877, 505–532. MR1057968 (91j:49052)

[Ste88] Peter Sternberg, *The effect of a singular perturbation on nonconvex variational problems*, Arch. Rational Mech. Anal. **101** (1988), no. 3, 209–260, DOI 10.1007/BF00253122. MR930124 (89h:49007)

[Tad03] Eitan Tadmor, *Entropy stability theory for difference approximations of nonlinear conservation laws and related time-dependent problems*, Acta Numer. **12** (2003), 451–512, DOI 10.1017/S0962492902000156. MR2249160 (2007g:35150)

[TN92] C. Truesdell and W. Noll, *The nonlinear field theories of mechanics*, 2nd ed., Springer-Verlag, Berlin, 1992. MR1215940 (94c:73002)

[vdW] J. D. van der Waals, *On the continuity of the gaseous and liquid states*, Studies in Statistical Mechanics, XIV, North-Holland Publishing Co., Amsterdam, 1988. Translated from the Dutch; Edited and with an introduction by J. S. Rowlinson. MR930454 (89d:01080)

[XS11] Yan Xu and Chi-Wang Shu, *Local discontinuous Galerkin methods for the Degasperis-Procesi equation*, Commun. Comput. Phys. **10** (2011), no. 2, 474–508, DOI 10.4208/cicp.300410.300710a. MR2799652 (2012h:65233)

WEIERSTRASS INSTITUTE, MOHRENSTRASSE 39, D-10117 BERLIN, GERMANY
Current address: University of Stuttgart, Institute of Applied Analysis and Numerical Simulation, Pfaffenwaldring 57, 70569 Stuttgart, Germany
E-mail address: jan.giesselmann@mathematik.uni-stuttgart.de

DEPARTMENT OF APPLIED MATHEMATICS, UNIVERSITY OF CRETE, GR-71409 HERAKLION, GREECE – AND – INSTITUTE FOR APPLIED AND COMPUTATIONAL MATHEMATICS, FOUNDATION FOR RESEARCH AND TECHNOLOGY-HELLAS, VASILIKA VOUTON P.O. BOX 1527, GR-71110 HERAKLION, GREECE
Current address: Department of Mathematics, University of Sussex, Falmer Campus, Brighton BN1 9QH, United Kingdom

SCHOOL OF MATHEMATICS, STATISTICS & ACTUARIAL SCIENCE, UNIVERSITY OF KENT, CANTERBURY, GB-CT2 7NF, ENGLAND UNITED KINGDOM
Current address: Department of Mathematics and Statistics, University of Reading, Whiteknights, P.O. Box 220, Reading RG6 6AX, United Kingdom
E-mail address: t.pryer@reading.ac.uk