Nonlinear partial differential equations in fluid dynamics: interfaces, microstructure, and stability.

Antoine Remond-Tiedrez

Submitted in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

in

Mathematical Sciences

Carnegie Mellon University
Pittsburgh, Pennsylvania
May 2020

Acknowledgements

Ian Tice, thank you for your guidance.

Robert Pego, Franziska Weber, and Ming Chen, thank you for your time.

David Gutman, thank you for your support.

Eric, Isabelle, Juliette, and Gabriel Rémond-Tiedrez and Kayla Bollinger, thank you for everything.

To my many wonderful companions in studies in the mathematics department, in the Graduate Student Assembly, and in the Eberly Center: I am glad that I walked alongside all of you for a while.

Introduction

In this thesis we report work completed in the study of nonlinear partial differential equations, and in particular those arising from the physical world and fluid dynamics. We study two two augmented versions of the incompressible Navier-Stokes equations, studying microstructure and interfaces in particular.

In joint work completed with Ian Tice [RTTa] we study *anisotropic* micropolar fluids, and prove that their stability depends on the shape of the microstructure. In joint work completed with Ian Tice [RTTb] we consider the implications of bending energies being stored at a free interface, proving a stability result.

In both instances, the stability analysis is rendered difficult by the presence of partial differential equations of *mixed type*. This means that the system studied contains equations of parabolic, hyperbolic, and elliptic type. To wrest these systems of mixed type under control we employ a nonlinear energy method.

Guide to the reader. There are four introductions in this thesis: the present one and one for each of the three chapters comprising the thesis. We highly recommend starting there for (1) broader context regarding the physical effects that are taken into account and the state of research in the area and for (2) high-level discussions of the technical content of each chapter.

Background

Nonlinear partial differential equations. The work reported here lies squarely in the area of nonlinear partial differential equations (PDE), and in particular in the area of nonlinear PDE arising from fluid dynamics. The research questions currently focusing most attention in the area of nonlinear PDE from fluid dynamics are old ones: we seek to establish local well-posedness of the equation and to then characterize long-time behaviour. Proving local well-posedness here means proving that solutions exist, are unique, and depend continuously on the initial data.

Once local well-posedness is established, the question becomes whether or not the unique solution exists for all time. If it does not, then can we characterize the blow-up scenarios? If it does, then can we characterize the long-time asymptotic behaviour of the solution? In particular it is often appealing to study the long-time behaviour of solutions whose initial data lives in a neighbourhood of equilibria, i.e. solutions which are stationary in time. This study of the stability of equilibria is particularly important because physical systems are typically expected to spend most of their time in configurations close to equilibria, and because long-time behaviour "at large", i.e. for arbitrary initial data, may be particularly difficult.

More specifically, we report here work on nonlinear PDE arising from augmented versions of the incompressible Navier-Stokes equations. The augmentations have consisted in incorporating one of two additional physical effects: the presence of an *interface* between the fluid and its surrounding medium, or the presence of additional structure within the fluid at a microscopic scale, so-called *microstructure*.

The incompressible Navier-Stokes equations. Before diving into greater detail into the consequences of the incorporation of these physical effects, let us briefly recall what the incompressible Navier-Stokes equations are. They consist of a system of equations whose unknowns are $u : [0, T) \times \Omega \to \mathbb{R}^3$ and $p : [0, T) \times \Omega \to \mathbb{R}$, where u is a vector field describing the fluid's velocity, p is a scalar field corresponding to the fluid's pressure, $T > 0$ is some (possibly infinite) time horizon, and the domain $\Omega \subseteq \mathbb{R}^n$ is some open set. The system consists of the equations

$$\begin{cases} \partial_t u + (u \cdot \nabla)\, u = \Delta u - \nabla p & \text{in } \Omega \text{ and} \quad (0.1a) \\ \nabla \cdot u = 0 & \text{in } \Omega, \quad (0.1b) \end{cases}$$

which are coupled with the initial condition $u\,(t = 0) = u_0$ for some given u_0. If Ω is a bounded set then the equations must be supplemented with appropriate boundary conditions, such as for example the classical no-slip boundary condition where we impose that $u = 0$ on the boundary $\partial\Omega$.

What do these equations mean? The equation (0.1a) states that linear momentum is conserved by the motion of the fluid and the equation (0.1b) is the incompressibility condition which states that the flow is locally volume-preserving (i.e. the volume of any subset of Ω remains constant throughout the motion of the fluid). Note that the right-hand side of (0.1a) may be written as (the negative of) the divergence of the stress tensor $S(p, u) = pI - \mathbb{D}u$, where $\mathbb{D}u$ is the symmetrized gradient of u, defined as twice the symmetric part of the gradient of u. Physically, the stress tensor encodes how the fluid reacts to external forces – this interpretation will be important later when discussing free boundary problems. Long-time behaviour of the incompressible Navier-Stokes is notoriously difficult to analyze. A heuristic explanation as to why, which highlights a key feature recurrent in nonlinear PDE from fluid dynamics, is that although sufficiently regular solutions of the incompressible Navier-Stokes equation satisfy the equation

$$\frac{\mathrm{d}}{\mathrm{d}t}\int_{\Omega}|u|^2+\int_{\Omega}|\nabla u|^2=0,$$

these conserved quantities are not sufficiently strong to wrest control over the nonlinearity $(u\cdot\nabla)\,u$, at least in dimension $n=3$.

Microstructure

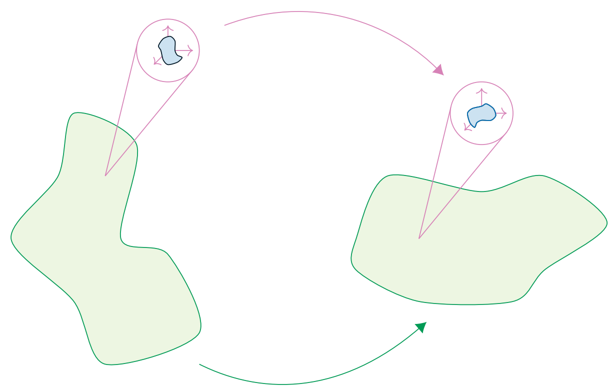


FIGURE 1. A micropolar continuum describes both the macroscopic motion of the continuum and the microscopic motion of the rigid microstructure. Micropolar continua are studied in Chapters 1 and 2

What do we mean by microstructure? Loosely speaking, fluids with microstructure correspond to a class of fluids for which a structure is present in the fluid, usually at a microscopic level, such that this structure has an impact on the overall dynamics of the fluid. A typical example of such a fluid is a liquid crystal, which is a fluid whose constituting particles are rigid rods. In particular, I have worked on *micropolar fluids*, introduced by Eringen [**Eri66**] which are fluids for which a rigid microstructure is postulated to be present at every point, where it is free to rotate, but not otherwise deform (see [**Eri99**, **Eri01**] for Eringen's reference 2-volume treatise on the matter) . This theory can be used to describe aerosols and colloidal suspensions such as those appearing in biological fluids [**Mau85**], blood flow [**Ram85**, **BBR$^+$08**, **MK08**], lubrication [**AK71**, **BŁ96**, **NS12**] and in particular the lubrication of human joints [**SSP82**], liquid crystals [**Eri66**, **LR04**, **GBRT13**], and ferromagnetic fluids [**NST16**].

To account for this rotating microstructure, the incompressible Navier-Stokes equations (0.1a)–(0.1b) are supplemented with two additional equations governing the dynamics of two additional unknowns.

- One unknown that needs to be taken into account is the moment of inertia of the microstructure, known as the micro-inertia, which essentially tracks the orientation of the microstructure. Its dynamics are governed by an additional equation, simply known as the conservation of micro-inertia.
- The other unknown to take into account is the angular velocity of the microstructure. The dynamics of the angular velocity are governed by the conservation of angular momentum.

It is important to note that the equations governing the dynamics of a micropolar fluid are derived from first principles. This means that the equations are derived from postulating that mass, micro-inertia, linear momentum, and angular momentum are conserved, and from postulating that the stress tensors depend linearly on the gradients of the fluid velocity and of the microstructure's angular velocity. Crucially, this is the same path as that which is followed to derive the Navier-Stokes equations (see for example [**Gur81**]), the only addition being the incorporation of the rigid microstructure.

Why do we care about microstructure? The interest in micropolar fluids is two-prong, coming from both physics and mathematics.

Physically, micropolar fluids are interesting because they are very common in the physical world. For example: milk, blood, and liquid crystals are all fluids which contain a microstructure (whether it be, respectively, fat molecules, hemoglobin, or the constituting molecules of the liquid crystal).

Mathematically, the study of *anisotropic* micropolar fluids is particularly enticing because, despite being vastly studied in the engineering and physics literature, anisotropic micropolar fluids are as of yet absent from the mathematical literature. Indeed, the mathematical endeavours to study micropolar fluids have so far been limited to the *isotropic* case, where the microstructure is taken to be spherical (see [Łuk99] for a reference on what is known about isotropic micropolar fluids). The study of anisotropic micropolar fluids is thus particularly enticing since it provides an opportunity for the development of new analytical techniques, especially in light of the difficulties engendered by the introduction of anisotropy.

Work reported here – Chapters 1 and 2. In joint work with Ian Tice we have studied the stability of anisotropic micropolar fluids under a constant microtorque (i.e. a constant torque acting on the microstructure) in a three-dimensional periodic domain, and where the microstructure is assumed to (essentially) have an axis of symmetry. The microstructure is thus rod-like or pancake-like depending on whether it is thin and long or wide and short, respectively. In this situation, a unique equilibrium exists corresponding to the microstructure rotating about its axis of symmetry, and the stability of this equilibrium depends on the shape of the molecule. We have proved that the equilibrium is unstable for rod-like microstructure [RTTa], and this result is report in Chapter 2.

In Chapter 1 we take care to carefully derive the equations of motion of micropolar fluids, following the path of rational continuum mechanics. Again, due to the scarcity of attention that anisotropic micropolar fluids are received in the mathematical literature we hope that this mathematically-minded introduction to the subject be of use to folks interested in studying this topic more robustly.

Interfaces and free boundary problems

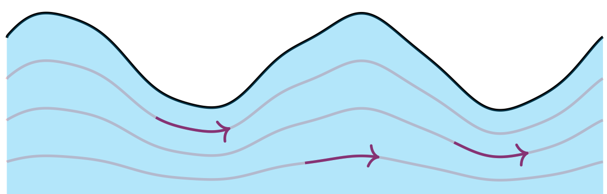


FIGURE 2. The geometry of the free boundary problem studied in Chapter 3

What do we mean by interfaces and free boundary problems? Free boundary problems arise when the domain on which the equations are solved changes with time, such that the dynamics of the domain are coupled to the unknowns. Physically, this occurs when a free interface between the fluid and its medium is present. Such situations are very common: the air–coffee interface in a cup of coffee, the elastic walls of blood vessels, the surface of the ocean, or the surface of a star are all examples of free boundaries on at least one side of which a fluid is present. Augmenting the incompressible Navier-Stokes equations (0.1a)–(0.1b) in order to account for a free boundary means supplementing that system with two equations: a dynamic boundary condition which accounts for the balance of stresses at the free boundary and a kinematic boundary condition which accounts for the fact that the free boundary is transported by the fluid's velocity normal to the interface.

An example of such a dynamic boundary condition would be $S\nu = -H\nu$ where ν denotes the unit normal to the free boundary and H denotes its mean curvature (i.e. the sum of its principal curvatures). If the free boundary is given as the graph of a function η then we may write the mean curvature as

$$H = \nabla \cdot \left(\frac{\nabla \eta}{\sqrt{1 + |\nabla \eta|^2}} \right) \tag{0.2}$$

This force $-H\nu$ arises as the surface force acting on the fluid due to surface tension. This is not surprising: surface tension appears because the fluid seeks to minimize the surface area of its fluid–medium interface, and the first variation of the surface area is precisely (the negative of) the mean curvature.

As one can now begin to see, introducing these additional boundary conditions introduces equations of mixed type into the system. Indeed: whilst the kinematic boundary condition is of hyperbolic type, by virtue of being a transport equation, the dynamic boundary condition is often of elliptic type.

Why do we care about interfaces and free boundary problems? Free boundary problems are particularly enticing because they are both physically and mathematically interesting. They are physically interesting, because they are frequently found in nature, and also because it is often the case that various physical effects are at play at interfaces. This is the case for example in biological models used to model the elastic lipid membranes surrounding living cells. They are also mathematically interesting for at least a couple of reasons.

- They are intrinsically appealing because they correspond to PDE whose domain changes and such that the dynamics of the domain is coupled with the unknowns.
- They are also of particular appeal given the current state of the research area. As evidenced by recent works [**Bea81**, **Bea84**, **BN85**, **GT13a**, **GT13b**, **NTY04**] who have made advances by developing new tools, free boundary problems are fertile ground for the development of new analytic techniques.

Work reported here – Chapter 3. In joint work with Ian Tice we have shown that, if the domain is three-dimensional and horizontally periodic with its free boundary given as the graph of a function, and if the physical effects taken into account at the interface are those coming from gravity, surface tension, and the presence of an elastic membrane at the interface, then the flat equilibrium is nonlinearly stable. This result is reported in Chapter 3.

Challenges

There are common challenges posed by the study of both free boundaries and micropolar fluids, and in particular when studying their stability.

First, let us recall that most nonlinear evolution PDE can be written as infinite-dimensional dynamical systems, where an evolution PDE is one of the form $\partial_t X = N(X)$ for some unknown function X and some operator N. By contrast with the finite-dimensional case, where it is known that linear stability implies nonlinear stability, this is not true in the infinite-dimensional case.

Moreover, both the study of free boundaries and of anisotropic micropolar fluids lead to the consideration of nonlinear PDE of *mixed type*. In order to solve equations of mixed type, whether it be in a moving domain (as is the case for free boundary problems) or in a fixed domain, there is no “off-the-shelf” theory that can be invoked to produce local well-posedness.

Finally: upon fixing their domains, free boundary problems give rise to *quasilinear* equations, and their analysis must therefore be carefully adapted to the nonlinearities, whilst some of the equations arising in the study of anisotropic micropolar fluids are *hypocoercive*, or equivalently can be thought of as equations that are degenerate parabolic, which means that their stability analysis is delicate since some components of the solution do not decay.

Nonlinear energy method. Overcoming these challenges is often achieved through a *nonlinear energy method*. As its name suggests, this method is inspired from traditional energy estimates. However it is important to contrast the two: whilst energy estimates are derived by using the *linear structure* of the equations considered, the nonlinear energy method leverages the *nonlinear structure* of the equations at hand. Put simply: energy estimates, obtained by integrating by parts, allow us to pass from the equations to “summary” equations involving functional norms of the solutions which provide bounds on these norms. Similarly, the nonlinear energy method relies on careful integration by parts (which must be respectful of the nonlinear structure) to obtain “summary” equations involving norm-like functionals which contain coefficients depending on the unknowns (this is due precisely to the nonlinear nature of the equations).

A sketch of the nonlinear energy method follows. Suppose we are considering a nonlinear evolution PDE of the form $\partial_t X = N(X)$ for some unknown function X and some operator N. If there exists non-negative functionals E and D of X satisfying the following "summary" equation,

$$\frac{\mathrm{d}}{\mathrm{d}t}E + D \leqslant 0,$$

which we call an energy-dissipation relation, then we are in business. This is referred to as a "summary equation" since it provides a succinct but coarse description of the dynamics: the energy E of the solutions is non-increasing in time, and spent through the dissipative mechanisms giving rise to the dissipation D. This relation immediately provides us with a priori estimates, as it tells us for example that the energy is non-decreasing in time. Moreover, we may deduce some decay information from the energy-dissipation relation provided that we can show a *coercivity* estimate of the form $E \leqslant CD$ (or $E \leqslant CD^\theta$ for some $0 < \theta < 1$ in the hypocoercive case), for some constant $C > 0$. Combining this estimate with the energy-dissipation relation tells us that the energy decays at an exponential rate (or at an algebraic rate dependent on θ in the hypocoercive case). The key challenge is due to the fact that the coercivity estimate does not hold for the natural energy and dissipation E and D above, but only holds for higher-order variants which are obtained by performing energy estimates on *differentiated* versions of the problem. Due to its nonlinear nature, the PDE usually does not behave well under differentiation, and hence this step introduces commutators. Taming these commutators is where modern tools come in, such as product and composition estimates in Sobolev spaces, interpolation theory, and high-low product estimates.

What is important to note about the nonlinear energy method is that it is robust and flexible. It is robust in the sense that it has been used in various areas, whether it be for problems arising from free boundaries in fluid dynamics or in the theory of self-gravitating stars. It is flexible in the sense that it can be used to prove stability and instability, even when the problem at hand is hypo-coercive and quasilinear.

Disclaimer

The work reported in Chapter 2 and Chapter 3 was first reported in [**RTTa**] and [**RTTb**], respectively. In particular, all work contained in those chapters is joint work with Ian Tice.

Contents

CHAPTER 1

A short treatise on micropolar continuum mechanics

ABSTRACT.

We derive the equations of motion of Newtonian incompressible homogeneous micropolar fluids by following the path of rational continuum mechanics. By contrast with classical fluids, micropolar fluids allows for the non-trivial behaviour of a rigid microstructure at the microscopic scale. This introduces an additional kinematic quantity, an additional conserved quantity, and an additional stress tensor responsible for the mediation of couples at the microscopic scale, namely the angular velocity and the microinertia of the microstructure and the couple stress tensor, respectively.

To be more precise, we derive the equations by postulating (1) the integral balance laws for conserved physical quantities such as mass, linear, and angular momentum, (2) the frame-invariance of the constitutive equations for the stress and couple-stress tensor, and (3) the satisfaction of the Onsager reciprocity relations.

Contents

Micropolar fluids were introduced by Eringen in [Eri66] as part of an effort to describe microcontinuum mechanics, which extend classical continuum mechanics by taking into account the effects of microstructure present in the medium. For viscous, incompressible continua, this results in a model in which the incompressible Navier-Stokes equations are coupled to an evolution equation for the rigid microstructure present at every point of the continuum. This theory can be used to describe aerosols and colloidal suspensions such as those appearing in biological fluids [Mau85], blood flow [Ram85, BBR+08, MK08], lubrication [AK71, BŁ96, NS12] and in particular the lubrication of human joints [SSP82], liquid crystals [Eri66, LR04, GBRT13], and ferromagnetic fluids [NST16].

In this chapter we carefully derive the equations governing micropolar fluids in the spirit of rational continuum mechanics, under the additional assumption that the fluid is incompressible and homogeneous.

Disclaimer. It is important to note that the derivation of the equations of motion for micropolar fluids is heavily inspired by the sections of [Eri99, Eri01] relevant to micropolar fluids. To a large degree, many portions of this chapter are more mathematically-minded reformulations of Eringen's original description of micropolar fluids.

1. Introduction

In this introduction we briefly sketch the derivation of the equations of motion of micropolar fluids, pointing out the relevant sections of this chapter where more details can be found regarding each step in the derivation. We will also highlight the differences between micropolar continua and classical continua throughout. This introduction will contain numerous links to later portion of this chapter, pointing to the precise definition of the notions discussed here.

The story of micropolar continuum mechanics, like that of classical continuum mechanics, begins with a *kinematic* description of the continuum. In the classical realm, this means that a continuum is fully determined by a flow map $\eta : [0, \infty) \to \mathbb{R}^n \to \mathbb{R}^n$ which determines the motion of the continuum. In the micropolar realm this description is supplemented by a microrotation map $Q : [0, \infty) \times \mathbb{R}^n \to SO(n)$ which assigns an orientation to every point in the micropolar continuum. This is illustrated in Figure 1. This kinematic description of continua and micropolar continua is detailed in Section 2, which also contains an elementary discussion of rigid motions and incompressible flows, two fundamental classes of continua. Once

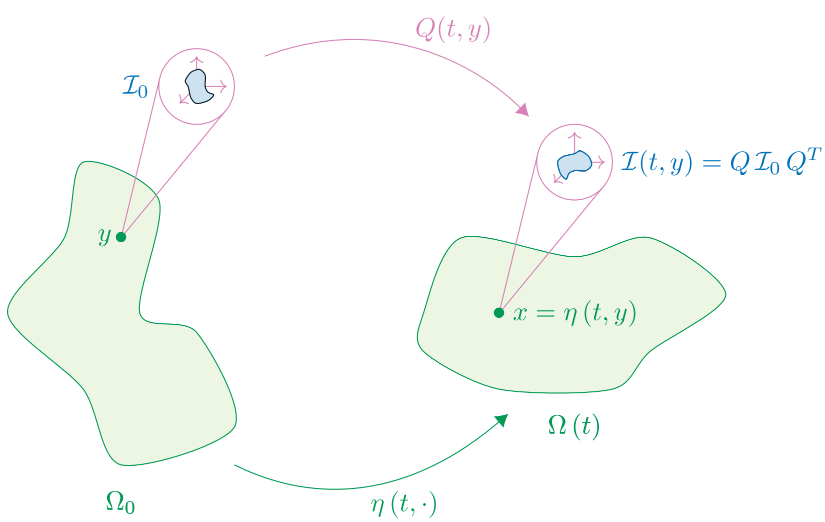


FIGURE 1. A depiction of how a subset $\Omega_0 \subseteq \mathbb{R}^n$ of the micropolar continuum behaves under the flow of η and Q. $\Omega\,(t) = \eta\,(t, \Omega_0)$ is the image of Ω_0 under the flow of η and $y \in \Omega_0$ is a point in Ω_0 at which the micropolar continuum has microinertia $\mathcal{I}_0$. At the point $x = \eta\,(t, y)$ the microinertia is $\mathcal{I}\,(t, y) = Q\,(t)\,\mathcal{I}_0\,Q^t\,(t)$ since the microinertia transforms as a 2-tensor under the flow of the microrotation Q.

a kinematic description of micropolar continua is established we throw physics into the mix. The objective, at this stage, is to postulate appropriate conservation laws for various physical quantities associated with the continuum. In order to do this we must first (1) carefully define these physical quantities and (2) carefully study them in the special case where the (classical) continuum is a rigid body. This is essential since the defining feature of micropolar continua is the presence of rigid microstructure. A good understanding of the physics of rigid bodies is therefore essential to ensure that the conservation laws which we will posit to hold for micropolar continua are physically sound. We define these physical quantities, namely mass, moment of inertia, linear momentum, and angular momentum in Section 3.

With these tools in hand we may now postulate appropriate conservation laws for micropolar continua. These conservation laws are formulated as integral balance laws and so the key point at the stage is to derive the corresponding local differential equations satisfied by the micropolar continuum. The formulation of the balance laws and the derivation of their local counterparts is carried out in Section 4. In particular, while classical continua are taken to conserve mass, linear momentum, and angular momentum, micropolar continua are posited to also conserve microinertia, which is defined to be the moment of inertia of the microstructure.

Note that throughout this treatment of micropolar fluids we deliberately avoid discussing boundary conditions. We make an exception in Section 4 to derive the natural boundary conditions that naturally arise from the conservation laws. This is done solely so that we may ultimately obtain a complete set of partial differential equations, and we do not focus on boundary effects in this chapter.

At this stage we already have in hand the complete set of equations satisfied by the micropolar continuum. However the balances of linear and angular momentum introduced two additional unknowns: the stress tensor T and the couple stress tensor M.

Physically, these tensors encode the response of the micropolar fluid to the forces and the torques induced by the neighbouring fluid. Note that only the stress tensor appears in classical fluids since in that case there is no microstructure that can support torques and hence no way for the fluid to apply torques to itself. Mathematically, these tensors render the system overdetermined, pending constitutive relations which determine T and M in terms of the dynamic variables. In the micropolar world there are two dynamic variables at play: the velocity u and angular velocity ω, which are essentially time-derivatives of the flow map and microrotation map, respectively.

To conclude we therefore impose three constraints on the stress tensor T and M: (1) they only depend on the dynamic variables and their gradients, and in a linear fashion (this is the Newtonian assumption), (2) their dependence on the dynamic variables is frame-invariant, i.e. independent of the frame of reference used to observe the fluid, and (3) they respect the Onsager reciprocity relations (which is a thermodynamical restriction). The constraints (1) and (2) are familiar from classical Newtonian fluids but (3) may not be, so we direct the reader's attention to Section 5.3 for a more detailed discussion of the Onsager reciprocity relations. The determination of the necessary forms of T and M given the constraints (1)–(3) is carried out in Section 5. We conclude Section 5 by recording the equations of motion of micropolar fluids in Corollary 5.21.

The last section of Chapter 1, namely Section 6, is a short appendix collecting various identities which are either well-known or elementary and which are used elsewhere in Chapter 1.

2. Kinematics

In this section we introduce the fundamental objects used to describe continua and micropolar continua, in Section 2.1 and Section 2.3 respectively. We also take the time to carefully discuss two important classes of continua, namely rigid motions and incompressible flows in Section 2.2.

2.1. Continua.

We begin with a discussion of continua, which are defined below in Definition 2.1. In this section we also record elementary results regarding derivatives of functions defined along the flow of a continuum.

Definition 2.1. (Continuum)
A *continuum* is a pair (Ω_0, η) where:
(1) Ω_0 is open, with $\partial\Omega_0$ Lipschitz, and is called a *reference configuration*.
(2) $\eta : [0,\infty) \times \Omega_0 \to \mathbb{R}^n$ is a map such that for every $t \geq 0$, $\eta_t := \eta(t, \cdot)$ is an orientation-preserving C^1-diffeomorphism onto its image, called a *flow map*.

We write $\Omega(t) := \eta_t(\Omega)$.

With the definition of a continuum in hand we can define what it means for functions and measures to be defined "along the flow". Put simply, and informally, a function is defined along the flow of a continuum (Ω, η) if its domain is $[0,\infty) \times \Omega(t)$. The precise definition is below in Definition 2.2.

DEFINITION 2.2. (Functions and measures defined 'along the flow')
Let (Ω_0, η) be a continuum.

(1) A map $f : [0,\infty) \times \Omega_0 \to X$ is a called a *Lagrangian function defined along the flow*.
(2) A collection of measures $\mu = (\mu_t)_{t \geqslant 0}$ such that for every $t \geqslant 0$, μ_t is a measure on Ω_0 is called a *Lagrangian measure defined along the flow*.
(3) We say that g is an *Eulerian function defined along the flow* if there exists a Lagrangian function f defined along the flow such that $g(t,x) = f\left(t, \eta_t^{-1}(x)\right)$, i.e. for every $t \geqslant 0$, $g(t,\cdot)$ is a map on $\Omega(t)$. We summarize this informally by writing $g = f \circ \eta^{-1}$.
(4) We say that ν is an *Eulerian measure defined along the flow* if there exists a Lagrangian measure μ defined along the flow such that $\nu_t = (\eta_t)_{\#}\mu_t$, i.e. for every $t \geqslant 0$, ν_t is a measure on $\Omega(t)$. We summarize this informally by writing $\nu = \eta_{\#}\mu$.

We now introduce the notion of Lagrangian and Eulerian coordinates. The motivation behind Lagrangian and Eulerian coordinates is that the former correspond to coordinates in the continuum's reference configuration Ω_0 whilst the latter correspond to coordinates in an observer's frame of reference where the complete history of the trajectory of each point of the continuum is not kept track of.

DEFINITION 2.3. (Lagrangian and Eulerian coordinates)
Let (Ω_0, η) be a continuum.

(1) $(t, y) \in [0,\infty) \times \Omega_0$ are called *Lagrangian coordinates*.
(2) $(t, x) \in [0,\infty) \times \Omega(t)$ are called *Eulerian coordinates*.

We continue this initial avalanche of definitions by introducing the velocity and acceleration of a continuum. These notions are absolutely fundamental for our purposes here: we seek equations of motion for micropolar fluids and one of these unknowns will be precisely the Eulerian velocity.

DEFINITION 2.4. (Velocity and acceleration)
Let (Ω_0, η) be a continuum.

(1) $v := \partial_t \eta : [0,\infty) \times \Omega_0 \to \mathbb{R}^3$ is called the *Lagrangian velocity*.
(2) $a := \partial_t v = \partial_t^2 \eta : [0,\infty) \times \Omega_0 \to \mathbb{R}^3$ is called the *Lagrangian acceleration*.
(3) $u : [0,\infty) \times \Omega(t)$, defined via, for every $t \geq 0$, $u_t := v_t \circ \eta_t^{-1}$, is called the *Eulerian velocity*.
(4) $b : [0,\infty) \times \Omega(t)$, defined via, for every $t \geq 0$, $b_t := a_t \circ \eta_t^{-1}$, is called the *Eulerian acceleration*.

In Lagrangian coordinates the acceleration a and the velocity v satisfy the familiar relation $a = \partial_t v$. This picture is slightly more complicated in Eulerian coordinates since the relationship between the Eulerian acceleration b and the Eulerian velocity u is given by $b = \partial_t u + (u \cdot \nabla) u$. Indeed we can verify from Proposition 2.6 below that

$$a = \partial_t v = ((\partial_t + u \cdot \nabla) u) \circ \eta$$

and hence $b = a \circ \eta^{-1} = (\partial_t + u \cdot \nabla) u$. This means that the Eulerian acceleration is the *material derivative* of the Eulerian velocity. Actually, the operators $f \mapsto \partial_t f + (u \cdot \nabla) f$ and its closely related cousin $f \mapsto \partial_t f + \nabla \cdot (f u)$, which we refer to as *material derivatives*, occur so often in continuum mechanics that they are given their own notation - see Definition 2.5 below.

DEFINITION 2.5. (Material derivatives)
Let T be a tensor field differentiable in both space and time and let u be a vector field. We define

$$D_t^u T := \partial_t T + (u \cdot \nabla) T \text{ and } \mathbb{D}_t^u := \partial_t T + \nabla \cdot (T \otimes u)$$

where we use the notation $(T \otimes u)_{i_1 \dots i_k j} := T_{i_1 \dots i_k} u_j$.

Now that we have introduce the fundamental objects of continuum mechanics we start recording some fundamental results associated with their derivatives. In particular in Proposition 2.6 below we record an identity for the temporal derivatives of functions and measures defined along the flow.

PROPOSITION 2.6. *(Derivatives along a flow)*
Let (Ω_0, η) be a continuum with Eulerian velocity field u.

(1) *(Derivatives of functions defined along the flow)*
For any $f : [0, \infty) \times \Omega(t) \to \mathbb{R}$*,*

$$\frac{\mathrm{d}}{\mathrm{d}t}(f \circ \eta) = \Big((\partial_t + u \cdot \nabla) f \Big) \circ \eta$$

(2) *(Derivatives of the volume form, i.e. of the Lebesgue measure, along the flow)*
For any Lebesgue-measurable $E_0 \subseteq \Omega_0$*, writing* $E(t) := \eta_t(E_0)$*, we have that*

$$\frac{\mathrm{d}}{\mathrm{d}t}\mathcal{L}^n(E(t)) = ((\nabla \cdot u)\, d\mathcal{L}^n)(E(t))$$

where the measure denoted $f d\mu$ *is defined via* $(f d\mu)(A) := \int_A f d\mu$*.*
Note that this boils down to the following equation, which is essentially a pointwise version of the equation above:

$$\frac{\mathrm{d}}{\mathrm{d}t} \det \nabla\eta = ((\nabla \cdot u) \circ \eta) \det \nabla\eta$$

Proof. (1) follows from an immediate computation upon recalling that $\partial_t \eta = u \circ \eta$:

$$\frac{\mathrm{d}}{\mathrm{d}t} f \circ \eta = \frac{\mathrm{d}}{\mathrm{d}t} f(\eta(y,t), t) = (\nabla f \circ \eta) \cdot \partial_t \eta + \partial_t f \circ \eta = ((\partial_t + u \cdot \nabla) f) \circ \eta$$

To derive (2), recall from Corollary 6.22 that $\det'|_{M_0}(A) = \det(M_0) \operatorname{tr}\left(M_0^{-1} A\right)$. Therefore

$$\frac{\mathrm{d}}{\mathrm{d}t} \det \nabla\eta = \det'|_{\nabla\eta}(\partial_t \nabla\eta) = \det(\nabla\eta) \operatorname{tr}\Big((\nabla\eta)^{-1} \partial_t \nabla\eta \Big)$$

where

$$\partial_t \nabla\eta \cdot (\nabla\eta)^{-1} = \nabla\partial_t\eta \cdot \left(\nabla\left(\eta^{-1}\right) \circ \eta\right) = \left(\left(\nabla\partial_t\eta \circ \eta^{-1}\right) \cdot \nabla\left(\eta^{-1}\right)\right) \circ \eta = \left(\nabla\left(\partial_t\eta \circ \eta^{-1}\right)\right) \circ \eta = \nabla u \circ \eta$$

and hence $\frac{\mathrm{d}}{\mathrm{d}t} \det \nabla\eta = \det(\nabla\eta) \operatorname{tr}(\nabla u \circ \eta) = ((\nabla \cdot u) \circ \eta) \det \nabla\eta$. So finally

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}t}\mathcal{L}^n(E(t)) = \frac{\mathrm{d}}{\mathrm{d}t}\int_{E(t)} d\mathcal{L}^n = \frac{\mathrm{d}}{\mathrm{d}t}\int_{E_0} \det(\nabla\eta) d\mathcal{L}^n = \int_{E_0} (\nabla \cdot u) \circ \eta \det(\nabla\eta) d\mathcal{L}^n = \int_{E(t)} (\nabla \cdot u)\, d\mathcal{L}^n \\ = ((\nabla \cdot u)\, d\mathcal{L}^n)(E(t))\end{aligned}$$

□

Using Proposition 2.6 above we are now equipped to prove a result which will be absolutely essential when it comes to deriving local versions of the integral balance laws. Indeed, Theorem 2.7 below tells us precisely how to "push time derivatives inside integrals defined over domains carried by the flow".

Theorem 2.7. *(Reynolds' transport theorem)*
Let $\mathcal{U}_0 \subseteq \mathbb{R}^n$ *be open and let* $\eta : [0, \infty) \times \mathcal{U}_0 \to \mathbb{R}^n$ *be a map such that, for every* $t \geq 0$*,* $\eta_t := \eta(t, \cdot)$ *is an orientation-preserving* C^1*-diffeomorphism. Define* $\mathcal{U}(t) := \eta_t(\mathcal{U}_0)$*, and* $u_t := \partial_t \eta_t \circ \eta_t^{-1}$ *for every* $t \geqslant 0$*. For any sufficiently regular* $f : [0, \infty) \times \mathcal{U}(t) \to \mathbb{R}$*,*

$$\frac{\mathrm{d}}{\mathrm{d}t}\int_{\mathcal{U}(t)} f = \int_{\mathcal{U}(t)} \partial_t f + \nabla \cdot (f u).$$

Proof. Equipped with Proposition 2.6, this is a direct computation

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}t}\int_{\mathcal{U}(t)} f = \frac{\mathrm{d}}{\mathrm{d}t}\int_{\mathcal{U}_0} (f \circ \eta) \det \nabla\eta = \int_{\mathcal{U}_0} ((\partial_t + u \cdot \nabla) f \circ \eta) \det \nabla\eta + f((\nabla \cdot u) \circ \eta) \det \nabla\eta \\ = \int_{\mathcal{U}(t)} \partial_t f + (\nabla f) \cdot u + f(\nabla \cdot u) = \int_{\mathcal{U}(t)} \partial_t f + \nabla \cdot (f u).\end{aligned}$$

□

We can interpret Proposition 2.6 and Theorem 2.7 above as telling us how to differentiate 0-forms and n-forms, i.e. smooth functions and volume forms respectively, in terms of the material derivatives introduced in Definition 2.5. More precisely: for a 0-form f and its associated n-form $f d\mathcal{L}^n$, Proposition 2.6 and Theorem 2.7 tell us that, if $\mathcal{U}(t)$ denotes a subset of $\Omega(t)$ carried by the flow (i.e. $\mathcal{U}(t) = \eta_t(\mathcal{U}_0)$ for some $\mathcal{U}_0 \subseteq \Omega_0$) then

$$\partial_t(f \circ \eta) = (D_t^u f) \circ \eta \text{ and } \partial_t\left(f d\mathcal{L}^n|_{\mathcal{U}(t)}\right) = (\mathbb{D}_t^u f)\, d\mathcal{L}^n|_{\mathcal{U}(t)}.$$

In particular, note that when $\nabla \cdot u$ vanishes (i.e. the flows is incompressible – see Definition 2.12 below) then the two material derivatives agree, i.e. $D_t^u = \mathbb{D}_t^u$, and therefore 0-forms and n-forms behave in the same way when differentiated in time along the flow. This is a manifestation of a feature of incompressible flows that we will prove below (in Proposition 2.15), namely that incompressible flows preserve the Lebesgue measure.

2.2. Important classes of continua: rigid motions and incompressible flows. In this section we introduce rigid motions and incompressible flows, which are two fundamental examples of continua. We also establish some of their properties, and in particular will establish the implications laid out in the diagram below:

$$\begin{tikzcd}
\begin{array}{c}\text{Isometric flow}\\ \text{Preserves the metric}\end{array} \arrow[r, leftrightarrow] \arrow[d] & \text{Rigid motion} \arrow[r, leftrightarrow] & u \in \ker \mathbb{D} \arrow[d] \\
\begin{array}{c}\text{Locally volume-preserving flow}\\ \text{Preserves the Lebesgue measure}\end{array} \arrow[rr, leftrightarrow] & & \nabla \cdot u = 0 \ \ \text{(incompressible flow)}
\end{tikzcd}$$

Rigid motions and rigid bodies, introduced in Definition 2.8 below, are important for several reasons.

(1) They are the simplest class of continua that one can study.
(2) They motivate and help to illustrate numerous definitions in the realm of micropolar continua, such as the definitions of linear momentum and angular momentum. Indeed, the microstructure of micropolar fluids is posited to be a microscopic rigid body, so a good understanding of rigid bodies is important in motivating the defining properties of micropolar media.
(3) They are necessary to define the notion of frame-invariance, which plays in essential role in the determination of the equations of motion of micropolar fluids.

Definition 2.8. (Rigid motion and rigid body)

(1) We say that $f : \mathbb{R}^n \to \mathbb{R}^n$ is a *rigid motion* if $f = R + z$ for some $R \in O(n)$ and $z \in \mathbb{R}^n$. Note that f is orientation-preserving if and only if $R \in SO(n)$.
(2) We say that a flow map $\eta : [0, \infty) \times \mathbb{R}^n \to \mathbb{R}^n$ is a *rigid motion* if η_t is a rigid motion for every $t \geq 0$.
(3) A continuum whose flow map is a rigid motion is called a *rigid body*.

The first property of rigid motions we establish is that rigid motions are precisely the isometries of $\mathbb{R}^n$. This also serves as a justification for the central important of rigid motions.

Proposition 2.9. *Rigid motions are precisely the isometries of Euclidean space, and moreover isometries of Euclidean space are necessarily bijective.*

Proof. First we show that rigid motions are isometries. Let $f = R + z$ be a rigid motion. Then, for every $x, y \in \mathbb{R}^n$, $|f(x) - f(y)| = |Rx - Ry| = |x - y|$ since R is orthogonal and hence

$$|R(x-y)| = (R(x-y) \cdot R(x-y))^{1/2} = \left(R^T R(x-y) \cdot (x-y)\right)^{1/2} = |x-y|$$

such that indeed f is an isometry.

Now we show that isometries must rigid motions. We proceed in several steps. In step 1 we use a polarization identity to show that isometries preserve inner products (i.e. angles). In step 2 we deduce that isometries must therefore be additive. In step 3 we note that isometries must be continuous and injective. In step 4 we deduce from the additivity and the continuity of the isometry that it must be linear. We finally conclude in step 5.

Step 1. Let f be an isometry of $\mathbb{R}^n$. We will use a polarization identity, which allows us to relate inner products and norms, to show that f also preserves the inner product. Suppose without loss of generality that $f(0) = 0$. Then, for any $x, y \in \mathbb{R}^n$,

$$\begin{aligned} x \cdot y = \frac{1}{2}\left(|x|^2 + |y|^2 - |x-y|^2\right) &= \frac{1}{2}\left(|x-0|^2 + |y-0|^2 - |x-y|^2\right) \\ &= \frac{1}{2}\left(|f(x) - f(0)|^2 + |f(y) - f(0)|^2 - |f(x) - f(y)|^2\right) \\ &= \frac{1}{2}\left(|f(x)|^2 + |f(y)|^2 - |f(x) - f(y)|^2\right) = f(x) \cdot f(y). \end{aligned}$$

Step 2. We now show that f is additive. For any $x, y, z \in \mathbb{R}^n$,

$$\begin{aligned}|(x+y)-z|^2 &= |x|^2 + 2x\cdot y + |y|^2 - 2\left(x\cdot z + y\cdot z\right) + |z|^2 \\ &= |f(x)|^2 + 2f(x)\cdot f(y) + |f(y)|^2 - 2\left(f(x)\cdot f(z) + f(y)\cdot f(z)\right) + |f(z)|^2 \\ &= |(f(x)+f(y)) - f(z)|\end{aligned}$$

and thus in particular if we pick $z = x+y$ we obtain that

$$f(x) + f(y) - f(x+y) = 0$$

Step 3. We now note that, since f is an isometry, it must be a continuous injection. For every $x \neq y$, $x, y \in \mathbb{R}^n$ $|f(x) - f(y)| = |x-y| > 0$ and hence $f(x) \neq f(y)$. Continuity is immediate since isometries are 1-Lipschitz.

Step 4. We now show that, since f is additive and continuous, it must be 1-homogeneous, and hence linear/. Let f be additive and continuous. Then, for any $x \in \mathbb{R}^n$, and for any $z \in \mathbb{Z}$, $f(zx) = zf(x)$, and hence $zf\left(\frac{x}{z}\right) = f\left(z\frac{x}{z}\right) = f(x)$. It follows that $f\left(\frac{x}{z}\right) = \frac{1}{z}f(x)$, and thus $f(qx) = qf(x)$ for any rational q. Finally, by continuity of f and density of $\mathbb{Q}$ in $\mathbb{R}$, we obtain that $f(\lambda x) = \lambda f(x)$ for any $\lambda \in \mathbb{R}$. So indeed f is 1-homogeneous. Since it is also additive, it must be linear.

Step 5. We may now conclude the argument and show that f is a rigid motion. We know that f is linear and injective. It is therefore bijective (since the dimensions of the domain and codomain of f coincide). So let us write $f(x) = Ax$ for some n-by-n matrix A. We note that, for arbitrary $x, y \in \mathbb{R}^n$,

$$x\cdot y = f(x)\cdot f(y) = Ax\cdot Ay = A^TAx\cdot y$$

and hence A is indeed an orthogonal matrix. □

We now introduce the symmetrized gradient. Due to its connections with rigid motions with frame-invariance, this differential operator appears throughout our treatment of micropolar continuum mechanics.

DEFINITION 2.10. (Symmetrized gradient)
For any differentiable vector field v we define its *symmetrized gradient*, denoted $\mathbb{D}v$, to be $(\mathbb{D}v)_{ij} = \partial_i v_j + \partial_j v_i$, i.e. $\mathbb{D}v = 2\,\mathrm{Sym}\,\nabla v$.

Having defined the symmetrized gradient we show that it can be used to characterize the Eulerian velocities of rigid motions.

PROPOSITION 2.11. *A flow map is a rigid motion if and only if its Eulerian velocity belongs to the kernel of the symmetrized gradient.*

PROOF. $\boxed{\Rightarrow}$ Suppose $\eta(t,y) = z(t) + R(t)\,y$ for some $z : [0,\infty) \to \mathbb{R}^n$ and some $R : [0,\infty) \to O(n)$. Note that $\eta^{-1}(t,x) = R(t)\,(x - z(t))$, and therefore:

$$u(t,x) = \left(\partial_t \eta \circ \eta^{-1}\right)(t,x) = \dot{z}(t) + \dot{R}(t)\,R(t)^{-1}\,(x - z(t))$$

i.e. $u(t,\cdot\,) = v + \Omega$ for

$$\begin{cases} v = \dot{z} - \dot{R}R^{-1}z \text{ and} \\ \Omega = \dot{R}R^{-1}. \end{cases} \tag{2.1}$$

Now note that, by Proposition 6.18, $\dot{R} = RB$ for some $B \in \Omega(n)$, and hence by Proposition 6.19 $\Omega = RBR^{-1} \in A(n)$. So finally, by Lemma 6.6, $u \in \ker \mathbb{D}$.

$\boxed{\Leftarrow}$ This direction of the proof amounts to solving the ODE (2.1), treating (Ω, v) as data and (R, z) as unknowns, with initial conditions $z(0) = 0$ and $R(0) = I$ (to ensure that $\eta_0 = \mathrm{id}$). So let $u \in \ker \mathbb{D}$, i.e by Lemma 6.6, $u(t,x) = v(t) + \Omega(t)\,x$ for some $v : [0,\infty) \to \mathbb{R}^n$ and some $\Omega : [0,\infty) \to \mathrm{Skew}\left(\mathbb{R}^{n\times n}\right)$. Then define, for $t \geq 0$,

$$z(t) = \int_0^t e^{\Omega(t-s)}v(s)\,ds \text{ and } R(t) = e^{\Omega t},$$

i.e. $z(t) = R(t)\int_0^t R(s)^{-1}v(s)\,ds$. Upon taking a derivative we see that

$$\dot{z} = RR^{-1}v + \dot{R}\int R^{-1}v = v + \dot{R}R^{-1}z \text{ and } \dot{R} = \Omega R$$

such that (2.1) holds indeed. □

Having discussed rigid motions we now introduce the second of the two classes of continua that we consider in this section, namely incompressible flows.

DEFINITION 2.12. (Incompressible flow)
We say that a flow map η is *incompressible* if its Eulerian velocity is divergence-free.

We now define what it means for a flow to be locally volume-preserving, a notion which will play with incompressible flows the role that isometries played with rigid motions.

DEFINITION 2.13. (Locally volume-preserving maps and flows)

(1) Let $E \subseteq \mathbb{R}^n$ be Lebesgue measurable. A map $f : E \to \mathbb{R}^n$ is said to be *locally volume-preserving* if $f_{\#}\mathcal{L}^n = \mathcal{L}^n$.
(2) A flow map η is said to be *locally volume-preserving* if, for every $t \geqslant 0$, η_n is locally volume-preserving, i.e. $(\eta_t)_{\#}\mathcal{L}^n = \mathcal{L}^n$.

Having introduced locally volume-preserving flows we record a few different equivalent characterizations of such flows. This will be helpful later when discussing the relationships between locally volume-preserving flows, incompressible flows, and isometries.

LEMMA 2.14. *(Alternative characterizations of locally volume-preserving flow maps)*
Let η be a flow map. The following are equivalent:

(1) *η is locally volume-preserving.*
(2) *For every Lebesgue-measurable set $E_0 \subseteq \Omega_0$, writing $E(t) := \eta_t(E_0)$, we have that*
$$\mathcal{L}^n(E(t)) = \mathcal{L}^n(E_0).$$
(3) $|\det \nabla\eta| \equiv 1$.

PROOF. $\boxed{(1) \Leftrightarrow (2)}$ This follows from the observation that, since η_t is a bijection for every $t \geq 0$,
$$\mathcal{L}^n(E_0) = \mathcal{L}^n\left(\eta_t^{-1}(E(t))\right) = \left((\eta_t)_{\#}\mathcal{L}^n\right)(E(t)).$$
$\boxed{(2) \Leftrightarrow (3)}$ This follows from the observation that
$$\mathcal{L}^n(E(t)) = \int_{E(t)} d\mathcal{L}^n = \int_{E_0} |\det \nabla\eta| d\mathcal{L}^n = \left(|\det \nabla\eta| d\mathcal{L}^n\right)(E_0).$$
□

We now show an analog of Proposition 2.9 in the realm of incompressible flows.

PROPOSITION 2.15. *A flow map is locally volume-preserving if and only if it is incompressible.*

PROOF. This follows from combining one of the alternative characterization of locally volume-preserving flows in Lemma 2.14 which says that a flow map η is locally volume-preserving if and only if $|\det \nabla\eta| \equiv 1$, with the computation of the time derivative of the volume form in Proposition 2.6, which tells us that $\frac{\mathrm{d}}{\mathrm{d}t} \det \nabla\eta = (\nabla \cdot u)\circ\eta \, \det \nabla\eta$. Note that we also need to use the fact that $\eta_0 = \mathrm{id}$, and hence $\det \nabla\eta_0 \equiv 1$. □

Finally we conclude this section by remarking on the relationship between locally volume-preserving isometric flow maps.

PROPOSITION 2.16. *Isometric flow maps are locally volume-preserving.*

PROOF. Since η is an isometry, it follows from Proposition 2.9 that it is a rigid motion, i.e.
$$\eta(t, x) = z(t) + R(t)\,x$$
for some $z : [0, \infty) \to \mathbb{R}^n$ and some $R : [0, \infty) \to O(n)$. In particular, $|\det \nabla\eta| = |\det R| = |\pm 1| = 1$, which by Lemma 2.14 tells us precisely that η is locally volume-preserving. □

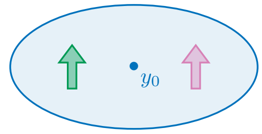
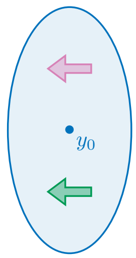
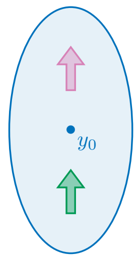
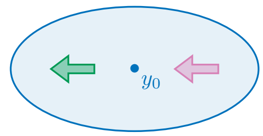
	Case 1	Case 2	Case 3
Configuration at time $t = 0$			
$\eta(t, \cdot)$	$e^{tR}(\cdot - y_0)$	$e^{tR}(\cdot - y_0)$	I
$Q(t, y)$	e^{tR}	I	e^{tR}
$u(t, x)$	Rx	Rx	0
$\frac{1}{2}\nabla \times u(t, x)$	e_3	e_3	0
$\omega(t, x)$	e_3	0	e_3
$\frac{1}{2}\nabla \times u - \omega$	0	e_3	$-e_3$
Configuration at time $t = \frac{\pi}{2}$			

Table 1. Three explicit examples of the motion of a micropolar continuum with the same initial configuration. These motions are chosen to be similar to emphasize that the microrotation Q is an *absolute* rotation. The figures shown correspond to cross-sections perpendicular to e_3, each colored arrow is a depiction of the orientation of the microstructure at that point, y_0 is some point in the micropolar continuum, and $R = e_2 \otimes e_1 - e_1 \otimes e_2$ corresponds to a (counter-clockwise) rotation by $\pi/2$ in the plane perpendicular to e_3.

2.3. Micropolar continua. In this section we follow a path similar to that which we took in Section 2.1 earlier: we define the fundamental kinematic objects used to describe micropolar continua. This begins with the definition of a micropolar continuum.

Definition 2.17. (Micropolar continuum and microrotation map)
A *micropolar continuum* is a triple (Ω_0, η, Q) where
(1) (Ω_0, η) is a continuum.
(2) $Q : \Omega_0 \times [0, \infty) \to SO(n)$ is called a *microrotation map*.

A word of warning: there are two ways to define the microrotation map and we have chosen here the convention that Q is *absolute*. Indeed, one may either define Q to be the rotation of the microstructure with respect to its immediate environment, in which case Q would be equal to the identity when the micropolar continuum undergoes rigid motions such as rotations, or one may define Q to be the identity at time $t = 0$ and to be the absolute rotation underwent by the micropolar continuum thereafter. We choose the latter convention. In order to illustrate the physical interpretation of the microrotation map Q, Table 1 contrasts the motions obtained for various simple expressions of η and Q

We now introduce two linear maps, ten and vec, which will play a fundamental role throughout.

Definition 2.18. (ten and vec)
We define ten : $R^3 \to \mathrm{Skew}(3)$ and vec : $\mathrm{Skew}(3) \to \mathbb{R}^3$ via: for every $v \in \mathbb{R}^3$ and every $A \in \mathrm{Skew}(3)$,

$$(\mathrm{ten}\, v)_{ij} := \epsilon_{iaj} v_a \text{ and } (\mathrm{vec}\, A)_i := \frac{1}{2}\epsilon_{aib} A_{ab}.$$

The linear maps ten and vec are essential since they allows us, in light of Proposition 6.13, to identify $\mathbb{R}^3$ with Skew(3), the space of 3-by-3 skew-symmetric matrices. Quantities like angular velocity and angular momentum, that would naturally take the form of a skew-symmetric matrix (since they arise from the rotational invariance of physical system and, as noted in Proposition 6.19, the tangent space to the space of orthogonal matrices is precisely the space of skew-symmetric matrices), will thus be treated as vectors.

We conclude this section with the analog of Definition 2.4 for the micropolar realm and introduce the dynamic quantities that can be used to describe the motion of the microstructure of micropolar continua.

Definition 2.19. (Angular velocity and angular velocity tensor)
Let (Ω_0, η, Q) be a micropolar continuum.
(1) $\Theta := (\partial_t Q)\, Q^{-1} : \Omega_0 \times [0,\infty) \to A\,(n)$ is called the *Lagrangian angular velocity tensor*.
(2) if $n = 3$, $\theta := \operatorname{vec}\Theta : \Omega_0 \times [0,\infty) \to \mathbb{R}^3$ is called the *Lagrangian angular velocity*.
(3) $\Omega : \Omega\,(t) \to A\,(n)$, defined via, for every $t \geqslant 0$, $\Omega_t := \Theta_t \circ \eta_t^{-1}$, is called the *Eulerian angular velocity tensor*
(4) if $n = 3$, $\omega : \Omega\,(t) \to \mathbb{R}^3$, defined via, for every $t \geqslant 0$, $\omega_t := \operatorname{vec}\Omega_t = \theta_t \circ \eta_t^{-1}$, is called the *Eulerian angular velocity*

We can motivate the definition of the angular velocity tensors as follows. If a rigid motion f maps maps a points y in the reference configuration to

$$x(t,y) = f(t,y) = b(t) + R(t)(y - b_0) \tag{2.2}$$

for $b_0 \in \mathbb{R}^n$, $b : [0,\infty) \to \mathbb{R}^n$, and $R : [0,\infty) \to O(n)$, then $\partial_t f(t,y) = \dot{b} + \dot{R}y$. In particular, since we may invert (2.2) to write $y(t,x) = R^T(x-b) + b_0$, we deduce that

$$\partial_t f(t, y(t,x)) = \dot{b} + \dot{R}R^T(x-b).$$

Expressing the time derivative of f in these coordinates is not merely a sleight of hand: those are precisely the coordinates in which we can measure f if we are not keeping track of the original position of each point x. Crucially: this expression motivates defining the angular velocity of the rigid motion as $\dot{R}R^T$, which is akin to how we defined the angular velocity tensors in Definition 2.19 above.

3. Physics and rigid bodies

In this section we introduce various physical quantities associated with continua and micropolar continua, such as mass and moment of inertia in Section 3.1 and linear momentum and angular momentum in Section 3.3. In Section 3.2 we take care to characterize the admissible moments of inertia, i.e. determining precisely which positive symmetric matrices are the moment of inertia of some continuum.

3.1. Mass and moments of inertia. In this section we introduce several concepts and quantities related to mass and moment of inertia. We then take care to compute the values of these quantities for some simple example, and we record how these quantities behave under rigid motions and Cartesian products. This will come in handy in Section 3.2 when we seek to characterize the admissible moments of inertia.

We being by defining the mass, moment of inertia, and other associated objects for a continuum. Recall that, physically-speaking, the mass and moment of inertia play very similar role. Each of these quantities is a phenomenological constant which encodes the inertia response of a body to exerted forces and torques.

Definition 3.1. (Mass, center of mass, and associated notions)
(1) Given a Borel measure ν on $\mathbb{R}^n$ we define, for every Borel set E,
   (a) its *mass*, denoted $\mathcal{M}\,(E)$, via $\mathcal{M}\,(E) := \nu\,(E)$,
   (b) its *center of mass*, denoted $\overline{x}$, via

$$\overline{x} := \frac{1}{\underbrace{\mathcal{M}\,(E)}_{\nu(E)}} \int_E x\, d\nu\,(x) = \fint_E x\, d\nu\,(x) = \mathbb{E}_\nu\,[x]$$

   (c) its *covariance matrix*, denoted V, via

$$V := \fint_E (x - \overline{x}) \otimes (x - \overline{x})\, d\nu\,(x) = \mathbb{E}_\nu\,[(x - \mathbb{E}_\nu\,[x]) \otimes (x - \mathbb{E}_\nu\,[x])] = \mathbb{V}_\nu\,[x]$$

   (d) if $n = 3$, its moment of inertia, denoted J, via

$$J := \int_E \left(|x - \overline{x}|^2 - (x - \overline{x}) \otimes (x - \overline{x}) \right) d\nu\,(x) = \mathcal{M}\,(E)\,((\operatorname{tr} V)\, I - V)$$

(2) We say that a measure is *finite* if the mass of its support is finite.
(3) We say that a measure has *finite second moment* if the covariance matrix of its support is finite.

(4) A Lagrangian measure μ defined along the flow such that, for every $t \geqslant 0$, μ_t is a finite Borel measure with finite second moment is also called a *Lagrangian mass measure*.
(5) If a Lagrangian mass measure μ is absolutely continuous with respect to the Lebesgue measure, meaning that for every t, $\mu_t \ll \mathcal{L}^n$, then we call the Lagrangian function σ defined along the flow via, for all $t \geqslant 0$, $\sigma_t := \frac{\mathrm{d}\mu_t}{\mathrm{d}\mathcal{L}^n}$ the *Lagrangian mass density* associated with μ.
(6) An Eulerian measure defined along the flow such that, for every $t \geqslant 0$, ν_t is a finite Borel measure with finite second moment is also called an *Eulerian mass measure*.
(7) If an Eulerian mass measure ν is absolutely continuous with respect to the Lebesgue measure, then we call the Eulerian function ρ defined along the flow via, for all $t \geqslant 0$, $\rho_t := \frac{\mathrm{d}\nu_t}{\mathrm{d}\mathcal{L}^n}$ the *Eulerian mass density* associated with ν.

Since the center of mass and covariance matrix of a finite measure with finite second moment are nothing more than the expectation and covariance matrix of the probability measure obtained by normalizing the measure, we will often refer to these quantities as *statistic functionals* of the measure.

Remark 3.2. The fact that the moment of inertia J has a form which is, at first sight, somewhat odd is worth remarking on. This particular form of J is a consequence of our insistence to identify Skew(3) with $\mathbb{R}^3$. Indeed: the natural space for the angular velocity tensor Ω to live in is Skew(3), which would mean that the moment of inertia $\mathcal{J}$ would be a linear map on Skew(3). It can be shown that this matrix-to-matrix moment of inertia would take the form

$$\mathcal{J}\Omega = \{V, \Omega\} = V\Omega + \Omega V,$$

where $\{\,\cdot\,,\,\cdot\,\}$ denotes the *anti-commutator* of two matrices and where V denotes the covariance matrix of the mass measure under consideration. $\mathcal{J}$ then gives rise to a linear map J on $\mathbb{R}^3$ by making the following diagram commute.

$$\begin{tikzcd}
\mathbb{R}^3 \arrow[r, shift left, "\mathrm{ten}"] \arrow[d, "J"'] & \mathrm{Skew}(3) \arrow[l, shift left, "\mathrm{vec}"] \arrow[d, "\mathcal{J}"] \\
\mathbb{R}^3 \arrow[r, shift left, "\mathrm{ten}"] & \mathrm{Skew}(3) \arrow[l, shift left, "\mathrm{vec}"]
\end{tikzcd}$$

More precisely, for any $\omega \in \mathbb{R}^3$, $J\omega = \operatorname{vec}\mathcal{J}(\operatorname{vec}\omega)$ and hence, using Lemma 6.1, the fact that, since $\mathcal{J} = \{V, \cdot\,\}$, $\mathcal{J}_{abcd} = V_{ac}\delta_{bd} + \delta_{ac}V_{bd}$, and the fact that V is symmetric,

$$\begin{aligned}
(J\omega)_i = \frac{1}{2}\epsilon_{aib}\mathcal{J}_{abcd}\epsilon_{cmd}\omega_m = \frac{1}{2}\epsilon_{aib}(V_{ac}\delta_{bd} + \delta_{ac}V_{bd})\epsilon_{cmd}\omega_m &= \frac{1}{2}(\epsilon_{aib}V_{ac}\epsilon_{cmb} + \epsilon_{aib}V_{bd}\epsilon_{amd})\omega_m \\
&= \frac{1}{2}(\delta_{ac}\delta_{im}V_{ac} - \delta_{am}\delta_{ic}V_{ac} + \delta_{im}\delta_{bd}V_{bd} - \delta_{id}\delta_{bm}V_{bd})\omega_m \\
&= \frac{1}{2}(V_{aa}\omega_i - V_{ai}\omega_a + V_{bb}\omega_i - V_{bi}\omega_b) = (\operatorname{tr} V)\omega_i - (V\omega))_i,
\end{aligned}$$

i.e. indeed $J = (\operatorname{tr} V)I - V$.

We now record a useful decomposition for the Eulerian velocity of a rigid body. Despite seeming quite innocuous, Proposition 3.3 below is quite important since it is later used to compute the linear and angular momentum of a rigid body. These computations are essential since they will in turn motivate the definition of the linear momentum and angular momentum densities.

Proposition 3.3. *(Decomposition of the Eulerian velocity of a rigid body)*
Let $(\mathcal{U}, \eta_0)$ be a rigid body. Then there exists constants $\bar{u}, \bar{\omega} \in \mathbb{R}^3$ such that, if u denotes the Eulerian velocity and $\bar{x}$ denotes the center of mass of the rigid body, then we can decompose u as $u = \bar{u} + \bar{\omega} \times (\,\cdot\, - \bar{x})$.

Proof. Since η is a rigid motion we know from Proposition 2.11 that $\mathbb{D}u = 0$. Lemma 6.6 then tells us that there exists constants $\tilde{u} \in \mathbb{R}^3$ and $\bar{\Omega} \in \mathrm{Skew}(3)$ such that $u(x) = \tilde{u} + \bar{\Omega}x$. So finally, for $\bar{\omega} := \operatorname{vec}\bar{\Omega}$ and $\bar{u} := \tilde{u} + \bar{\Omega}\bar{x}$ we have that $u = \bar{u} + \bar{\omega} \times (\,\cdot\, - \bar{x})$. □

We now compute that mass, center of mass, covariance matrix, and moment of inertia of several simple rigid bodies. These computations serve several purposes.

(1) They serve a pedagogical purpose by showing us what these physical quantities look like for simple cases. For example, these computations tell us that the roles of eigenspaces corresponding to zero and non-zero eigenvalues are flipped when translating between the covariance matrix V and the moment of inertia J. In more geometric terms we can phrase this as follows: the geometric extent of a rigid body is concentrated along the orthogonal complement of the kernel of V, which corresponds to the kernel of J. This is particularly evident when considering Example 3.4 below.
(2) These examples will later be used as elementary building blocks to construct mass measures with arbitrarily prescribed moments of inertia.

EXAMPLE 3.4. (Idealized dumbbell)
Consider the measure $\nu = \frac{m}{2}\delta_{\frac{le_n}{2}} + \frac{m}{2}\delta_{\frac{-le_n}{2}}$.

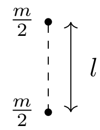


FIGURE 2. Idealized dumbbell

The center of mass and covariance matrix are given by

$$\bar{x} = \frac{1}{m}\Big(\frac{le_n}{2}\frac{m}{2} + \frac{-le_n}{2}\frac{m}{2}\Big) = 0 \text{ and } V = \frac{1}{m}\Big(\left(\frac{le_n}{2}\otimes\frac{le_n}{2}\right)\frac{m}{2} + \left(\frac{-le_n}{2}\otimes\frac{-le_n}{2}\right)\frac{m}{2}\Big) = \frac{l^2}{4}e_n\otimes e_n$$

and therefore the moment of inertia is

$$J = m\Big((\operatorname{tr}V)\, I - V\Big) = m\left(\frac{l^2}{4}I - \frac{l^2}{4}e_n\otimes e_n\right) = \frac{ml^2}{4}\left(I - e_n\otimes e_n\right).$$

EXAMPLE 3.5. (Idealized n-dumbbell)
Consider the measure $\nu = \frac{1}{n}\sum_{i=1}^{n}\frac{m}{2}\left(\delta_{\frac{le_i}{2}} + \delta_{\frac{-le_i}{2}}\right)$.

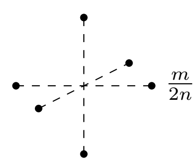


FIGURE 3. Idealized 3-dumbbell

The center of mass and covariance matrix are given by

$$\begin{cases} \bar{x} = \frac{1}{m}\left(\frac{1}{n}\sum_{i=1}^{n}\frac{le_i}{2}\frac{m}{2} + \frac{-le_i}{2}\frac{m}{2}\right) = \frac{l}{4n}\sum_{i=1}^{n}(e_i - e_i) = 0 \text{ and} \\ V = \frac{1}{m}\left(\frac{1}{n}\sum_{i=1}^{n}\left(\frac{le_i}{2}\otimes\frac{le_i}{2}\right)\frac{m}{2} + \left(\frac{-le_i}{2}\otimes\frac{-le_i}{2}\right)\frac{m}{2}\right) = \frac{l^2}{4n}\sum_{i=1}^{n}e_i\otimes e_i = \frac{l^2}{4n}I \end{cases}$$

and therefore the moment of inertia is

$$J = m\left((\operatorname{tr}V)\, I - I\right) = m\left(\frac{l^2}{4}I - \frac{l^2}{4n}I\right) = \frac{ml^2}{4}\frac{n-1}{n}I$$

EXAMPLE 3.6. (Rod)
Consider the measure $\nu = \rho\mathcal{H}^1 \llcorner \frac{l}{2}[-e_1, e_1]$ where $[-e_1, e_1] := \{\theta(-e_1) + (1-\theta)e_1 \mid \theta\in[0,1]\}$. The mass and center of mass of the rod are given by

$$M = \int_{-\frac{l}{2}e_1}^{\frac{l}{2}e_1}\rho\, d\mathcal{H}^1(x) = \int_{-\frac{l}{2}}^{\frac{l}{2}}\rho\, ds = \rho l \text{ and } \bar{x} = \fint_{-\frac{l}{2}e_1}^{\frac{l}{2}e_1}\rho x\, d\mathcal{H}^1(x) = \frac{1}{\rho l}\left(\int_{-\frac{l}{2}}^{\frac{l}{2}}\rho s\, ds\right)e_1 = 0.$$

The covariance matrix is given by

$$V = \fint_{-\frac{l}{2}e_1}^{\frac{l}{2}e_1}\rho x\otimes x\, d\mathcal{H}^1(x) = \left(\fint_{-\frac{l}{2}e_1}^{\frac{l}{2}e_1}\rho s^2\, ds\right)e_1\otimes e_1 = \frac{1}{\rho l}\frac{\rho l^3}{3}e_1\otimes e_1 = \frac{l^2}{3}e_1\otimes e_1$$

and hence the angular moment of inertia is $J = \frac{l^2}{3}\left(I - e_1 \otimes e_1\right)$.

Example 3.7. (Sphere)
Consider the measure $\nu = \mathcal{L}^n \llcorner B\left(0, r\right)$. When evaluating integrals below, we will often use the following fact: given a map $\Phi : \mathbb{R}^n \to \mathbb{R}^n$ preserving the Lebesgue measure, if the integrand f is odd under Φ, i.e. $f \circ \Phi = -f$ and E is some (Lebesgue-measurable) set invariant under Φ, then $\int_E f\, dx = 0$.

First we compute the center of mass: by symmetry,

$$\bar{x} = \fint_{B(0,r)} x\, dx = 0.$$

Now we compute the covariance matrix and moment of inertia. By symmetry,

$$V_{ii} = \left(\fint_{B(0,r)} x \otimes x\, dx\right)_{ii} = \fint_{B(0,r)} x_i^2 dx = \frac{1}{n}\left(\fint_{B(0,r)} x_1^2\, dx + \cdots + \fint_{B(0,r)} x_n^2\, dx\right) = \frac{1}{n}\fint_{B(0,r)} |x|^2\, dx$$

and

$$V_{ij} = \left(\fint_{B(0,r)} x \otimes x\, dx\right)_{ij} = \fint_{B(0,r)} x_i x_j\, dx = 0 \text{ if } i \neq j.$$

Now let us write $\alpha_n := \mathcal{L}^n\left(B\left(0,1\right)\right)$ and note that then $\mathcal{H}^{n-1}\left(\partial B\left(0,1\right)\right) = n\alpha_n$. We can then compute:

$$\frac{1}{n}\fint_{B(0,r)} |x|^2\, dx = \frac{1}{n\mathcal{L}^n\left(B\left(0,r\right)\right)}\int_0^r \left(\int_{\partial B(0,s)} |x|^2\, d\mathcal{H}^{n-1}\left(x\right)\right) ds = \frac{1}{n\alpha_n r^n}\int_0^r \mathcal{H}^{n-1}\left(\partial B\left(0,s\right)\right) s^2\, ds$$
$$= \frac{\mathcal{H}\left(\partial B\left(0,1\right)\right)}{n\alpha_n r^n}\int_0^r s^{(n-1)+2}\, ds = \frac{n\alpha_n}{n\alpha_n r^n}\frac{r^{n+2}}{n+2} = \frac{r^2}{n+2}$$

Therefore the covariance matrix is $V = \frac{r^2}{n+2} I$. So finally the moment of inertia is

$$J = \mathcal{L}^n\left(B\left(0,r\right)\right)\left(\left(\operatorname{tr} V\right) I - V\right) = \alpha_n r^n\left(\frac{nr^2}{n+2} I - \frac{r^2}{n+2} I\right) = \frac{\alpha_n\left(n-1\right)}{n+2} r^{n+2} I.$$

Example 3.8. (Cylinder)
Consider the measure $nu = \mathcal{L}^3 \llcorner C$ where

$$C = B_2\left(r\right) \times \left(-l, l\right) = \left\{\left(x_1, x_2, x_3\right) \in \mathbb{R}^3 \;\middle|\; x_1^2 + x_2^2 < r^2 \text{ and } |x_3| < l\right\}.$$

Then the mass is

$$M = \mathcal{L}^3\left(C\right) = \mathcal{L}^2\left(B_2\left(r\right)\right)\mathcal{L}^1\left(\left(-l, l\right)\right) = 2\pi r^2 l$$

and by symmetry of C the center of mass is

$$\bar{x} = \fint_C x\, dx = \frac{1}{M}\int_{-l}^{l} \underbrace{\int_{B_2(r)} x\, d\left(x_1, x_2\right) dx_3}_{=0} = 0.$$

Now let us compute the covariance matrix. Observe that by symmetry of C, $V_{ij} = \fint_C x_i \otimes x_j\, dx = 0$ when $i \neq j$, and that

$$V_{11} = V_{22} = \frac{1}{M}\int_{-l}^{l} dx_3 \int_{B_2(r)} x_1^2 d\left(x_1, x_2\right) = \frac{2l}{2\pi r^2 l}\frac{1}{2}\int_{B_2(r)}\left(x_1^2 + x_2^2\right) d\left(x_1, x_2\right) = \frac{1}{2\pi r^2}\int_0^r s^2\left(2\pi s\right) ds = \frac{r^2}{4}$$

whilst

$$V_{33} = \frac{1}{M}\int_{-l}^{l} x_3^2\left(\pi r^2\right) dx_3 = \frac{\pi r^2}{2\pi r^2 l}\frac{2l^3}{3} = \frac{l^2}{3}.$$

We have thus computed the covariance matrix to be

$$V = \begin{pmatrix} \frac{r^2}{4} & 0 & 0 \\ 0 & \frac{r^2}{4} & 0 \\ 0 & 0 & \frac{l^2}{3} \end{pmatrix}.$$

So finally, the moment of inertia is

$$J = M\left(\left(\operatorname{tr} V\right) I - V\right) = 2\pi r^2 l \left(\left(\frac{r^2}{2} + \frac{l^2}{3}\right) I - \begin{pmatrix} \frac{r^2}{4} & 0 & 0 \\ 0 & \frac{r^2}{4} & 0 \\ 0 & 0 & \frac{l^2}{3} \end{pmatrix}\right) = 2\pi r^2 l \begin{pmatrix} \frac{r^2}{4} + \frac{l^2}{3} & 0 & 0 \\ 0 & \frac{r^2}{4} + \frac{l^2}{3} & 0 \\ 0 & 0 & \frac{r^2}{2} \end{pmatrix}$$

$$= \frac{\pi r^2 l}{6} \begin{pmatrix} 3r^2 + 4l^2 & 0 & 0 \\ 0 & 3r^2 + 4l^2 & 0 \\ 0 & 0 & 6r^2 \end{pmatrix} = \frac{M}{12}\left(3r^2 I + 4l^2 I_2 + 3r^2 e_3 \otimes e_3\right)$$

where $I_2 := e_1 \otimes e_1 + e_2 \otimes e_2$. In particular, if $4l^2 = 3r^2$, then $J = \frac{M}{2} r^2 I$. It is worth contrasting this with the moment of inertia of a sphere of mass M and radius R, which is $\frac{2M}{5} r^2 I$.

Recall that the computations above are helpful since they provide building blocks that can later be used to construct more complicated mass measures – this will be essential in Section 3.2 when we characterize admissible moments of inertia. The other tool we need is to be able to say how statistical functionals associated with various rigid bodies behave under various transformation. First we record how statistical functionals behave under transformation by rigid motions.

Proposition 3.9. *(Transformation of statistical functionals of a measure under rigid motions) Let ν be a finite Borel measure with finite second moment and let $T : \mathbb{R}^n \to \mathbb{R}^n$ be a rigid motion.*

(1) *If T is a translation, i.e. $T(x) = x + z$ for some $z \in \mathbb{R}^n$, then the mass, center of mass, and covariance matrix of the measure transform as*

$$(M, \bar{x}, V) \to (M, \bar{x} + z, V)$$

i.e. $(M, \bar{x} + z, V)$ are the statistical functionals of $T_{\#}\nu$. In particular the moment of inertia is invariant under translations, i.e. $J \to J$.

(2) *If T is a rotation, i.e. $T(x) = Rx$ for some $R \in O(n)$, then the mass, center of mass, and covariance matrix of the measure transform as*

$$(M, \bar{x}, V) \to \left(M, R\bar{x}, RVR^T\right)$$

i.e. $\left(M, R\bar{x}, RVR^T\right)$ are the statistical functionals of $T_{\#}\nu$. In particular the moment of inertia transforms as $J \to RJR^T$.

Proof. (1) Denote by μ the push-forward of ν under T, i.e. $\mu := T_{\#}\nu$ such that

$$d\mu(x + z) = d\nu(x).$$

Mass: performing the change of variables $y = x + z$ we obtain

$$\int_{E+z} d\mu(y) = \int_E d\mu(x + z) = \int_E d\nu(x) = M.$$

Center of mass: we perform the same change of variable $y = x + z$ to compute

$$\fint_{E+z} y\, d\mu(y) = \fint_E (x + z)\, d\nu(x) = \bar{x} + z.$$

Covariance matrix: since $\bar{y} = \bar{x} + z$, performing the change of variables $y = x + z$ yields $y - \bar{y} = x - \bar{x}$ and hence

$$\fint_{E+z} (y - \bar{y}) \otimes (y - \bar{y})\, d\mu(y) = \fint_E (x - \bar{x}) \otimes (x - \bar{x})\, d\nu(x) = V.$$

(2) Denote by μ the push-forward of ν under T, i.e. $\mu := T_{\#}\nu$ such that $d\mu(Rx) = d\nu(x)$. To compute the mass, we perform the change of variables $y = Rx$, noting that $dy = dx$ since $R \in O(n)$ and hence $|\det R| = 1$. Therefore we have

$$\int_{RE} d\mu(y) = \int_E d\mu(Rx) = \int_E d\nu(x) = M.$$

Now we compute the center of mass, performing the same change of variables:

$$\fint_{RE} y\, d\mu(y) = \fint_E Rx\, d\nu(x) = R\bar{x}.$$

Finally we compute the covariance matrix:

$$\fint_{RE} (y-\bar{y}) \otimes (y-\bar{y})\, d\mu(y) = \fint_{E} R(x-\bar{x}) \otimes R(x-\bar{x})\, d\nu(x) = RVR^T.$$

□

Proposition 3.9 gives us a particularly simple rule for the transformation of statistical functionals under rigid motions that preserve the center of mass. This is recorded in Corollary 3.10 below.

Corollary 3.10. *Under the transformation $T(x) = R(x-\bar{x}) + \bar{x}$ for some $R \in O(n)$, the mass, center of mass, and covariance matrix of a measure transform as*

$$(M, \bar{x}, V) \to \left(M, \bar{x}, RVR^T\right)$$

In particular the moment of inertia transforms as $J \to RJR^T$.

We continue establishing properties of statistical functionals of mass measures under various transformations. Having established how they transform under rigid motions we now record how they behave with respect to Cartesian products.

Proposition 3.11. *(Statistical functionals of product measures) Let ν_1, ν_2 be finite Borel measures with finite second moment on $\mathbb{R}^{n_1}, \mathbb{R}^{n_2}$ respectively, with masses M_1, M_2, centers of masses $\bar{x}_1, \bar{x}_2$, and covariance matrices V_1, V_2. The measure $\nu := \nu_1 \times \nu_2$ on $\mathbb{R}^{n_1+n_2}$ has mass $M_1 M_2$, center of mass $(\bar{x_1}, \bar{x}_2)$, and covariance matrix $\begin{pmatrix} V_1 & 0 \\ 0 & V_2 \end{pmatrix}$.*

Proof. First we compute the mass: $\nu(\mathbb{R}^{n_1+n_2}) = \nu(\mathbb{R}^{n_1} \times \mathbb{R}^{n_2}) = \nu_1(\mathbb{R}^{n_1})\, \nu_2(\mathbb{R}^{n_2}) = M_1 M_2$. Now we compute the covariance matrix, writing $x = (x_1, x_2) \in \mathbb{R}^{n_1} \times \mathbb{R}^{n_2} = R^{n_1+n_2}$ to obtain

$$\fint_{\mathbb{R}^{n_1+n_2}} x d\nu(x) = \frac{1}{M_1 M_2} \int_{\mathbb{R}^{n_1} \times \mathbb{R}^{n_2}} (x_1, x_2)\, d\nu_1(x_1)\, d\nu_2(x_2)$$

$$= \frac{1}{M_2} \int_{\mathbb{R}^{n_2}} \left(\frac{1}{M_1} \int_{\mathbb{R}^{n_1}} (x_1, x_2)\, d\nu_1(x_1) \right) d\nu_2(x_2) = \frac{1}{M_2} \int_{\mathbb{R}^{n_2}} (\bar{x}_1, x_2)\, d\nu_2(x_2) = (\bar{x}_1, \bar{x}_2).$$

Finally, to compute the covariance matrix, we first note that since covariance matrices are invariant under translation, we may without loss of generality assume that $\bar{x} = 0$. Then

$$\begin{aligned} \fint_{\mathbb{R}^{n_1+n_2}} x \otimes x d\nu(x) &= \fint_{\mathbb{R}^{n_2}} \fint_{\mathbb{R}^{n_1}} (x_1, x_2) \otimes (x_1, x_2)\, d\nu_1(x)\, d\nu_2(x) \\ &= \fint_{\mathbb{R}^{n_2}} \left(\fint_{\mathbb{R}^{n_1}} \begin{pmatrix} x_1 \otimes x_1 & x_1 \otimes x_2 \\ x_2 \otimes x_1 & x_2 \otimes x_2 \end{pmatrix} d\nu_1(x) \right) d\nu_2(x) \\ &= \fint_{\mathbb{R}^{n_2}} \begin{pmatrix} V_1 & 0 \otimes x_2 \\ x_2 \otimes 0 & x_2 \otimes x_2 \end{pmatrix} d\nu_2(x_2) = \begin{pmatrix} V_1 & 0 \\ 0 & V_2 \end{pmatrix}. \end{aligned}$$

□

3.2. Admissible moments of inertia. The goal of this section is to characterize precisely when, given a positive symmetric matrix J, it is the moment of inertia of *some* mass measure.

We begin with preliminary results from linear algebra used to formulate the conditions on the eigenvalues of J that will characterize admissible moments of inertia. First we record a result telling us to relate the spectrum of a covariance matrix with the spectrum of its moment of inertia.

Lemma 3.12. *Let $V \in \mathbb{R}^{n \times n}$ be symmetric and positive. Then $S := (\operatorname{tr} V) I - V$ is symmetric and positive. Moreover S and V have the same eigenspaces and, if $\lambda = (\lambda_1, \ldots, \lambda_n)$ denotes the eigenvalues of V and $\mu = (\mu_1, \ldots, \mu_n)$ denotes the corresponding (which is well-defined by item 1 above) eigenvalues of S, then $\mu = (C - I)\lambda$, where $C \in \mathbb{R}^{n \times n}$ such that $C_{ij} = 1$ for all i, j.*

Proof. Clearly S is symmetric since V is. Now, since V is symmetric and positive, it has positive eigenvalues and is diagonal in some orthonormal basis $\{v_i\}_{i=1}^n$, i.e. $V = \sum_i \lambda_i v_i \otimes v_i$ for some $\lambda_i \geqslant 0$. Therefore, for any $\xi \in \mathbb{R}^n$,

$$S\xi \cdot \xi = \left(\sum_i \lambda_i \right) |\xi|^2 - \sum_i \lambda_i (v_i \cdot \xi)^2 = \sum_i \lambda_i \left(|\xi|^2 - (v_i \cdot \xi)^2 \right)$$

where indeed, since $|v_i| = 1$ for each i, $|\xi|^2 - (v_i \cdot \xi)^2 = |\xi|^2|v_i|^2 - (v_i \cdot \xi)^2 \geqslant 0$ by Cauchy-Schwarz. Finally observe that

$$Sv_i = \left(\sum_j \lambda_j\right) v_i - \lambda_i v_i = \underbrace{\left(\sum_{j\neq i} \lambda_j\right)}_{=\mu_i} v_i$$

i.e. indeed S and V share the same eigenspaces, with moreover $\mu = (C - I)\,\lambda$. ☐

Since Lemma 3.12 established that there was a linear relationship between the spectrum of a covariance matrix and its associated moment of inertia we now observe, in Lemma 3.13 below, that this relationship is invertible and compute its inverse.

Lemma 3.13. *Let $C \in \mathbb{R}^{n\times n}$ such that $C_{ij} = 1$ for all i, j. Then $C - I$ is invertible, with*

$$(C - I)^{-1} = \frac{1}{n-1}C - I.$$

Proof. Since C and I commute, it is sufficient to show that $(C - I)\left(\frac{1}{n-1}C - I\right) = I$. The key observation is that $C^2 = nC$. We may now compute directly that

$$(C - I)\left(\frac{1}{n-1}C - I\right) = \frac{1}{n-1}C^2 - C - \frac{1}{n-1}C + I = \left(\frac{n}{n-1} - 1 - \frac{1}{n-1}\right) C + I = I.$$

☐

We now have the tools to establish the first of the two main results of this section, namely providing a necessary condition for a positive symmetric matrix to be the moment of inertia of some mass measure.

Proposition 3.14. *(Admissible moments of inertia – necessity)*
Let ν be a finite Borel measure on $\mathbb{R}^n$ with finite second moment. Its moment of inertia is a positive symmetric n-by-n matrix, and moreover if we denote by $\mu \in \mathbb{R}^n$ its eigenvalues, then

$$\mu_i \leqslant \frac{1}{n-1}\sum_{j=1}^{n} \mu_j \text{ for all } i = 1, \dots, n \tag{1}$$

or equivalently

$$\mu \in \mathbb{R}_+\nabla^{n-1} = \left\{ sx \;\middle|\; s \in \mathbb{R}_+, x \in \nabla^{n-1} \right\} \tag{2}$$

where $\mathbb{R}_+ := [0,\infty)$ and $\nabla^{n-1} = 1 - \Delta^{n-1}$, for Δ^{n-1} denoting the $(n-1)$-simplex such that

$$\nabla^{n-1} := \left\{ x \in \mathbb{R}^n \;\middle|\; x_i = 1 - \theta_i = \sum_{j\neq i} \theta_j \text{ for some } \theta_j \geqslant 0 \text{ with } \sum_j \theta_j = 1 \right\}.$$

Proof. Suppose without loss of generality that ν has unit mass. Let V denote the covariance matrix of ν and let J denote the corresponding moment of inertia, i.e. $J = (\operatorname{tr} V)\, I - V$. It follows from Lemma 3.12 that J is symmetric, positive, has the same eigenspaces as V, and that if λ and $\mu \in \mathbb{R}^n$ denote the eigenvalues of V and J respectively, then $\mu = (C - I)\,\lambda$ where $C \in \mathbb{R}^n$ such that $C_{ij} = 1$ for all i, j. Note that since V is symmetric and positive, it follows that $\lambda \in \mathbb{R}^n_+$ and hence $\mu \in (C - I)\,\mathbb{R}^n_+$. Now Lemma 3.13 says that $C - I$ is invertible with inverse $\frac{1}{n-1}C - I$ and therefore $\left(\frac{1}{n-1}C - I\right)\mu \in \mathbb{R}^n_+$, i.e.

$$\frac{1}{n-1}\sum_{j=1}^{n} \mu_j - \mu_i \geqslant 0 \text{ for all } i$$

so (1) holds. Now let us show that (1) $\Rightarrow$ (2). To do so, simply define

$$s := \frac{1}{n-1}\sum_j \mu_j \quad \text{and} \quad \theta := \mathbb{1} - \frac{\mu}{s}$$

where $\mathbb{1}_i = 1$ for all i, and observe that then

- $s \geqslant 0$ since J is positive and hence it's eigenvalues μ_i are positive,

- $\theta \geqslant 0$ since, by (1), $s \geqslant 0$, and
- $\sum_i \theta_i = n - \frac{1}{s}\sum_i \mu_i = n - (n-1) = 1$.

So indeed $\mu = s(\mathbb{1} - \theta) \in \mathbb{R}_+ \nabla^{n-1}$. Finally we show that (2) $\Rightarrow$ (1). This is immediate since by assumption $\mu = s(\mathbb{1} - \theta)$ for some $s \geqslant 0$ and some $\theta \in \Delta^{n-1}$, and therefore $\sum_i \mu_i = s\,(n-1)$, i.e. $s = \frac{1}{n-1}\sum_i \mu_i$. We can thus conclude that, since $\theta_i \leqslant 1$, $\mu_i = s\,(1-\theta_i) \leqslant s = \frac{1}{n-1}\sum_i \mu_i$. □

The remainder of this section is now devoted to proving that the necessary condition for a positive symmetric matrix to be the moment of inertia of some mass measure recorded in Proposition 3.14 above is actually sufficient. To do so means being able to, given any positive symmetric matrix J satisfying one of the equivalent conditions provided in Proposition 3.14, constructing a mass measure whose moment of inertia is J. Our strategy to do so is to use the simples examples considered in Section 3.1 as building blocks that we can piece together using the results also proved in Section 3.1.

We begin by noting that, using Example 3.6 we may construct a mass measure corresponding to any *one-dimensional* covariance matrix.

Lemma 3.15. *(Existence of mass measures with prescribed one-dimensional covariance matrix) For any $M > 0$ and any $\lambda \geqslant 0$ there exists a Borel measure ν on $\mathbb{R}^n$ with mass M, center of mass 0 and covariance matrix $\lambda e_1 \otimes e_1$.*

Proof. If $\lambda = 0$ then define a 'point-mass' measure $nu := M\delta_0$ such that indeed

$$\nu\left(\mathbb{R}^n\right) = M,\ \mathbb{E}_\nu\left[x\right] = 0,\ \text{ and } \mathbb{V}_\nu\left[x\right] = \mathbb{E}_\nu\left[x \otimes x\right] = 0.$$

If $\lambda > 0$ then define $\nu := \rho\mathcal{H}^1 \llcorner \frac{l}{2}\left[-e_1, e_1\right]$ where $l := \sqrt{3\lambda}$ and $\rho := \frac{M}{l}$. We may thus compute, as in example 3.6, and obtain

$$\nu\left(\mathbb{R}^n\right) = \rho l = M,\ \mathbb{E}_\nu\left[x\right] = 0,\ \text{ and } \mathbb{V}_\nu\left[x\right] = \frac{l^2}{3} e_1 \otimes 3_1 = \lambda e_1 \otimes e_1.$$

□

Using Lemma 3.15 immediately above and Proposition 3.11 we now show that we can construct mass measures corresponding to arbitrary *diagonal* covariance matrices.

Lemma 3.16. *(Existence of mass measures with prescribed diagonal covariance matrix) For any $M > 0$ and any $\lambda_1, \ldots, \lambda_n \geqslant 0$ there exists a Borel measure ν on $\mathbb{R}^n$ with mass M, center of mass 0 and covariance matrix $\operatorname{diag}\left(\lambda_1, \ldots, \lambda_n\right)$.*

Proof. This follows immediately from combining Lemma 3.15 and Proposition 3.11. For any $M > 0$ and $\lambda_1, \ldots, \lambda_n \geqslant 0$, by Lemma 3.15 there are Borel measures ν_i, $i = 1 \ldots, n$, on $\mathbb{R}^1$ with masses $M^{1/n}$, centers of mass 0 and covariance matrices $\lambda_i \left(\in \mathbb{R}^{1\times 1} = \mathbb{R}\right)$. Therefore, by Proposition 3.9 the measure $\nu := \nu_1 \times \cdots \times \nu_n$ has the desired statistical functionals. □

Combing Lemma 3.16 with Lemma 3.12 allows us to deduce that we can construct mass measures corresponding to arbitrary diagonal *moment of inertia*.

Lemma 3.17. *(Existence of mass measures with prescribed diagonal moment of inertia) For any $M > 0$ and any $\mu \in \mathbb{R}_+\nabla^{n-1}$ there exists a Borel measure ν on $\mathbb{R}^n$ with finite second moment, mass M, center of mass 0 and moment of inertia $\operatorname{diag}\left(\mu_1, \ldots, \mu_n\right)$.*

Proof. Let $\lambda := \frac{1}{M}\left(\frac{1}{n-1}C - I\right)\mu$. Note that part (1) of Proposition 3.14 tells us precisely that $\lambda \in \mathbb{R}^n_+$. By Lemma 3.16 we therefore know that there exists a Borel measure ν with mass M, center of mass 0 and covariance matrix $\operatorname{diag}\left(\lambda_1, \ldots, \lambda_n\right)$. Since the moment of inertia is $J = M\left(\left(\operatorname{tr} V\right) I - V\right)$, Lemma 3.12 tells us that J is diagonal with eigenvalues $M\left(C - I\right)\lambda$. Finally, since Lemma 3.13 tells us that $\left(C - I\right)^{-1} = \frac{1}{n-1}C - I$, it follows that the eigenvalues of J are indeed μ. □

We now have all the tools to prove the sufficiency of the conditions for admissibility of a moment of inertia introduced in Proposition 3.14. We do so to conclude this section.

PROPOSITION 3.18. *(Admissible moments of inertia – sufficiency)*
Let $M > 0$ and let J be a positive symmetric n-by-n matrix with eigenvalues $\mu_1, \ldots, \mu_n$ satisfying

$$\mu_i \leqslant \frac{1}{n-1} \sum_{j=1}^{n} \mu_j \text{ for every } i.$$

Then there exist a Borel measure ν with mass M, center of mass 0, and angular moment of inertia J.

PROOF. Since J is symmetric and positive, there exists a rotation matrix $R \in O(n)$ such that

$$J = R \operatorname{diag}(\mu_1, \ldots, \mu_n) R^T.$$

Since, by Proposition 3.14, $\mu \in \mathbb{R}_+ \nabla^{n-1}$ it follows from Lemma 3.17 that there exists a Borel measure $\tilde{\nu}$ with mass M, center of mass 0 and moment of inertia $\operatorname{diag}(\mu_1, \ldots, \mu_n)$. Finally Corollary 3.10 tells us that after applying the transformation $T(x) := R(x - \overline{x}) + \overline{x}$ the measure $\nu := T_{\#}\tilde{\nu}$ has mass M, center of mass 0 and moment of inertia $R \operatorname{diag}(\mu_1, \ldots, \mu_n) R^T = J$. ☐

Putting Proposition 3.14 and Proposition 3.18 together, we have proved the theorem below.

THEOREM 3.19. *(Admissible moments of inertia)*
Let J be a positive symmetric n-by-n matrix. There exists a finite Borel measure ν on $\mathbb{R}^n$ with moment of inertia J if and only if the eigenvalues $\mu = (\mu_1, \ldots, \mu_n)$ of J satisfy

$$\mu_i \leqslant \frac{1}{n-1} \sum_{j=1}^{n} \mu_j \text{ for every } i.$$

3.3. Linear and angular momentum. In this section we define the linear and angular momentum associated with a rigid body and compute their values for a rigid body.

DEFINITION 3.20. (Linear and angular momentum)
Let (Ω_0, η) be a continuum with Eulerian mass measure ν. For every Borel set $E \subseteq \Omega(t)$ we define

(1) its *linear momentum* $\mathcal{P}(E) := \int_E u d\nu_t$ and
(2) if $n = 3$, its *angular momentum about a point* $z \in \mathbb{Z}^n$ is defined to be

$$\mathcal{L}_z(E) := \int_E (x - z) \times u_t(x) \, d\nu_t(x).$$

In particular, its angular momentum about its center of mass is simply called its *angular momentum* and denoted $\mathcal{L}(E)$.

Having defined linear and angular momentum, we compute their values for a rigid body.

PROPOSITION 3.21. *(Linear and angular momentum of a rigid body)*
Let $(\mathcal{U}_0, \eta)$ be a rigid body. Then, using the notation of Proposition 3.3 for $\bar{u}$ and $\bar{\omega}$ and using M and J to denote respectively the mass and moment of inertia of $\mathcal{U}(t)$, we have

$$\mathcal{P}(\mathcal{U}(t)) = M\bar{u} \text{ and } \mathcal{L}(\mathcal{U}(t)) = J \cdot \bar{\omega}.$$

PROOF. In light of Proposition 3.3 we may immediately compute the linear momentum to be

$$\mathcal{P}(\mathcal{U}(t)) = \int_{\mathcal{U}(t)} u d\nu_t = \int_{\mathcal{U}(t)} (\bar{u} + \omega \times (x - \bar{x})) \, d\nu_t = M\bar{u} + \omega \times \underbrace{\left(\int_{\mathcal{U}(t)} x d\nu_t - M\bar{x} \right)}_{=0} = M\bar{u}.$$

Using the 'vectorized' version of the epsilon-delta identity in Lemma 6.2 we may now compute the angular momentum to be

$$\begin{aligned} \mathcal{L}(\mathcal{U}(t)) = \int_{\mathcal{U}(t)} (x - \bar{x}) \times u d\nu_t &= \int_{\mathcal{U}(t)} ((x - \bar{x}) \times \bar{u} + (x - \bar{x}) \times (\omega \times (x - \bar{x}))) \, d\nu_t \\ &= 0 + \int_{\mathcal{U}(t)} \Big(|x - \bar{x}|^2 \omega - (\omega \cdot (x - \bar{x})) (x - \bar{x}) \Big) \, d\nu_t = J \cdot \omega. \end{aligned}$$

☐

Proposition 3.21 comes in handy later since it motivates the definition of a linear momentum density and a angular momentum density.

4. Conservation laws

The goal of this section is two-fold:
(1) introduce the integral balance laws corresponding to the conservation of mass, linear and angular momentum, and energy, and
(2) derive the corresponding local versions of these balance laws.

To be more precise, we discuss the conservation of mass in Section 4.1, the conservation of microinertia in Section 4.2, the conservation of linear and angular momentum in Section 4.3. We conclude Section 4 with a brief discussion of boundary conditions in Section 4.5.

4.1. Mass. In this section we define what it means for a mass measure to be conserved and we derive the associated local conservation law. We also use the conserved mass of rigid body to obtain a useful characterization of the kinematic and dynamics variables describing a rigid body.

Definition 4.1. (Conservation of mass)
Let (Ω_0, η) be a continuum. An Eulerian mass measure ν satisfying $\nu_t = (\eta_t)_{\#}\nu_0$ is called a *conserved Eulerian mass measure.*

Having defined what it means for mass to be conserved we derive the associated local conservation law, which turns out to be the well-known continuity equation.

Proposition 4.2. *(Local conservation of mass - continuity equation)*
Let (Ω_0, η) be a continuum with Eulerian velocity u. If a conserved Eulerian mass measure is absolutely continuous with respect to the Lebesgue measure, then its Eulerian mass density ρ satisfies

$$\partial_t \rho + \nabla \cdot (\rho u) = 0.$$

In other words: as a consequence of conservation of mass, the Eulerian mass density satisfies the continuity equation.

Proof. We simply use Reynolds' transport theorem, i.e. Theorem 2.7, to take a time derivative of the equation of conservation of mass for an arbitrary set. Indeed, if we let $\mathcal{U}_0 \subseteq \Omega_0$ be any open set, then

$$0 = \frac{\mathrm{d}}{\mathrm{d}t}\mathcal{M}\left(\mathcal{U}_0\right) = \frac{\mathrm{d}}{\mathrm{d}t}\mathcal{M}\left(\mathcal{U}\left(t\right)\right) = \frac{\mathrm{d}}{\mathrm{d}t}\int_{\mathcal{U}(t)} \rho = \int_{\mathcal{U}(t)} \partial_t \rho + \nabla \cdot \left(\rho u\right).$$

Therefore, since $\mathcal{U}_0$ was an arbitrary subset of the Lagrangian domain and since η_t are diffeomorphisms (such that we can go back and forth between Lagrangian and Eulerian coordinates), the integrand above must vanish everywhere in the Eulerian domain. □

We now record a result relating the fact that a flow is locally volume preserving with the absolute continuity of its conserved mass measure. This result will be used in Proposition 4.5 below, which is itself used to justify the definition of micropolar continua.

Proposition 4.3. *(Locally volume-preserving flows preserve absolute continuity of the Eulerian mass measure)*
Let η be a locally volume-preserving flow map and let ν be the conserved Eulerian mass measure. Suppose that ν is initially absolutely continuous (with respect to the Lebesgue measure). Then ν is absolutely continuous for all time, and moreover we have an explicit representation for the Eulerian mass density in terms of the initial density and the flow map, namely: $\rho_t = \rho_0 \circ \eta_t^{-1}$.

Proof. Since η is locally volume-preserving and ν is conserved by the flow, both sides of the 'inequality' $\nu_0 \ll \mathcal{L}^n$ are preserved under pushforwards along the flow, i.e.: $\nu_t = (\eta_t)_{\#}\nu_0 \ll (\eta_t)_{\#}\mathcal{L}^n = \mathcal{L}^n$. Moreover

$$\left(\rho_t d\mathcal{L}^n\right)(E) = d\nu_t\left(E\right) = d\left(\eta_t^{-1}\left(E\right)\right) = \int_{\eta_t^{-1}(E)} \rho_0\left(y\right)dy \overset{(*)}{=} \int_E \rho_0\left(\eta_t^{-1}\left(x\right)\right)dx = \left(\left(\rho_0 \circ \eta_t^{-1}\right)d\mathcal{L}^n\right)(E)$$

i.e. indeed $\rho_t = \rho_0 \circ \eta_t^{-1}$, where $(*)$ holds since η is locally volume-preserving, and hence by Lemma 2.14 $|\det \nabla \eta| = 1$. □

We now record another result having to do with the densities of conserved mass measures which allows us to translate between the Lagrangian and Eulerian mass densities.

PROPOSITION 4.4. *(Relating Eulerian and Lagrangian mass densities)*
Let (Ω_0, η) be a continuum, let μ be a Lagrangian mass measure with mass density σ and let ν be an Eulerian mass measure with mass density ρ. If $\nu_t = (\eta_t)_{\#}\mu_t$ for every $t \geqslant 0$ then

$$\sigma_t = (\rho_t \circ \eta_t) \det \nabla \eta_t$$

for every $t \geqslant 0$.

PROOF. Let $v_0 \in \Omega_0$, let $r > 0$, and let $t \geqslant 0$. On one hand

$$\mu_t \left(B\left(y_0, r\right)\right) = \int_{B(y_0,r)} \sigma_t\left(y\right) dy$$

and on the other hand, upon using the change of variables $x = \eta_t\left(y\right)$ we see that

$$\nu_t \left(\eta_t \left(B\left(y_0, r\right)\right)\right) = \int_{\eta_t(B(y_0,r))} \rho_t\left(x\right) dx = \int_{B(y_0,r)} \left(\rho_t \circ \eta_t\right)\left(y\right) \det \nabla \eta_t\left(y\right) dy.$$

Since $\nu_t = (\eta_t)_{\#}\mu_t$, we have in particular that $\nu_t\left(\eta_t\left(B\left(y_0, r\right)\right)\right) = \mu_t\left(B\left(y_0, r\right)\right)$, and hence

$$\int_{B(y_0,r)} \sigma_t\left(y\right) dy = \int_{B(y_0,r)} \left(\rho_t \circ \eta_t\right)\left(y\right) \det \nabla \eta_t\left(y\right) dy.$$

So finally, dividing both sides of the equation immediately above by $\mathcal{L}^n\left(B\left(y_0, r\right)\right)$ and sending $r \downarrow 0$ we obtain:

$$\sigma_t\left(y_0\right) = \lim_{r \downarrow 0} \fint_{B(y_0,r)} \sigma_t\left(y\right) dy = \lim_{r \downarrow 0} \fint_{B(y_0,r)} \left(\rho_t \circ \eta_t\right)\left(y\right) \det \nabla \eta_t\left(y\right) dy = \left(\rho_t \circ \eta_t\right)\left(y_0\right) \det \nabla \eta_t\left(y_0\right).$$

□

We conclude this section with a characterization of the kinematic and dynamic descriptors of a rigid body, i.e. its flow map η and its Eulerian velocity u respectively, provided this rigid body has a conserved mass measure.

PROPOSITION 4.5. *(Canonical representation of rigid motions via their Eulerian mass measures)*
If η is a rigid body with conserved Eulerian mass measure ν, then

$$\eta(t, y) = \bar{x}(t) + R(t)(y - \bar{x}(0)) \text{ and } u(t, x) = \bar{u}(t) + \Omega(t)(x - \bar{x}(t))$$

where $\bar{x} : [0, \infty) \to \mathbb{R}^n$, $V : [0, \infty) \to \mathrm{Sym}(b)$, and $R : [0, \infty) \to O(n)$ are given by, for every $t \geqslant 0$,

$$\bar{x}(t) = \mathbb{E}_{\nu_t}[x] = \fint_{\mathcal{U}(t)} x d\nu_t(x),\ V(t) = \mathbb{V}_{\nu_t}[x] = \fint_{\mathcal{U}(t)} (x - \bar{x}(t)) \otimes (x - \bar{x}(t)) d\nu_t(x), \text{ and } R(t) = V(t)^{1/2} V(0)^{1/2},$$

and $\bar{u} : [0, \infty) \to \mathbb{R}^n$ and $\Omega : [0, \infty) \to \mathrm{Skew}(n)$ are given by, for every $t \geqslant 0$,

$$\bar{u}(t) = \dot{\bar{x}}(t) \text{ and } \Omega(t) = \dot{R}(t)R(t)^T = \left(V(t)^{1/2}\right)' V(t)^{-1/2}.$$

PROOF. First we show that $\eta = \bar{x} + R\left(\cdot - \bar{x}\left(0\right)\right)$. Since η is a rigid motion, we have by Proposition 2.11 that $\eta\left(t, y\right) = z\left(t\right) + R\left(t\right) y$ for some $z : [0, \infty) \to \mathbb{R}^n$ and some $R : [0, \infty) \to O\left(n\right)$. The key computation is:

$$\frac{1}{\mathcal{M}\left(\mathcal{U}\left(t\right)\right)} \int_{\mathcal{U}(t)} x \rho_t\left(x\right) dx \overset{(1)}{=} \frac{1}{\mathcal{M}\left(\mathcal{U}\left(t\right)\right)} \int_{\mathcal{U}_0} \eta\left(y, t\right) \rho_t\left(\eta\left(y, t\right), t\right) dy \overset{(2)}{=} \frac{1}{\mathcal{M}\left(\mathcal{U}_0\right)} \int_{\mathcal{U}_0} \left(z\left(t\right) + R\left(t\right) y\right) \rho_0\left(y\right) dy$$

which relies in (1) on the fact that η is a rigid motion and so by Lemma 2.14 $|\det \nabla \eta| = 1$, and in (2) on the fact that $\rho_t = \rho_0 \circ \eta_t$ by Proposition 4.3. So finally, where the computation above comes into play in $(*)$, we obtain that:

$$\overline{x} = \mathbb{E}_{\nu_t}\left(x\right) \overset{(*)}{=} \mathbb{E}_{\nu_0}\left(z + Ry\right) = z + R\,\mathbb{E}_{\nu_0}\left(y\right) = z + R\,\overline{x}\left(0\right).$$

Now we show that $R = V^{1/2} V_0^{-1/2}$. To do this, we first compute how to express the covariance matrix at a time t in terms of the initial covariance matrix. Using the same change of variable as above, and observing that $x - \bar{x} = \eta - \bar{x} = R\left(y - \bar{x}_0\right)$, we obtain that, using Lemma 6.5,

$$\begin{aligned} V = \mathbb{V}_{\nu_t}\Big(\left(x - \bar{x}\right) \otimes \left(x - \bar{x}\right)\Big) &= \mathbb{V}_{\nu_0}\Big(R\left(y - \bar{x}_0\right) \otimes R\left(y - \bar{x}_0\right)\Big) \\ &= R\,\mathbb{V}_{\nu_0}\Big(\left(y - \bar{x}_0\right) \otimes \left(y - \bar{x}_0\right)\Big) R^T = R V_0 R^T. \end{aligned}$$

Now all that is left to do is to solve the conjugacy equation $V = RV_0R^T$ for R. We make the educated guess $R := V^{1/2}V_0^{-1/2}$, noting that the square root of V is well-defined since V is symmetric and positive-definite, and observe that then:

$$RV_0R^T = V^{1/2}\left(V_0^{-1/2}V_0\left(V_0^{-1/2}\right)^T\right)\left(V^{1/2}\right)^T = V^{1/2}V^{1/2} = V$$

where we have used that for S symmetric and positive-definite, $\left(S^{1/2}\right)^T = \left(S^T\right)^{1/2} = S^{1/2}$.

Now we show that $\Omega = \dot{R}R^{-1}$. Since the flow map is

$$\eta_t(y) = \overline{x}(t) + R(t)(y - \overline{x}(0))$$

it follows that the inverse flow map is

$$\eta_t^{-1}(x) = \overline{x}(0) + R(t)^{-1}(x - \overline{x}(t)).$$

We can therefore compute the Lagrangian and Eulerian velocities to be

$$v(y,t) = \frac{\mathrm{d}}{\mathrm{d}t}\eta_t(y) = \dot{\overline{x}} + \dot{R}(t)(y - \overline{x}(0))$$

and

$$u(x,t) = v\left(\eta_t^{-1}(x), t\right) = \dot{\overline{x}}(t)\overline{u}(t) + \dot{R}(t)R(t)^{-1}(x - \overline{x}(t)) = \bar{u}(t) + \Omega(t)(x - \bar{x}(t)).$$

Finally we show that $\Omega = \left(V^{1/2}\right)' V^{-1/2}$. We proceed as above, but using V instead of R. Indeed, since

$$\eta_t = \bar{x} + V^{1/2}V_0^{-1/2}(\cdot - \bar{x}_0)$$

it follows that

$$\eta_t^{-1} = \bar{x}_0 + \left(V^{1/2}V_0^{-1/2}\right)^{-1}(\cdot - \bar{x}).$$

Therefore

$$v = \bar{x} + \left(V^{1/2}V_0^{-1/2}\right)' \text{ and } u = \dot{\bar{x}} + \left(V^{1/2}V_0^{1/2}\right)'\left(V^{1/2}V_0^{-1/2}\right)^{-1}(\cdot - \bar{x}) = \dot{\bar{x}} + (V^{1/2})'V^{-1/2}(\cdot - \bar{x}).$$

□

Proposition 4.5 above helps us motivate the definition of a micropolar continuum provided in Definition 2.17. Indeed, Proposition 4.5 tells us that a rigid motion way be fully characterized by the behaviour of its center of mass $\bar{x}(t)$ and some rotation matrix $R(t)$, and therefore it stands to reason that we would define the motion of a micropolar continuum to be determined by a flow map η and a microrotation map Q: η plays the role of $\bar{x}$ since it tracks the translational motion of the microstructure, and Q plays the role of R since it tracks the rotational motion of the microstructure.

4.2. Microinertia. In this section we define the microinertia of a micropolar continuum, define what it means for microinertia to be conserved, and finally derive the associate local conservation law. First we recall that we have characterized the space of admissible moments of inertia.

Definition 4.6. (Admissible moments of inertia)
We denote by $\mathcal{I}(n)$ the set of *admissible moments of inertia*, i.e. in light of Theorem 3.19,

$$\mathcal{I}(n) := \left\{ J \in \mathbb{R}^{n\times n} \;\middle|\; J \geqslant 0, J = J^T, \text{ and its eigenvalues } \mu_1, \ldots, \mu_n \text{ satisfy } \mu_i \leqslant \frac{1}{n-1}\sum_{j=1}^{n}\mu_j \text{ for all } i \right\}.$$

We now define the microinertia of a micropolar continuum, which is nothing more than a function defined along the flow taking values in the space of admissible moments of inertia.

Definition 4.7. (Microinertia)
Let (Ω_0, η, Q) be a micropolar continuum.

(1) A Lagrangian function defined along the flow with codomain $\mathcal{I}(n)$, i.e. a map

$$i : \Omega_0 \times [0,\infty) \to \mathcal{I}(n),$$

is also called a *Lagrangian microinertia density*.

(2) Given a Lagrangian mass measure μ and a Lagrangian microinertia density i, the Lagrangian measure defined along the flow $\gamma := i\mu$ is also called a *Lagrangian microinertia measure*, and we say that γ is *subordinate* to μ to mean that γ is absolutely continuous with respect to μ. Moreover, if μ has an associated Lagrangian mass density σ (i.e. $\mu \ll \mathcal{L}^n$ and $\sigma = \frac{\mathrm{d}\mu}{\mathrm{d}\mathcal{L}^n}$) then $I := i\sigma$ is called a *Lagrangian microinertia*.

(3) An Eulerian function defined along the flow with codomain $\mathcal{I}(n)$ is also called an *Eulerian microinertia density*.

(4) Given an Eulerian mass measure ν and an Eulerian microinertia density j, the Eulerian measure defined along the flow λ, defined via $\lambda_t := j_t\nu_t$ for all $t \geqslant 0$, is also called an *Eulerian microinertia measure*, and we say that λ is *subordinate* to ν to mean that λ_t is absolutely continuous with respect to ν_u for every $t \geqslant 0$. Moreover, if ν has an associated Eulerian mass density ρ (i.e. $\nu_t \ll \mathcal{L}^n$ and $\rho_t = \frac{\mathrm{d}\nu_t}{\mathrm{d}\mathcal{L}^n}$ for all $t \geqslant 0$) then $J := j\rho$ is called an *Eulerian microinertia*.

In Definition 4.7 above we make a careful distinction between the microinertia density and the microinertia. However in the sequel we will almost exclusively devote our attention to incompressible flows, in which case the microinertia density and the microinertia are the same up to a constant factor of the mass density ρ. Nonetheless, it is important to make this distinction in order for the foundation of micropolar continuum mechanics laid so far to also be useful when it is used to investigate *compressible* micropolar fluids.

We now define what it means for the microinertia density to be conserved. This definition is inspired from the way in which the moment of inertia transforms under rigid motions that preserve the center of mass, as recorded in Corollary 3.10.

Definition 4.8. (Conservation of a microinertia density)
Let (Ω_0, η, Q) be a micropolar continuum

(1) A Lagrangian microinertia density i satisfying $i = Qi_0Q^T$, i.e.

$$i\,(t,y) = Q\,(t,y)\,i\,(0,y)\,Q^T\,(t,y) \qquad \text{for all } (t,y) \in [0,\infty) \times \Omega_0$$

is called a *conserved Lagrangian microinertia density*.

(2) An Eulerian microinertia density j is called a *conserved Eulerian microinertia density* if $j_t = i_t \circ \eta_t^{-1}$ for all $t \geqslant 0$ for some conserved Lagrangian microinertia density i.

Having defined what it means for microinertia density to be conserved we derive the associate local conservation law.

Proposition 4.9. *(Local conservation of a microinertia density)*
Let (Ω_0, η, Q) be a micropolar continuum with Eulerian velocity u, Eulerian angular velocity tensor Ω and conserved Eulerian microinertia density j. Then

$$\partial_t j + (u \cdot \nabla) j - [\Omega, j] = 0.$$

Proof. Let i be the Lagrangian microinertia density corresponding to j, i.e. $i = j \circ \eta$. Then, by conservation of microinertia,

$$j \circ \eta = i = Qi_0Q^T.$$

Upon taking a time derivative, using Proposition 2.6 to compute $\frac{\mathrm{d}}{\mathrm{d}t}(j \circ \eta)$ and denoting by Θ the Lagrangian angular velocity tensor $\Theta := \Omega \circ \eta = (\partial_t Q)\,Q^{-1}$ we have that

$$\begin{aligned}((\partial_t + u \cdot \nabla)\,j) \circ \eta = \partial_t Q i_0 Q^T + Q i_0 \partial_t Q^T &= \left(\partial_t Q Q^{-1}\right)\left(Q i_0 Q^T\right) + \left(Q i_0 Q^T\right)\left(Q^{-T}\partial_t Q^T\right) \\ &= \Theta\,(j \circ \eta) + (j \circ \eta)\,\Theta^T.\end{aligned}$$

So finally, precomposing by η^{-1} on both sides and recalling that Θ and Ω are anti-symmetric we obtain that

$$\partial_t j + (u \cdot \nabla)\,j = \left(\Theta \circ \eta^{-1}\right) j + j\left(\Theta \circ \eta^{-1}\right)^T = \Omega j - j\Omega.$$

□

It is worth noting that the differential operator $\partial_t + u \cdot \nabla - [\Omega, \cdot]$, which appears in Proposition 4.9 and will appear throughout the sequel whenever the microinertia is involved, is an analog of the well-known advection operator $\partial_t + u \cdot \nabla$. Indeed: while the advection operator $\partial_t + u \cdot \nabla$ takes into account the change in a quantity defined along the flow due to the advection by the flow map , a heuristic explanation made precise in Proposition 2.6, the operator $\partial_t + u \cdot \nabla - [\Omega, \cdot\,]$ also takes into account the rotation due to the microrotation map. We will thus refer to this operator as an *advection-rotation* operator.

We now note that, as a consequence of Proposition 4.2 and Proposition 4.9 which provide us with the local conservations of mass and microinertia density respectively, we may now derive the local conservation law satisfied by the microinertia.

COROLLARY 4.10. *(Local conservation of the microinertia)*
Let (Ω_0, η, Q) be a micropolar continuum with Eulerian velocity u and Eulerian angular velocity tensor Ω. Let ρ be an Eulerian mass density which is conserved and let j be a conserved Eulerian microinertia density. Then the microinertia $J := j\rho$ satisfies

$$\partial_t J + \nabla \cdot (J \otimes u) - [\Omega, J] = 0.$$

PROOF. In light of Proposition 4.2 which says that $D_t^u j = [\Omega, j]$, Proposition 4.9 which says that $\mathbb{D}_t^u \rho = 0$, and Lemma 6.7 which provides a "Leibniz Rule" for material derivatives, this result is the consequence of a direct computation:

$$\partial_t J + \nabla \cdot (J \otimes u) = \mathbb{D}_t^u J = \mathbb{D}_t^u (j\rho) = (D_t^u j)\rho + j(\mathbb{D}_t^u) = [j, \Omega]\,\rho = [J, \Omega]\,.$$

□

4.3. Momenta and stresses. In this section we define the linear momentum measure and angular momentum measure associated with a conserved mass measure, state what it means for linear and angular momentum to be conserved – introducing the stress and couple stress tensor in the process, and conclude with the derivation of the associated local conservation law. First we define the linear momentum measure and angular momentum measure.

DEFINITION 4.11. (Linear momentum measure, linear momentum density)
Let (Ω_0, η) be a continuum, let ν be an Eulerian mass measure, and let u denote the Eulerian velocity of the continuum. We call νu a *linear momentum measure.* Moreover, if ρ is the Eulerian mass density of ν, i.e. $d\nu = \rho d\mathcal{L}^n$, then we call ρu a *linear momentum density.*

DEFINITION 4.12. (Angular momentum measure, angular momentum density)
Let $n = 3$, let (Ω_0, η, Q) be a micropolar continuum, let ν be an Eulerian mass measure, let j be an Eulerian microinertia density and let ω be the Eulerian angular velocity of the continuum. We call $j\omega\nu$ a *angular momentum measure.* Moreover, if ρ is the Eulerian mass density of ν, i.e. $d\nu = \rho d\mathcal{L}^n$, and $J := \rho i$ is the associated Eulerian microinertia then we call $J\omega$ a *angular momentum density.*

With Definition 4.11 and Definition 4.12 in hand we are now ready to define what it means for linear and angular momentum to be conserved. Prior to doing so we introduce the notion of a *physical* micropolar continuum. This is simply a convenient way to combine together *kinematic* and *physical* information. The kinematic information, namely the flow map and the microrotation map, is already built in the definition of a micropolar continuum, and so when defining a physical micropolar continuum we add in physical information, namely postulating the existence of a conserved mass measure and a conserved microinertia density.

DEFINITION 4.13. (Physical micropolar continuum)
A *physical micropolar continuum* is a tuple $(\Omega_0, \eta, Q, \nu, j)$ such that
(1) (Ω_0, η, Q) is a micropolar continuum,
(2) ν is a conserved Eulerian mass measure, and
(3) j is a conserved Eulerian microinertia density.

At last we are equipped to define what it means for linear and angular momentum to be conserved. Note that this introduces the notion of a stress tensor and a couple stress tensor. The postulation that such tensors exist is a core tenet of rational continuum mechanics. Physically, these stress and couple stress tensors are manifestations of Newton's third law: "Every action creates an equal and opposite reaction", which in this context means that they encode how the fluid reacts to forces and torques induces by the neighbouring fluid.

DEFINITION 4.14. (Balance of momenta, stress and couple stress tensor, and external force and torque)
Let $n = 3$, let $(\Omega_0, \eta, Q, \nu, j)$ be a physical micropolar continuum, let
- u denote the Eulerian velocity,

- ω denote the Eulerian angular velocity,
- ρ denote the Eulerian mass density, i.e. $d\nu = \rho d\mathcal{L}^3$,
- $J = \rho j$ denote the Eulerian microinertia,
- T, M be Eulerian functions defined along the flow with codomain $\mathbb{R}^{3\times 3}$ called the *stress tensor* and *couple stress tensor* respectively,
- f, τ be Eulerian functions defined along the flow with codomain $\mathbb{R}^3$ called the *external force* and *external torque* respectively.

If, for every $\mathcal{U}_0 \subset\subset \Omega_0$,

$$\frac{\mathrm{d}}{\mathrm{dt}} \int_{\mathcal{U}(t)} \rho u \, dx = \int_{\partial \mathcal{U}(t)} Tn \, dx + \int_{\mathcal{U}(t)} f \, dx$$

and

$$\frac{\mathrm{d}}{\mathrm{dt}} \left(\int_{\mathcal{U}(t)} x \times \rho u \, dx + \int_{\mathcal{U}(t)} J\omega \, dx \right) = \int_{\partial \mathcal{U}(t)} x \times (Tn) \, dx + \int_{\partial \mathcal{U}(t)} Mn \, dx + \int_{\mathcal{U}(t)} x \times f \, dx + \int_{\mathcal{U}(t)} \tau \, dx$$

where $\mathcal{U}(t) := \eta_t(\mathcal{U}_0)$ for every $t \geqslant 0$, then we say that the physical micropolar continuum is *governed by* (T, M) *subject to* (f, τ). Moreover, the two integral equations above are referred to as the *balance of linear momentum* and the *balance of angular momentum* respectively.

Note that in the definition of the balances of linear and angular momentum above we restrict the integral balances to subsets of the continuum satisfying $\mathcal{U}_0 \subset\subset \Omega_0$, i.e. staying away from the boundary of the continuum. This is done because the stress tensor T and the couple stress tensor M are manifestations of Newton's third law. Indeed, as mentioned earlier, these tensors encode the fact that the continuum reacts to forces and torques induced at a point by the continuum present in a neighbourhood of that point. In particular, if $\partial\mathcal{U}_0 \cap \partial\Omega_0 \neq$ then the balances of linear and angular momentum must be modified slightly from their respective versions provided in Definition 4.14 above. This is done in Section 4.5.

Lemmas 4.15 and 4.16 below are useful intermediate computations which are consequences of the local conservation of mass and microinertia respectively, as well as the "Leibniz Rule" for material derivatives recorded in Lemma 6.7. This two lemmas allow us to compute the material derivatives of the linear momentum density ρu and of the angular momentum density $J\omega$. These computations come in very handy when deriving the local version of the balance of linear and angular momentum in Proposition 4.17.

Lemma 4.15. *(Material derivative of the linear momentum density)*
Let (Ω_0, η) be a continuum with Eulerian velocity u. If a conserved Eulerian mass measure is absolutely continuous with respect to the Lebesgue measure, then the following identity holds:

$$\partial_t(\rho u) + \nabla \cdot (\rho u \otimes u) = \rho \left(\partial_t u + (u \cdot \nabla) u \right).$$

Proof. This follows from the local conservation of mass (which says that $\mathbb{D}_t^u \rho = 0$) and Lemma 6.7:

$$\partial_t(\rho u) + \nabla \cdot (\rho u \otimes u) = \mathbb{D}_t^u(\rho u) = (\mathbb{D}_t^u \rho) u + \rho (\mathrm{D}_t^u u) = 0 + \rho \left(\partial_t u + (u \cdot \nabla) u \right).$$

$\square$

Lemma 4.16. *(Material derivative of the angular momentum density)*
Let (Ω_0, η, Q) be a micropolar continuum with Eulerian velocity u and Eulerian angular velocity ω. Let ρ be an Eulerian mass density which is conserved, let j be a conserved Eulerian microinertia density, and let $J := j\rho$ denote the microinertia. The following identity holds:

$$\partial_t(J\omega) + \nabla \cdot (J\omega \otimes u) = J \left(\partial_t \omega + (u \cdot \nabla) \omega \right) + \omega \times J\omega.$$

Proof. This follows from Corollary 4.10 and Lemma 6.7. Indeed, local conservation of microinertia says that $\mathbb{D}_t^u J = [\Omega, J]$ where $\Omega = \operatorname{ten} \omega$ is the Eulerian angular velocity tensor, and hence we obtain that, since $(\operatorname{ten} \omega)\omega = \omega \times \omega = 0$,

$$(\mathbb{D}_t^u J)\,\omega = [\Omega, J]\,\omega = (\operatorname{ten} \omega) J\omega - J(\operatorname{ten} \omega)\omega = \omega \times (J\omega).$$

So finally:

$$\partial_t(J\omega) + \nabla \cdot ((J\omega) \otimes u) = \mathbb{D}_t^u (J\omega) = (\mathbb{D}_t^u J)\,\omega + J\,(\mathrm{D}_t^u \omega) = \omega \times (J\omega) + J\left(\partial_t \omega + (u \cdot \nabla) \omega \right)$$

$\square$

We now come to the main result of this section, and one of the main results of this chapter, as we derive the local version of the balance of linear and angular momentum.

PROPOSITION 4.17. *(Local conservation of linear and angular momentum)*
Let $n = 3$, let $(\Omega_0, \eta, Q, \nu, j)$ be a physical micropolar continuum governed by (T, M) subject to (f, τ), and let

- *u denote the Eulerian velocity,*
- *ω denote the Eulerian angular velocity,*
- *ρ denote the Eulerian mass density, i.e. $d\nu = \rho d\mathcal{L}^3$, and*
- *$J = \rho j$ denote the Eulerian microinertia.*

Then

$$\rho\left(\partial_t u + (u\cdot\nabla)\,u\right) = \nabla\cdot T + f \text{ and } J\left(\partial_t\omega + (u\cdot\nabla)\,\omega\right) + \omega\times(J\omega) = 2\operatorname{vec}T + \nabla\cdot M + \tau.$$

PROOF. The key idea is to use the Reynolds and divergence theorems (c.f. Theorem 2.7) to write all terms of the balance of linear and angular momentum as integrals over $\mathcal{U}(t)$, no longer differentiated in time, and to then use the conservation of mass and the conservation of micro-inertia to simplify the resulting equations.

First we deal with linear momentum. The balance of linear momentum is

$$\frac{\mathrm{d}}{\mathrm{d}t}\int_{\mathcal{U}(t)}\rho u = \int_{\partial\mathcal{U}(t)} T\cdot n + \int_{\mathcal{U}(t)} f.$$

By the Reynolds transport theorem

$$\frac{\mathrm{d}}{\mathrm{d}t}\int_{\mathcal{U}(t)}\rho u = \int_{\mathcal{U}(t)}\partial_t(\rho u) + \nabla\cdot(\rho u\otimes u)$$

and by the divergence theorem

$$\int_{\partial\mathcal{U}(t)} T\cdot n = \int_{\mathcal{U}(t)}\nabla\cdot T.$$

We may thus write the balance of linear momentum as

$$\int_{\mathcal{U}(t)}\partial_t(\rho u) + \nabla\cdot(\rho u\otimes u) = \int_{\mathcal{U}(t)}\nabla\cdot T + f.$$

Since $\mathcal{U}_0 \subset\subset \Omega_0$ is arbitrary and η_t is a diffeomorphism it follows that the integral equation immediately above holds for arbitrary subsets $\mathcal{U}(t)\subseteq\Omega(t)$ and hence the following PDE holds pointwise:

$$\partial_t(\rho u) + \nabla\cdot(\rho u\otimes u) = \nabla\cdot T + f \tag{4.1}$$

In particular, note that local conservation of mass, and more specifically Lemma 4.15, tells us that

$$\partial_t(\rho u) + \nabla\cdot(\rho u\otimes u) = \rho\left(\partial_t u + (u\cdot\nabla)\,u\right).$$

So finally:

$$\rho\left(\partial_t u + (u\cdot\nabla)\,u\right) = \nabla\cdot T + f. \tag{4.2}$$

Now we deal with angular momentum. We begin by recording some preliminary computations:

$$\nabla\cdot((x\times\rho u)\otimes u) = x\times\nabla\cdot(\rho u\otimes u)\,, \tag{4.3}$$
$$x\times(Tn) = (x\times T)\,n, \text{ and} \tag{4.4}$$
$$\nabla\cdot(x\times T) = x\times(\nabla\cdot T) + 2\operatorname{vec}T, \tag{4.5}$$

where we define $(v\times v)_{ij} := \epsilon_{iab}v_a T_{bj}$. Note that, in particular, it follows from (4.3) that

$$\mathbb{D}_t^u(x\times\rho u) = x\times\mathbb{D}_t^u(\rho u)\,. \tag{4.6}$$

Indeed, (4.3), (4.4), and (4.5) follow from direct computations:

$$\begin{aligned}(\nabla\cdot((x\times\rho u)\otimes u))_i &= \partial_j\left(\epsilon_{ikl}x_k\rho u_l u_j\right) = \epsilon_{ikl}x_k\partial_j\left(\rho u_l u_j\right) + \epsilon_{ikl}\delta_{jk}\rho u_l u_j\\ &= (x\times\nabla\cdot(\rho u\otimes u))_i + \epsilon_{ijl}\rho u_l u_j\end{aligned}$$

where $\epsilon_{ijl}u_l u_j = 0$ since ϵ_{ijl} is anti-symmetric with respect to (j,l) but $u_l u_j$ is symmetric with respect to (j,l),

$$(x \times (Tn))_i = \epsilon_{ijk} x_j T_{kl} n_l = (x \times T)_{il} n_l = ((x \times T)\, n)_i ,$$

and finally

$$(\nabla \cdot (x \times T))_i = \partial_j \left(\epsilon_{ikl} x_k T_{lj}\right) = \epsilon_{ikl} x_k \partial_j T_{lj} + \epsilon_{ikl} \delta_{jk} T_{lj} = (x \times (\nabla \cdot T))_i + (2 \operatorname{vec} T)_i .$$

Now recall that the balance of angular momentum is

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(\int_{\mathcal{U}(t)} x \times \rho u + \int_{\mathcal{U}(t)} J\omega\right) = \int_{\partial\mathcal{U}(t)} x \times (Tn) + \int_{\partial\mathcal{U}(t)} Mn + \int_{\mathcal{U}(t)} x \times f + \int_{\mathcal{U}(t)} \tau .$$

By the Reynolds transport theorem, (4.6), Lemma 4.15, and Lemma 4.16 we obtain that

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}t}\left(\int_{\mathcal{U}(t)} x \times \rho u + J\omega\right) &= \int_{\mathcal{U}(t)} \mathbb{D}_t^u (x \times \rho u) + \mathbb{D}_t^u (J\omega) \\ &= \int_{\mathcal{U}(t)} x \times \mathbb{D}_t^u (\rho u) + \mathbb{D}_t^u (J\omega) = \int_{\mathcal{U}(t)} x \times (\rho\, \mathrm{D}_t^u\, u) + J\, (\mathrm{D}_t^u\, \omega) + \omega \times (J\omega)\end{aligned}$$

whilst using (4.4), the divergence theorem, and (4.5) yields

$$\begin{aligned}\int_{\partial\mathcal{U}(t)} x \times (Tn) + Mn &= \int_{\partial\mathcal{U}(t)} (x \times T + M)\, n = \int_{\mathcal{U}(t)} \nabla \cdot (x \times T + M) \\ &= \int_{\mathcal{U}(t)} x \times (\nabla \cdot T) + 2 \operatorname{vec} T + \nabla \cdot M .\end{aligned}$$

We may thus write the balance of angular momentum as

$$\int_{\mathcal{U}(t)} x \times (\rho\, \mathrm{D}_t^u\, u) + J\, (\mathrm{D}_t^u\, \omega) + \omega \times (J\omega) = \int_{\mathcal{U}(t)} x \times (\nabla \cdot T + f) + 2 \operatorname{vec} T + \nabla \cdot M + \tau .$$

Since $\mathcal{U}_0 \subset\subset \Omega_0$ is arbitrary and η_t is a diffeomorphism it follows that the integral equation immediately above holds for arbitrary subsets $\mathcal{U}(t) \subseteq \Omega(t)$ and hence the following PDE holds pointwise:

$$x \times (\rho\, \mathrm{D}_t^u\, u) + J\, (\mathrm{D}_t^u\, \omega) + \omega \times (J\omega) = x \times (\nabla \cdot T + f) + 2 \operatorname{vec} T + \nabla \cdot M + \tau .$$

In particular, since (4.2) says that

$$\rho\, \mathrm{D}_t^u\, u = \nabla \cdot T + f$$

it follows that

$$J\, (\mathrm{D}_t^u\, \omega) + \omega \times (J\omega) = 2 \operatorname{vec} T + \nabla \cdot M + \tau$$

or, in more expansive but standard notation,

$$J\, (\partial_t \omega + (u \cdot \nabla)\, \omega) + \omega \times (J\omega) = 2 \operatorname{vec} T + \nabla \cdot M + \tau .$$

□

4.4. Energy. Having introduced the conservation of mass, microinertia, and linear and angular momentum, we conclude Section 4 by considering another conserved quantity: energy. We proceed with the process: we define the kinetic energy associated with a continuum, we compute its expression for a rigid body and use that to define an appropriate notion of the conservation of energy for micropolar continua, and finally we derive a local version of the balance of energy.

Definition 4.18. (Kinetic energy of a flow map)
Let (Ω_0, η) be a continuum with Eulerian velocity u and conserved mass measure ν. For any Borel subset $E \subseteq \Omega(t)$ we define the *kinetic energy* of E to be $K(E) := \frac{1}{2}||u||^2_{L^2_\nu(E)}$, i.e.

$$K(E) = \frac{1}{2}\int_E |u(x,t)|^2 d\nu_t(x) .$$

As alluded to earlier, now that we have defined the kinetic energy associated with a continuum we compute its expression for a rigid body. This makes use of the decomposition of the Eulerian velocity of a rigid body established in Proposition 3.3.

Proposition 4.19. *(Kinetic energy of 3-dimensional rigid motions)*
Let $n = 3$, let (Ω_0, η) be a rigid body, and let $\bar{u}$ and $\bar{\omega}$ be as in Proposition 3.3. Then the kinetic energy of the rigid body is a quadratic form

$$K : \mathbb{R}^3 \times \mathbb{R}^3 \to [0, \infty)$$
$$(\bar{u}, \bar{\omega}) \mapsto \frac{1}{2}M|\bar{u}|^2 + \frac{1}{2}J : (\bar{\omega} \otimes \bar{\omega})$$

where M and J are the mass and moment of inertia of the rigid body, respectively.

Proof. First we show that $\bar{u}$ and $\bar{\omega} \times (\cdot - \bar{x})$ are orthogonal in L^2_ν. This is a direct computation:

$$\Big(\bar{u}, \bar{\omega} \times (\cdot - \bar{x})\Big)_{L^2_\nu} = M \fint_{\mathcal{U}(t)} \bar{u} \cdot \bar{\omega} \times (x - \bar{x})\, d\nu_t(x) = M\bar{u} \cdot \bar{\omega} \times \underbrace{\left(\fint_{\mathcal{U}(t)} x d\nu_t(x) - \bar{x}\right)}_{=0} = 0$$

where $M := \mathcal{M}(\mathcal{U}(t))$. Since $\bar{u}$ and $\bar{\omega} \times (\cdot - \bar{x})$ are orthogonal in L^2_ν, we know that

$$K(\bar{u}, \bar{\omega}) = \frac{1}{2}\tilde{I} : (\bar{u} \otimes \bar{u}) + \frac{1}{2}\tilde{J} : (\bar{\omega} \otimes \bar{\omega})$$

for some $\tilde{I}, \tilde{J} \in \mathbb{R}^{3x3}$. Let us compute $\tilde{I}$ and $\tilde{J}$:

$$K(\bar{u}, \bar{\omega}) = \frac{1}{2}||u||^2_{L^2_\nu} = \frac{1}{2}||\bar{u}||^2_{L^2_\nu} + \frac{1}{2}||\bar{\omega} \times (\cdot - \bar{x})||^2_{L^2_\nu}$$

where

$$||\bar{u}||^2_{L^2_\nu} = \int_{\mathcal{U}(t)} |\bar{u}|^2 d\nu_t = M|\bar{u}|^2$$

i.e. $\tilde{I} = MI$, and where

$$||\bar{\omega} \times (\cdot - \bar{x})||^2_{L^2_\nu} = \int_{\mathcal{U}(t)} |\bar{\omega} \times (x - \bar{x})|^2 d\nu_t(x) = \int_{\mathcal{U}(t)} \Big(|\bar{\omega}|^2|x - \bar{x}|^2 - |\bar{\omega} \cdot (x - \bar{x})|^2\Big) d\nu_t(x)$$
$$= \int_{\mathcal{U}(t)} \Big(|x - \bar{x}|^2 I - (x - \bar{x}) \otimes (x - \bar{x})\Big) d\nu_t(x) : (\bar{\omega} \otimes \bar{\omega}) = M((\mathrm{tr}\, V) I - V) : (\bar{\omega} \otimes \bar{\omega}) = J : (\bar{\omega} \otimes \bar{\omega})$$

where we have used Proposition 6.3 to expand $|\bar{\omega} \times (x - \bar{x})|^2$. So indeed $\tilde{J} = J$. □

Note that Proposition 4.19 helps motivate the definition of the kinetic energy density in Definition 4.20 below where we define what it means for energy to be conserved for a micropolar continuum.

Definition 4.20. (Purely mechanical micropolar continuum, balance of energy, and related notions)
Let $n = 3$. We say that a physical micropolar continuum $(\Omega_0, \eta, Q, \nu, j)$ governed by (T, M) subject to (f, τ), where

- u denote the Eulerian velocity,
- ω denote the Eulerian angular velocity,
- ρ denote the Eulerian mass density, i.e. $d\nu = \rho d\mathcal{L}^3$,
- $J = \rho j$ denote the Eulerian microinertia,

is *purely mechanical* if there exist Eulerian functions defined along the flow with codomain $\mathbb{R}_+$, denoted ϵ and δ, called respectively the *mechanical energy density* and the *thermodynamic dissipation density*, such that

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(\int_{\mathcal{U}(t)} (\epsilon + K)\right) + \int_{\mathcal{U}(t)} \delta = \int_{\partial\mathcal{U}(t)} (Tn) \cdot u + (Mn) \cdot \omega + \int_{\mathcal{U}(t)} f \cdot u + \tau \cdot \omega$$

where n is the outer unit normal and where $K := \frac{1}{2}\rho|u|^2 + \frac{1}{2}J : \omega \otimes \omega$ is called the *kinetic energy density*. Moreover, the integral equation above is referred to as the *balance of energy.*

We are now ready to conclude this section by obtaining the local version of the balance of energy introduce in Definition 4.20 above.

Theorem 4.21. *(Local conservation of energy)*
Let $n = 3$, let $(\Omega_0, \eta, Q, \nu, j)$ be a purely mechanical physical micropolar continuum governed by (T, M) subject to (f, τ), and let

- u *denote the Eulerian velocity,*
- ω *denote the Eulerian angular velocity,*
- Ω *denote the Eulerian angular velocity tensor,*
- ϵ *denote the mechanical energy density.*

Then

$$\partial_t \epsilon + \nabla \cdot (\epsilon u) \leqslant (\nabla u - \Omega) : T + \nabla \omega : M.$$

PROOF. We begin by observing that, since the thermodynamic dissipation density δ is non-negative, we can re-write the balance of energy as the inequality

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(\int_{\mathcal{U}(t)} (\epsilon + K)\right) \leqslant \int_{\partial\mathcal{U}(t)} (T \cdot n) \cdot u + (M \cdot n) \cdot \omega + \int_{\mathcal{U}(t)} f \cdot u + \tau \cdot \omega,$$

which will hereafter be referred to as the *energy inequality*.

We now seek to rewrite all integrals appearing in the energy inequality in terms of non-time-differentiated integrals over the bulk domain $\mathcal{U}(t)$. In particular, we use the Reynolds transport theorem to see that

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(\int_{\mathcal{U}(t)} (\epsilon + K)\right) = \int_{\mathcal{U}(t)} \partial_t (\epsilon + K) + \nabla \cdot ((\epsilon + K) u)$$

and we use the divergence theorem to see that

$$\int_{\partial\mathcal{U}(t)} (T \cdot n) \cdot u + (M \cdot n) \cdot \omega = \int_{\mathcal{U}(t)} \nabla \cdot (u \cdot T) + \nabla \cdot (\omega \cdot M)$$

since indeed, for any differentiable matrix and vector fields A and v and any sufficiently regular open set $\mathcal{U}$,

$$\int_{\partial\mathcal{U}} (An) \cdot v = \int_{\partial\mathcal{U}} A_{ij} n_j v_i = \int_{\mathcal{U}} \partial_j (A_{ij} v_i) = \int_{\mathcal{U}} \nabla \cdot (v \cdot A).$$

For simplicity, we now define the *power density* $P := \nabla \cdot (u \cdot T) + \nabla \cdot (\omega \cdot M) + f \cdot u + \tau \cdot \omega$. We can thus write the energy inequality as

$$\int_{\mathcal{U}(t)} \partial_t (\epsilon + K) + \nabla \cdot ((\epsilon + K) u) \leqslant \int_{\mathcal{U}(t)} \nabla \cdot (u \cdot T) + \nabla \cdot (\omega \cdot M) + f \cdot u + \tau \cdot \omega = \int_{\mathcal{U}(t)} P$$

i.e.

$$\int_{\mathcal{U}(t)} \partial_t \epsilon + \nabla \cdot (\epsilon u) \leqslant \int_{\mathcal{U}(t)} P - (\partial_t K + \nabla \cdot (K u)).$$

Since $\mathcal{U}_0 \subseteq \Omega_0$ is arbitrary and η_t is a diffeomorphism it follows that the integral equation immediately above holds for arbitrary subsets $\mathcal{U}(t) \subseteq \Omega(t)$ and hence the following differential inequality holds pointwise:

$$\partial_t \epsilon + \nabla \cdot (\epsilon u) \leqslant P - (\partial_t K + \nabla \cdot (K u)).$$

To conclude the proof, we simply compute $P - (\partial_t K + \nabla \cdot (K u)) = P - \mathbb{D}_t^u K$. The key observation is that

$$\mathbb{D}_t^u K = (\rho \, \mathrm{D}_t^u \, u) \cdot u + (J D_t^u \omega + \omega \times J\omega) \cdot \omega. \tag{4.7}$$

Indeed, this follows from writing $K = \left(\frac{1}{2}\rho u\right) \cdot u + (J\omega) \cdot \omega$, using Lemma 6.7 and simplifying the result using Lemma 4.15, Lemma 4.16, and the symmetry of J:

$$\mathbb{D}_t^u \left(\frac{1}{2}\rho u \cdot u\right) = \mathbb{D}_t^u \left(\frac{1}{2}\rho u\right) \cdot u + \frac{1}{2}\rho u \cdot (\mathrm{D}_t^u \, u) = \frac{1}{2}\rho \, (\mathrm{D}_t^u \, u) \cdot u + \frac{1}{2}\rho u \cdot (\mathrm{D}_t^u \, u) = (\rho \, \mathrm{D}_t^u \, u) \cdot u$$

and

$$\begin{aligned}\mathbb{D}_t^u \left(\frac{1}{2} J\omega \cdot \omega\right) = \mathbb{D}_t^u \left(\frac{1}{2} J\omega\right) \cdot \omega + \frac{1}{2} J\omega \cdot \mathrm{D}_t^u \, \omega &= \left(\frac{1}{2} J \left(\mathrm{D}_t^u \, \omega\right) + \frac{1}{2}\omega \times J\omega\right) \cdot \omega + \frac{1}{2} J\omega \cdot \mathrm{D}_t^u \, \omega \\ &= \frac{1}{2}\left(\left(J + J^T\right)\left(\mathrm{D}_t^u \, \omega\right) + \omega \times J\omega\right) \cdot \omega = J \, \mathrm{D}_t^u \, \omega \cdot \omega.\end{aligned}$$

So finally, using (4.7), Proposition 4.17, Lemma 6.8, Proposition 6.13, and Lemma 6.17 we obtain that

$$\begin{aligned} P - (\partial_t K - \nabla \cdot (Ku)) &= \nabla \cdot (u \cdot T + \omega \cdot M) + f \cdot u + \tau \cdot M - \mathbb{D}_t^u K \\ &= \nabla u : T + \nabla \omega : M + (\nabla \cdot T + f) \cdot u + (\nabla \cdot M + \tau) \cdot \omega \\ &\quad - (\rho \, \mathrm{D}_t^u \, u) \cdot u - (J \, \mathrm{D}_t^u \, \omega + \omega \times J\omega) \cdot \omega \\ &= \nabla u : T + \nabla \omega : M - (2 \operatorname{vec} T) \cdot \omega \\ &= (\nabla u - \Omega) : T + \nabla \omega : M \end{aligned}$$

such that indeed

$$\partial_t \epsilon + \nabla \cdot (\epsilon u) \leqslant (\nabla u - \Omega) : T + \nabla \omega : M.$$

□

4.5. Boundary conditions. In this section we only briefly discuss boundary conditions associated with micropolar continua. The only purpose of the boundary conditions detailed here is to be such that the equations of motion for micropolar fluids ultimately derived are complete, and so we only consider so-called *natural* boundary conditions. That is not to say that there is not much to be done when it comes to discussing appropriate boundary conditions for micropolar continua in various contexts, but such a discussion is simply not the focus here.

This section follows the usual path: we define a version of the balance of linear and angular momentum taking into account boundary effects and derive the local version of this balance law.

DEFINITION 4.22. (External boundary force and torque)
Let $n = 3$, let $(\Omega_0, \eta, Q, \nu, j)$ be a physical micropolar continuum governed by (T, M) subject to (f, τ), and let

- u denote the Eulerian velocity,
- ω denote the Eulerian angular velocity,
- ρ denote the Eulerian mass density, i.e. $d\nu = \rho d\mathcal{L}^3$, and
- $J = \rho j$ denote the Eulerian microinertia.

Let f_b and τ_b Eulerian functions defined along the flow with codomain $\mathbb{R}^3$ called the *external boundary force* and *external boundary torque* respectively. If, for every $\mathcal{U}_0 \subseteq \Omega_0$,

$$\frac{\mathrm{d}}{\mathrm{d}t} \int_{\mathcal{U}(t)} \rho u \, dx = \int_{\partial \mathcal{U}(t) \setminus \partial \Omega(t)} Tn \, dx + \int_{\partial \mathcal{U}(t) \cap \partial \Omega(t)} f_b \, dx + \int_{\mathcal{U}(t)} f \, dx$$

and

$$\begin{aligned} \frac{\mathrm{d}}{\mathrm{d}t} \left(\int_{\mathcal{U}(t)} x \times \rho u \, dx + \int_{\mathcal{U}(t)} J\omega \, dx \right) &= \int_{\partial \mathcal{U}(t) \setminus \partial \Omega(t)} x \times (Tn) \, dx + \int_{\partial \mathcal{U}(t) \setminus \partial \Omega(t)} Mn \, dx \\ &\quad + \int_{\partial \mathcal{U}(t) \cap \partial \Omega(t)} x \times f_b \, dx + \int_{\partial \mathcal{U}(t) \cap \partial \Omega(t)} \tau_b \, dx + \int_{\mathcal{U}(t)} x \times f \, dx + \int_{\mathcal{U}(t)} \tau \, dx \end{aligned}$$

where $\mathcal{U}(t) := \eta_t (\mathcal{U}_0)$ for every $t \geqslant 0$, then we say that the physical micropolar continuum is *subject to the boundary effects* (f_b, τ_b). Moreover, the two integral equations above are referred to as the *balance of linear momentum for boundary flows* and the *balance of angular momentum for boundary flows* respectively.

We now derive the local version of the balance of linear and angular momentum introduced above in Definition 4.22, which are called the *natural boundary conditions.*

PROPOSITION 4.23. *(Natural boundary conditions)*
Let $n = 3$, let $(\Omega_0, \eta, Q, \nu, j)$ be a physical micropolar continuum governed by (T, M) subject to (f, τ) and the boundary effects (f_b, τ_b), and let

- *u denote the Eulerian velocity,*
- *ω denote the Eulerian angular velocity,*
- *ρ denote the Eulerian mass density, i.e. $d\nu = \rho d\mathcal{L}^3$, and*
- *$J = \rho j$ denote the Eulerian microinertia.*

Then

$$Tn = f_b \text{ and } Mn = \tau_n \text{ on } \partial \Omega (t),$$

where n denotes the outer unit normal to $\Omega\,(t)$. These two equations are called the natural boundary conditions *associated with the* boundary effects (f_b, τ_b).

Proof. The key is to write

$$\int_{\partial\mathcal{U}(t)\setminus\partial\Omega(t)} = \int_{\partial\mathcal{U}(t)} - \int_{\partial\mathcal{U}(t)\cap\partial\Omega(t)}$$

where the integrand is either Tn or $x \times Tn + Mn$. Combining this with the balances of linear and angular momentum for boundary flows tells us that

$$\frac{\mathrm{d}}{\mathrm{d}t}\int_{\mathcal{U}(t)} \rho u - \int_{\partial\mathcal{U}(t)} Tn - \int_{\mathcal{U}(t)} f = \int_{\partial\Omega(t)\cap\partial\Omega(t)} (f_b - Tn) \tag{4.8}$$

and

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(\int_{\mathcal{U}(t)} x \times \rho u + \int_{\mathcal{U}(t)} J\omega\right) - \int_{\partial\mathcal{U}(t)} x \times Tn - \int_{\partial\mathcal{U}(t)} Mn - \int_{\mathcal{U}(t)} x \times f - \int_{\mathcal{U}(t)} \tau$$
$$= \int_{\partial\Omega(t)\cap\partial\mathcal{U}(t)} x \times (f_f - Tn) + \int_{\partial\Omega(t)\cap\partial\mathcal{U}(t)} (\tau_b - Mn)\,. \tag{4.9}$$

Proceeding as in the proof of Proposition 4.17 tells us that the left-hand side of (4.8) is

$$\int_{\mathcal{U}(t)} \mathbb{D}_t^u\,(\rho u) - \nabla \cdot T - f$$

which, by Proposition 4.17, vanishes. Similarly, the left-hand side of (4.9) may be written as

$$\int_{\mathcal{U}(t)} \mathbb{D}_t^u\,(x \times \rho u + J\omega) - \nabla \cdot (x \times J\omega) - \nabla \cdot M - x \times f - \tau,$$

which also vanishes. So finally we deduce from (4.8) and the arbitrariness of $\mathcal{U}_0$ that $f_b = Tn$ on $\partial\Omega\,(t)$. Plugging this into (4.9) and once again using the fact that $\mathcal{U}_0$ is arbitrary tells us that $\tau_b = Mn$ on $\partial\Omega\,(t)$. ☐

5. Constitutive equations

To begin this section let us comment on where we stand with respect to the derivation of the equations of motions for micropolar fluids. Combing Corollary 4.10 and Proposition 4.17 tells us what the equations of motion are. To close the system all that we have to do is specify how the stress tensor T and the couple stress tensor M depend on the dynamic variables (u, ω). This is precisely what we do in this section.

First we discuss frame-invariance in Section 5.1, then we record some results on the representation of frame-invariant linear maps in Section 5.2. We discuss the Onsager reciprocity relations in Section 5.3 and we conclude this section by putting it all together to derive the equations of motion of homogeneous incompressible Newtonian micropolar fluids in Section 5.4.

5.1. Frame-invariance. In this section we introduce the notions of frame-invariance for tensor-valued functions, and in particular for functions defined along the flow under the notion of *similarity* of continua and micropolar continua. We then compute how various kinematic quantities behave under similarity..

Definition 5.1. (Similar continua)
Given two continua (Ω_0, η) and $(\tilde{\Omega}_0, \tilde{\eta})$ we say that they are *similar* if there exists a time-dependent orientation-preserving rigid motion $f : [0,\infty) \times \mathbb{R}^n \to \mathbb{R}^n$ such that $\Omega_0 = \tilde{\Omega}_0$ and $\tilde{\eta}_t = f_t \circ \eta_t$ for every $t \geqslant 0$. Moreover we say that f *maps η to $\tilde{\eta}$.*

Note that since rigid motions form a group under composition, similarity of continua is an equivalence relation on the set of continua. The notion of similar continua is important since constitutive equations, which postulate which quantities the stress tensor and couple stress tensor may depend on, must *pass to the quotient* induced by the equivalence relation of similarity. In other words: the constitutive equations for the stress and couple stress tensor must be well-defined on *equivalence classes* of continua. This is made precise in Definition 5.5 under the name of *frame-invariance.*

Of course, the similarity of continua can be interpreted physically: two continua (Ω_0, η) and $(\Omega_0, \tilde{\eta})$ are similar precisely when they correspond to the same system, but viewed by different observers.

Having defined the similarity of continua we now compute how various kinematic quantities behave under similarity.

Proposition 5.2 (The behaviour of kinematic quantities of continua under similarity). *Let (Ω_0, η) and $(\Omega_0, \tilde{\eta})$ be similar continua where the rigid motion f which maps η to $\tilde{\eta}$ is given by*

$$f(t,x) = b(t) + R(t)(x - b_0) \tag{5.1}$$

for $b_0 \in \mathbb{R}^n$, $b : [0,\infty) \to \mathbb{R}^n$, and $R : [0,\infty) \to SO(n)$. Then
(1) $\partial_t \tilde{\eta} = \dot{b} + \dot{R}(\eta - b_0) + R\partial_t \eta$,
(2) $\nabla \tilde{\eta} = R\nabla\eta$,
(3) $\tilde{u} \circ f = \dot{b} + \dot{R}(\,\cdot\, - b_0) + Ru$, *and*
(4) $(\nabla\tilde{u}) \circ f = \dot{R}R^T + R\nabla u R^T$,
where u and $\tilde{u}$ denote the Eulerian velocities of (Ω_0, η) and $(\Omega_0, \tilde{\eta})$ respectively.

Proof. Throughout this proof we will abuse notation to various degrees. We will omit the explicit dependence of the various functions on their variables, except when that dependence is essential to the computation being carried out. We will also abusively write $f \circ \eta$ to denotes $f(t, \eta(t,x))$ wherever it is helpful to do so for the brevity and clarity of the argument.

We begin by computing the derivatives of f. It follows immediately from (5.1) that

$$\partial_t f = \dot{b} + \dot{R}(\,\cdot\, - b_0) \text{ and } \nabla f = R. \tag{5.2}$$

We are now equipped to compute the derivatives of $\tilde{\eta}$. Since $\tilde{\eta}(t) = f(t, \eta(t))$ it follows that

$$\partial_t \tilde{\eta} = (\partial_t f) \circ \eta + ((\nabla f) \circ \eta)\partial_t \eta \text{ and } \nabla\tilde{\eta} = ((\nabla f) \circ \eta)\nabla\eta$$

such that, in light of (5.2),

$$\partial_t \tilde{\eta} = \dot{b} + R(\eta - b_0) + R\partial_t \eta \text{ and } \nabla\tilde{\eta} = R\nabla\eta,$$

i.e. (1) and (2) hold.

We now compute the Eulerian velocity $\tilde{u}$. Since $\tilde{u}$ is defined as $\tilde{u} := \partial_t \tilde{\eta} \circ \tilde{\eta}^{-1}$ we must first compute the inverse of $\tilde{\eta}$. Since $\tilde{\eta} = f \circ \eta$ this is immediate:

$$\tilde{\eta}^{-1} = \eta^{-1} \circ f^{-1}. \tag{5.3}$$

Note that this really means that, for every $t \geqslant 0$,

$$\tilde{\eta}_t^{-1} = \eta_t^{-1} \circ f_t^{-1}.$$

We may now compute $\tilde{u}$. Using (1) and (5.3) we see that

$$\tilde{u} = \partial_t \tilde{\eta} \circ \tilde{\eta}^{-1} = \left(\dot{b} + R(\eta - b_0) + R\partial_t\eta\right) \circ \left(\eta^{-1} \circ f^{-1}\right) = \left(\dot{b} + R(\,\cdot\, - b_0) + Ru\right) \circ f^{-1}$$

from which (3) follows.

Finally we compute $(\nabla\tilde{u}) \circ f$. We introduce $(\nabla\tilde{u}) \circ f$ by differentiating both sides of (3), which yields

$$((\nabla\tilde{u}) \circ f)\nabla f = \dot{R} + R\nabla u. \tag{5.4}$$

Using (5.2) we may apply R^T to both sides of (5.4) and obtain that

$$(\nabla\tilde{u}) \circ f = \dot{R}R^T + R\nabla u R^T,$$

i.e. (4) holds. □

We now define an analog of Definition 5.1 for the micropolar realm by defining the similarity of micropolar continua.

Definition 5.3. (Similar micropolar continua)
Given two micropolar continua (Ω_0, η, Q) and $(\tilde{\Omega}_0, \tilde{\eta}, \tilde{Q})$ we say that they are *similar* if there exists a time-dependent orientation-preserving rigid motion $f = b + R(\,\cdot\, - b_0)$ such that $\tilde{\Omega}_0 = \Omega_0$, $\tilde{\eta}_t = f_t \circ \eta_t$, and $\tilde{Q}_t = R_t Q_t$ for every $t \geqslant 0$. Moreover we say that f *maps* (η, Q) *to* $(\tilde{\eta}, \tilde{Q})$.

A justification for why we define the similarity of micropolar continua in this way is provided in Figure 4. Having defined similarity for micropolar continua we now proceed as we did for classical continua and compute how various kinematic quantities behave under similarity.

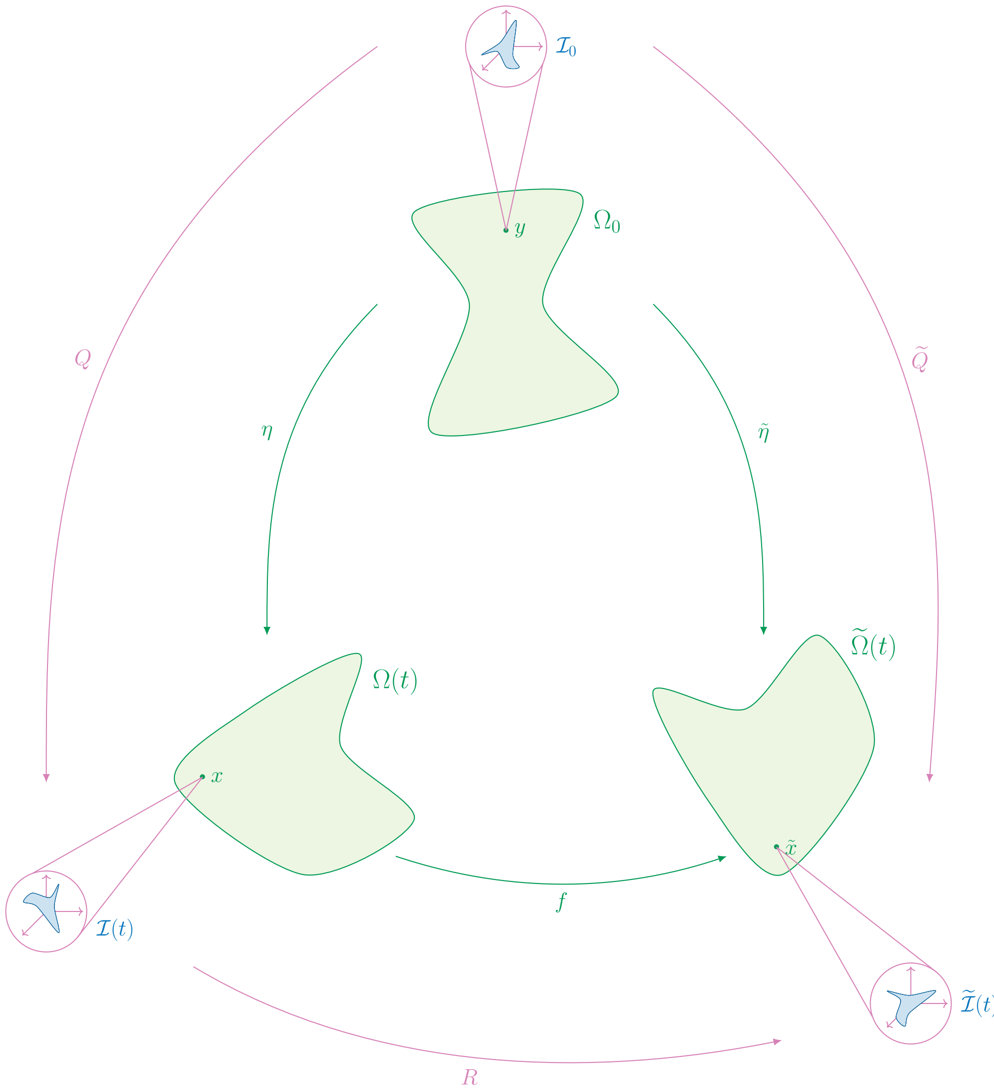


FIGURE 4. Two micropolar continua (Ω_0, η, Q) and $(\Omega_0, \tilde{\eta}, \tilde{Q})$ are similar when there exists a rigid motion $f = b + R(\cdot - b_0)$ such that this diagram commutes, i.e. $\tilde{\eta} = f \circ \eta$ and $\tilde{Q} = RQ$.

PROPOSITION 5.4. *(The behaviour of kinematic quantities of micropolar continua under similarity) Let (Ω_0, η, Q) and $(\Omega_0, \tilde{\eta}, \tilde{Q})$ be similar micropolar continua where $f = b + R(\cdot - b_0)$ maps (η, Q) to $(\tilde{\eta}, \tilde{Q})$. Then*

(1) *Items (1)–(4) of Proposition 5.2 hold,*
(2) $\partial_t \tilde{Q} = \dot{R}Q + R\partial_t Q$,
(3) $\tilde{\Omega} \circ f = \dot{R}R^T + R\Omega R^T$,
(4) $\tilde{\omega} \circ f = \operatorname{vec}\left(\dot{R}R^T\right) + R\omega$, *and*
(5) $(\nabla\tilde{\omega}) \circ f = R\nabla\omega R^T$.

PROOF. We know that items (1)–(4) of Proposition 5.2 hold since the continua underlying similar micropolar continua must themselves be similar, so here (Ω_0, η) and $(\Omega_0, \tilde{\eta})$ are similar continua. Note that (2) is immediate since $\tilde{Q} = RQ$. To obtain (3) we compute that

$$\begin{aligned}\tilde{\Omega} = ((\partial_t \tilde{Q})\tilde{Q}^T) \circ \tilde{\eta}^{-1} = ((\dot{R}Q + R\partial_t Q)(Q^T R^T)) \circ (\eta^{-1} \circ f^{-1}) &= (\dot{R}R^T + R(\partial_t Q)Q^T R^T) \circ (\eta^{-1} \circ f^{-1}) \\ &= (\dot{R}R^T + R\Omega R^T) \circ f^{-1}\end{aligned}$$

from which (3) follows. Since f is orientation-preserving we know that $\det R = 1$, and hence we may use Lemma 6.14 and apply vec to both sides of the chain of equalities immediately above to deduce (4). To derive (5) we proceed as we did to obtain item (4) of Proposition 5.2: we differentiate (4), which tells us that

$$((\nabla\tilde{\omega}) \circ f)\nabla f = R\nabla\omega$$

and use the fact that $\nabla f = R$ to obtain (5). □

We conclude this section with the definition of the notion of *frame-invariance*.

Definition 5.5. (Frame-invariance)

(1) (Vector functions) We say that $F : \mathbb{R}^n \to \mathbb{R}^n$ is *frame-invariant* if $F(RV) = RF(v)$ for every $v \in \mathbb{R}^n$ and every $R \in SO(n)$.
(2) (Tensors functions) Let $X, Y \subseteq \mathbb{R}^{n\times n}$ be closed under conjugation by orientation-preserving orthogonal matrices. We say that $F : X \to Y$ is *frame-invariant* if $F(RMR^T) = RF(M)R^T$ for every $M \in X$ and every $R \in SO(n)$.
(3) (Functions defined along the flow) Let $\mathcal{T}$ be a function which maps continua to k-tensor fields along that flow. We say that $\mathcal{T}$ is *frame-invariant* if for every similar continua (Ω_0, η) and $(\tilde{\Omega}_0, \tilde{\eta})$, where f which maps η to $\tilde{\eta}$ is given by
$$f(t, x) = b(t) + R(t)(x - b_0)$$
for some $b_0 \in \mathbb{R}^n$, $b : [0, \infty) \to \mathbb{R}^n$, and $R : [0, \infty) \to \mathbb{R}^n$, if we write
$$T := \mathcal{T}(\Omega_0, \eta) \text{ and } \widetilde{T} := \mathcal{T}(\tilde{\Omega}_0, \tilde{\eta})$$
then
$$\widetilde{T}_{j_1, \cdots, j_k} \circ f = R_{j_1 i_1} R_{j_2 i_2} \dots R_{j_k i_k} T_{i_1, \cdots, i_k}.$$
Note that this definition applies *mutatis-mutandis* to functions mapping micropolar continua to k-tensor fields along that flow.

With the notion of frame-invariance in hand, we can look back at our computations from Proposition 5.2 and Proposition 5.4 to identity the kinematic quantities which are frame-invariant. We record this below.

Corollary 5.6. *(Identification of the frame-invariant kinematic quantities)*
Let (Ω_0, η, Q) be a micropolar continuum, let u denotes its Eulerian velocity, and let ω be its Eulerian angular velocity. Then $\nabla u - \omega$ and $\nabla\omega$ are frame-invariant.

Proof. The frame-invariance of $\nabla\omega$ is precisely item (5) of Proposition 5.4. The frame-invariance of $\nabla u - \Omega$ follows from item (4) of Proposition 5.2 and item (3) of Proposition 5.4 since upon subtracting the former from the latter we see that

$$\nabla\tilde{u} - \tilde{\Omega} = (\dot{R}R^T + R\nabla u R^T) - (\dot{R}R^T + R\Omega R^T) = R(\nabla u - \Omega)R^T.$$

□

5.2. Representation of frame-invariant linear maps. In this section we record several results on the representation of frame-invariant linear maps, which are inspired by analogous results in [**Gur81**, **Wan70a**, **Wan70b**, **Smi71**].

We make the technical assumption that the dimension n be odd, which is not concerning for our purposes here since we ultimately wish to consider the case $n = 3$. Nonetheless, it seems that the results below should hold in arbitrary dimensions. The source of this technical restriction lies in the fact that we consider the angular velocity and angular momentum to be in $\mathbb{R}^3$ (and not in Skew(3)). Crucially: the identification of $\mathbb{R}^3$ and Skew(3) is made via ten and vec, which are not invariant under actions by $O(3)$, but only invariant under actions by $SO(3)$. A key tool in obtaining the representation formulae below is to consider, for an appropriately chosen unit vector v, the transformation $2v \otimes v - I$. This matrix is always in $O(n)$ but is only in $SO(n)$ when n is odd, and this is precisely the source of our technical restriction.

We begin with a couple of lemmas that will come in handy when discussing the frame-invariance of linear maps whose domain lies within the space of symmetric matrices. First we note relate the commutativity of symmetric matrices to the invariance of eigenspaces.

Lemma 5.7. *(Commuting symmetric operators keep eigenspaces invariant)*
Let S and T be real symmetric matrices. S and T commute if and only if the eigenspaces of S are invariant under T.

Proof. Suppose first that S and T commute and let $V \subseteq \mathbb{R}^n$ be an eigenspace of S with eigenvalue λ. Then, for every $x \in V$, $STx = TSx = \lambda Tx$ such that $Tx \in V$ (since Tx is an eigenvector of S with eigenvalue λ) and so indeed the eigenspaces of S are invariant under S.

Suppose now that the eigenspaces of S are invariant under T and let us write $S = \sum_i \lambda_i v_i \otimes v_i$ (such a decomposition exists since S is symmetric). By assumption we know that, for every i, Tv_i belongs to the eigenspace of v_i such that $S(Tv_i) = \lambda_i Tv_i = TSv_i$ and thus S and T commute on each of the eigenspaces of S. Since the union of the eigenspaces of S constitutes all of $\mathbb{R}^n$ we conclude that S and T commute. ☐

We now note that frame-invariant functions whose domain is the space of symmetric matrices preserve eigenspaces.

Lemma 5.8. *(Frame-invariant functions preserve eigenspaces)*
Let n be odd and let $F : \mathrm{Sym}(n) \to \mathbb{R}^{n\times n}$ be frame-invariant. Then, for every $S \in \mathrm{Sym}(n)$, eigenvectors of S are eigenvectors of $F(S)$.

Proof. Let $S \in \mathrm{Sym}(n)$ and let $v \in \mathbb{R}^n$ be an eigenvector of S. Let us define $R := 2v \otimes v - I$, which is an orthogonal transformation since it is symmetric and satisfies

$$(2v \otimes v - I)^2 = 4|v|^2 - 4v \otimes v - I = I.$$

Moreover, R is orientation preserving. Indeed: let us complete v to a basis $\{v, w_1, \ldots, w_{n-1}\}$ of $\mathbb{R}^n$. Then $Rv = v$ and $Rw_i = -w_i$ such that every element of that basis is an eigenvector of R. Therefore, since n is odd, $\det R = 1 \cdot (-1)^{n-1} = 1$.

Geometrically, we may describe R as the reflection through the line spanned by v. Indeed, as already mentioned above when showing that R is orientation-preserving:

$$Rv = v \text{ and } Rw = -w \text{ if } w \perp v. \tag{5.5}$$

In particular: for any eigenvector w of S distinct from v we have that $Rw = -w$, so the eigenspaces of S are invariant under R. As a consequence, we deduce from Lemma 5.7 that S and R commute. Combining this with the frame-invariance of F tells us that

$$RF(S)R^T = F(RSR^T) = F(S),$$

i.e. $F(S)$ and R also commute. So finally:

$$R(F(S)v) = F(S)(Rv) = F(S)v$$

which, by (5.5), may only occur if $F(X)v \in \operatorname{span}\{v\}$. So indeed we may conclude that v is an eigenvector of $F(S)$. ☐

As an immediate consequence of Lemma 5.8 we obtain the following corollary.

Corollary 5.9. *Let $F : \mathrm{Sym}(n) \to \mathbb{R}^{n\times n}$ be frame-invariant. Then $\operatorname{im} F \subseteq \mathrm{Sym}(n)$.*

Proof. Let $S \in \mathrm{Sym}(n)$ and let us write $S = \sum_i \lambda_i v_i \otimes v_i$ where the v_i's form an orthonormal basis of $\mathbb{R}^n$. By Lemma 5.8 we know that there exists μ_i's such that $F(S)$ has eigenpairs $\{(\mu_i, v_i)\}$, and therefore:

$$F(S) = \sum_i \mu_i v_i \otimes v_i$$

such that indeed $F(S)$ is symmetric. ☐

We are now ready to begin to establish a slew of representation formulae for frame-invariant linear maps. We begin with vector-to-vector frame-invariant linear maps.

Lemma 5.10. *(Representation formula for linear frame-invariant maps from $\mathbb{R}^n$ to itself)*
Let n be odd and let $F : \mathbb{R}^n \to \mathbb{R}^n$ be linear and frame-invariant. Then F is a scalar multiple of the identity.

Proof. Since F is linear we know that there exists a matrix $M \in \mathbb{R}^{n\times n}$ such that $F(v) = Mv$ for every $v \in \mathbb{R}^n$. The fact that F is frame-invariant then tells us that M commutes with all orientation-preserving orthogonal matrices.

So let us consider a particular orthogonal matrix: let $x \in \mathbb{R}^n$ be a unit vector and note that $I - 2x \otimes x$ is orthogonal since it is symmetric and satisfies

$$(2x \otimes x - I)^2 = 4|x|^2 x \otimes x - 4x \otimes x + I = I.$$

Moreover, R is orientation-preserving. To see this, note that if we complete $\{x\}$ to a basis $\{x, v_1, \dots, v_{n-1}\}$ of $\mathbb{R}^n$ then $Rx = x$ and $Rv_i = -v_i$, therefore each of the basis vectors is an eigenvectors of R. So finally: $\det R = 1 \cdot (-1)^{n-1} = 1$ since n is odd.

Therefore M commutes with $2x \otimes x - I$, and hence with $x \otimes x$, such that

$$(Mx) \otimes x = M(x \otimes x) = (x \otimes x)M = x \otimes (M^T x).$$

Applying both sides of this equality to x we see that, since x is a unit vector,

$$Mx = (M^T x \cdot x)x = (Mx \cdot x)x.$$

In particular, since M commutes with all orientation-preserving orthogonal matrices we may pick $R \in SO(n)$ such that $x = Re_1$ and observe that then

$$MRe_1 \cdot Re_1 = R^T RMe_1 \cdot e_1 = Me_1 \cdot e_1$$

and hence $Mx = (Me_1 \cdot e_1)x$ for all unit vectors x. So indeed M is a multiple of the identity since $M = (Me_1 \cdot e_1)I$. □

We now prove a representation formula for linear frame-invariant maps from $\mathrm{Sym}(n)$ to itself.

Lemma 5.11. *(Representation formula for linear frame-invariant maps from* $\mathrm{Sym}(n)$ *to itself) Let* n *be odd and let* $F : \mathrm{Sym}(n) \to \mathrm{Sym}(n)$ *be linear and frame-invariant. Then there exist* $c_1, c_2 \in \mathbb{R}$ *such that*

$$F(S) = c_1(\operatorname{tr} S)I + c_2 S \text{ for every } S \in \mathrm{Sym}(n).$$

Proof. First we will show that the claim holds on the set of rank-1 symmetric matrices with unit norm. Let $v \in \mathbb{R}^n$ be a unit vector and let us consider the symmetric matrix $v \otimes v$. Lemma 5.8 tells us that, since F is frame-invariant, $F(v \otimes v)$ has two eigenspaces: span $\{v\}$ and its orthogonal complement. So either the eigenspaces of $F(v \otimes v)$ are the same as those of $v \otimes v$ or the sole eigenspace of $F(v \otimes v)$ is $\mathbb{R}^n$. Either way we have that

$$F(v \otimes v) = \tilde{c}_1(v) v \otimes v + \tilde{c}_2(v)(I - v \otimes v) = c_1(v) v \otimes v + c_2(v) I$$

for some $c_1, c_2 : \mathbb{S}^{n-1} \to \mathbb{R}$.

Now let $w \in \mathbb{R}^n$ be another unit vector and let R be the orientation-preserving orthogonal transformation which takes v to w, i.e. $Rv = w$. Then, by the frame-invariance of F,

$$\begin{aligned} 0 = F(R(v \otimes v)R^T) - F(w \otimes w) &= RF(v \otimes v)R^T - F(w \otimes w) \\ &= R(c_1(v) v \otimes v + c_2(v) I)R^T - (c_1(w) w \otimes w + c_2(w) I) \\ &= (c_1(v) - c_1(w)) w \otimes w + (c_2(v) - c_2(w)) I. \end{aligned}$$

So finally, since $w \otimes w$ and I are linearly independent we know that c_1 and c_2 are constants on $\mathbb{S}^{n-1}$, i.e.

$$F(v \otimes v) = c_1 v \otimes v + c_2 I, \tag{5.6}$$

for every $v \in \mathbb{S}^{n-1}$. We have thus just shown that the representation formula holds on the set of symmetric matrices of rank-1 with unit norm (up to switching c_1 and c_2).

To conclude we leverage the linearity of F. Let S be an arbitrary real symmetric matrix and let us decompose it as

$$S = \sum_i \lambda_i v_i \otimes v_i$$

where each of the v_i has unit norm. Then, by (5.6) and the linearity of F we deduce that

$$F(S) = \sum_i \lambda_i (c_1 v_i \otimes v_i + c_2 I) = c_1 S + c_2 (\operatorname{tr} S) I$$

which is precisely the representation formula we sought (up to interchanging c_1 and c_2). □

We now prove a result similar to Lemma 5.11 but considering frame-invariant linear maps whose domain now lies within the space of skew-symmetric matrices. Moreover we restrict our attention to dimension $n = 3$ in order to be able to use ten and vec and thus streamline the argument.

LEMMA 5.12. *(Representation formula for linear frame-invariant maps from* Skew(3) *to* $\mathbb{R}^{3\times 3}$*)* *Let $n = 3$ and let $F : \mathrm{Skew}(3) \to \mathbb{R}^{3\times 3}$ be linear and frame-invariant. Then F is a scalar multiple of the identity.*

PROOF. Let $A \in \mathrm{Skew}(3)$ with $|A|^2 = A : A = 2$ and let $v \in \mathbb{R}^3$ be defined as $v := \operatorname{vec} A$. Note that we have chosen the seemingly odd normalization $|A|^2 = 2$ since it ensures that v is a unit vector. Indeed, by Lemma 6.17,

$$|v|^2 = \operatorname{vec} A : \operatorname{vec} A = \frac{1}{2} A : \operatorname{ten} \operatorname{vec} A = \frac{1}{2}|A|^2 = 1.$$

It will be helpful throughout to recall the geometric interpretation of the action of A: A is a (counter-clockwise) rotation by $\frac{\pi}{2}$ in the plane orthogonal to v and annihilates vectors colinear with v.

Before beginning the proof in earnest we record some useful computations about A whose outcomes are obvious in light of the geometric interpretation of the action of A. First, note that

$$A(v \otimes v) = (v \otimes v)A = 0. \tag{5.7}$$

Indeed, for any $w \in \mathbb{R}^3$,

$$A(v \otimes v)w = (v \cdot w)Av = (v \cdot w)(v \times w) = 0 \text{ and } (v \otimes v)Aw = (v \cdot Aw)v = (v \cdot (v \times w))v = 0.$$

Second, observe that

$$-A^2 = I - v \otimes v = \operatorname{proj}_{\{v\}^\perp}. \tag{5.8}$$

Indeed: $-A^2 v = -A(v \times v) = 0$ whilst, for any $w \perp v$,

$$-A^2 w = -v \times (v \times w) = -(v \cdot w)v + (v \cdot v)w = w.$$

Step 1. It will be very convenient for the remainder of this argument to consider the transformation $R := A + v \otimes v$. In particular we will show that R is orthogonal, orientation-preserving, and commutes with $F(A)$. The orthogonality of R follows from (5.7) and (5.8) since

$$RR^T = (A + v \otimes v)(-A + v \otimes v) = -A^2 + v \otimes v = I$$

and, similarly, $R^T R = I$. We can understand the action of R geometrically by comparing it to the action of A: R acts in the same way as A on vectors orthogonal to v but is the identity on the span of v. In particular, this tells us that R can de decomposed as the direct sum of two orientation-preserving transformations, so R itself is orientation-preserving. Now let us show that R and $F(A)$ commute. Observe that, by (5.7) and (5.8),

$$RAR^T = (A + v \otimes v)A(-A + v \otimes v) = -A^3 = A(I - v \otimes v) = A$$

and hence, since F is frame-invariant,

$$F(A) = F(RAR^T) = RF(A)R^T$$

such that indeed $F(A)$ and R commute.

Step 2. We now show that the actions of A and $F(A)$ agree on $\operatorname{span}\{v\}$ and $\operatorname{span}\{v\}^\perp$. To do so, observe that we may characterize the action of R^2 as follows:

$$\begin{cases} R^2 x = x \iff x \in \operatorname{span}\{v\} \text{ and} \\ R^2 x = -x \iff x \perp v. \end{cases} \tag{5.9}$$

Indeed:

$$\begin{aligned} R^2 - I = (A + v \otimes v)^2 - I = A^2 + v \otimes v - I = -(I + v \otimes v) + v \otimes v - I &= -2(I - v \otimes v) \\ &= -2 \operatorname{proj}_{\{v\}^\perp} \end{aligned}$$

and thus $R^2 + I = 2v \otimes v = 2 \operatorname{proj}_{\operatorname{span}\{v\}}$.

Since R and $F(A)$ commute we may then compute that, for any $w \perp v$,

$$\begin{cases} R^2 F(A)v = F(A)R^2 v = F(A)v \text{ and} \\ R^2 F(A)w = F(A)R^2 w = -F(A)w \end{cases}$$

such that, by (5.9),

$$F(A)v \in \operatorname{span}\{v\} \text{ and } F(A)w \perp v. \tag{5.10}$$

Now let us pick $w_1, w_2 \in \mathbb{R}^3$ to be unit vectors orthogonal to v such that (v, w_1, w_2) forms an *oriented* orthonormal basis of $\mathbb{R}^3$, i.e.

$$v \times w_1 = w_2, w_1 \times w_2 = v, \text{ and } w_2 \times v = w_1.$$

Note that, since $w_1, w_2 \perp v$, we have that

$$Rw_1 = Aw_1 = v \times w_1 = w_2 \text{ and } Rw_2 = Aw_2 = v \times w_2 = -w_1.$$

Therefore, since R and $F(A)$ commute,

$$RF(A)w_1 = F(A)Rw_1 = F(A)w_2 \text{ and } RF(A)w_2 = F(A)Rw_2 = -F(A)w_1$$

and hence, in light of (5.10), $(F(A)v, F(A)w_1, F(A)w_2)$ forms an oriented orthogonal basis of $\mathbb{R}^3$.

This allows us to characterize the behaviour of $F(A)$ on $\operatorname{span}\{v\}^\perp$: $F(A)$ takes the orthonormal basis $\{w_1, w_2\}$ of $\operatorname{span}\{v\}^\perp$, which complements v to an oriented orthonormal basis (v, w_1, w_2) of $\mathbb{R}^3$, to an orthogonal basis $\{F(A)w_1, F(A)w_2\}$ of $\operatorname{span}\{v\}^\perp$, which complements v to an oriented orthogonal basis $(F(A)v, F(A)w_1, F(A)w_2)$ of $\mathbb{R}^3$. So indeed the actions of $F(A)$ and A agree on $\operatorname{span}\{v\}^\perp$. Combining this observation with (5.10) allows us to conclude that

$$F(A) = c_1(A) v \otimes v + c_2(A)A$$

for some scalars $c_1(A), c_2(A) \in \mathbb{R}$.

Step 3. We know show that F is constant on the set of skew-symmetric matrices of norm $\sqrt{2}$. Let A and B be 3-by-3 skew-symmetric matrices with $|A|^2 = |B|^2 = 2$ and let $v := \operatorname{vec} A$ and $w := \operatorname{vec} B$. Then, by Step 2, there exist scalars $c_1(A)$, $c_2(A)$, $c_1(B)$, and $c_2(B) \in \mathbb{R}$ such that

$$F(A) = c_1(A)v \otimes v + c_2(A)A \text{ and } F(B) = c_1(B)w \otimes w + c_2(B)B.$$

Now let $Q \in SO(3)$ such that $w = Qv$, and hence, by Lemma 6.14, $B = QAQ^T$. Then, by frame-invariance of A we have that

$$c_1(B)w \otimes w + c_2(B)B = F(B) = QF(A)Q^T = Q(c_1(A)v \otimes v + c_2(A)A)Q^T = c_1(A)w \otimes w + c_2(A)B.$$

Since $w \otimes w$ and B are linearly independent we deduce that $c_1(A) = c_1(B)$ and $c_2(A) = c_2(B)$, so indeed F is constant on $\left\{A \in \operatorname{Skew}(3) : |A|^2 = 2\right\}$.

Step 4. We conclude by linearity of F. Since there exist scalars $c_1, c_2 \in \mathbb{R}$ such that

$$F(A) = c_1 \operatorname{vec} A \otimes \operatorname{vec} A + c_2 A$$

for every 3-by-3 skew-symmetric matrix A satisfying $|A|^2 = 2$ it follows that, for any $B \in \operatorname{Skew}(3)$,

$$F(B) = \frac{|B|}{\sqrt{2}} F\left(\frac{\sqrt{2}}{|B|}B\right) = \frac{c_1\sqrt{2}}{|B|} \operatorname{vec} B \otimes \operatorname{vec} B + c_2 B.$$

In particular, leveraging once again the linearity of F we note $F(B) - F(C) - F(B - C)$ must vanish for all $B, C \in \operatorname{Skew}(3)$, from which we deduce that $c_1 = 0$. So finally:

$$F(B) = c_2 B \text{ for every } B \in \operatorname{Skew}(3).$$

□

We conclude this section with a representation formulae for general frame-invariant linear maps from $\mathbb{R}^{n\times n}$ it itself (i.e. we do not specify, as we did in the previous two results, that the domain is contained in either the space of symmetric matrices or the space of skew-symmetric matrices). Once again we restrict our attention to dimension $n = 3$.

Proposition 5.13. *(Representation formula for linear frame-invariant maps from $\mathbb{R}^{3\times 3}$ to itself)* *Let $n = 3$ and let $F : \mathbb{R}^{3\times 3} \to \mathbb{R}^{3\times 3}$ be linear and frame-invariant. Then there exists $c_1, c_2, c_3 \in \mathbb{R}$ such that*

$$F(M) = c_1(\operatorname{tr} M)I + c_2 \operatorname{Sym}(M) + c_3 \operatorname{Skew}(M) \text{ for every } M \in \mathbb{R}^{n\times n}.$$

Proof. Equipped with Corollary 5.9, Lemma 5.11, Lemma 5.12, and the observation that $\mathrm{Sym} + \mathrm{Skew} = \mathrm{id}$, the only details that remain to be checked are that Sym and Skew are frame-invariant. This follows immediately from the observation that, for any $M \in \mathbb{R}^{3\times 3}$ and any $R \in SO(3)$, $(RMR^T)^T = R(M^T)R^T$. Therefore

$$\mathrm{Sym}\left(RMR^T\right) = \frac{1}{2}\left(RMR^T + RM^TR^T\right) = R\,\mathrm{Sym}(M)R^T$$

and

$$\mathrm{Skew}\left(RMR^T\right) = \frac{1}{2}\left(RMR^T - RM^TR^T\right) = R\,\mathrm{Skew}(M)R^T.$$

Using the linearity of F we may then decompose

$$F = F \circ \mathrm{Sym} + F \circ \mathrm{Skew},$$

where

- $F \circ \mathrm{Sym} : \mathrm{Sym}(3) \to \mathbb{R}^{3\times 3}$ is linear and frame-invariant, and so by Corollary 5.9 and Lemma 5.11 above we know that $F \circ \mathrm{Sym} = c_1 \,\mathrm{tr}(\,\cdot\,) + c_2 \,\mathrm{id}$ for some scalars $c_1, c_2 \in \mathbb{R}$ and
- $F \circ \mathrm{Skew} : \mathrm{Skew}(3) \to \mathbb{R}^{3\times 3}$ is linear and frame-invariant, and so by Lemma 5.12 above we know that $F \circ \mathrm{Skew} = c_3 \,\mathrm{id}$ for some scalar $c_3 \in \mathbb{R}$.

Putting it all together we see that indeed

$$F = F \circ \mathrm{Sym} + F \circ \mathrm{Skew} = c_1 \,\mathrm{tr}(\,\cdot\,) \,+\, c_2 \,\mathrm{Sym} + c_3 \,\mathrm{Skew}\,.$$

□

5.3. Onsager reciprocity relations. The Onsager reciprocity relations were proposed by Onsager [**Ons31a**, **Ons31b**] to provide a theoretical justification for the fact that, in some physical systems, the irreversibility of an underlying microscopic process leads to symmetry properties of some macroscopic observables. This fact has since been extensively verified experimentally [**Mil60**]. As pointed out in [**MRP16**], the fact that the symmetry arises at the macroscopic level as a consequence of irreversibility at a microscopic level is a purely mathematical feature, as shown in Theorem 5.14 below. Note that the statement of Theorem 5.14 is taken from [**MRP16**] and its proof is taken from [**dGM62**].

It is also worth pointing out that the Onsager reciprocity relations need not be invoked when the equations of motion of classical fluids are derived. This is explained in more detail in Remark 5.19 below.

Theorem 5.14. *(Onsager reciprocity relations)*
Let X_t be a Markov process in $\mathbb{R}^n$ with transition kernel $P_t(dx|x_0)$ and invariant measure $\mu(dx)$. Define the expectation $z_t(x_0)$ of X_t given that $X_0 = x_0$, i.e.

$$z_t(x_0) = \mathbb{E}_{x_0} X_t = \int x P_t(dx|x_0).$$

Assume that
1. *μ is reversible, i.e. $\mu(dx_0)P_t(dx|x_0) = \mu(dx)P_t(dx_0|x)$ for every $x, x_0 \in \mathbb{R}^n$ and every $t > 0$,*
2. *μ is Gaussian with mean zero and covariance matrix G, and*
3. *$t \mapsto z_t(x_0)$ satisfies the equation $\dot{z}_t = -Az_t$ for some nonnegative matrix A.*

Then $M := AG$ is symmetric.

Proof. It follows immediately from item 3 that, for every $t > 0$ and every $x_0 \in \mathbb{R}^n$,

$$z_t(x_0) = e^{-At}x_0. \tag{5.11}$$

Taking the outer product of (5.11) with x_0 and integrating with respect to $\mu(dx_0)$ tells us that

$$\int x_0 \otimes z_t(x_0)\mu(dx_0) = \int x_0 \otimes e^{-At}x_0\mu(dx_0).$$

By definition of z, reversibility of μ, and (5.11) we note that we may rearrange the left-hand side to see that

$$\begin{aligned}\int x_0 \otimes z_t(x_0)\mu(dx_0) &= \int x_0 \otimes \left(\int x P_t(dx|x_0)\right)\mu(dx_0) = \int\int x_0 \otimes x P_t(dx|x_0)\mu(dx_0)\\ &= \int\int x_0 \otimes x P_t(dx_0|x)\mu(dx) = \int \left(\int x_0 P_t(dx_0|x)\right)\otimes x\mu(dx)\\ &= \int z_t(x)\otimes x\mu(dx) = \int e^{-At}x\otimes x\mu(dx).\end{aligned}$$

In particular, since μ is a centered Gaussian with covariance matrix G we note that $\int y \otimes y\mu(dx) = G$ and hence

$$\int e^{-At}x\otimes x\mu(dx) = e^{-At}G \text{ and } \int x_0 \otimes e^{-At}x_0\mu(dx_0) = \left(\int x_0\otimes x_0\mu(dx_0)\right)e^{-At} = Ge^{-At}$$

such that

$$e^{-At}G = Ge^{-At}$$

holds for all $t > 0$. Differentiating in time and setting $t = 0$, we deduce that

$$AG = GA^T$$

such that indeed $M = AG$ is symmetric. □

5.4. Micropolar fluids: definition and derivation of their constitutive equations. This section is the conclusion of this chapter, where we define micropolar fluids and derive their equations of motion. First we define various classes of micropolar continua. In particular, note that we will only consider homogeneous incompressible continua in the sequel.

Definition 5.15. (Homogeneity, incompressibility, and isotropy of micropolar continua)
Let $(\Omega_0, \eta, Q, \nu, j)$ be a physical micropolar continuum. We say that it is

- *homogeneous* if $d\nu = \rho d\mathcal{L}^3$ for some constant $\rho > 0$,
- *incompressible* if the flow map η is incompressible, and
- *isotropic* if $j = I$, i.e. if the microinertia density is constant and equal to the identity matrix.

We are now ready to define a micropolar fluid.

Definition 5.16. (Homogeneous incompressible micropolar fluid and homogeneous incompressible Newtonian micropolar fluid)
A *homogeneous incompressible micropolar fluid* is a homogeneous incompressible physical micropolar continuum governed by (T, M) for which there exist a scalar function p defined along the flow, called the *pressure*, as well as functions $\hat{T}, \hat{M} : \mathbb{R}^{3\times3}\times\mathbb{R}^3\times\mathbb{R}^{3\times3}\to\mathbb{R}^{3\times3}$ and $\hat{\epsilon} : \mathbb{R}^{3\times3}\times\mathbb{R}^3\times\mathbb{R}^{3\times3}\to\mathbb{R}$, which are independent of the flow, such that

$$T = -pI + \hat{T}(\nabla u, \omega, \nabla\omega),\ M = \hat{M}(\nabla u, \omega, \nabla\omega),\ \text{ and } \epsilon = \hat{\epsilon}(\nabla u, \omega, \nabla\omega)$$

where u denotes the Eulerian velocity, ω denotes the Eulerian angular velocity, and ϵ denotes the mechanical energy density. Moreover, if both $\hat{T}$ and $\hat{M}$ are linear then the homogeneous incompressible micropolar fluid is said to be *Newtonian*.

Now that we have defined micropolar fluids we seek to use frame-invariance and the balance of energy to establish precisely in what ways the stress tensor and couple stress tensor depend on the dynamic variables and their gradients. First we note that, for incompressible micropolar fluids, the mechanical energy density must be constant.

Proposition 5.17. *(Constancy of the mechanical energy in incompressible fluids)*
Consider a homogeneous incompressible micropolar fluid governed by (T, M). If the fluid is purely mechanical then its mechanical energy density ϵ must be constant.

Proof. As proven in Theorem 4.21 and in light of the Eulerian velocity u being divergence-free (by incompressibility), the inequality

$$\partial_t\epsilon + u\cdot\nabla\epsilon - T : (\nabla u - \Omega) = M : \nabla\omega \leqslant 0$$

holds. To make the computation more palatable, let us write $\hat{\epsilon} = \hat{\epsilon}(A, v, B)$. Then, since incompressibility and Lemma 6.7 tells us that D_t^u obeys the Leibniz Rule, we may compute that

$$(\partial_A \hat{\epsilon})(\nabla u, \omega, \nabla\omega) : D_t^u \nabla u + (\partial_v \hat{\epsilon})(\nabla u, \omega, \nabla\omega) : D_t^u \omega + (\partial_B \hat{\epsilon})(\nabla u, \omega, \nabla\omega) : D_t^u \nabla\omega$$
$$-T : (\nabla u - \Omega) - M : \nabla\omega \leqslant 0. \tag{5.12}$$

Since the inequality (5.12) must hold for all possible flows, the key observation is that for any n-by-n matrices A, B, C, D and any vectors $e, f \in \mathbb{R}^n$ we may construct a flow such that, at some space-time point (t, x),

$$D_t^u \nabla u = A,\ \nabla u = B,\ D_t^u \nabla\omega = C,\ \nabla\omega = D,\ D_t^u \omega = e, \text{ and } D\omega = f.$$

Therefore, since $D_t^u \nabla u$, $D_t^u \omega$, and $D_t^u \nabla\omega$ appear linearly in (5.12), we may violate the inequality *unless* $\partial_A \hat{\epsilon} = \partial_B \hat{\epsilon} = 0$ and $\partial_v \hat{\epsilon} = 0$, in which case ϵ is indeed constant. □

It is worth noting that the constancy of the mechanical energy proven in Proposition 5.17 above is a feature of incompressibility. Indeed, this result is a mathematical manifestation of the physical observation that incompressible continua are incapable of exerting mechanical work.

Our goal remains to establish the dependence of the stress tensor and the couple stress tensor on the dynamic variables and their gradients. In particular, to do so we need to invoke the Onsager reciprocity relations, which are phrased in terms of the inequality obtained in Theorem 4.21. The corresponding postulate is stated below.

Definition 5.18. (Dissipation inequality and Onsager reciprocity relations)
Consider a homogeneous incompressible micropolar fluid governed by (T, M) that is purely mechanical. The inequality

$$T : (\nabla u - \Omega) + M : \nabla\omega \geqslant 0$$

is called the *dissipation inequality*. In particular, for $\mathcal{Y} := (\nabla u - \Omega, \nabla\omega)$ and $\mathcal{J}(\mathcal{Y}) := (T, M)$ the dissipation inequality may be written as

$$\mathcal{Y} \cdot \mathcal{J}(\mathcal{Y}) \geqslant 0.$$

We say that the micropolar fluid obeys the *Onsager reciprocity relations* if $\nabla \mathcal{J}$ is symmetric, i.e. $\partial_i \mathcal{J}_j = \partial_j \mathcal{J}_i$ for every i, j.

Note that in the context of continuum mechanics the dissipation inequality is also known as the Clausius-Duhem inequality.

Remark 5.19. Note that, when we derive the equations of motion of classical fluids (e.g. the Euler or Navier-Stokes) equations by arguments from rational continuum mechanics we do not appeal to the Onsager reciprocity relations since the symmetry is guaranteed by the dissipation inequality and frame-invariance.

Indeed, for classical fluids the Clausius-Duhem inequality reads

$$T(\mathbb{D}u) : \nabla u = T(\mathbb{D}u) : \mathbb{D}u \geqslant 0.$$

Under the Newtonian assumption which postulates that the stress tensor T is linear, we deduce from frame-invariance arguments (and Lemma 5.11 in particular) that $T = c_1 \operatorname{tr}(\,\cdot\,) I + c_2 \operatorname{id}$ for some constants c_1, c_2. In particular T is necessarily symmetric.

We are now ready to state and prove the necessary form of the constitutive equations relating the stress tensor and the couple stress tensor to the dynamic variables and their gradients.

Theorem 5.20. *(Constitutive equations of homogeneous incompressible Newtonian micropolar fluids)*
Consider a homogeneous incompressible micropolar fluid governed by (T, M) that is purely mechanical. If $\hat{T}$ and $\hat{M}$ are frame-invariant and the micropolar fluid satisfies Onsager's reciprocity relations then there exist universal constants $\mu, \kappa, \alpha, \beta, \gamma \geqslant 0$ such that

$$T = \mu \mathbb{D}u + \kappa \operatorname{ten}\left(\frac{1}{2}\nabla \times u - \omega\right) - pI \text{ and } M = \alpha(\nabla \cdot \omega) I + \beta \mathbb{D}^0 \omega + \gamma \operatorname{ten} \nabla \times \omega.$$

Proof. In light of Corollary 5.6 the frame-invariance of $\hat{T} = \hat{T}(\nabla u, \omega, \nabla\omega)$ and $\hat{M}$ tells us that we may write

$$\hat{T} = \hat{T}(\nabla u - \Omega, \nabla\omega) \text{ and } \hat{M} = \hat{M}(\nabla u - \Omega, \nabla\omega).$$

Since $\hat{T}$ and $\hat{M}$ are linear we may write $\hat{L} := (\hat{T}, \hat{M})$ for some linear operator $\hat{L}$ which, by Onsager's reciprocity relations, is symmetric. This means that $\hat{L}$ is completely determined by the quadratic form it generates.

Adapting the arguments from Section 5.1 and Wang's paper on the representation of frame-invariant (or, in the terminology he employs, isotropic) functions we deduce that, since $\hat{T}$ and $\hat{M}$ are linear and frame-invariant and since u is divergence-free,

$$\hat{T}(\nabla u - \Omega, \nabla \omega) = c_1 \mathbb{D}u + c_2(\operatorname{Skew} \nabla u - \Omega) + c_3(\nabla \cdot \omega)I + c_4 \mathbb{D}\omega + c_5 \operatorname{Skew} \nabla \omega$$

and

$$\hat{M}(\nabla u - \Omega, \nabla \omega) = c_6(\nabla \cdot u)I + c_7 \mathbb{D}u + c_8(\operatorname{Skew} \nabla u - \Omega) + c_9(\nabla \cdot \omega)I + c_{10}\mathbb{D}\omega + c_{11} \operatorname{Skew} \nabla \omega$$

for some constants $c_1, c_2, \ldots, c_{11} \in \mathbb{R}$.

The quadratic form generated by $\hat{L}$ is therefore

$$\begin{aligned} \hat{L}(\nabla u - \Omega, \nabla \omega) \cdot (\nabla u - \Omega, \nabla \omega) &= \hat{T} : (\nabla u - \Omega) + \hat{M} : \nabla \omega \\ = c_1|\mathbb{D}u|^2 + c_2|\operatorname{Skew} \nabla u - \Omega|^2 &+ c_9|\nabla \cdot \omega|^2 + c_{10}|\mathbb{D}\omega|^2 + c_{11}|\operatorname{Skew} \nabla \omega|^2 \end{aligned}$$

and so we deduce that $c_3 = c_4 = c_5 = c_6 = c_7 = c_8 = 0$. In particular this tells us that $\hat{T} = \hat{T}(\nabla u - \Omega)$ and that $\hat{M} = \hat{M}(\nabla \omega)$.

Finally we seek to leverage the dissipation inequality

$$\hat{T} : (\nabla u - \Omega) + \hat{M} : \nabla \omega \geqslant 0$$

to obtain sign conditions on the coefficients appearing in T and M. To make this process easier we group the terms in $\hat{T}$ and $\hat{M}$ according to the orthogonal decomposition (with respect to the Frobenius inner product)

$$\mathbb{R}^{n \times n} \cong \mathbb{R}I \oplus \operatorname{Dev}(n) \oplus \operatorname{Skew}(n). \tag{5.13}$$

Recalling the identities

$$\mathbb{D}v = \mathbb{D}^0 v + \frac{2}{3}(\nabla \cdot v)I \text{ and } \operatorname{Skew} \nabla v = \frac{1}{2} \operatorname{ten} \nabla \times v$$

from Lemma 6.11 and Lemma 6.16 respectively, we write

$$\hat{T}(\nabla u - \Omega) = c_2 \mathbb{D}u + c_3 \operatorname{ten}\left(\frac{1}{2}\nabla \times u - \omega\right)$$

and

$$\hat{M}(\nabla \omega) = (c_9 + (2/3)c_{10})(\nabla \cdot \omega)I + c_{10}\mathbb{D}^0\omega + (c_{11}/2) \operatorname{ten} \nabla \times \omega.$$

Defining $\mu = c_2$, $\kappa = c_3$, $\alpha = c_9 + (2/3)c_{10}$, $\beta = c_{10}$, and $\gamma = c_{11}/2$ and employing Lemma 6.17, the dissipation inequality now read

$$\mu|\mathbb{D}u|^2 + 2\kappa|(1/2)\nabla \times u - \omega|^2 + \alpha|\nabla \cdot \omega|^2 + \beta|\mathbb{D}^0\omega|^2 + 2\gamma|\nabla \times \omega|^2 \geqslant 0. \tag{5.14}$$

This inequality must hold for arbitrary flows. Since we can construct flows with arbitrary values of $\nabla u - \Omega$ and $\nabla \omega$ we deduce from (5.14) and the orthogonality of the decomposition (5.13) that $\mu, \kappa, \alpha, \beta, \gamma \geqslant 0$. ☐

Finally, we may now state and prove the main result of this chapter which establishes the equations of motion for homogeneous incompressible Newtonian micropolar fluids.

COROLLARY 5.21. *(Equations of motions for homogeneous incompressible Newtonian micropolar fluids) Let $(\mathbb{R}^3, \eta, Q)$ be a homogeneous incompressible Newtonian micropolar fluid subject to (f, τ) and the boundary effects (f_b, τ_b) such that*

- *the fluid is purely mechanical,*
- *the stress tensor T and couple stress tensor M are frame-invariant, and*
- *the Onsager reciprocity relations are satisfied.*

Then the Eulerian velocity u, pressure p, Eulerian microinertia J, and Eulerian angular velocity ω satisfy

$$\begin{cases} \rho(\partial_t u + u\cdot\nabla u) = (\mu+\kappa/2)\Delta u + \kappa\nabla\times\omega - \nabla p + f \ \textit{in}\ \Omega(t)\,, \\ \nabla\cdot u = 0 \ \textit{in}\ \Omega(t)\,, \\ J(\partial_t\omega + u\cdot\nabla\omega) + \omega\times J\omega = \kappa\nabla\times u - 2\kappa\omega + (\alpha+\beta/3-\gamma)\nabla(\nabla\cdot\omega) + (\beta+\gamma)\Delta\omega + \tau \ \textit{in}\ \Omega(t)\,, \\ \partial_t J + u\cdot\nabla J - [\Omega, J] = 0 \ \textit{in}\ \Omega(t)\,, \\ \mu(\mathbb{D}u)n + \kappa((1/2)\nabla\times u - \omega)\times n - pn = f_b \ \textit{on}\ \partial\Omega(t)\,, \ \textit{and} \\ \alpha(\nabla\cdot\omega)n + \beta(\mathbb{D}^0\omega)n + \gamma(\nabla\times\omega)\times n = \tau_b \ \textit{on}\ \partial\Omega(t) \end{cases}$$

for some mass density $\rho > 0$, dissipation coefficients $\mu, \kappa, \alpha, \beta, \gamma \geqslant 0$, and where n denotes the outer unit normal to $\Omega(t)$.

Proof. It follows from the incompressibility, Corollary 4.10, Proposition 4.17, that (u, ω, J) satisfy

$$\begin{cases} \rho(\partial_t u + u\cdot\nabla u) = \nabla\cdot T + f, \nabla\cdot u = 0, \\ J(\partial_t\omega + u\cdot\nabla\omega) + \omega\times J\omega = 2\operatorname{vec} T + \nabla\cdot M + \tau, \text{ and} \\ \partial_t J + \nabla\cdot(J\otimes u) - [\Omega, J]\,. \end{cases}$$

In particular since u is divergence-free we see immediately that

$$\partial_t J + \nabla\cdot(J\otimes u) - [\Omega, J] = \partial_t J + u\cdot\nabla J - [\Omega, J]\,.$$

Now we deduce from Theorem 5.20 that

$$T = \mu\mathbb{D}u + \kappa\operatorname{ten}\left(\frac{1}{2}\nabla\times u - \omega\right) - pI \text{ and } M = \alpha(\nabla\cdot\omega)I + \beta\mathbb{D}^0\omega + \gamma\operatorname{ten}\nabla\times\omega$$

such that, in light of Lemma 6.16, Lemma 6.12, and the incompressibility,

$$\begin{aligned} \nabla\cdot T &= \mu\nabla\cdot\mathbb{D}u + \kappa\nabla\cdot\operatorname{ten}\left(\frac{1}{2}\nabla\times u - \omega\right) - \nabla p \\ &= \mu\Delta u - \kappa\nabla\times\left(\frac{1}{2}\nabla\times u - \omega\right) - \nabla p \\ &= \left(\mu + \frac{\kappa}{2}\right)\Delta u + \kappa\nabla\times\omega - \nabla p \end{aligned}$$

whilst

$$2\operatorname{vec} T = 2\operatorname{vec}\left(\kappa\operatorname{ten}\left(\frac{1}{2}\nabla\times u - \omega\right)\right) = \kappa\nabla\times u - 2\kappa\omega,$$

and therefore

$$\begin{aligned} \nabla\cdot M &= \alpha\nabla(\nabla\cdot\omega) + \beta\nabla\cdot\left(\mathbb{D}^0\omega\right) + \gamma\nabla\cdot\operatorname{ten}\nabla\times\omega \\ &= \alpha\nabla(\nabla\cdot\omega) + \beta\left(1 - \frac{2}{n}\right)\nabla(\nabla\cdot\omega) + \beta\Delta\omega - \gamma\nabla\times\nabla\times\omega \\ &= \left(\alpha + \frac{\beta}{3}\right)\nabla(\nabla\cdot\omega) + \beta\Delta\omega - \gamma\nabla(\nabla\cdot\omega) + \gamma\Delta\omega \\ &= \left(\alpha + \frac{\beta}{3} - \gamma\right)\nabla(\nabla\cdot\omega) + (\beta+\gamma)\Delta\omega. \end{aligned}$$

To conclude we combine Proposition 4.23 and Theorem 5.20 to obtain the boundary conditions. □

6. Appendix

In this section we record various results that are either well-known or elementary, but which are nonetheless useful elsewhere in this chapter. In Section 6.1 we record identities from calculus and linear algebra, in Section 6.2 we record results related to ten and vec, and in Section 6.3 we record some elementary results related to matrix groups and Lie groups.

6.1. Identities from linear algebra and calculus. First we record the well-known $\epsilon - \delta$ identity.

Lemma 6.1. *(Epsilon-delta identities)*
The following identities hold: $\epsilon_{ija}\epsilon_{kla} = \delta_{ik}\delta_{jl} - \delta_{il}\delta_{jk}$ *and* $\epsilon_{iab}\epsilon_{jab} = 2\delta_{ij}$.

We then record a vectorized version of the $\epsilon - \delta$ identity.

Lemma 6.2. *(Vectorized epsilon-delta identities) Let* $a, b, c \in \mathbb{R}^3$. *Then* $a \times (b \times c) = (a \cdot c)b - (a \cdot b)c$.

Proof. This is nothing more than the 'vectorized' version of an epsilon-delta identity of Lemma 6.1:

$$(a \times (b \times c))_i = \epsilon_{ijk} a_j (b \times c)_k = \epsilon_{ijk}\epsilon_{lmk} a_j b_l c_m = (\delta_{il}\delta_{jm} - \delta_{im}\delta_{jl})\, a_j b_l c_m = ((a \cdot c)\, b - (a \cdot b)\, c)_i.$$

□

We now record a characterization of orthogonal projections in $\mathbb{R}^3$ using cross-products.

Proposition 6.3. *Let* $e, v \in \mathbb{R}^3$ *with* e *a unit vector. We can write the orthogonal decomposition of* v *with respect to* e *conveniently as*

$$v = (v \cdot e)\, e - e \times (e \times v)$$

such that

$$|v|^2 = |v \cdot e|^2 + |e \times v|^2.$$

More generally, for $v, w \in \mathbb{R}^3$,

$$v = \left(v \cdot \frac{w}{|w|}\right) \frac{w}{|w|} - \frac{w}{|w|} \times \left(\frac{w}{|w|} \times v\right) \text{ and } |v|^2|w|^2 = |v \cdot w|^2 + |v \times w|^2.$$

Proof. It suffices to show that $-e \times (e \times \cdot\,) = (I - e \otimes e)$ since the latter is precisely the orthogonal projection unto the orthogonal complement of e. This identity follows from Lemma 6.1:

$$(-e \times (e \times v))_i = -\epsilon_{ijk}\epsilon_{lmk} e_j e_l v_m = -(\delta_{il}\delta_{jm} - \delta_{im}\delta_{jl}) e_j e_l v_m = -(e_j e_i v_j - e_j e_j v_i) = (v - (v \cdot e) e)_i$$

i.e. indeed $-e \times (e \times v) = (I - e \otimes e)v$. □

We now record the frame-invariance of the cross product, which is not surprising since cross products are characterized in terms of lengths and angles, both of which are preserved by orientation-preserving orthogonal transformations.

Lemma 6.4. *(Frame-invariance of the cross product)*
Let Q *be an orthogonal transformation of* $\mathbb{R}^3$, *i.e.* $Q \in O(3)$. *Then, for all* $v, w \in \mathbb{R}^3$

$$Q(v \times w) = (\det Q)(Qv) \times (Qw).$$

Proof. The key observation is that for any $u, v, w \in \mathbb{R}^3$, $\det(Qu \,|\, Qv \,|\, Qw) = (\det Q) \det(u \,|\, v \,|\, w)$ since $\det Q = 1$ if Q is orientation-preserving and $\det Q = -1$ if Q is orientation-reversing. So, for all $u, v, w \in \mathbb{R}^3$,

$$u \cdot Q(v \times w) = (Q^T u) \cdot (v \times w) = \det(Q^T u \,|\, v \,|\, w) = (\det Q^T) \det(u \,|\, Qv \,|\, Qw) = (\det Q)\, u \cdot ((Qv) \times (Qw))$$

such that indeed the identity holds. □

We now record an elementary computation dealing with outer products and matrix multiplication.

Lemma 6.5. *Let* $A \in \mathbb{R}^{r \times s}$, $B \in \mathbb{R}^{t \times u}$, $v \in \mathbb{R}^s$ *and* $w \in \mathbb{R}^t$. *Then* $(Av) \otimes (wB) = A\,(v \otimes w)\,B$.

Proof. This is the result of an immediate computation:

$$\Big((Av) \otimes (wB)\Big)_{ij} = A_{ik} v_k w_l B_{lj} = A_{ik}\,(v \otimes w)_{kl}\, B_{lj} = \Big(A\,(v \otimes w)\,B\Big)_{ij}.$$

□

We prove here a characterization of the kernel of the symmetrized gradient.

Lemma 6.6. *(Characterization of the kernel of the symmetrized gradient)*
Let $\mathcal{U} \subseteq R^n$ *be a bounded open set with Lipschitz boundary. Then*

$$\ker \mathbb{D} \cong \mathbb{R}^n \times \mathrm{Skew}\,(n)$$

i.e. for every $v \in H^1(\mathcal{U})$, $\mathbb{D}v = 0$ *if and only if* $v(x) = \bar{v} + \bar{\Omega}x$ *for some* $\bar{v} \in \mathbb{R}^n$ *and some* $\bar{\Omega} \in \mathrm{Skew}\,(n)$.

PROOF. If $v(x) = \bar{v} + \bar{\Omega}x$ with $(v) \in \mathbb{R}^n$ and $\bar{\Omega} \in \operatorname{Skew}(n)$, then $\mathbb{D}v = \operatorname{Sym}(\nabla v) = \operatorname{Sym}(\Omega) = 0$. Conversely, suppose that $v \in \ker \mathbb{D}$, and define

$$\bar{v} := \fint v \text{ and } \bar{\Omega} := \fint \operatorname{Skew}(\nabla v)$$

such that $\left(x \mapsto \bar{v} + \bar{\Omega}x\right) = \operatorname{proj}_{\ker \mathbb{D}} v$. Let $w := v - (\bar{v} + \bar{\Omega}\,\cdot)$ and observe that $\mathbb{D}w = \mathbb{D}v = 0$. Crucially, note that

$$\fint w = 0 \text{ and } \fint \operatorname{Skew}(\nabla w) = 0.$$

Using the fundamental theorem of calculus to expand w about its average we deduce that $w = 0$. So indeed $v(x) = \bar{v} + \bar{\Omega}x$. □

We prove a "Leibniz Rule" for the material derivatives introduced in Definition 2.5.

LEMMA 6.7. *('Leibniz rule' for material derivatives)*
Let u be a vector field and let T and S be differentiable tensor fields. Then $\mathbb{D}_t^u (T \cdot S) = (\mathbb{D}_t^u T) \cdot S + T \cdot (\mathrm{D}_t^u S)$.

PROOF. For better readability despite the number of indices involved, we will write I instead of $i_1 \dots i_k$ and J instead of $j_1 \dots j_l$. We then compute:

$$\mathbb{D}_t^u (T \cdot S) = \partial_t (T \cdot S) + \nabla \cdot ((T \cdot S)\, u)$$

where

$$\begin{aligned}(\nabla \cdot ((T \cdot S)\, u))_{IJ} = \partial_a (T_{Ib} S_{bJ} u_a) = \partial_a (T_{Ib} u_a) S_{bJ} + T_{Ib} u_a \partial_a S_{bJ} &= (\nabla \cdot (T \otimes u))_{Ib} S_{bJ} + T_{Ib} ((u \cdot \nabla) S)_{bJ} \\ &= ((\nabla \cdot (T \otimes u)) \cdot S + T \cdot ((u \cdot \nabla) S))_{IJ}\end{aligned}$$

and hence

$$\mathbb{D}_t^u (T \cdot S) = (\partial_t T) \cdot S + T \cdot (\partial_t S) + (\nabla \cdot (T \otimes u)) \cdot S + T \cdot ((u \cdot \nabla) S) = (\mathbb{D}_t^u T) \cdot S + T \cdot (\mathrm{D}_t^u S).$$

□

We record a differential identity useful when deriving the local version of the balance of energy.

LEMMA 6.8. *Let M and v be differentiable matrix and vector fields respectively, i.e. $M : \mathbb{R}^n \to \mathbb{R}^{k \times n}$ and $v : \mathbb{R}^n \to \mathbb{R}^n$. Then $\nabla \cdot (v \cdot M) = (\nabla \cdot M) \cdot v + M : \nabla v$.*

PROOF. We compute: $\nabla \cdot (v \cdot M) = \partial_i (v_j M_{ji}) = M_{ji} \partial_i v_j + (\partial_i M_{ji}) v_j = M : \nabla v + (\nabla \cdot M) \cdot v$. □

Here we define the *deviatoric part* of a matrix. Due to the fact that this plays well with an orthogonal decomposition of $\mathbb{R}^{n \times n}$, this comes in handy when establishing sign conditions on the coefficients that arise in the constitutive equations for the stress and couple stress tensors.

DEFINITION 6.9. (Deviatoric part)
We define $\operatorname{Dev} : \mathbb{R}^{n \times n} \to \mathbb{R}^{n \times n}$ via, for every $M \in \mathbb{R}^{n \times n}$, $\operatorname{Dev} M := \operatorname{Sym}(M) - \frac{1}{n} \operatorname{tr} M I$, i.e. $\operatorname{Dev} M$ is the traceless symmetric part of M.

We also define the related notion of the deviatoric gradient.

DEFINITION 6.10. (Deviatoric gradient)
Let v be a differentiable vector field on $\mathbb{R}^n$. Its *deviatoric gradient*, denoted $\mathbb{D}^0 v$, is defined to be the deviatoric part of the symmetrized gradient of v, i.e. $\mathbb{D}^0 v := \operatorname{Dev}(\mathbb{D}v)$.

Now we take note of how to relate the deviatoric gradient to the symmetrized gradient.

LEMMA 6.11. *(Relating the symmetrized gradient and the deviatoric gradient)*
For any differentiable n-dimensional vector field v, i.e. $v \in C^1(\mathbb{R}^n, \mathbb{R}^n)$, $\mathbb{D}^0 v = \mathbb{D}v - \frac{2}{n}(\nabla \cdot v) I$.

We conclude this section with elementary identities involving various first order differential operators.

LEMMA 6.12. *(Identities involving the curl, divergence, symmetrized gradient, and deviatoric gradient). For any twice-differentiable n-dimensional vector field v, i.e. $v \in C^2(\mathbb{R}^n, \mathbb{R}^n)$, the following hold:*
(1) $\nabla \times (\nabla \times v) = \nabla (\nabla \cdot v) - \Delta v$ *(when $n = 3$)*
(2) $\nabla \cdot (\mathbb{D}v) = \nabla (\nabla \cdot v) + \Delta v$

(3) $\nabla \cdot \left(\mathbb{D}^0 v\right) = \left(1 - \frac{2}{n}\right) \nabla \left(\nabla \cdot v\right) + \Delta v$

Proof. These identities follow from direct computations. To obtain (1) we use Lemma 6.1:

$$(\nabla \times (\nabla \times v))_i = \epsilon_{kij} \partial_j \left(\epsilon_{klm} \partial_l v_m\right) = \left(\delta_{il}\delta_{jm} - \delta_{im}\delta_{jl}\right)\left(\partial_j \partial_l v_m\right) = \partial_i \left(\partial_j v_j\right) - \partial_j \partial_j v_i = \left(\nabla \left(\nabla \cdot v\right) - \Delta v\right)_i.$$

(2) and (3) then follow from direct computations:

(2) $\left(\nabla \cdot (\mathbb{D} v)\right)_i = \partial_j (\mathbb{D} v)_{ij} = \partial_j \left(\partial_j v_i + \partial_i v_j\right) = \partial_j \partial_j v_i + \partial_i \left(\partial_j v_j\right) = \left(\Delta v + \nabla \left(\nabla \cdot v\right)\right)_i$ and

(3) $\nabla \cdot \left(\mathbb{D}^0 v\right) = \nabla \cdot \left(\operatorname{Dev}\left(\mathbb{D} v\right)\right) = \nabla \cdot \left(\mathbb{D} v - \frac{2}{n} \left(\nabla \cdot v\right) I\right) = \Delta v + \left(1 - \frac{2}{n}\right) \nabla \left(\nabla \cdot v\right).$

∎

6.2. Skew-symmetric matrices in three dimensions. In this section we obtain various results that have to do with ten and vec. The first result is the most important one, showing that ten and vec are linear isomorphisms and can thus indeed be used to identify Skew(3) and $\mathbb{R}^3$.

Proposition 6.13. *(Isomorphism between 3-by-3 skew-symmetric matrices and 3-dimensional vectors)* $\operatorname{vec} : \operatorname{Skew}(3) \to \mathbb{R}^3$ *is an isomorphism, whose inverse is* ten*, such that for every* $\Omega \in \operatorname{Skew}(3)$*,* $\Omega = \operatorname{vec}(\Omega) \times$*, i.e. the action of skew-symmetric matrices is equivalent to the action of a vectors via the cross-product.*

Proof. It suffices to show that vec and ten from Definition 2.18 satisfy

$$\operatorname{vec} \circ \operatorname{ten} = \operatorname{id}_{\operatorname{Skew}(3)} \text{ and } \operatorname{ten} \circ \operatorname{vec} = \operatorname{id}_{\mathbb{R}^3}$$

and that, for every $\Omega \in \operatorname{Skew}(3)$ and every $v \in \mathbb{R}^3$, $\Omega v = \operatorname{vec}(\Omega) \times v$. The first two identities follow from the epsilon-delta identities in Lemma 6.1. Indeed we may compute that

$$\operatorname{vec}\left(\operatorname{ten}(\omega)\right)_i = \frac{1}{2} \epsilon_{aib} \operatorname{ten}(\omega)_{ab} = \frac{1}{2} \epsilon_{aib} \epsilon_{ajb} \omega_j = \frac{1}{2} \left(2\delta_{ij}\right) \omega_j = \omega_i$$

and

$$\operatorname{ten}\left(\operatorname{vec}(\Omega)\right)_{ij} = \epsilon_{iaj} \left(\frac{1}{2} \epsilon_{kal} \Omega_{kl}\right) = \frac{1}{2} \left(\delta_{ik}\delta_{jl} - \delta_{il}\delta_{jk}\right) \Omega_{kl} = \operatorname{Skew}(\Omega)_{ij} = \Omega_{ij}$$

Similarly, a direct computation shows that

$$\left((\operatorname{vec} \Omega) \times v\right)_i = \epsilon_{ijk} (\operatorname{vec} \Omega)_j v_k = \frac{1}{2} \epsilon_{ijk} \epsilon_{ajb} \Omega_{ab} v_k = \frac{1}{2} \left(\delta_{ia}\delta_{kb} - \delta_{ib}\delta_{ka}\right) \Omega_{ab} v_k = \frac{1}{2} \left(\Omega_{ik} - \Omega_{ki}\right) v_k$$

i.e. $(\operatorname{vec} \Omega) \times v = (\operatorname{Skew} \Omega) v = \Omega v$. ∎

We now record the fact that ten and vec are frame-invariant.

Lemma 6.14. *(Frame-invariance of* ten *and* vec*)*
For any $v \in \mathbb{R}^3$*, any* 3-by-3 *skew-symmetric matrix* A*, and any orthogonal transformation* Q*,*

$$\operatorname{ten}(Qv) = (\det Q)\, Q (\operatorname{ten} v) Q^T \text{ and } \operatorname{vec} QAQ^T = (\det Q) \operatorname{vec} A.$$

Proof. This is an immediate consequence of Lemma 6.4 since it allows us to compute that, for any $w \in \mathbb{R}^3$,

$$\operatorname{ten}(Qv) w = (Qv) \times w = (\det Q)\, Q (v \times (Q^T w)) = (\det Q)\, Q (\operatorname{ten} v) Q^T w$$

and

$$\operatorname{vec}(QAQ^T) \times w = QAQ^T w = Q(\operatorname{vec} A \times Q^T w) = (\det Q)(Q \operatorname{vec} A) \times w.$$

∎

Here we record how ten and vec relate to Skew, the linear operator which isolates the skew-symmetric part of a matrix.

Lemma 6.15. *(Relation between* ten*,* vec*, and* Skew*) The following identity holds:* $\operatorname{ten} \circ \operatorname{vec} = \operatorname{Skew}$*.*

Proof. This identity is nothing more than the classical epsilon-delta identity (c.f. Lemma 6.1) in coordinate-invariant-form. Indeed, let A be a real n-by-n matrix. Then using Lemma 6.1 and Proposition 6.13 we compute that

$$\operatorname{ten}\left(\operatorname{vec} A\right)_{ij} = \epsilon_{iaj} (\operatorname{vec} A)_a = \frac{1}{2} \epsilon_{iaj} \epsilon_{paq} A_{pq} = \frac{1}{2} \left(\delta_{ip}\delta_{jq} - \delta_{iq}\delta_{jp}\right) A_{pq} = \frac{1}{2} \left(A_{ij} - A_{ji}\right) = (\operatorname{Skew} A)_{ij}.$$

∎

We now derive the relationship between ten, vec, and various first order differential operators.

Lemma 6.16. *(Relation of* ten, vec, Skew*, and differential operators)*
For any 3-dimensional differentiable vector field v, i.e. $v:\mathbb{R}^3\to\mathbb{R}^3$, the following identities hold:

$$\operatorname{vec}\nabla v=\frac{1}{2}\nabla\times v,\ \operatorname{Skew}\nabla v=\frac{1}{2}\operatorname{ten}\nabla\times v,\ \text{ and } \nabla\cdot(\operatorname{ten}v)=-\nabla\times v.$$

Proof. Recall that $(\nabla\times v)_i=\epsilon_{ijk}\partial_j v_k$. We may thus compute that

$$(\operatorname{vec}\nabla v)_i=\frac{1}{2}\epsilon_{aib}(\nabla v)_{ab}=\frac{1}{2}\epsilon_{iba}\partial_b v_a=\frac{1}{2}\nabla\times v \text{ and } (\nabla\cdot(\operatorname{ten}v))_i=\partial_j(\operatorname{ten}v)_{ij}=\partial_j\left(\epsilon_{iaj}v_a\right)=-\nabla\times v$$

such that the first and third identities hold. The second identity follows from combining the first identity and Lemma 6.15 since $\operatorname{Skew}\nabla v=\operatorname{ten}\operatorname{vec}\nabla v=\frac{1}{2}\operatorname{ten}\nabla\times v$. □

We conclude this section with a result relating ten, vec, and inner products.

Lemma 6.17. *(Relation between* ten, vec*, and inner products)*
For any $v\in\mathbb{R}^3$ and any $M\in\mathbb{R}^{3\times 3}$, $(\operatorname{vec}M)\cdot v=\frac{1}{2}M:(\operatorname{ten}v)$.

Proof. We compute: $(\operatorname{vec}M)\cdot v=(\operatorname{vec}M)_i v_i=\frac{1}{2}\epsilon_{aib}M_{ab}v_i=\frac{1}{2}M_{ab}(\operatorname{ten}v)_{ab}=\frac{1}{2}M:(\operatorname{ten}v)$. □

6.3. Matrix groups, and Lie groups. In this section we prove some elementary results concerning matrix groups and Lie groups. We begin by showing that the space of skew-symmetric matrices is closed under conjugacy by orthogonal matrices.

Proposition 6.18. *(Closure of skew-symmetric matrices under conjugacy by orthogonal matrices)*
$\operatorname{Skew}(n)$ *is closed under conjugacy by $O(n)$, i.e. for any $A\in\operatorname{Skew}(n)$ and $R\in O(n)$, $RAR^{-1}\in\operatorname{Skew}(n)$.*

Proof. This is immediate: $\left(RAR^{-1}\right)^T=R^{-T}A^TR^T=R(-A)R^{-1}=-\left(RAR^{-1}\right)$. □

We now record a result dealing with matrix groups, identifying the tangent space to the Lie group of orthogonal matrices.

Proposition 6.19. *(Tangents to matrix groups, or a glimpse into Lie algebras)*
(1) $T_I\,O(n)=\operatorname{Skew}(n)$*, i.e. the Lie algebra of orthogonal matrices consists precisely of the algebra of skew-symmetric matrices.*
(2) *For any $R\in O(n)$, $T_R\,O(n)=R\operatorname{Skew}(n)$.*

Proof. (1) Consider $R:\mathbb{R}\to O(n)$ such that $R(0)=I$. Upon differentiating $I=RR^T$ and evaluating at 0, we obtain that: $0=\dot R(0)R(0)^T+R(0)\dot R(0)^T=\dot R(0)+\dot R(0)^T$ i.e. indeed $\dot R(0)\in\operatorname{Skew}(n)$. Conversely, for any $A\in\operatorname{Skew}(n)$, define $R(t):=e^{At}$ and observe that

$$\left(e^{At}\right)^Te^{At}=e^{A^Tt}e^{At}=e^{\left(A^T+A\right)t}=e^0=I \text{ and } \left(e^{At}\right)'|_{t=0}=Ae^0=A$$

such that indeed $\operatorname{Skew}(n)\subseteq T_I\,O(n)$ and so (1) holds.

(2) Fix $R_0\in O(n)$ and consider $R:\mathbb{R}\to O(n)$ such that $R(0)=R_0$. Note that $L_{R_0^{-1}}\circ R$ is a path along $O(n)$ going through the identity at time zero, and hence by the above:

$$d\left(L_{R_0^{-1}}\circ R\right)|_{t=0}=:A\in\operatorname{Skew}(n)$$

Now, since the exterior derivative 'commutes' with composition, we have that

$$A=d\left(L_{R_0^{-1}}\circ R\right)=dL\circ dR=R_0^{-1}\circ\dot R$$

i.e. $\dot R=R_0A\in R_0\operatorname{Skew}(n)$.
Conversely, for any $R_0\in O(n)$ and $A\in\operatorname{Skew}(n)$, define $R(t):=R_0e^{At}$ such that

$$\left(R_0e^{At}\right)^TR_0e^{At}=e^{A^Tt}R_0^TR_0e^{At}=e^{\left(A^T+A\right)t}=I \text{ and } \left(R_0e^{At}\right)'|_{t=0}=R_0Ae^0=R_0A$$

such that indeed $R_0\operatorname{Skew}(n)\subseteq T_{R_0}O(n)$ and thus (2) holds.

□

We now prove results about Lie groups and the determinant in particular. The goal of these results is to justify in a clean way the formula for the derivative of the determinant. First we prove that, in order to compute the derivative of a Lie group homomorphism (such as the determinant), it suffices to compute its derivative at the identity.

Proposition 6.20. *(Differential of a Lie group homomorphism)*
Let G, H be Lie groups and let $F : G \to H$ be a Lie group homomorphism. Recall that for any $g \in G$, L_g denotes left-multiplication, i.e. $L_g(h) = gh$ for any $h \in G$. Then, for any $g \in G$,

$$dF_g = dL_{F(g)} \circ dF_e \circ dL_{g^{-1}}$$

or, written in all its gory detail: $dF|_g = dL^{(H)}_{F(g)}\Big|_{e_H} \circ dF|_{e_G} \circ dL^{(G)}_{g^{-1}}\Big|_g$.
Moral: To compute the differential of a Lie group homomorphism it is enough to know how to compute its differential at the identity and the differentials of the left-multiplication operators.

$$\begin{tikzcd} T_gG \arrow[r, dashed, "dF_g"] \arrow[d, "dL_{g^{-1}}"] & T_{F(g)}H \\ T_eG \arrow[r, "dF_e"] & T_eH \arrow[u, "dL_{F(g)}"'] \end{tikzcd}$$

Proof. Since F is a group homomorphism, the following holds for any $g \in G$:

$$F = L_{F(g)} \circ F \circ L_{g^{-1}}$$

Therefore, upon applying the exterior derivative and noting that it 'commutes' with composition, we obtain the desired result. ☐

Following Proposition 6.20, since we are seeking a formula for the derivative of the determinant we compute its derivative at the identity.

Lemma 6.21. *(Derivative of the determinant at the identity)*

$$\det'|_I = \operatorname{tr}$$

Proof. Let H be an arbitrary n-by-n matrix which we write as $H = (h_1|h_2|\dots|h_n)$. Then

$$\begin{aligned}\det(I + \epsilon H) &= \det(e_1 + \epsilon h_1|\dots|e_n + \epsilon h_n)\\ &= \det(e_1|\dots|e_n) + \epsilon\left(\det(h_1|\dots|e_n) + \dots + \det(e_1|\dots|h_n)\right) + \mathcal{O}\left(\epsilon^2\right)\\ &= 1 + \epsilon \underbrace{(h_{11} + \dots + h_{nn})}_{\operatorname{tr} H} + \mathcal{O}\left(\epsilon^2\right)\end{aligned}$$

and hence

$$\det'|_I(H) = \lim_{\epsilon \to 0} \frac{\det(I + \epsilon H) - \det I}{\epsilon} = \operatorname{tr} H$$

☐

Finally we conclude this section by using Proposition 6.20 and Lemma 6.21 to establish a formula for the derivative of the determinant.

Corollary 6.22. *(Derivative of the determinant)*
Let A and M_0 be n-by-n matrices, with M_0 invertible. Then

$$\det'|_{M_0}(A) = \det(M_0) \operatorname{tr}\left(M_0^{-1} A\right)$$

Proof. Since the determinant is a Lie group homomorphism from $GL(n)$ to $\mathbb{R}$ (as a multiplicative group), we know by Proposition 6.20 that it is enough to compute the derivative of the determinant at the identity as well as the derivative of the left-multiplications. By Lemma 6.21, $\det'|_I = \operatorname{tr}$, and since here both left-multiplications are linear (and hence equal to their derivatives), we obtain that:

$$\det'|_{M_0}(A) = \left(L_{\det(M_0)} \circ \operatorname{tr} \circ L_{M_0^{-1}}\right)(A) = \det(M_0) \det\left(M_0^{-1} A\right)$$

☐

CHAPTER 2

Anisotropic micropolar fluids subject to a uniform microtorque: the unstable case

Abstract.

We study a three-dimensional, incompressible, viscous, micropolar fluid with anisotropic microstructure on a periodic domain. Subject to a uniform microtorque, this system admits a unique nontrivial equilibrium. We prove that this equilibrium is nonlinearly unstable. Our proof relies on a nonlinear boostrap instability argument which uses control of higher-order norms to identify the instability at the L^2 level.

Contents

1. Introduction

1.1. Brief discussion of the model. We restrict our attention to problems in which the microinertia plays a significant role, and so in this chapter we only consider *anisotropic* micropolar fluids for which the microinertia tensor is not isotropic, i.e. J has at least two distinct eigenvalues. In fact, we study micropolar fluids whose microstructure has an *inertial axis of symmetry*, which means that the microinertia J has a repeated eigenvalue. More concretely: there are some physical constants $\lambda, \nu > 0$ which depend on the microstructure such that, at every point, J is a symmetric matrix with spectrum $\{\lambda, \lambda, \nu\}$. This is in some sense the intermediate case between the case of isotropic microstructure where the microinertia has a repeated eigenvalue of multiplicity three and the "fully" anisotropic case where the microstructure has three distinct eigenvalues.

The equations of motion related to these quantities in the periodic spatial domain $\mathbb{T}^3 = \mathbb{R}^3/(2\pi\mathbb{Z})^3$, subject to an external microtorque τe_3, read:

$$\begin{cases} \partial_t u + (u \cdot \nabla)\, u = \tilde{\mu}\Delta u + \kappa \nabla \times \omega - \nabla p & \text{on } (0,T) \times \mathbb{T}^3, \quad (1.1a) \\ \nabla \cdot u = 0 & \text{on } (0,T) \times \mathbb{T}^3, \quad (1.1b) \\ J\left(\partial_t \omega + (u \cdot \nabla)\, \omega\right) + \omega \times J\omega = \kappa \nabla \times u - 2\kappa\omega + (\tilde{\alpha} - \tilde{\gamma})\, \nabla\left(\nabla \cdot \omega\right) + \tilde{\gamma}\Delta\omega + \tau e_3 & \text{on } (0,T) \times \mathbb{T}^3, \quad (1.1c) \\ \partial_t J + (u \cdot \nabla)\, J = [\Omega, J] & \text{on } (0,T) \times \mathbb{T}^3, \quad (1.1d) \end{cases}$$

where $[\,\cdot\,,\,\cdot\,]$ denotes the matrix commutator, $\tilde{\mu}$, κ, $\tilde{\alpha}$, and $\tilde{\gamma}$ are physical constants related to viscosity, τ denotes the magnitude of the microtorque, and Ω is the 3-by-3 antisymmetric matrix identified with ω via the identity $\Omega v = \omega \times v$ for every $v \in \mathbb{R}^3$.

We have chosen to consider the situation in which external forces are absent and the external microtorque is constant, namely equal to τe_3 for some fixed $\tau > 0$. Note that the choice of e_3 as the direction of the microtorque may be made without loss of generality since the equations are equivariant under proper rotations, in the sense that if (u, p, ω, J) is a solution of (1.1a)–(1.1d) then, for any $\mathcal{R} \in SO\,(3)$, $\left(u, p, \mathcal{R}\omega, \mathcal{R}J\mathcal{R}^T\right)$ is a solution of (1.1a)–(1.1d) *provided* that the external torque τe_3 is replaced by $\tau\mathcal{R}e_3$.

It is worth noting that this system is equivariant under Galilean transformations. More precisely: if (u, p, ω, J) is a sufficiently regular solution of (1.1a)–(1.1d) then $u_{\text{avg}} := \fint_{\mathbb{T}^3} u$ is constant in time and

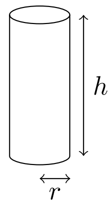
$$(0,T) \times \mathbb{T}^3 \ni (t, y) \mapsto (u - u_{\text{avg}}, p, \omega, J)\,(t, y + u_{\text{avg}})$$

also satisfies (1.1a)–(1.1d). We may therefore assume without loss of generality that u has average zero at all times. Similarly, since the pressure only appears in the equations with a gradient, we are free to posit that p has average zero for all times.

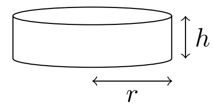
There are two ways to motivate our choice to have no external forces and a constant microtorque. On one hand, it is reminiscent of certain chiral active fluids constituted of self-spinning particles which continually pump energy into the system [BSAV17], as our constant microtorque does. On the other hand, this choice of an external force − external microtorque pair is motivated by the dearth of analytical results on anisotropic micropolar fluids. It is indeed natural, as a first step in the study of non-trivial equilibria of anisotropic micropolar fluids, to consider a simple external force − external microtorque pair yielding non-trivial equilibria for the angular velocity ω and the microinertia J. The simplest nonzero such pair is precisely our choice of $(0, \tau e_3)$.

Let us now turn to the aforementioned equilibrium. Due to the uniform microtorque, the system admits a nontrivial equilibrium. At equilibrium the fluid velocity is quiescent ($u_{eq} = 0$), the pressure is null ($p_{eq} = 0$), the angular velocity is aligned with the microtorque ($\omega_{eq} = \frac{\tau}{2\kappa} e_3$), and the inertial axis of symmetry of the microstructure is aligned with the microtorque such that the microinertia is $J_{eq} = \text{diag}(\lambda, \lambda, \nu)$.

Physically-motivated heuristics (which again we postpone until Section 2) suggest that the stability of this equilibrium depends on the 'shape' of the microstructure. The heuristics suggest that if the microinertia is inertially oblong, i.e. if $\lambda > \nu$, then the equilibrium is unstable, and that if the microinertia is inertially oblate, i.e. if $\nu > \lambda$, then the equilibrium is stable. This nomenclature is justified by the fact that for rigid bodies with an axis of symmetry and a uniform mass density, the notions of being oblong (or oblate), which essentially means that the body is longer (respectively shorter) along its axis of symmetry than it is wide across it, and being inertially oblong (respectivelly inertially oblate) coincide. Examples of inertially oblong and oblate rigid bodies are provided in Figure 1. This chapter deals with the instability of inertially oblong microstructure. In future work we will study the stability of inertially oblate microstructure.



(A) This rigid body is inertially oblong if $h^2 > 6r^2$.



(B) This rigid body is inertially oblate if $h^2 < 6r^2$.

FIGURE 1. Two rigid bodies with uniform density which possess an inertial axis of symmetry.

1.2. Statement of the main result. The main thrust of this chapter is to prove that if the microstructure is inertially oblong, then the equilibrium is nonlinearly unstable in L^2. A precise statement of the theorem may be found in Theorem 5.2, but an informal statement of the result is the following.

THEOREM 1.1 (L^2 instability of the equilibrium). *Suppose that the microstructure is inertially oblong, i.e. suppose that $\lambda > \nu$, and let $X_{eq} = (u_{eq}, \omega_{eq}, J_{eq}) = (0, \frac{\tau}{2\kappa} e_3, \operatorname{diag}(\lambda, \lambda, \nu))$ be the equilibrium solution of* (1.1a)–(1.1d). *Then X_{eq} is nonlinearly unstable in L^2.*

Here the notion of nonlinear instability is the familiar one from dynamical systems: there exists a radius $\delta > 0$ and a sequence of initial data $\{X_n^0\}_{n=0}^{\infty}$, converging to X_{eq} in L^2, such that the solutions to (1.1a)–(1.1d) starting from X_n^0 exit the ball $B(X_{eq}, \delta)$ in finite time, depending on n.

Note that in Theorem 1.1 the pressure has disappeared from consideration. This is because the pressure plays only an auxiliary role in the equations and may be eliminated from (1.1a) by projecting onto the space of divergence-free vector fields.

2. Background, preliminaries, and discussion

2.1. Previous work. Micropolar fluids have been extensively studied by the continuum mechanics community over the last fifty years and an exhaustive literature review is beyond the scope of this chapter. We restrict our attention to the mathematics literature here, in which case, to the best of our knowledge all results relate to *isotropic* microstructure, where the microinertia J is a scalar multiple of the identity. In that case the precession term $\omega \times J\omega$ from (1.1c) vanishes and the entire equation (1.1d) trivializes. Note that in two dimensions the micro-inertia is a scalar, and therefore all micropolar fluids are isotropic.

In two dimensions the problem is globally well-posed, as per [**Łuk01**] where global well-posedness and qualitative results on the long-time behaviour are obtained. Some quantitative information on long-time behaviour is also known in two dimensions: for example, decay rates are obtained in [**DC09**]. The situation is more delicate in three dimensions, which is an unsurprising assertion in the setting of viscous fluids. The first discussion of well-posedness in three dimensions is due to Galdi and Rionero [**GR77**]. Łukaszewicz then obtained weak solutions in [**Łuk90**] and uniqueness of strong solutions in [**Lu89**]. More recent work has established global well-posedness for small data in critical Besov spaces [**CM12**] and in the space of pseudomeasures [**FVR07**], as well as derived blow-up criteria [**Yua10**]. There is also an industry devoted to the study of micropolar fluids when one or more of the viscosity coefficients vanishes: we refer to [**DZ10**] for an illustrative example.

Various extensions of the incompressible micropolar fluid model considered here have been studied. For example, compressible models [**LZ16**], models coupled to heat transfer [**Tar06, KLŁ19**], and models with coupled magnetic fields [**AS74, RM97**] have all been studied. Again, to the best of our knowledge all of these works consider *isotropic* micropolar fluids.

2.2. Equilibria. In this section we describe the two classes of equilibria which arise as particular solutions of (1.1a)–(1.1d). A critical piece of this description is the following energy-dissipation relation:

$$\frac{\mathrm{d}}{\mathrm{d}t} \int_{\mathbb{T}^3} \frac{1}{2}|u|^2 + \frac{1}{2} J(\omega - \omega_{eq}) \cdot (\omega - \omega_{eq}) - \frac{1}{2} J\omega_{eq} \cdot \omega_{eq} \tag{2.1}$$
$$= -\int_{\mathbb{T}^3} \frac{\mu}{2}|\mathbb{D}u|^2 + 2\kappa \left| \frac{1}{2} \nabla \times u - (\omega - \omega_{eq}) \right|^2 + \alpha |\nabla \cdot \omega|^2 + \frac{\beta}{2} |\mathbb{D}^0 \omega|^2 + 2\gamma |\nabla \times \omega|^2.$$

It is obtained by testing (1.1a) and (1.1c) against u and $\omega - \omega_{eq}$ respectively and integrating by parts. For a full derivation, see Appendix 6.3. With the relation (2.1) in hand we may define two classes of equilibria.

DEFINITION 2.1. We say that a solution (u, p, ω, J) of (1.1a)–(1.1d) is an *equilibrium* if $\partial_t (u, p, \omega, J) = 0$ and we say that it is an *energetic equilibrium* if $\frac{\mathrm{d}}{\mathrm{d}t}\mathcal{E}_{\mathrm{rel}} = 0$ where the relative energy $\mathcal{E}_{\mathrm{rel}}$ is given as in (2.1) by

$$\mathcal{E}_{\mathrm{rel}}(u, p, \omega, J) = \int_{\mathbb{T}^3} \frac{1}{2}|u|^2 + \frac{1}{2}J(\omega - \omega_{eq}) \cdot (\omega - \omega_{eq}) - \frac{1}{2}J\omega_{eq} \cdot \omega_{eq}. \tag{2.2}$$

There are two reasons why one might study the energetic equilibria introduced in Definition 2.1: (1) they arise naturally as the stationary points of a Lyapunov functional and (2) we believe that they play an essential role in characterizing the long-time behaviour of the system.

We justify (1) now and postpone the justification of (2) until after the identification of the various equilibria is carried out in Proposition 2.2. Since the relative energy $\mathcal{E}_{\mathrm{rel}}$ is both non-increasing in time and bounded below we may indeed view it as a Lyapunov functional. The observation that $\frac{\mathrm{d}}{\mathrm{d}t}\mathcal{E}_{\mathrm{rel}} \leqslant 0$ follows immediately from (2.1) and the boundedness from below of $\mathcal{E}_{\mathrm{rel}}$ follows from the fact that the spectrum of the microinertia J is invariant over time.

More precisely: the conservation of microinertia for a *homogeneous* micropolar fluid means that there exists some reference microinertia J_{ref} to which $J(t, x)$ is similar at all times $0 \leqslant t < T$ and at every point $x \in \mathbb{T}^3$. Denoting by $\lambda_{\max}$ the largest eigenvalue of J_{ref} it follows that the only non-positive term in $\mathcal{E}_{\mathrm{rel}}$ is bounded below: $-J\omega_{eq} \cdot \omega eq \geqslant -\lambda_{\max}|\omega_{eq}|^2$, and hence $\mathcal{E}_{\mathrm{rel}}$ itself is bounded below.

We now identify all of the (sufficiently regular) equilibria which belong to each class as defined in Definition 2.1. Recall that we are considering a homogeneous micropolar fluid whose microstructure has an inertial axis of symmetry, which means that there are physical constants $\lambda, \nu > 0$ such that the microinertia has spectrum $\{\lambda, \lambda, \nu\}$. In particular this microinertia tensor is *physical* precisely when $2\lambda \geqslant \nu \geqslant 0$. We will assume thereafter that strict inequalities hold, i.e. $2\lambda > \nu > 0$. This assumptions means that the microstructure is not degenerate, in the sense that it corresponds to a genuinely three-dimensional rigid body (as opposed to a degenerate rigid body which would be lower-dimensional, e.g. because it is flat in one or more directions).

PROPOSITION 2.2. *Let (u, p, ω, J) be a sufficiently regular solution of (1.1a)–(1.1d) where u has average zero.*

1. *If (u, p, ω, J) is an equilibrium then $u = 0$, $p = 0$, $\omega = \omega_{eq} = \frac{\tau}{2\kappa}e_3$, and $J = \operatorname{diag}(\lambda, \lambda, \nu) = \lambda I_2 \oplus \nu$.*
2. *If (u, p, ω, J) is an energetic equilibrium then either it is an equilibrium or $u = 0$, $p = 0$, $\omega = \omega_{eq}$, and $J = e^{t\frac{\tau}{2\kappa}R}\bar{J}(0)e^{-t\frac{\tau}{2\kappa}R} \oplus \lambda$ where $R = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$ and where the spectrum of $\bar{J}(0)$ is $\{\lambda, \nu\}$. Here '$\oplus$' denotes the direct sum of two linear operators, see Section 2.6 to recall the precise definition.*

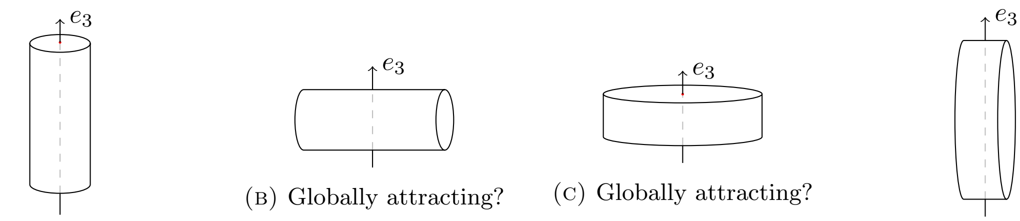
In simpler words Proposition 2.2 says that for both equilibria and energetic equilibria the microstructure rotates in the direction of the imposed microtorque, with one crucial difference: the unique equilibrium corresponds to the inertial axis of symmetry of the microstructure being *aligned* with the microtorque, giving rise to a *constant* microinertia, whilst the energetic equilibria consist of an orbit where the inertial axis of symmetry rotates in the plane *perpendicular* to the microtorque, giving rise to a *periodic* microinertia (with period $4\pi\kappa/\tau$).

PROOF OF PROPOSITION 2.2. Since equilibria are energetic equilibria we suppose that (u, p, ω, J) is an energetic equilibrium. It follows from the energy-dissipation relation (2.1) that the dissipation vanishes, i.e.

$$\int_{\mathbb{T}^3} \frac{\mu}{2}|\mathbb{D}u|^2 + 2\kappa\left|\frac{1}{2}\nabla \times u - (\omega - \omega_{eq})\right|^2 + \alpha|\nabla \cdot \omega|^2 + \frac{\beta}{2}|\mathbb{D}^0\omega|^2 + 2\gamma|\nabla \times \omega|^2 = 0.$$

In particular: ω is constant and u has constant curl. Coupling this with the fact that u is divergence-free we deduce that u is harmonic. Since u has average zero, it follows that $u = 0$, and hence that $p = 0$ (recall that we require p to have average zero) and $\omega = \omega_{eq}$.

So now we know from (1.1c) that the precession term $\omega \times J\omega = \left(\frac{\tau}{2\kappa}\right)^2 e_3 \times Je_3$ vanishes, and hence J has the block form $J = \bar{J} \oplus J_{33}$ for some 2-by-2 matrix $\bar{J}$. The conservation of microinertia (1.1d) now becomes the ODE $\partial_t J = [\operatorname{ten}\omega_{eq}, J] = \frac{\tau}{2\kappa}\left[R, \bar{J}\right] \oplus 0$ which may be solved explicitly to yield $\bar{J}(t) = e^{t\frac{\tau}{2\kappa}R}\bar{J}(0)e^{-t\frac{\tau}{2\kappa}R}$ and $J_{33}(t) = J_{33}(0)$.



(A) Unstable. (B) Globally attracting? (C) Globally attracting? (D) Unstable?

FIGURE 2. Depictions of the microstructure for the equilibrium (A, C) and an energetic equilibrium (B, D) corresponding to both the oblong (A, B) and oblate cases (C, D). B and C are conjectured to be globally attracting for the oblong and oblate cases respectively, D is conjectured to be be unstable for the oblate case, and we prove in Theorem 1.1 that A is unstable.

There are now two cases to consider: either $\bar{J}$ has a repeated eigenvalue λ or $\bar{J}$ has distinct eigenvalues λ and ν. Since $e^{t\frac{\tau}{2\kappa}R}\bar{J}(0)e^{-t\frac{\tau}{2\kappa}R}$ is constant in time if and only if $\bar{J}(0)$, and hence $\bar{J}(t)$, has a repeated eigenvalue, the result follows. $\square$

As the next section suggests, we believe that the global attractors of (1.1a)–(1.1d) may be characterized in terms of the equilibrium and the orbit of energetic equilibria. This is summarized in the conjecture below, which is the second reason why energetic equilibria are worthy of attention.

CONJECTURE 2.3.

(1) *If the microstructure is inertially oblong, i.e. $\lambda > \nu$, then the orbit of energetic equilibria identified in Proposition 2.2 is the global attractor of the system* (1.1a)–(1.1d).
(2) *If the microstructure is inertially oblate, i.e. $\lambda < \nu$, then the equilibrium identified in Proposition 2.2 is the global attractor of the system* (1.1a)–(1.1d).

A depiction of the equilibrium and the energetic equilibria configurations of the microstructure can be found in Figure 2, where we also label each configuration with its relevant conjectured long-time behaviour.

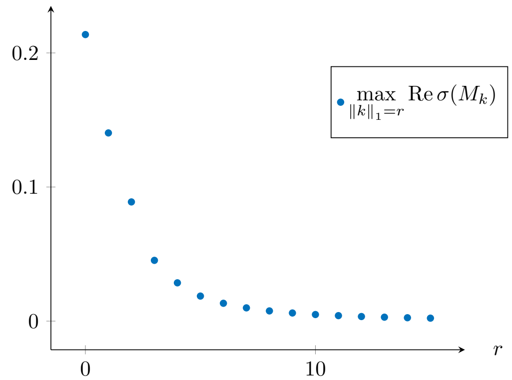
2.3. Heuristics for the long-time behaviour. In this section we briefly discuss heuristics for the long-term behaviour of the system (1.1a)–(1.1d). The central element of the reasoning that follows is the energy-dissipation relation (2.1). As remarked in Section 2.2, this relation tells us that the relative energy $\mathcal{E}_{\text{rel}}$ defined in (2.2) is non-increasing in time and bounded below. Let us therefore, for the sake of this discussion, assume that $\mathcal{E}_{\text{rel}}$ approaches its absolute minimum as time approaches $+\infty$. In particular this means that each term in $\mathcal{E}_{\text{rel}}$ approaches its absolute minimum, from which we deduce that u approaches zero, ω approaches ω_{eq} (since J is strictly positive-definite at time $t = 0$ and hence strictly positive-definite for all time), and $-J_{33}$ approaches $-\lambda_{\max}$ for $\lambda_{\max}$ denoting the maximum eigenvalue of J, i.e. $\lambda_{\max} = \max(\lambda, \nu)$.

This last observation is precisely where the dichotomy between inertially oblong and inertially oblate microstructure comes in. If the microstructure is inertially oblong, i.e. $\lambda > \nu$, then J_{33} approaches λ which means that $\bar{J}$ must consist of the *distinct* eigenvalues λ, ν, and hence the global attractor is conjectured to be the orbit of energetic equilibria. If the microstructure is inertially oblate, i.e. $\nu > \lambda$, then J_{33} approaches ν and hence $\bar{J}$ has repeated eigenvalues equal to λ, such that the global attractor is conjectured to be the equilibrium.

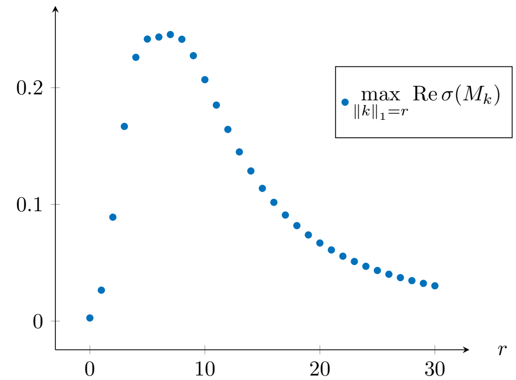
2.4. Heuristics for the origin of the instability. In this section we discuss heuristics for the origin of the instability of the system (1.1a)–(1.1d). Beyond being helpful heuristics that physically motivate the instability of the system, the ideas presented below actually form the core of our proof of the nonlinear instability.

We begin with another energy-dissipation relation, which is associated with the linearization of the problem (1.1a)–(1.1d) about its equilibrium. This relation is

$$\frac{\mathrm{d}}{\mathrm{d}t}\mathcal{E}_{\text{lin}} := \frac{\mathrm{d}}{\mathrm{d}t}\int_{\mathbb{T}^3}\left(\frac{1}{2}|u|^2 + \frac{1}{2}J_{eq}\omega\cdot\omega - \frac{1}{2}\frac{1}{\lambda-\nu}\left(\frac{\tau}{2\kappa}\right)^2|a|^2\right) = -\mathcal{D}\left(u, \omega - \omega_{eq}\right) \tag{2.3}$$



(A) Physical parameters: $\lambda = 3.2$, $\nu = 0.6$, $\mu = 4.3$, $\kappa = 3.3$, $\alpha = 0.9$, $\beta = 6.8$, $\gamma = 0.4$, $\tau = 4.4$.



(B) Physical parameters: $\lambda = 3.6$, $\nu = 1.2$, $\mu = 2.4$, $\kappa = 0.4$, $\alpha = 5.3$, $\beta = 3.1$, $\gamma = 1.7$, $\tau = 20$.

FIGURE 3. An illustration of the fact that the instability is not exclusively due to the zero mode: depending on the physical parameter regime the eigenvalue with largest real part may or may not occur when $k = 0$. Here M_k denotes the symbol of the linearization of (1.1a)–(1.1d) about the equilibrium.

where $a = (J_{31}, J_{32}) = (J_{13}, J_{23})$ and where the dissipation $\mathcal{D}$ is given as in (2.1) by

$$\mathcal{D}(u,\omega) = \int_{\mathbb{T}^3} \frac{\mu}{2}|\mathbb{D}u|^2 + 2\kappa\left|\frac{1}{2}\nabla \times u - \omega\right|^2 + \alpha|\nabla\cdot\omega|^2 + \frac{\beta}{2}|\mathbb{D}^0\omega|^2 + 2\gamma|\nabla\times\omega|^2.$$

Note that only part of the micro-inertia J appears in (2.3), namely $a = (J_{31}, J_{32})$ which corresponds to the products of inertia which describe the moment of inertia about the e_1-axis and e_2-axis, respectively, when the microstructure rotates about the e_3-axis. This is due to the fact that, as explained in detail in Section 3.1, the linearized problem can de decomposed into blocks which do not interact with one another. In particular the block governing the dynamics of u, ω, and a is the only block which produces non-trivial dynamics, and it is this block which gives rise to (2.3).

Since the integrand of $\mathcal{E}_{\text{lin}}$ in (2.3), viewed as a quadratic form on (u, ω, a), has negative directions precisely when the microstructure is inertially oblong, i.e. when $\lambda > \nu$, this suggests that the equilibrium is unstable in that case.

We actually know a little bit more about the instability mechanism. If we denote by $M(k)$, where $k \in \mathbb{Z}^3$, the symbol of the linearized operator about the equilibrium, then we can compute the spectrum of $M(0)$ explicitly and see that is has exactly two unstable eigenvalues, which come as a conjugate pair. An important point to note here is that the only nonzero components of the eigenvectors corresponding to this conjugate pair are the components corresponding to a and $\bar{\omega}$, which denotes the horizontal components of ω, i.e. $\bar{\omega} = (\omega_1, \omega_2)$. It is thus precisely a and $\bar{\omega}$ that are at the origin of the instability.

This is particularly interesting since $M(0)$ is precisely (up to neglecting its components depending on u) the linearization of the ODE

$$\begin{cases} J\dfrac{d\omega}{dt} + \omega \times J\omega = \tau e_3 - 2\kappa\omega \\ \dfrac{dJ}{dt} = [\Omega, J] \end{cases}$$

about its equilibrium $(\omega_{eq}, J_{eq}) = \left(\frac{\tau}{2\kappa}e_3, \operatorname{diag}(\lambda,\lambda,\nu)\right)$, where here ω and J are only time-dependent. This ODE describes the rotation of a damped rigid body subject to a uniform torque, which tells us that instability of the system (1.1a)–(1.1d) stems precisely from the instability of this ODE.

Finally note that, although this ODE plays a key role in explaining the instability mechanism, it does not fully characterize it. To understand what we mean by this, recall that the linearization of the ODE about its equilibrium describes the evolution of the zero Fourier mode of the linearized PDE. However, the nonzero Fourier modes play a nontrivial role in the instability mechanism. Indeed numerics show that, depending on the physical regime, the most unstable mode (i.e. that giving rise to the eigenvalue with the largest positive real part) may or may not be the zero mode. This is shown in Figure 3.

2.5. Summary of techniques and plan of chapter. Our technique for proving Theorem 1.1 is to employ the nonlinear bootstrap instability framework first introduced by Guo-Strauss [**GS95a**], which is not so much a black-box theorem as it is a strategy for proving instability. In broad strokes, the idea is to construct a maximally unstable solution to the linearized equations and then employ a nonlinear energy method to prove that this solution is nonlinearly stable, i.e. the nonlinear dynamics stay close to the linear growing mode, which then leads to instability.

An essential feature of the Guo-Strauss bootstrap instability framework is that it does *not* require the presence of a spectral gap, as is required for other standard methods used to prove nonlinear instability (see for example [**FSV97**]). This is crucial for us since it is quite delicate to obtain spectral information about the problem at hand, as discussed in more detail below. In particular, note that Proposition 3.9 tells us that a pair of conjugate eigenvalues of the linearized operator approach the imaginary axis as the wavenumber approaches infinity. As an immediate consequence, we may thus deduce that there is no spectral gap.

In order to implement the bootstrap instability strategy we need four ingredients. The first is the maximally unstable linear growing mode. This is a solution to the linearized equations (linearized around the equilibrium) that grows exponentially in time (when measured in various Sobolev norms) at a rate that is maximal in the sense that no other solution to the linearized equations grows more rapidly. The second is a scheme of nonlinear energy estimates that allows us to obtain control of high-regularity norms of solutions to the nonlinear problems in terms of certain low-regularity norms. This is the bootstrap portion of the argument. The third is a low-regularity estimate of the nonlinearity in terms of the square of the high-regularity energy, valid at least in a *small energy regime*. Finally, we need a local existence theory for the nonlinear problem that is capable of producing solutions to which the bootstrap estimates apply. With these ingredients in hand, we can then prove that the nonlinear solution stays sufficiently close to the growing linearized solution that it must leave a ball of fixed radius within a timescale computed in terms of the data.

In Section 3 we construct the maximally unstable solution to the linearized equations. A principal difficulty is encountered immediately upon linearizing: the resulting (spatial) differential operator is not self-adjoint. This is due entirely to the anisotropy of the microstructure, and in particular to the term $\omega \times J\omega$ in (1.1c); indeed, in the case of isotropic microstructure this term vanishes and the linearized operator becomes self-adjoint. The lack of self-adjointness means we have far fewer tools at our disposal, and in particular it means that we cannot employ variational methods to find the maximal growing mode.

Since we work on the torus and the linearization is a constant coefficient problem, we are naturally led to seek the maximal solution in the form of a growing Fourier mode solution. This leads to an ODE in $\mathbb{C}^8$ of the form $\partial_t \hat{X}_k = \hat{\mathcal{B}}_k \hat{X}_k$, where $k \in \mathbb{Z}^3$ is the wavenumber and $\hat{\mathcal{B}}_k \in \mathbb{C}^{8\times 8}$ is not Hermitian. Without the precision tools associated to Hermitian matrices, we are forced to naively study the degree eight characteristic polynomial of $\hat{\mathcal{B}}_k$, which, due to the appearance of the physical parameters α, β, γ, κ, μ, τ, λ, ν, in addition to the wave number k, is an unmitigated mess. Numerics (see Figure 3) suggest that for any $k \in \mathbb{Z}^3$ the spectrum consists of a conjugate pair of unstable eigenvalues, a zero eigenvalue (coming from the incompressibility condition), and five stable eigenvalues. However, due to the inherent complexity of $\hat{\mathcal{B}}_k$ and its characteristic polynomial, we were unable to prove this, except in the case $k = 0$.

Failing at the direct approach of simply factoring the characteristic polynomial of $\hat{\mathcal{B}}_k$, we instead employ an indirect approach based on isolating the highest order (in terms of the wavenumber k) part of the characteristic polynomial and deriving its asymptotic form as $|k| \to \infty$. For this it's convenient to parameterize the matrices in terms of $k \in \mathbb{R}^3$ rather than $\mathbb{Z}^3$. Using this idea, the special form of the highest-order term, and the implicit function theorem, we are then able to prove the existence of an unstable conjugate pair of eigenvalues, smoothly parameterized by $k \in \mathbb{R}^3$ in a neighborhood of infinity. Remarkably, since the neighborhood of infinity contains all but finitely many lattice points from $\mathbb{Z}^3$, we conclude from this argument that for *all but finitely many* wavenumbers $\hat{\mathcal{B}}_k$ is unstable. Combining this with a number of delicate spectral estimates and an application of Rouché's theorem, we are then able to find $k_* \in \mathbb{Z}^3$ with the largest growth rate. From this and a Fourier synthesis we then construct the desired maximal growing mode.

The lack of self-adjointness is also an issue when we seek to use spectral information about $\hat{\mathcal{B}}_k$ to obtain bounds on the corresponding matrix exponential $e^{t\hat{\mathcal{B}}_k}$. These bounds are required to obtain the bounds on the semigroup generated by the linearization that verify that our growing mode is actually maximal among all linear solutions. We only know that $e^{t\hat{\mathcal{B}}_k}$ is similar to its diagonal matrix up to a change of basis matrix whose norm *depends on* k. Circumventing this issue requires a good understanding of the decay of the spectrum of the symmetric part of $\hat{\mathcal{B}}_k$ as k becomes large, and the precise workaround is discussed at the beginning of the proof of Proposition 3.11.

In Section 4 we derive the nonlinear bootstrap energy estimates and the nonlinearity estimate. Here the primary difficulty is related to rewriting the problem in a way that prevents time derivatives from entering the nonlinearity. If we were to naively rewrite (1.1c) by writing $J\partial_t\omega = J_{eq}\partial_t\omega + (J - J_{eq})\,\partial_t\omega$ and considering the term $(J - J_{eq})\,\partial_t\omega$ as a remainder term, then we would then not be able to close the estimates due to this time derivative being present as part of the nonlinear remainder. Instead we must multiply (1.1c) by $J_{eq}J^{-1}$, which solves the time derivative problem but significantly worsens the form of the remaining terms in the nonlinearity. In spite of this, we are able to derive the appropriate estimates needed for the bootstrap argument.

We delay the development of the final ingredient, the local existence theory, until Appendix 6.1. Our local existence theory is built on a nonlinear Galerkin scheme that employs the Fourier basis for the finite dimensional approximations. To solve the resulting nonlinear, but finite dimensional, ODE we borrow many of the nonlinear estimates from Section 4.

Section 5 combines the four ingredients to prove our instability result. This culminates in Theorem 5.2, the main result of the chapter. Finally, in Appendix 6.2 we record a number of auxiliary results that are used throughout the main body of the chapter.

2.6. Notation. We say a constant C is universal if it only depends on the various parameters of the problem, the dimension, etc., but not on the solution or the data. The notation $\alpha \lesssim \beta$ will be used to mean that there exists a universal constant $C > 0$ such that $\alpha \leqslant C\beta$.

Let us also record here some basic notation for linear algebraic operations. For any $w \in \mathbb{R}^n$ we denote by $P_{\parallel}(w)$ and $P_{\perp}(w)$ the orthogonal projections onto the span of w and its orthogonal complement, respectively. More precisely: for any nonzero w, $P_{\parallel}(w) = \frac{w\otimes w}{|w|^2}$ and $P_{\perp}(w) = I - \frac{w\otimes w}{|w|^2}$, whilst $P_{\parallel}(0) = 0$ and $P_{\perp}(0) = I$. For any $v \in \mathbb{R}^2$ and $w \in \mathbb{R}^3$ we write $\bar{w} = (w_1, w_2)$, $\bar{w}^{\perp} = (-w_2, w_1)$, $\tilde{v} = (v_1, v_2, 0)$, and $\tilde{v}^{\perp} = (-v_2, v_1, 0)$. Finally, let X_1, X_2, Y_1, and Y_2 be normed vector spaces, let $L_1 \in \mathcal{L}(X_1, Y_1)$, and let $L_2 \in \mathcal{L}(X_2, Y_2)$. The *direct sum* of L_1 and L_2, denoted $L_1 \oplus L_2$, is the bounded linear operator from $X_1 \times X_2$ to $Y_1 \times Y_2$ defined via, for every $(f_1, f_2) \in X_1 \times X_2$, $(L_1 \oplus L_2)(f_1, f_2) := (L_1 f_1, L_2 f_2)$.

3. Analysis of the linearization

To begin we record the precise form of the linearization of (1.1a)–(1.1d) about the equilibrium solution $(u_{eq}, p_{eq}, \omega_{eq}, J_{eq}) = (0, 0, \frac{\tau}{2\kappa}e_3, \operatorname{diag}(\lambda, \lambda, \nu))$ and introduce notation which allows us to write the linearized problem in a compact form. Then in Section 3.1 we note that the linearized operator has a natural block structure with only *one* block which gives rise to non-trivial dynamics. It is this component whose spectrum we study in detail in Section 3.2. The results from Section 3.2 are then used to construct the semigroup associated with the linearization in Section 3.3 and to construct a maximally unstable solution to the linearized problem in Section 3.4.

The linearization is

$$\begin{cases} \partial_t u = (\mu + \kappa/2)\,\Delta u + \kappa\nabla\times\omega - \nabla p, & (3.1a)\\ J_{eq}\partial_t\omega = -\left(\omega\times J_{eq}\omega_{eq} + \omega_{eq}\times J\omega_{eq} + \omega_{eq}\times J_{eq}\omega\right) & \\ \qquad\qquad +\kappa\nabla\times u - 2\kappa\omega + (\alpha + \beta/3 - \gamma)\,\nabla\nabla\cdot\omega + (\beta+\gamma)\,\Delta\omega, \text{ and} & (3.1b)\\ \partial_t J = [\Omega_{eq}, J] + [\Omega, J_{eq}] & (3.1c)\end{cases}$$

subject to $\nabla\cdot u = 0$ which, for $X = (u, \omega, J)$, $D = I_3 \oplus J_{eq} \oplus I_{\mathrm{Mat}(3)}$ (where $I_{\mathrm{Mat}(3)}$ denotes the identity function on the space of 3-by-3 matrices), $\Lambda(p) = (-\nabla p, 0, 0)$, and an appropriate linear operator $\widetilde{\mathcal{L}}$ can be written more succintly as

$$\partial_t DX = \widetilde{\mathcal{L}}X + \Lambda(p) \text{ subject to } \nabla\cdot u = 0. \tag{3.2}$$

3.1. The block structure. The linearization (3.1a)–(3.1c) can be decomposed into blocks which do not interact with one another. Notably, only one of these blocks gives rise to non-trivial dynamics, so we will identify this block before studying its spectrum in Section 3.2. More precisely: writing

$$J = \begin{pmatrix} \bar{J} & a \\ a^T & J_{33} \end{pmatrix},$$

the linearization becomes

$$\begin{cases} \partial_t u = (\mu + \kappa/2)\,\Delta u + \kappa \nabla \times \omega - \nabla p, & \text{(3.3a)} \\ J_{eq}\partial_t \omega = \kappa \nabla \times u - 2\kappa\omega + (\alpha + \beta/3 - \gamma)\,\nabla\nabla\cdot\omega + (\beta+\gamma)\,\Delta\omega - (\lambda - \nu)\,\dfrac{\mathfrak{t}}{2\kappa}\tilde{\omega}^{\perp} - \left(\dfrac{\mathfrak{t}}{2\kappa}\right)^2 \tilde{a}^{\perp}, & \text{(3.3b)} \\ \partial_t a = (\lambda - \nu)\,\bar{\omega}^{\perp} + \dfrac{\mathfrak{t}}{2\kappa} a^{\perp}, & \text{(3.3c)} \\ \partial_t \bar{J} = \dfrac{\tau}{2\kappa}\left[R, \bar{J}\right], \text{ and} & \text{(3.3d)} \\ \partial_t J_{33} = 0 & \text{(3.3e)} \end{cases}$$

subject to $\nabla \cdot u = 0$, where R is the 2-by-2 matrix given by $R = e_2 \otimes e_1 - e_1 \otimes e_2$. In particular, if we write $Y = (u, \omega, a)$ and $\bar{D} = I_3 \oplus J_{eq} \oplus I_2$ then (3.3a), (3.3b), and (3.3c) can be written as $\partial_t \bar{D} Y = \widetilde{\mathcal{M}} Y + \Lambda(p)$ subject to $\nabla \cdot u = 0$ for an appropriate operator $\widetilde{\mathcal{M}}$. In particular, since $\widetilde{\mathcal{M}}$ commutes with the application of the Leray projector to u it suffices to study $\partial_t \bar{D} Y = \widetilde{\mathcal{M}}\,\bar{\mathbb{P}}\,Y$, where $\bar{\mathbb{P}} := \mathbb{P}_L \oplus I_3 \oplus I_2$ for $\mathbb{P}_L$ denoting the Leray projector. Recall that the Leray projector is the projection onto divergence-free vector fields, which on the 3-torus can be written explicitly as $\mathbb{P}_L = -\nabla \times \Delta^{-1}\,\nabla\times$ (see Lemma 6.26).

So finally, for $\mathcal{B} := \bar{D}^{-1}\widetilde{\mathcal{M}}\,\bar{\mathbb{P}}$ we have that $\mathcal{L} := D^{-1}\widetilde{\mathcal{L}}\,\mathbb{P}$, where $\mathbb{P} := \mathbb{P}_L \oplus I_3 \oplus I_{\mathrm{Mat}(3)}$, can be written as $\mathcal{L} = \mathcal{B} \oplus \frac{\tau}{2\kappa}\,[R, \cdot\,] \oplus 0$. Note that using this notation we may write the linearized problem (3.2), after Leray projection, as

$$\partial_t X = \mathcal{L} X. \tag{3.4}$$

This is a particularly convenient formulation since it is amenable to attack via semigroup theory.

What matters for the purpose of the spectral analysis carried out in the following section is that the equations governing the non-trivial dynamics of the problem can be written as $\partial_t Y = \mathcal{B} Y$. The punchline is that it suffices to study the spectrum of $\mathcal{B}$, which is precisely what we do in Section 3.2 below.

3.2. Spectral analysis. In this subsection we study the spectrum of the operator $\mathcal{B}$ introduced in the preceding section. Since our domain is the torus it is natural to consider the symbol $\hat{\mathcal{B}}$ of this operator, which gives a matrix in $\mathbb{C}^{8\times 8}$ for each wavenumber $k \in \mathbb{Z}^3$. However, it will be more convenient for us to parameterize these with a continuous wavenumber $k \in \mathbb{R}^3$; for each such k we define $\hat{\mathcal{B}}_k \in \mathbb{C}^{8\times 8}$ according to

$$\hat{\mathcal{B}}_k := \begin{pmatrix} -\left(\mu + \frac{\kappa}{2}\right)|k|^2 P_{\perp}(k) & i\kappa k\times & 0 \\ J_{eq}^{-1}\left(i\kappa k\times\right) P_{\perp}(k) & -2\kappa J_{eq}^{-1} - \tilde{\alpha}|k|^2 J_{eq}^{-1} P_{\parallel}(k) - \tilde{\gamma}|k|^2 J_{eq}^{-1} P_{\perp}(k) - \left(1 - \frac{\nu}{\lambda}\right)\frac{\tau}{2\kappa} R_{33} & -\frac{1}{\lambda}\left(\frac{\tau}{2\kappa}\right)^2 R_{32} \\ 0 & (\lambda - \nu)\,R_{23} & \frac{\tau}{2\kappa} R_{22} \end{pmatrix}, \tag{3.5}$$

where $P_{\parallel}$ and $P_{\perp}$ are as defined in Section 2.6, and

$$R_{22} = R = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}, \; R_{23} = \begin{pmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \end{pmatrix}, \; R_{32} = \begin{pmatrix} 0 & -1 \\ 1 & 0 \\ 0 & 0 \end{pmatrix}, \text{ and } R_{33} = \begin{pmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$

Note here that we have abused notation by writing $i\kappa k\times$ as a place-holder to indicate the matrix corresponding to the linear map $z \mapsto i\kappa k \times z$.

It is somewhat tricky to extract useful spectral information from $\hat{\mathcal{B}}_k$ directly. Instead, we introduce a sort of similarity transformation $M_k := Q_k \hat{\mathcal{B}}_k \bar{Q}_k$ in such a way that M_k is a real matrix, i.e. $M_k \in \mathbb{R}^{8\times 8}$ for each $k \in \mathbb{R}^3$, which carries the spectral information of $\hat{\mathcal{B}}_k$. Here the matrices $Q_k, \bar{Q}_k \in \mathbb{C}^{8\times 8}$ are defined by

$$Q_k := T(k) \oplus J_{eq}^{1/2} \oplus s R_{22} \text{ and } \bar{Q}_k := T(k) \oplus J_{eq}^{-1/2} \oplus \left(-s^{-1}\right) R_{22},$$

where $T(k) := \frac{ik\times}{|k|}$ if $k \neq 0$, $T(0) := 0$, and $s := \frac{-1}{\sqrt{\lambda - \nu}}\frac{\mathfrak{t}}{2\kappa}$. Unfortunately, Q_k and $\bar{Q}_k$ are not quite invertible, so this isn't exactly a similarity transformation. When $k \neq 0$, this is due to the fact that $(k, 0, 0)$ belongs to the kernels of both operators, a fact that is ultimately related to the divergence-free condition for u, which reads $k \cdot \hat{u}_k = 0$ on the Fourier side. In principle we could remove the kernel and restore invertibility, but the resulting 7-by-7 matrices are less convenient to work with. As such, we will stick with the 8-by-8 setup and find a work-around for the invertibility issue. Ultimately we will prove in Propositions 3.10 and 3.11 that we can gain good spectral information about M_k, and it will follow from Definition 3.1 and Lemmas 3.2 and 6.8 that the spectrum of $\hat{\mathcal{B}}_k$ coincides with that of M_k. Note that for all these k-dependent matrices we will write equivalently M_k or $M(k)$.

An important observation is that the matrix $M_k \in \mathbb{R}^{8\times 8}$ may be decomposed into its symmetric part $S_k \in \mathbb{R}^{8\times 8}$ and its antisymmetric part $A \in \mathbb{R}^{8\times 8}$ such that A is *independent of* k. More precisely

$$S_k = \begin{pmatrix} -\left(\mu + \frac{\kappa}{2}\right)|k|^2 P_\perp(k) & \kappa|k|P_\perp(k)\, J_{eq}^{-1/2} & 0 \\ \kappa|k|J_{eq}^{-1/2}P_\perp(k) & -2\kappa J_{eq}^{-1} - \tilde{\alpha}|k|^2 J_{eq}^{-1/2}P_\parallel(k)\, J_{eq}^{-1/2} - \tilde{\gamma}|k|^2 J_{eq}^{-1/2}P_\perp(k)\, J_{eq}^{-1/2} & \phi I_{32} \\ 0 & \phi I_{23} & 0 \end{pmatrix} \quad (3.6)$$

and

$$A = 0 \oplus cR_{33} \oplus dR_{22}, \quad (3.7)$$

where

$$\phi = \sqrt{1-\frac{\nu}{\lambda}}\,\frac{\mathfrak{t}}{2\kappa},\ c = \left(\frac{\nu}{\lambda}-1\right)\frac{\mathfrak{t}}{2\kappa},\ \text{and } d = \frac{\mathfrak{t}}{2\kappa}. \quad (3.8)$$

Note that M_k is written out explicitly in all its gory details in Appendix 6.4.

We now turn to the issue of proving that the spectra of $\hat{\mathcal{B}}_k$ and M_k coincide. To do this we will need to use the notion of linear maps acting on quotient spaces. Here we quotient out by the spaces V_k defined as $V_0 := \operatorname{span}\left\{(v,0,0) \mid v \in \mathbb{R}^3\right\}$ as well as, for any nonzero $k \in \mathbb{R}^3$, $V_k := \operatorname{span}(k,0,0)$.

Definition 3.1 (Linear maps acting on quotient spaces). Let $A \in \mathbb{C}^{n\times n}$ and let V be a subspace of $\mathbb{C}^n$. We say that A *acts on* $\mathbb{C}^n/V$ if and only if $\ker A = V$ and $\operatorname{im} A \subseteq V^\perp$, where $V^\perp$ is the orthogonal complement relative to the standard Hermitian structure on $\mathbb{C}^n$.

We refer to Lemma 6.8 for the key property of linear maps acting on quotient spaces which we will use in the sequel, namely conditions under which two matrix representations of such maps are equivalent, even when the 'change of basis' matrices involved are not invertible. We now prove that the matrices we are dealing with here do satisfy the hypotheses of Lemma 6.8.

Lemma 3.2. *For any $k \in \mathbb{R}^3$, $\hat{\mathcal{B}}_k$, Q_k, and $\bar{Q}_k$ act on $\mathbb{C}^8/V_k$ and $Q_k\bar{Q}_k = \bar{Q}_kQ_k = \operatorname{proj}_{V_k^\perp}$.*

Proof. First we consider $\hat{\mathcal{B}}_k$ for $k \neq 0$. Since $\hat{\mathcal{B}}_k^\dagger(k,0,0) = \hat{\mathcal{B}}_k(k,0,0) = 0$, where $\dagger$ denotes the conjugate transpose, we know that $\operatorname{im}\hat{\mathcal{B}}_k \subseteq V_k$ and that $V_k \subseteq \ker\hat{\mathcal{B}}_k$, so we only have to show that $\ker\hat{\mathcal{B}}_k \subseteq V_k$. Let $y = (v,\theta,b) \in \ker\hat{\mathcal{B}}_k$. The third row of (3.5) tells us that $b = \frac{2\kappa(\lambda-\nu)}{\mathfrak{t}}\bar{\theta}$ and hence

$$0 = \bar{D}\hat{\mathcal{B}}_k y\cdot y = -\mu|k|^2|v_\perp|^2 - 2\kappa\left|\frac{1}{2}ik\times v - \theta\right|^2 - \tilde{\alpha}|k|^2\left|\theta_\parallel\right|^2 - \tilde{\gamma}|k|^2|\theta_\perp|^2.$$

Therefore $\theta = v_\perp = 0$, and hence also $b = 0$, such that indeed $y = \left(v_\parallel,0,0\right) \in V_k$. So indeed $\hat{\mathcal{B}}_k$ acts on $\mathbb{C}^8/V_k$.

Now we consider $\hat{\mathcal{B}}_0$, proceeding essentially as we did above for the case $k \neq 0$. Since $\hat{\mathcal{B}}_0^\dagger(v,0,0) = \hat{\mathcal{B}}_0(v,0,0) = 0$ for any $v \in \mathbb{R}^3$ it follows that $\operatorname{im}\hat{B}_0 \subseteq V_0$ and that $V_0 \subseteq \ker\hat{\mathcal{B}}_0$. Now let $y = (v,\theta,b) \in \ker\hat{\mathcal{B}}_0$ and observe that, as above, $b = \frac{2\kappa(\lambda-\nu)}{\mathfrak{t}}\bar{\theta}$ and that hence $0 = \bar{D}\hat{\mathcal{B}}_0 y\cdot y = -2\kappa|\theta|^2$. Therefore $\theta = 0$ and $b = 0$ such that indeed $y = (v,0,0) \in V_0$. So $\ker\hat{\mathcal{B}}_0 \subseteq V_0$ and thus indeed $\hat{\mathcal{B}}_0$ acts on $\mathbb{C}^8/V_0$.

We now turn our attention to Q_k and $\bar{Q}_k$. Since $(k,0,0)^T Q_k = (k,0,0)^T\bar{Q}_k = (k\cdot T(k)) \oplus 0 \oplus 0 = 0$ for any nonzero $k \in \mathbb{R}^3$ and since $(v,0,0)\cdot Q_0 = (v,0,0)\cdot\bar{Q}_0 = v\cdot T(0) \oplus 0 \oplus 0 = 0$, we may deduce that $\operatorname{im} Q_k$, $\operatorname{im}\bar{Q}_k \subseteq V_k^\perp$ for *all* $k \in \mathbb{Z}^3$. Now observe that, since $J_{eq}^{1/2}$ and R_{22} are invertible, we deduce that $\ker Q_k = \ker\bar{Q}_k = (\ker T(k)) \oplus 0 \oplus 0$. Therefore, since $\ker T(k) = \operatorname{span}\{k\}$ when k is nonzero and since $\ker T(0) = \mathbb{R}^3$, we have that indeed $\ker Q_k = \ker\bar{Q}_k = V_k$ for *all* $k \in \mathbb{Z}^3$, i.e. Q_k and $\bar{Q}_k$ act on $\mathbb{C}^8/V_k$ for all $k \in \mathbb{Z}^3$.

Finally observe that, since $R_{22}^2 = -I_2$, it follows that $Q_k\bar{Q}_k = \bar{Q}_kQ_k = T(k)^2 \oplus I_3 \oplus I_2$, where $T(0)^2 = 0$ and $T(k)^2 = \frac{(ik\times)^2}{|k|^2} = \operatorname{proj}_{\operatorname{span}\{k\}^\perp}$ for $k \neq 0$. Note that we have used the ε-δ identity $\varepsilon_{aij}\varepsilon_{akl} = \delta_{ik}\delta_{jl} - \delta_{il}\delta_{jk}$ to deduce that $(k\times)^2 = -|k|^2\operatorname{proj}_{\operatorname{span}\{k\}^\perp}$. So indeed $Q_k\bar{Q}_k = \bar{Q}_kQ_k = \operatorname{proj}_{V_k^\perp}$. $\square$

We now record how M_k behaves under transformations of the form $k \mapsto -k$ and $k = (\bar{k},k_3) \mapsto (\bar{H}\bar{k},k_3)$ for $\bar{H}$ an orthogonal map. This comes in handy when constructing the maximally unstable solution in Section 3.4.

Lemma 3.3 (Equivariance and invariance of M). *Let H be a* horizontal rotation, *i.e. $H \in \mathbb{R}^{3\times 3}$ such that $H = \bar{H} \oplus 1$ for some 2-by-2 orthogonal matrix $\bar{H}$. We call $\widetilde{H} := H \oplus H \oplus \bar{H}$ the* joint horizontal rotation *associated with H.*

(1) *M is equivariant under horizontal rotations, i.e. for any $k \in \mathbb{R}^3$ and any horizontal rotation H, $M(Hk) = \widetilde{H} M(k) \widetilde{H}^T$ and*
(2) *M is even, i.e. for any $k \in \mathbb{R}^3$, $M(-k) = M(k)$.*

Proof. Note that $k \mapsto P_{\parallel}(k), P_{\perp}(k)$ are both even and equivariant under horizontal rotations, i.e., for any horizontal rotation H, $P_{\parallel}(Hk) = HP_{\parallel}(k) H^T$ and similarly for $P_{\perp}$, whilst $k \mapsto |k|$ is even and invariant under horizontal rotations. We can therefore write

$$S(k) = \begin{pmatrix} A(k) & B(k) & 0 \\ C(k) & -2\kappa J_{eq}^{-1} + D(k) & \phi I_{32} \\ 0 & \phi I_{23} & 0 \end{pmatrix}$$

for some A, B, C, D which are equivariant under horizontal rotations and even. It follows immediately that M is even. Now let H be a horizontal rotation. Since $\bar{H} I_{23} H = I_{23}$, $H I_{32} \bar{H} = I_{32}$, and since H commutes with J_{eq}^{-1} one may readily compute that $S(Hk) = \widetilde{H} S(k) \widetilde{H}^T$. Finally, since two-dimensional rotations (i.e. elements of $O(2)$) commute with one another, $A = \widetilde{H} A \widetilde{H}^T$ and so indeed M is equivariant under horizontal rotations. $\square$

We now obtain some fairly crude bounds on the spectrum of M_k in Lemmas 3.5, 3.6, and 3.7. These bounds are nonetheless essential in the proofs of Propositions 3.10 and 3.11. As a first step in obtaining these bounds we identify the quadratic form associated with S_k, the symmetric part of M_k, in Lemma 3.4.

Lemma 3.4 (Quadratic form associated with S_k). *For any $y = (v, \theta, b) \in \mathbb{R}^3 \times \mathbb{R}^3 \times \mathbb{R}^2 = \mathbb{R}^8$ and any $k \in \mathbb{R}^3$,*

$$S(k)\, y \cdot y = -\mu |k|^2 |v_\perp|^2 - 2\kappa \left| \frac{1}{2} |k| v_\perp - J_{eq}^{-1/2} \theta \right|^2 - \tilde{\alpha} |k|^2 \left| \left(J_{eq}^{-1/2} \theta \right)_{\parallel} \right|^2 - \tilde{\gamma} |k|^2 \left| \left(J_{eq}^{-1/2} \theta \right)_{\perp} \right|^2 + 2\phi \bar{\theta} \cdot b.$$

where, for any $w \in \mathbb{R}^3$, $w_\parallel := \operatorname{proj}_k w$ and $w_\perp := (I - \operatorname{proj}_k) w$, and ϕ is as in (3.8).

Proof. This follows immediately from the definition of S in (3.6). $\square$

We now use Lemma 3.4 to obtain upper bounds on the eigenvalues of S.

Lemma 3.5 (Spectral bounds on S_k). *For any $k \in \mathbb{R}^3$, it holds that* $\max \sigma(S_k) \leqslant \min \left(\phi, \frac{C_\sigma}{|k|^2} \right)$, *where $C_\sigma := \frac{\phi^2 \lambda}{\min(\tilde{\alpha}, \tilde{\gamma})}$ and ϕ is as in (3.8).*

Proof. Let $k \in \mathbb{R}^3$ and let $y = (v, \theta, b) \in \mathbb{R}^3 \times \mathbb{R}^3 \times \mathbb{R}^2$. By Lemma 3.4

$$S(k)\, y \cdot y \leqslant -\tilde{\alpha} |k|^2 \left| \left(J_{eq}^{-1/2} \theta \right)_{\parallel} \right|^2 - \tilde{\gamma} |k|^2 \left| \left(J_{eq}^{-1/2} \theta \right)_{\perp} \right|^2 + 2\phi \bar{\theta} \cdot b \tag{3.9}$$

from which it follows that $S(k)\, y \cdot y \leqslant \phi \left(|\bar{\theta}|^2 + |b|^2 \right)$ and hence that $\max \sigma(S_k) \leqslant \phi$. Now observe that

$$-\tilde{\alpha} |k|^2 \left| \left(J_{eq}^{-1/2} \theta \right)_{\parallel} \right|^2 - \tilde{\gamma} |k|^2 \left| \left(J_{eq}^{-1/2} \theta \right)_{\perp} \right|^2 \leqslant -\min(\tilde{\alpha}, \tilde{\gamma}) |k|^2 \left| J_{eq}^{-1/2} \theta \right|^2 \leqslant -\frac{1}{\lambda} \min(\tilde{\alpha}, \tilde{\gamma}) |k|^2 |\bar{\theta}|^2. \tag{3.10}$$

Combining (3.9) and (3.10) tells us that, for $k \neq 0$,

$$S(k)\, y \cdot y \leqslant -\frac{\phi^2 |k|^2}{C_\sigma} |\bar{\theta}|^2 + 2\phi \bar{\theta} \cdot b = -\frac{\phi^2 |k|^2}{C_\sigma} \left| \bar{\theta} - \frac{C_\sigma}{\phi |k|^2} b \right|^2 + \frac{C_\sigma}{|k|^2} |b|^2 \leqslant \frac{C_\sigma}{|k|^2} |y|^2$$

from which we deduce that $\max \sigma(S_k) \leqslant \frac{C_\sigma}{|k|^2}$. $\square$

The bounds on S from Lemma 3.5 coupled with elementary considerations from linear algebra allow us to deduce bounds on the real parts of the eigenvalues of M_k.

Lemma 3.6 (Bounds on the real parts of eigenvalues of M_k). *For any $k \in \mathbb{R}^3$, and with ϕ as in (3.8), it holds that* $\max \operatorname{Re} \sigma(M_k) \leqslant \phi$.

PROOF. This follows immediately from Lemmas 3.5 and 6.9. ◻

To conclude this batch of spectral estimates we obtain bounds on the imaginary parts of the eigenvalues of M_k as a corollary of the Gershgorin disk theorem (Theorem 6.10).

LEMMA 3.7 (Bounds on the imaginary parts of eigenvalues of M_k). *For any $k \in \mathbb{R}^3$ it holds that* $\max|\operatorname{Im}\sigma(M_k)| \leqslant \frac{\sqrt{7}\mathfrak{t}}{2\kappa}$.

PROOF. This follows from Corollary 6.11 since

$$||A||_2^2 = 2\left(c^2+d^2\right) = \left(\frac{\mathfrak{t}}{2\kappa}\right)^2 \left(1-2\nu\left(\nu-2\lambda\right)\right) \leqslant \left(\frac{\mathfrak{t}}{2\kappa}\right)^2.$$

◻

We now record some useful facts about the characteristic polynomial p of M_k. Computing p was done by using a computer algebra system, and we thus record M_k in Appendix 6.4 in a form which can readily be used for computer-assisted algebraic manipulations.

Upon computing p we observe that it is a polynomial in k of degree 10 and that it only depends on even powers of $|\bar{k}|$ and k_3. Therefore we may write

$$p\left(x,k\right) = \sum_{q=0}^{5} r_q\left(x,|\bar{k}|,k_3\right) \tag{3.11}$$

where each r_q is a polynomial in $\left(x,|\bar{k}|,k_3\right)$ which is *homogeneous of degree* $2q$ *in* $\left(|\bar{k}|,k_3\right)$. In particular:

$$r_5\left(x,|\bar{k}|,k_3\right) = C_0 x\left(x^2+d^2\right)|k|^{10} \text{ and } r_4\left(x,|\bar{k}|,k_3\right) = |k|^6\left(t_1\left(x\right)|\bar{k}|^2 + t_2\left(x\right)k_3^2\right) \tag{3.12}$$

where

$$t_i\left(x\right) = x^2\left(-C_{i,0} + C_{i,1}x + C_{i,2}\left(x^2+d^2\right)\right) \tag{3.13}$$

and

$$C_0 = \left(\mu+\kappa/2\right)^2\left(\alpha+4\beta/3\right)\left(\beta+\gamma\right)^2/\left(\nu\lambda^2\right),$$
$$C_{1,0} = \left(\alpha+5\beta/3+\gamma\right)\left(\beta+\gamma\right)\left(\mu+\kappa/2\right)^2\phi/\left(\nu\lambda\right),\ C_{2,0} = 2\left(\alpha+4\beta/3\right)\left(\beta+\gamma\right)\left(\mu+\kappa/2\right)\phi/\left(\nu\lambda\right),$$
$$C_{1,1} = C_{2,1} = 2\kappa\left(\mu+\kappa/2\right)\left(\beta+\gamma\right)\left(2\mu\left(\alpha+4\beta/3\right) + \left(\mu+\kappa/2\right)\left(\beta+\gamma\right)\right)/\left(\nu\lambda^2\right),$$
$$C_{1,2} = \left(\mu+\kappa/2\right)\left(\beta+\gamma\right)\left(2\left(\alpha+4\beta/3\right)\left(\beta+\gamma\right) + \left(\mu+\kappa/2\right)\left(\left(\alpha+5\beta/3+\gamma\right)\lambda + \left(\alpha+4\beta/3\right)\nu\right)\right)/\left(\nu\lambda^2\right) \text{ and}$$
$$C_{2,2} = 2\left(\mu+\kappa/2\right)\left(\beta+\gamma\right)\left(\left(\alpha+4\beta/3\right)\left(\beta+\gamma\right) + \left(\mu+\kappa/2\right)\left(\left(\alpha+4\beta/3\right)\lambda + \left(\beta+\gamma\right)\nu/2\right)\right)/\left(\nu\lambda^2\right).$$

The exact dependence of these constants on the various physical parameters is not of concern here, since all that matters is that all these constants are strictly positive, i.e. $C_0, C_{i,j} > 0$ for all i,j.

We now use Rouché's Theorem (c.f. Theorem 6.16) and our explicit expressions for the leading factors (with respect to $|k|$) of the characteristic polynomial p of M_k to control the number of eigenvalues remaining within bounded neighbourhoods of the origin as $|k|$ becomes large. This is stated precisely in Lemma 3.8 below, which is another ingredient of the proof of Proposition 3.10.

LEMMA 3.8 (Isolation of some eigenvalues of M for large wavenumbers). *For any $R > \frac{\mathfrak{t}}{2\kappa}$ there exist $K_I > 0$ such that for any $k \in \mathbb{R}^3$, if $|k| > K_I$ then there are precisely three eigenvalues of M_k in an open ball of radius R about the origin.*

PROOF. Let $k \in \mathbb{R}^3$ be nonzero, let $p\left(\cdot,k\right)$ denote the characteristic polynomial of M_k, and let us write $s := p - r_5$ for r_5 as in (3.12). The key observations are that r_5 has precisely three roots in B_R when $R > \frac{\mathfrak{t}}{2\kappa}$ and that s is lower-order in k than r_5. The result then follows from Rouché's Theorem since r_5 dominates s for large $|k|$.

More precisely, let $R > d = \frac{\mathfrak{t}}{2\kappa}$ and let $\tilde{r}_5\left(x\right) := C_0 x\left(x^2+d^2\right)$ for C_0 as in (3.12) such that $r_5\left(x,k\right) = \tilde{r}_5\left(x\right)|k|^{10}$. Since $\tilde{r}_5$ is a polynomial whose roots are away from ∂B_R, since $s\left(x,k\right)$ is a polynomial of degree 8 in k, and since ∂B_R is compact, it follows that $C_r := \inf_{\partial B_R}|\tilde{r}_5| > 0$ and that $C_s := \sup_{\substack{x\in\partial B_R\\ k\neq 0}} \frac{s(x,k)}{|k|^8} < \infty$. So pick $K_I := \sqrt{\frac{C_s}{C_r}}$ and observe that for any $k \in \mathbb{Z}^3$, if $|k| > K_I$ then, on ∂B_R,

$$|r_5\left(\cdot,k\right)| = |\tilde{r}_5||k|^{10} > C_r|k|^8 K_I^2 \geqslant \frac{C_r}{C_s}K_I^2|s\left(\cdot,k\right)| = |s\left(\cdot,k\right)|. \tag{3.14}$$

Since $r_5(\cdot,k)$ has three roots in B_R, namely 0 and $\pm\frac{\mathfrak{t}}{2\kappa}$, we may use (3.14) to deduce from Theorem 6.16 that $p(\cdot,k)$ has three roots in B_R. ☐

In Proposition 3.9 below we use the Implicit Function Theorem to identify the trajectories of some unstable eigenvalues of M_k when $|k|$ is large. In particular we will see in the proof of Proposition 3.10 that, combining this result with earlier results from this section, we may deduce that these eigenvalues are the *most* unstable eigenvalues of M_k for large k. Here we say that an eigenvalue is unstable when it has strictly positive real part.

PROPOSITION 3.9 (Trajectories of some eigenvalues of M for large wavenumbers). *There exists $K_T > 0$ and a function $z : \{k\in\mathbb{R}^3 : |k| > K_T\}\to\mathbb{C}$, which is continuously differentiable in the real sense (i.e. after identifying $\mathbb{C}$ with $\mathbb{R}^2$ in the canonical way), such that*

(1) *for every $k\in\mathbb{R}^3$, if $|k|>K_T$ then*
   (a) *$z(k)$ and $\overline{z}(k)$ are eigenvalues of $M(k)$ and*
   (b) *$\operatorname{Re} z(k) > 0$, and*
(2) *$z(k)\to\frac{i\mathfrak{t}}{2\kappa}$ as $|k|\to\infty$.*

PROOF. Recall that $d=\frac{\mathfrak{t}}{2\kappa}$ and let $p(\cdot,k)$ denote the characteristic polynomial of M_k. We proceed in three steps: first we define s to be essentially $|\varepsilon|^5 p\left(\cdot,\varepsilon^{-1/2}\right)$ (such that the study of s about zero is equivalent to the study of p about infinity) and verify that we may apply the Implicit Function Theorem to s about $(x,\varepsilon)\sim(id,0)$, second we deduce from explicit computations of p (namely (3.12)) that, for small nonzero ε, s has two roots with strictly positive real parts, and third we write $k\sim\varepsilon^{-1/2}$ to turn our result from step 2 about $\varepsilon\sim 0$ into a result about $k\sim\infty$ which allows us to conclude that, for large $|k|$, p has two roots with strictly positive real part.

Step 1: Recall (from (3.11) and the preceding discussion) that p only depends on $|\bar{k}|$ and k_3, so we may write $p(x,k)=\tilde{p}\left(x,|\bar{k}|,k_3\right)$. Now define, for any $x\in\mathbb{C}$ and any $\varepsilon=(\varepsilon_h,\varepsilon_v)\in\mathbb{R}^2_{>0}$, $s(x,\varepsilon) := |\varepsilon|_1^5\,\tilde{p}\left(x,\frac{\left(\sqrt{\varepsilon_h},\sqrt{\varepsilon_v}\right)}{|\varepsilon|_1}\right)$, where $|\cdot|_1$ denotes the l^1 norm. It follows from (3.11) that $s(x,\varepsilon)=\sum_{q=0}^5 u_q(x,\varepsilon)$ for $u_{5-q}(x,\varepsilon) := |\varepsilon|_1^5\, r_q\left(x,\frac{\sqrt{\varepsilon_h}}{|\varepsilon|_1},\frac{\sqrt{\varepsilon_h}}{|\varepsilon|_1}\right)$. Since the only dependence of r_q on k is through $\left(|\bar{k}|,k_3\right)$, i.e. since $r_q\left(x,|\bar{k}|,k_3\right)=\tilde{r}_q\left(x,|\bar{k}|^2,k_3^2\right)$ for some $\tilde{r}_q$, we may write $r_q\left(x,|\bar{k}|,k_3\right)=C_q(x)\bullet\left(|\bar{k}|^2,k_3^2\right)^{\otimes q}$ for some polynomial C_q. In particular, it follows that

$$u_2(x,\varepsilon)=C_3(x)\bullet\frac{\varepsilon^{\otimes 3}}{|\varepsilon|_1},\ u_3(x,\varepsilon)=|\varepsilon|_1\,C_2(x)\bullet\varepsilon^{\otimes 2},\ u_4(x,\varepsilon)=|\varepsilon|_1^3\,C_1(x)\,,\ \text{and}\ u_5(x,\varepsilon)=|\varepsilon|_1^5\,C_0(x)$$

such that, for $q\geqslant 2$, $u_q(x,0)=0$ and both $\partial_x u_q(x,0)=0$ and $\nabla_\varepsilon u_q(x,0)=0$. Moreover we may compute, using (3.12), that

$$u_0(x,\varepsilon)=C_0x\left(x^2+d^2\right)=:u_0(x)\ \text{and}\ u_1(x,\varepsilon)=\left(t_1(x),t_2(x)\right)\cdot\varepsilon=:\bar{u}_1(\varepsilon)\,. \tag{3.15}$$

So finally, for $v := s-(u_0+u_1)=\sum_{q=2}^5 u_q$, we have that $s(x,\varepsilon)=u_0(x)+\bar{u}_1(x)\cdot\varepsilon+v(x,\varepsilon)$ where $v(x,0)=0$ and both $\partial_x v(x,0)=0$ and $\nabla_\varepsilon v(x,0)=0$. In particular, note that $s(id,0)=u_0(id)=0$ and that $\partial_x s(id,0)=u_0'(id)=-2C_0d^2\neq 0$.

Step 2: By step 1 we may apply Theorem 6.17 to s about id to deduce that there exists a number $\xi>0$ and a function $w : B^+_{1,\xi}\to\mathbb{C}$ which is continuously differentiable in the real sense, where $B^+_{1,\xi}$ is the intersection of the first quadrant and the l^1-ball of radius ξ, i.e. $B^+_{1,\xi} := \{(\varepsilon_h,\varepsilon_v)\mid\varepsilon_h,\varepsilon_v>0\text{ and }\varepsilon_h+\varepsilon_v<\xi\}$, such that $w(0)=id$, $s(w(\varepsilon),\varepsilon)=0$ for every $\varepsilon\in B^+_{1,\xi}$, and $\nabla_\varepsilon w(0)=\frac{-\nabla_\varepsilon s(id,0)}{\partial_x s(id,0)}$. Moreover we may compute from (3.13) and (3.15) that $\nabla_\varepsilon w(0)=\frac{1}{2C_0}\begin{pmatrix}C_{1,0}+iC_{1,1}d\\ C_{2,0}+iC_{2,1}d\end{pmatrix}$, such that $\operatorname{Re}\nabla_\varepsilon w(0)\in\mathbb{R}^2_{>0}$. It follows that there exists $0<\sigma<\xi$ such that $\operatorname{Re} w(\varepsilon)>0$ for all $\varepsilon\in B^+_{1,\sigma}$.

Step 3: Pick $K_T := 1/\sqrt{\sigma}$ and define z via, for every $k\in\mathbb{R}^3$ such that $|k|>K_T$, $z(k) := w(\varepsilon(k))$ for $\varepsilon(k) := \frac{1}{|k|^4}\left(|\bar{k}|^2,k_3^2\right)$. Note that z is well-defined on $\{k\in\mathbb{R}^3 : |k|>K_T\}$ since, for every $k\in\mathbb{R}^3$, $|k|>K_T \iff |\varepsilon(k)|=1/|k|^2<\sigma$. Now observe that, for every $k\in\mathbb{R}^3$ such that $|k|>K_T$, $\tilde{p}(z(k),k)=\frac{1}{|\varepsilon|_1^5}s(w(\varepsilon(k)),\varepsilon(k))=0$, i.e. indeed $z(k)$ is a root of $p(\cdot,k)$ and hence an eigenvalue of M_k. Since M_k is a matrix with real entries, we may deduce that $\bar{z}(k)$ is also an eigenvalue of M_k. Moreover it follows from step 2 above that $\operatorname{Re} z(k) > 0$ for every $|k| > K_T$. Finally, note that since $w(0) = id$, since w is continuous, and since $\varepsilon(k)$ is continuous away from $k = 0$, we may conclude that $z(k) \to id$ as $k \to \infty$. $\square$

We now have all the ingredients in hand to prove one of the two key results of this section, namely Proposition 3.10. This result tells us that there exists a *most* unstable eigenvalue of M_k, i.e. an eigenvalue with largest strictly positive real part.

PROPOSITION 3.10 (Maximally unstable eigenvalues). *There exist $k_* \in \mathbb{Z}^3$ and $w_* \in \mathbb{C}$ with strictly positive real part such that*

(1) *w_* is an eigenvalue of $M(k_*)$ and*
(2) *for every $k \in \mathbb{Z}^3$ and every eigenvalue w of $M(k)$, $\operatorname{Re} w \leqslant \operatorname{Re} w_*$.*

We define $\eta_* := \operatorname{Re} w_*$.

PROOF. The key observations are that: (i) by combining Proposition 3.9 and Lemmas 3.6 and 3.7, we can show that for $|k|$ large enough, the eigenvalues whose trajectory can be obtained via the implicit function theorem in Proposition 3.9 are the most unstable eigenvalues (i.e those with the largest real part) and that (ii) by Proposition 3.9 we know that $\operatorname{Re} z(k) \to 0$ as $|k| \to \infty$. We prove the first observation in step 1 below, and in step 2 we use the first step and the second observation to conclude.

Step 1: We show that there exists $K_* > 0$ such that, for every $|k| > K_*$, $\operatorname{Re} z(k) = \max_{w \in \sigma(M(k))} \operatorname{Re} w$. Pick $R > \phi^2 + 7d^2$ and note that since $R > d = \frac{\mathfrak{t}}{2\kappa}$ we may pick $K_I = K_I(R)$ as in Lemma 3.8. Let $K_* := \max(K_I, K_T)$ for K_T as in Proposition 3.9, let H denote the half-slab $\left\{ w \in \mathbb{C} \,\middle|\, \operatorname{Re} z \leqslant \phi, |\operatorname{Im} z| \leqslant \sqrt{7} d \right\}$, and let $B_R \subseteq \mathbb{C}$ denote the open ball of radius R about the origin.

Let $k \in \mathbb{Z}^3$ such that $|k| > K_*$. By Lemmas 3.6 and 3.7 we know that all the eigenvalues of $M(k)$ are in H, and by Lemma 3.8 we know that exactly three eigenvalues of $M(k)$ are in $B_R \cap H$. Moreover, by Proposition 3.9 we know that the three eigenvalues of $M(k)$ in $B_R \cap H$ are precisely 0 (since $M(k)(k,0,0) = 0$), $z(k)$, and $z\bar{(k)}$, for z as in Proposition 3.9.

In particular, since $R > \phi^2 + 7d^2$ such that no points in the half-slab H have larger real parts than all points in $B_R \cap H$, it follows that indeed the eigenvalues of $M(k)$ with largest real part are $z(k)$ and $z\bar{(k)}$.

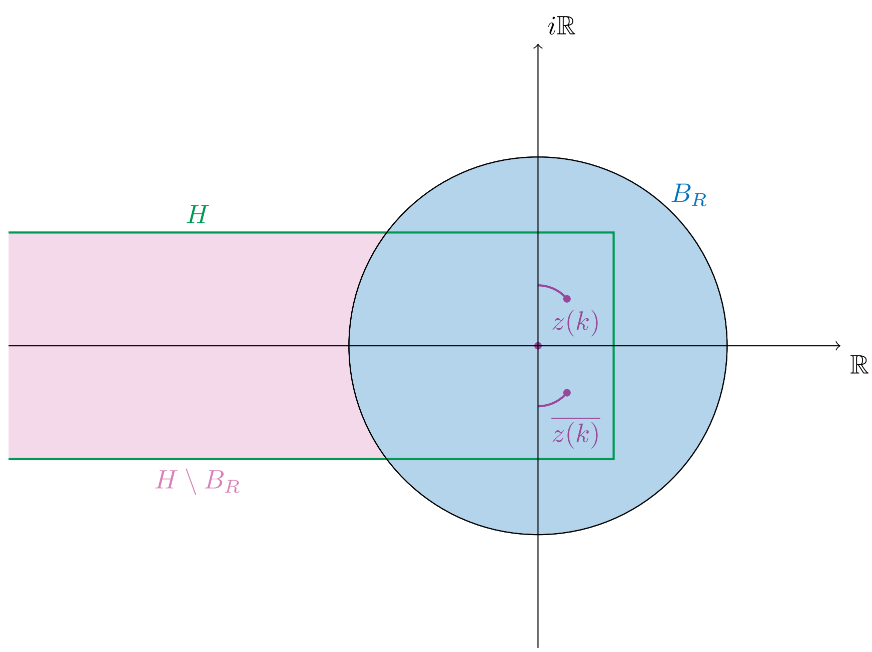


FIGURE 4. A pictorial summary of step 1 of the proof of Proposition 3.10.

Step 2: We want to show that the supremum

$$\sup_{k \in \mathbb{Z}^3} \max_{w \in \sigma(M(k))} \operatorname{Re} w$$

is strictly positive and attained. It is clearly strictly positive since for any $k \in \mathbb{Z}^3$ such that $|k| > K_T$ it follows from Proposition 3.9 that $z(k)$ is an eigenvalue of $M(k)$ with strictly positive real part. To see that this supremum is attained, we write for simplicity

$$s(E) := \sup_{k \in E} \max_{w \in \sigma(M(k))} \operatorname{Re} w$$

for any $E \subseteq \mathbb{Z}^3$. We thus want to show that $s(\mathbb{Z}^3)$ is attained. On one hand, by step 1, the supremum $s\left(\left\{k \in \mathbb{Z}^3 : |k| > K_*\right\}\right)$ is achieved. Indeed, we may pick the eigenvalue $z(k_{\text{crit}})$ of $M(k_{\text{crit}})$ corresponding to any k_{crit} such that $|k_{\text{crit}}|$ is equal to the smallest integer strictly larger than K_* which can be written as a sum of squares of integers. On the other hand the supremum $s\left(\left\{k \in \mathbb{Z}^3 : |k| \leqslant K_*\right\}\right)$ is attained since it is taken over a finite set. Since $\mathbb{Z}^3$ is the union of $\left\{k \in \mathbb{Z}^3 : |k| > K_*\right\}$ and $\left\{k \in \mathbb{Z}^3 : |k| \leqslant K_*\right\}$ we may conclude that the supremum $s(\mathbb{Z}^3)$ is attained. $\square$

We conclude this section with the second of its two key results: Proposition 3.11. This result is essential in the construction of the semigroup associated with the linearized operator. This construction is performed in Section 3.3 below.

PROPOSITION 3.11 (Uniform bound on the matrix exponentials). *Let η_* be as in Proposition 3.10. There exists $C_S > 0$ such that for every $k \in \mathbb{Z}^3$ and every $t > 0$, $|e^{tM_k}| \leqslant C_S\left(1 + t^8\right) e^{\eta_* t}$. As a consequence, for every $\varepsilon > 0$ there exists $C_S(\varepsilon) > 0$ such that for every $k \in \mathbb{Z}^3$ and every $t > 0$, $|e^{t\hat{\mathcal{B}}_k}| \leqslant C_S(\varepsilon)\, e^{(\eta_* + \varepsilon)t}$.*

PROOF. Naively, one may seek to use the bound from Corollary 6.14 to control e^{tM_k}. However, this bounds only holds up to a constant *dependent on* k. To circumvent this issue, we observe that alternatively one may bound e^{tM_k} using its symmetric part (as per Lemma 6.9). Coupling this observation with the fact that we have an upper bound which decays as $|k|^{-2}$ for the spectrum of S_k, namely Lemma 3.5, we see that for sufficiently large $|k|$ the exponential e^{tM_k} grows at most like $e^{\eta_* t}$. It thus suffices to use Corollary 6.14 for the finitely many modes with non-large $|k|$, in which case the dependence of the constant on k is harmless.

More precisely: let $K_S := \sqrt{\frac{C_\sigma}{\eta_*}}$ where C_σ is as in Lemma 3.5, write $C(k) := C(M_k)$ for $C(M)$ as in Corollary 6.14, and let $C_S := \max\left(1, \max_{|k| < K_S} C(k)\right) > 0$. Then, for every $k \in \mathbb{Z}^3$, if $|k| \geqslant K_S$ then $\frac{C_\sigma t}{|k|^2} \leqslant \frac{C\sigma t}{K_S^2} = \eta_* t$ and hence, by Lemmas 6.9 and 3.5, $\left|\left|e^{tM_k}\right|\right|_{\mathcal{L}(l^2,\, l^2)} \leqslant e^{\frac{C_\sigma t}{|k|^2}} \leqslant e^{\eta_* t}$, and if $|k| < K_S$ then by Corollary 6.14, the choice of C_S, and Proposition 3.10

$$\left|\left|e^{tM_k}\right|\right|_{\mathcal{L}(l^2,\, l^2)} \leqslant C(k)\left(1 + t^8\right) e^{(\max \operatorname{Re} \sigma(M_k))t} \leqslant C_S\left(1 + t^8\right) e^{\eta_* t}$$

from which the first part of the result follows. To obtain the second part we simply use the fact that polynomials of arbitrarily large degree can be controlled by exponentials of arbitrarily slow growth, i.e. the fact that for every $j \in \mathbb{N}$ and every $\varepsilon > 0$ there exists $C = C(j, \varepsilon) > 0$ such that, for every $t \geqslant 0$, $1 + t^j \leqslant C e^{\varepsilon t}$. $\square$

3.3. The semigroup. In this section we proceed in a standard fashion and use Proposition 3.11 to construct the semigroup associated with the linearized problem as recorded after Leray projection in (3.4).

PROPOSITION 3.12 (Semigroup for the linearization). *Let η_* be as in Proposition 3.10. For every $t \geqslant 0$ we define the operator $e^{t\mathcal{B}}$ on $L^2\left(\mathbb{T}^3, \mathbb{R}^8\right)$ via the Fourier multiplier $\left(e^{t\mathcal{B}}\right)^\wedge(k) := e^{t\hat{\mathcal{B}}_k}$ for every $k \in \mathbb{Z}^3$ and we define $e^{t\mathcal{L}}$ as $e^{t\mathcal{L}} := e^{t\mathcal{B}} \oplus e^{t\left[\bar{\Omega}_{eq},\, \cdot\right]} \oplus 1$, i.e. for every $\left(f, \bar{J}, J_{33}\right) \in L^2\left(\mathbb{T}^3, \mathbb{R}^8\right) \times L^2\left(\mathbb{T}^3, \mathbb{R}^{2\times 2}\right) \times L^2\left(\mathbb{T}^3, \mathbb{R}\right)$, $e^{t\mathcal{L}}\left(f, \bar{J}, J_{33}\right) := \left(e^{t\mathcal{B}} f, e^{t\left[\bar{\Omega}_{eq},\, \cdot\right]} \bar{J}, J_{33}\right)$.*

Then $\left(e^{t\mathcal{L}}\right)_{t \geqslant 0}$ is a semigroup on L^2 and for every $\varepsilon > 0$ it is an $(\eta_* + \varepsilon)$-contractive semigroup with domain containing $H^2\left(\mathbb{T}^3, \mathbb{R}^6\right) \times L^2\left(\mathbb{T}^3, \mathbb{R}^2\right) \times L^2\left(\mathbb{T}^3, \mathbb{R}^{2\times 2}\right) \times L^2\left(\mathbb{T}^3, \mathbb{R}\right) =: \mathfrak{D}$ and generator $\mathcal{L}$.

Moreover, for every $\varepsilon > 0$ there exists a constant $C_S(\varepsilon) > 0$ such that, for every $p, q, r \geqslant 0$ and every $t \geqslant 0$, $e^{t\mathcal{L}}$ is a bounded operator on $H^{p,q,r} := H^p\left(\mathbb{T}^3, \mathbb{R}^8\right) \times H^q\left(\mathbb{T}^3, \mathbb{R}^{2\times 2}\right) \times H^r\left(\mathbb{T}^3, \mathbb{R}\right)$ such that for any $\left(f, \bar{J}, J_{33}\right) \in H^{p,q,r}$, $\left|\left|e^{t\mathcal{L}}\left(f, \bar{J}, J_{33}\right)\right|\right|^2_{H^{p,q,r}} \leqslant C_S^2(\varepsilon)\, e^{2(\eta_* + \varepsilon)t} \left|\left|\left(f, \bar{J}, J_{33}\right)\right|\right|^2_{H^{p,q,r}}$, where

$$\left|\left|\left(f, \bar{J}, J_{33}\right)\right|\right|^2_{H^{p,q,r}} := ||f||^2_{H^p} + \left|\left|\bar{J}\right|\right|^2_{H^q} + ||J_{33}||^2_{H^r}.$$

Finally: the semigroup propagates incompressibility. More precisely: let

$$X_0 = \left(u_0, \omega_0, a_0, \bar{J}_0, (J_{33})_0\right) \in L^2\left(\mathbb{T}^3, \mathbb{R}^3\right) \times L^2\left(\mathbb{T}^3, \mathbb{R}^3\right) \times L^2\left(\mathbb{T}^3, \mathbb{R}^2\right) \times L^2\left(\mathbb{T}^3, \mathbb{R}^{2\times 2}\right) \times L^2\left(\mathbb{T}^3, \mathbb{R}\right)$$

and let $X(t,\cdot) = \left(u, \omega, a, \bar{J}, J_{33}\right)(t,\cdot) := e^{t\mathcal{L}} X_0$ *for all* $t > 0$. *If* u_0 *is incompressible, in a distributional sense, then* $u(t,\cdot)$ *is incompressible for all time* $t > 0$.

Proof. **Step 1:** We begin by constructing the semigroup $e^{t\mathcal{B}}$. Note that, in this proof, all matrix norms are norms in $\mathcal{L}\left(l^2, l^2\right)$. To construct this semigroup we will use Proposition 6.15 and must therefore verify that (i) for every $\varepsilon > 0$ there exists $C_S(\varepsilon) > 0$ such that for every $k \in \mathbb{Z}^3$ and every $t > 0$, $\left|\left|e^{t\hat{\mathcal{B}}_k}\right|\right| \leqslant C_S(\varepsilon)\, e^{(\eta_* + \varepsilon)t}$ and that (ii) there exists $C_D > 0$ such that for every $(v, \theta, b) \in \mathbb{R}^3 \times \mathbb{R}^3 \times \mathbb{R}^2$, $\left|\hat{\mathcal{B}}_k(v, \theta, b)\right| \leqslant C_D \left(\langle k\rangle^4 \left(|u|^2 + |\omega|^2\right) + |a|^2\right)$. Note that (ii) follows immediately from the expression provided for $\hat{\mathcal{B}}$ in (3.5). To obtain (i) we note that it follows from Lemmas 3.2 and 6.8 that

$$\hat{\mathcal{B}}_k^n = \left(\bar{Q}_k M_k Q_k\right)^n = \bar{Q}_k M_k^n Q_k \text{ for every } n \geqslant 1$$

whilst $\hat{B}_k^0 = \mathrm{id} = \mathrm{proj}_{V_k} + \mathrm{proj}_{V_k^\perp} = \mathrm{proj}_{V_k} + \bar{Q}_k M_k^0 Q_k$. Therefore

$$e^{t\hat{\mathcal{B}}_k} = \mathrm{proj}_{V_k} + \bar{Q}_k e^{tM_k} Q_k \tag{3.16}$$

where

$$\frac{1}{2}\left(||Q_k||^2 + ||\bar{Q}_k||^2\right) \leqslant \left|\left|\frac{ik\times}{|k|}\right|\right|^2 + \frac{1}{2}\left(\left|\left|J_{eq}^{1/2}\right|\right|^2 + \left|\left|J_{eq}^{-1/2}\right|\right|^2\right) + \frac{1}{2}\left(s + s^{-1}\right)||R||^2 \leqslant C_b \tag{3.17}$$

for some $C_b > 0$ *independent* of k. We may thus deduce from (3.16), (3.17), and Proposition 3.10 that (i) holds.

With (i) and (ii) in hand we apply Proposition 6.15 and obtain that $e^{t\mathcal{B}}$ is a semigroup on L^2 which is $(\eta_* + \varepsilon)$-contractive on all H^r spaces, for $r \geqslant 0$, with domain $H^2\left(\mathbb{T}^3, \mathbb{R}^3 \times \mathbb{R}^3\right) \times L^2\left(\mathbb{T}^3, \mathbb{R}^3\right)$ and generator $\mathcal{B}$.

Step 2: Now we construct the full semigroup $e^{t\mathcal{L}}$. First observe that, since $\left[\bar{\Omega}_{eq}, \cdot\right]$ is a finite-dimensional linear operator, $\left(e^{t\left[\bar{\Omega}_{eq}, \cdot\right]}\right)_{t\geqslant 0}$ is a semigroup on $\mathbb{R}^{2\times 2}$ and moreover

$$\text{the domain of } \left(e^{t\left[\bar{\Omega}_{eq}, \cdot\right]}\right)_{t\geqslant 0} \text{ is } \mathbb{R}^{2\times 2} \text{ and its generator is } \left[\bar{\Omega}_{eq}, \cdot\right]. \tag{3.18}$$

Moreover, Lemma 6.20 tells us that $\left[\bar{\Omega}_{eq}, \cdot\right]$ is antisymmetric, and thus it follows from Lemma 6.12 that $\left(e^{t\left[\bar{\Omega}_{eq}, \cdot\right]}\right)_{t\geqslant 0}$ is a contractive semigroup, i.e.

$$\left|\left|e^{t\left[\bar{\Omega}_{eq}, \cdot\right]}\right|\right|_{\mathcal{L}(l^2, l^2)} \leqslant 1. \tag{3.19}$$

From (3.19) and step 1 it follows that $e^{t\mathcal{L}} = e^{t\mathcal{N}} \oplus e^{t\left[\bar{\Omega}_{eq}, \cdot\right]} \oplus 1$ is a direct sum of semigroups which are, for every $\varepsilon > 0$, $(\eta_* + \varepsilon)$-contractive (since contractive semigroups are η-contractive for any $\eta > 0$ and since $1 = e^0$ is the trivial semigroup, which is contractive), and is hence $(\eta_* + \varepsilon)$-contractive itself. Moreover, it follows from the observation (3.18) and step 1 that the domain and generator of $e^{t\mathcal{L}}$ are as claimed. Finally the $H^{p,q,r}$ estimates follow immediately from (3.19) and the H^r estimates of step 1, upon observing that since, for each $t > 0$, $e^{t\left[\bar{\Omega}_{eq}, \cdot\right]}$ is a linear operator *independent of the spatial variable* x, it commutes with partial derivatives and with the Fourier transform.

Step 3: We now prove that incompressibility is propagated. Let us write $Y(t,\cdot) := (u, \omega, a)(t,\cdot)$. The key observation is that, as a consequence of Lemma 3.2, $\partial_t\left((k,0,0)\cdot\hat{Y}_k\right) = (k,0,0)\cdot\hat{\mathcal{B}}_k\hat{Y}_k = 0$ for every $k \in \mathbb{Z}^3$. In particular, if $\nabla \cdot_0 u = 0$ then indeed

$$(\nabla\cdot u)(t,\cdot) = \sum_{k\in\mathbb{Z}^3} (k,0,0)\cdot\hat{Y}_k(t) = \sum_{k\in\mathbb{Z}^3} (k,0,0)\cdot\hat{Y}_k(0) = \nabla\cdot u_0 = 0.$$

□

3.4. A maximally unstable solution. In this section we construct a maximally unstable solution of the linearized problem (3.4). Recall that (3.4) is obtained from the linearized problem by Leray projection. In particular, since (3.4) is invariant under the transformation $u \mapsto u + C$ for any constant C, the component corresponding to u in this maximally unstable solution will have average zero (this is as expected in light of the Galilean equivariance of the original system of equations). Note that, just as Proposition 3.12 is essentially a semigroup version of Proposition 3.11, Proposition 3.13 below is essentially a semigroup version of Proposition 3.10.

Proposition 3.13 (Maximally unstable solution). *Let η_* be as in Proposition 3.10. There is a smooth function $Y : [0,\infty) \times \mathbb{T}^3 \to \mathbb{R}^8$ such that $\partial_t Y = \mathcal{B} Y$ and $||Y(t,\cdot)||_{H^r(\mathbb{T}^3,\mathbb{R}^8)} = e^{\eta_* t} ||Y(0,\cdot)||_{H^r(\mathbb{T}^3,\mathbb{R}^8)}$ for every $t \geqslant 0$ and every $r \geqslant 0$. Moreover, if we write $Y = (u, \omega, a) \in \mathbb{R}^3 \times \mathbb{R}^3 \times \mathbb{R}^2$, then $\nabla \cdot u = 0$, and for every $t \geqslant 0$ and every $r \geqslant 0$*

$$
\begin{aligned}
||u(t,\cdot)||_{H^r(\mathbb{T}^3,\mathbb{R}^3)} &= e^{\eta_* t} ||u(0,\cdot)||_{H^r(\mathbb{T}^3,\mathbb{R}^3)}, \\
||\omega(t,\cdot)||_{H^r(\mathbb{T}^3,\mathbb{R}^3)} &= e^{\eta_* t} ||\omega(0,\cdot)||_{H^r(\mathbb{T}^3,\mathbb{R}^3)}, \quad \textit{and} \\
||a(t,\cdot)||_{H^r(\mathbb{T}^3,\mathbb{R}^3)} &= e^{\eta_* t} ||a(0,\cdot)||_{H^r(\mathbb{T}^3,\mathbb{R}^3)}.
\end{aligned}
$$

Proof. Let $k_* \in \mathbb{Z}^3$ and $w_* \in \mathbb{C}$ be as in Proposition 3.10 and recall that $\eta_* := \operatorname{Re} w_*$. It follows from Lemma 3.2 and Lemma 6.8 that, for any $k \in \mathbb{Z}^3$, $\hat{\mathcal{B}}_k$ and M_k are similar, so in particular w_* is an eigenvalue of $\hat{\mathcal{B}}_k$ and thus there exists $v_* \in \mathbb{C}^8$ such that $\hat{\mathcal{B}}(k_*) v_* = w_* v_*$. Now define, for every $t \geqslant 0$ and every $x \in \mathbb{T}^3$, $Y(t,x) := v_* e^{w_* t + i k_* \cdot x} + v_*^\dagger e^{w_*^\dagger t - i k_* \cdot x}$ where, for any complex number w, we denote its complex conjugate by $w^\dagger$. For a complex matrix A we will write, in this proof only, $A^\dagger$ to denote its entry-wise complex conjugate (and not its conjugate transpose).

Observe that $Y^\dagger = Y$ and hence Y is real-valued. Note that since $\hat{\mathcal{B}}_k = Q_k M_k \bar{Q}_k$ (which follows from Lemmas 3.2 and 6.8), since M_k has real entries and is even in k (i.e. $M_{-k} = M_k$), and since $Q(k)^\dagger = Q(-k)$ and $\bar{Q}(k)^\dagger = \bar{Q}(-k)$, we obtain that $\hat{B}(k)^\dagger = \hat{\mathcal{B}}(-k)$ and hence $\left(w_*^\dagger,\, v_*^\dagger\right)$ is an eigenvalue-eigenvector pair for $\hat{\mathcal{B}}(-k_*)$. Therefore

$$
\partial_t Y = w_* v_* e^{w_* t + i k_* \cdot x} + w_*^\dagger v_*^\dagger e^{w_*^\dagger t - i k_* \cdot x} = \hat{\mathcal{B}}(k_*) v_* e^{w_* t + i k_* \cdot x} + \hat{\mathcal{B}}(-k_*) v_*^\dagger e^{w_*^\dagger t - i k_* \cdot x} = \mathcal{B} Y. \tag{3.20}
$$

Now we argue that $u := (Y_1, Y_2, Y_3)$ is divergence-free. Observe that if $k_* = 0$ then Y is constant in the spatial variable $x \in \mathbb{T}^3$ and thus u is constant and hence divergence-free. Now consider the case $k_* \neq 0$. Note that we have proved in Lemma 3.2 that, for all $k \in \mathbb{Z}^3$, $\operatorname{im} \hat{\mathcal{B}}_k \subseteq V_k^\perp$ and hence $(k, 0, 0) \cdot v = 0$ for any eigenvector v of $\hat{\mathcal{B}}_k$. We may thus compute:

$$
\nabla \cdot u = \sum_{k \in \mathbb{Z}^3} k \cdot \hat{u}(k) = (k_*, 0, 0) \cdot \hat{Y}(k_*) + (-k_*, 0, 0) \cdot \hat{Y}(-k_*) = 0. \tag{3.21}
$$

Finally, observe that for any $j = 1, \ldots, 8$, $Y_j(t,x) = (v_*)_j e^{w_* t + i k_* \cdot x} + (v_*^\dagger)_j e^{w_*^\dagger t - i k_* \cdot x}$ and hence, proceeding as above yields

$$
||Y_j(t,\cdot)||_{H^r}^2 = \langle k_* \rangle^{2r} |(v_*)_j|^2 \, |e^{\operatorname{Re} w_* t}|^2 + \langle k_* \rangle^{2r} |(v_*^\dagger)_j|^2 \, |e^{\operatorname{Re} w_*^\dagger t}|^2 = 2 \langle k_* \rangle^{2r} |v_*| \, e^{2\eta_* t} = e^{2\eta_* t} \, ||Y_j(0,\cdot)||_{H^r}^2.
$$

We can thus conclude that, for $u = (Y_1, Y_2, Y_3)$, $\omega = (Y_4, Y_5, Y_6)$, and $a = (Y_7, Y_8)$,

$$
||u(t,\cdot)||_{H^r}^2 = e^{2\eta_* t} ||u(0,\cdot)||_{H^r}^2, \quad ||\omega(t,\cdot)||_{H^r}^2 = e^{2\eta_* t} ||\omega(0,\cdot)||_{H^r}^2, \quad \text{and} \quad ||a(t,\cdot)||_{H^r}^2 = e^{2\eta_* t} ||a(0,\cdot)||_{H^r}^2.
$$

□

4. Nonlinear energy estimates

In this section we perform the nonlinear energy estimates necessary to carry out the bootstrap instability argument in Section 5. First we record the precise form of the nonlinearities and introduce, in Definitions 4.1 and 4.2, notation used in the remainder of the chapter. In Section 4.1 we obtain bounds on the nonlinearity in L^2. We record the energy-dissipation relations satisfied by solutions of (1.1a)–(1.1d) and their derivatives in Section 4.2. In Section 4.3 we estimate the interaction terms appearing in the relations obtained in the preceding section. Finally we use the results of Sections 4.2 and 4.3 in Section 4.4 to obtain a chain of energy inequalities from which we deduce the key bootstrap energy inequality.

Writing the problem compactly using the same notation as that which was used in (3.2) and defining $Z := X - X_{eq}$ and $q := p - p_{eq}$ we may write the original problem (1.1a)–(1.1d) as

$$\partial_t DZ = \widetilde{\mathcal{L}}Z + \Lambda(q) + N(Z) \text{ subject to } \nabla \cdot u = 0. \tag{4.1}$$

For simplicity we will abuse notation in this section and write the components of the perturbative unknown Z as $Z = (u, \omega, J)$. This does conflict with the notation used in Section 3 for X. However confusion may be avoided by noting that all the unknowns appearing in this section are *perturbative*, i.e. (u, ω, J) will always denote the components of Z. We also abuse notation and, in this section, write $p = q$.

Using this notation we have that $N = (N_1, N_2, N_3)$ for

$$N_1(Z) = -(u \cdot \nabla) u,\ N_3(Z) = [\Omega, J] - (u \cdot \nabla) J, \tag{4.2}$$

and

$$\begin{aligned} N_2(Z) = &-J_{eq}(u \cdot \nabla)\omega - \left(I + JJ_{eq}^{-1}\right)^{-1}(\omega \times J\omega + \omega_{eq} \times J\omega + \omega \times J_{eq}\omega + \omega \times J\omega_{eq}) \\ &- JJ_{eq}^{-1}\left(I + JJ_{eq}^{-1}\right)^{-1}\big(\kappa\nabla \times u - 2\kappa\omega + (\tilde{\alpha} - \tilde{\gamma})\nabla(\nabla \cdot \omega) + \tilde{\gamma}\Delta\omega \\ &- \omega \times J_{eq}\omega_{eq} - \omega_{eq} \times J\omega_{eq} - \omega_{eq} \times J_{eq}\omega\big) \end{aligned} \tag{4.3}$$

Note that Z being a solution of (4.1) is equivalent to Z being a solution of

$$\partial_t Z = \mathcal{L}Z + \Lambda(p) + D^{-1}N(Z) \text{ subject to } \nabla \cdot u = 0, \tag{4.4}$$

for $\mathcal{L}$ as in (3.4). The fact that both of these formulations are equivalent is very handy since (4.1) is particularly convenient for energy estimates whilst semigroup theory may be readily applied to (4.4).

Definition 4.1. Let $\mathfrak{B} := \left\{ A \in \mathbb{R}^{n\times n} \,\middle|\, ||A||_{\text{op}} < 1 \right\}$ and define $m(A) := (I + A)^{-1}$ for any $A \in \mathfrak{B}$. Note that m is well-defined by Corollary 6.19.

Definition 4.2 (Small energy regime). Since $n = 3$ there exists $C_0 > 0$ such that $||J||_\infty \leqslant C_0 ||J||_{H^4}$ for every $J \in H^4\left(\mathbb{T}^3, \mathbb{R}^{3\times 3}\right)$. We define $\delta_0 := \min\left(\frac{1}{2}, \frac{1}{2C_0\left|\left|J_{eq}^{-1}\right|\right|_\infty}\right)$.

4.1. Estimating the nonlinearity. In this section we record some preliminary results in Lemmas 4.3 and 4.4 and then estimate the nonlinearity in L^2 in Proposition 4.5.

First we record for convenience some elementary consequences of the Sobolev embeddings. In particular Lemma 4.3 tells us that in the small energy regime Z, ∇Z, and $\nabla^2 Z$ are L^∞-multipliers, which simplifies many of the estimates below. It is precisely because the estimates are easier to perform when $\nabla^2 Z$ is in L^∞ that we have chosen to close the estimates in H^4.

Lemma 4.3. *Let $Z \in H^4\left(\mathbb{T}^3, \mathbb{R}^3 \times \mathbb{R}^{2\times 2} \times \mathbb{R}\right)$.*

(1) *There exists $C > 0$ independent of Z such that $||Z||_{L^\infty} + ||\nabla Z||_{L^\infty} + \left|\left|\nabla^2 Z\right|\right|_{L^\infty} \leqslant C||Z||_{H^4}$.*
(2) *For any polynomial p with no zeroth-order term there exists $C(p) > 0$ such that if $||Z||_{H^4} \leqslant 1$ then $p(||Z||_{H^4}) \leqslant C(p)||Z||_{H^4}$.*

Proof. (1) follows from the Sobolev embedding $H^2\left(\mathbb{T}^3\right) \hookrightarrow L^\infty\left(\mathbb{T}^3\right)$ and (2) is immediate. □

The result below ensures, when combined with Corollary 6.19, that the nonlinearities written in (4.2) and (4.3) are well-defined. Note that the only subtlety in ensuring that the nonlinearities are well-defined comes from the presence of $\left(I + JJ_{eq^{-1}}\right)^{-1} = m\left(JJ_{eq}^{-1}\right)$. This terms owes its appearance to our choice to write (1.1c) in a form such that the left-hand side is $J_{eq}\partial_t\omega$, and not $J\partial_t\omega$. The former is more convenient since it makes it possible to use semigroup theory.

Lemma 4.4. *Let δ_0 be as in the small energy regime (c.f. Definition 4.2). If $||Z||_{H^4} \leqslant \delta_0$ then $\left|\left|JJ_{eq}^{-1}\right|\right|_\infty \leqslant \frac{1}{2}$ and $\left|\left|m\left(JJ_{eq}^{-1}\right)\right|\right|_\infty \leqslant 2$.*

Proof. If $||Z||_{H^4} \leqslant \delta_0$ then $\left|\left|JJ_{eq}^{-1}\right|\right|_\infty \leqslant ||J||_\infty \left|\left|J_{eq}^{-1}\right|\right|_\infty \leqslant C_0||J||_{H^4}\left|\left|J_{eq}^{-1}\right|\right|_\infty \leqslant C_0\delta_0\left|\left|J_{eq}^{-1}\right|\right|_\infty \leqslant \frac{1}{2}$ and hence, by Corollary 6.19, $\left|\left|m\left(JJ_{eq}^{-1}\right)\right|\right|_\infty \leqslant \frac{1}{1 - \left|\left|JJ_{eq}^{-1}\right|\right|_\infty} \leqslant 2$. □

We now prove the main result of this section, namely the L^2 bound on the nonlinearity.

Proposition 4.5 (Estimate of the nonlinearity). *Let δ_0 be as in the small energy regime (c.f. Definition 4.2). There exists $C_N > 0$ such that if $||Z||_{H^4} \leqslant \delta_0$ then $||N(Z)||_{L^2} \leqslant C_N ||Z||^2_{H^2}$.*

Proof. Recall that $N = (N_1, N_2, N_3)$ is recorded in (4.2)–(4.3). In particular, one immediately obtains that $||N_1||_{L^2} + ||N_3||_{L^2} \lesssim ||Z||_{L^2} ||Z||_{H^1} \lesssim ||Z||^2_{H^2}$. Dealing with N_2 is only slightly more delicate. Considering $m\left(JJ_{eq}^{-1}\right)$ as a fixed L^∞ multiplier we see that all terms in N_2 are quadratic or cubic in Z (more precisely: the only cubic term is $-\left(I + JJ_{eq}^{-1}\right)^{-1}(\omega \times J\omega)$). We can thus use the generalized Hölder inequality as well as the Sobolev embeddings $H^1\left(\mathbb{T}^3\right) \hookrightarrow L^6\left(\mathbb{T}^3\right) \hookrightarrow L^p\left(\mathbb{T}^3\right)$ for all $p \in [1,6]$ and $H^2\left(\mathbb{T}^3\right) \hookrightarrow L^\infty\left(\mathbb{T}^3\right)$ to obtain that $||N_2||_{L^2} \lesssim ||Z||^2_{H^2} + ||Z||^3_{H^2} \lesssim (1+\delta_0)\,||Z||^2_{H^2}$. $\square$

Remark 4.6. The operator which must be estimated in the bootstrap instability argument is actually $\mathbb{P}N$ (and not merely N as is done in Proposition 4.5 above), where $\mathbb{P} = \mathbb{P}_L \oplus \mathrm{id} \oplus \mathrm{id}$ for $\mathbb{P}_L$ denoting the Leray projector. However, since $\hat{\mathbb{P}}_L(k) = \mathrm{proj}_{(\mathrm{span}\,k)^\perp} = I - \frac{k\otimes k}{|k|^2}$ for every $k \in \mathbb{Z}^3$, i.e. since $\mathbb{P}_L$ is a bounded Fourier multiplier, it follows that it is bounded on L^2.

4.2. The energy-dissipation identities. In this section we begin by recording the energy-dissipation relation and then remark on the coercivity of the dissipation.

Proposition 4.7 (The energy-dissipation relation). *If Z solves (4.1) then for any multi-index $\alpha \in \mathbb{N}^3$*

$$\frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}t}\left|\left|\sqrt{D}\left(\partial^\alpha Z\right)\right|\right|^2_{L^2} + \mathcal{D}\left(\partial^\alpha u, \partial^\alpha \omega\right) = B\left(\partial^\alpha \bar{\omega}, \partial^\alpha a\right) + \int_{\mathbb{T}^3} \partial^\alpha N(Z)\cdot \partial^\alpha Z$$

where

$$\mathcal{D}(u,\omega) := \int_{\mathbb{T}^3} \frac{\mu}{2}|\mathbb{D}u|^2 + 2\kappa\left|\frac{1}{2}\nabla\times u - \omega\right|^2 + \alpha|\nabla\cdot\omega|^2 + \frac{\beta}{2}|\mathbb{D}^0\omega|^2 + 2\gamma|\nabla\times\omega|^2$$

and

$$B(\bar{\omega}, a) := \left(2(\lambda-\nu) + \left(\frac{\mathfrak{t}}{2\kappa}\right)^2\right)\int_{\mathbb{T}^3}\bar{\omega}^\perp\cdot a.$$

Proof. To compute the energy-dissipation relation we take a derivative ∂^α of (4.1), multiply by Z, and integrate over the torus. Note that due to incompressibility $\int_{\mathbb{T}^3}\partial^\alpha\Lambda(p)\cdot\partial^\alpha Z = \int_{\mathbb{T}^3} -\left(\nabla\partial^\alpha p\right)\cdot\partial^\alpha u = 0$. Now we compute $\int_{\mathbb{T}^3}\widetilde{\mathcal{L}}Z\cdot Z$. Observe that for T and M denoting the stress and couple stress tensors, if we write $\tilde{T}$ for the trace-free part of T, i.e. $\tilde{T} = T + pI$, then we have that

$$\begin{aligned}\int_{\mathbb{T}^3}\left((\mu+\kappa/2)\,\Delta u + \kappa\nabla\times\omega\right)\cdot u + \int_{\mathbb{T}^3}\left(\kappa\nabla\times u - 2\kappa\omega + (\alpha+\beta/3-\gamma)\,\nabla\left(\nabla\cdot\omega\right) + (\beta+\gamma)\,\Delta\omega\right)\cdot\omega \\ = \int_{\mathbb{T}^3}\left(\nabla\cdot\tilde{T}\right)\cdot u + \left(2\,\mathrm{vec}\,\tilde{T} + \nabla\cdot M\right)\cdot\omega = -\int_{\mathbb{T}^3}\tilde{T} : (\nabla u - \Omega) + M : \nabla\omega = -\mathcal{D}(u,\omega)\,. \end{aligned}\tag{4.5}$$

Moreover, we may compute

$$\omega_{eq}\times J\omega_{eq} = \left(\frac{\mathfrak{t}}{2\kappa}\right)^2\tilde{a}^\perp,\ \ \omega_{eq}\times J_{eq}\omega = \frac{\lambda\mathfrak{t}}{2\kappa}\tilde{\omega}^\perp,\ \text{ and }\ [\Omega, J_{eq}] = (\lambda-\nu)\begin{pmatrix}0 & 0 & -\omega_2\\ 0 & 0 & \omega_1\\ \omega_2 & -\omega_1 & 0\end{pmatrix}$$

such that

$$\int_{\mathbb{T}^3} -\left(\omega\times J_{eq}\omega_{eq} + \omega_{eq}\times J\omega_{eq} + \omega_{eq}\times J_{eq}\omega\right)\cdot\omega + \int_{\mathbb{T}^3}\left([\Omega_{eq}, J] + [\Omega, J_{eq}]\right) : J = B(\bar{\omega}, a)\tag{4.6}$$

where we have used that $[\Omega_{eq}, J] : J = 0$ (c.f. Lemma 6.20). Combining (4.5) and (4.6), we obtain that $\int_{\mathbb{T}^3}\widetilde{\mathcal{L}}Z\cdot Z = -\mathcal{D}(u,\omega) + B(\bar{\omega}, a)$, and hence we may conclude that

$$\begin{aligned}\frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}t}\left|\left|\sqrt{D}\left(\partial^\alpha Z\right)\right|\right|^2_{L^2} &= \int_{\mathbb{T}^3}\partial_t\left(D\partial^\alpha Z\right)\cdot\partial^\alpha Z = \int_{\mathbb{T}^3}\widetilde{\mathcal{L}}\,\partial^\alpha Z\cdot\partial^\alpha Z + \int_{\mathbb{T}^3}\partial^\alpha N(Z)\cdot\partial^\alpha Z \\ &= -\mathcal{D}\left(\partial^\alpha u, \partial^\alpha\omega\right) + B\left(\partial^\alpha(\omega), \partial^\alpha a\right) + \int_{\mathbb{T}^3}\partial^\alpha N(Z)\cdot\partial^\alpha Z.\end{aligned}$$

$\square$

Besides the interaction term $\int_{\mathbb{T}^3} \partial^\alpha N(Z)\cdot\partial^\alpha Z$, the only term appearing in the energy-dissipation relation which does not have a sign is the term $B(\partial^\alpha\bar\omega, \partial^\alpha a)$. We refer to this term as the *unstable term* since, as detailed in Section 2.4 the instability originates from $\bar\omega$ and a. In Lemma 4.8 below we estimate this term in a manner which allows us to absorb a high-order contribution into the dissipation and leaves us with a lower-order term which is controlled by the energy.

LEMMA 4.8 (Bounds on the unstable term). *For any $\sigma > 0$ there exists $C_\sigma > 0$ such that for any sufficiently regular (ω, a) and any nonzero multi-index α,*

$$|B(\partial^\alpha\bar\omega, \partial^\alpha a)| \leqslant \sigma\left|\left|\partial^{\alpha+1}\bar\omega\right|\right|_{L^2}^2 + C_\sigma\left|\left|\partial^{\alpha-1}a\right|\right|_{L^2}^2$$

where we write $\alpha \pm 1 := \alpha \pm e_i$ for some i such that α_i nonzero.

PROOF. This follows immediately from integrating by parts and applying an ε-Cauchy inequality: if we define $C := 2(\lambda-\nu) + \left(\frac{\mathfrak{t}}{2\kappa}\right)^2$ then, for any $\varepsilon > 0$,

$$|B(\partial^\alpha\bar\omega, \partial^\alpha a)| = C\left|\int_{\mathbb{T}^3}\partial^{\alpha+1}\bar\omega^\perp\cdot\partial^{\alpha-1}a\right| \leqslant \varepsilon\left|\left|\partial^{\alpha+1}\bar\omega^\perp\right|\right|_{L^2}^2 + \frac{C^2}{4\varepsilon}\left|\left|\partial^{\alpha-1}a\right|\right|_{L^2}^2.$$

□

We now prove that the dissipation is coercive, since the velocity u has average zero.

LEMMA 4.9 (Coercivity of the dissipation over linear velocities of average zero). *There exists a constant $C_D > 0$ such that for every $(u,\omega) \in H^1\left(\mathbb{T}^3, \mathbb{R}^3\times\mathbb{R}^3\right)$, if $⨍ u = 0$ then $\mathcal{D}(u,\omega) \geqslant C_D\left(||u||_{H^1}^2 + ||\omega||_{H^1}^2\right)$.*

PROOF. Since u has average zero, it follows from Propositions 6.21 and 6.22 that

$$||u||_{H^1}^2 \lesssim ||\mathbb{D}u||_{L^2}^2 \lesssim D(u,\omega). \tag{4.7}$$

To see that the dissipation also controls the H^1 norm of ω we observe that, by (4.7),

$$||\omega||_{L^2}^2 \lesssim \int_{\mathbb{T}^3}\left|\frac{1}{2}\nabla\times u - \omega\right|^2 + \int_{\mathbb{T}^3}\left|\frac{1}{2}\nabla\times u\right|^2 \lesssim D(u,\omega) + ||u||_{H^1}^2 \lesssim D(u,\omega)$$

whilst, by Lemma 6.23, $||\nabla\omega||_{L^2}^2 = \int_{\mathbb{T}^3}|\nabla\cdot\omega|^2 + \int_{\mathbb{T}^3}|\nabla\times\omega|^2 \lesssim D(u,\omega)$, such that indeed $||\omega||_{H^1}^2 \lesssim D(u,\omega)$.

□

Recall that, due to the Galilean equivariance of (1.1a)–(1.1d) solutions of that system can be assumed without loss of generality to have an Eulerian velocity with average zero. Since $u_{eq} = 0$ it follows that we can assume that the perturbative velocity u has average zero as well, and hence the coercivity result proven in Lemma 4.9 applies.

4.3. Estimating the interactions. In this section we introduce notation which makes it easier to write down the Faà di Bruno formula for the chain rule, use this notation to record useful bounds on m (defined in Definition 4.1), and finally we estimate the interactions arising from the energy-dissipation relations satisfied by derivatives of solutions to (1.1a)–(1.1d) in Proposition 4.16.

DEFINITION 4.10 (Integer partitions and derivatives). Let $k \in \mathbb{N}$.

- Let $i_1 \geqslant i_2 \geqslant \cdots \geqslant i_l \geqslant 1$ be integers such that $k = i_1 + i_2 + \cdots + i_l$. The sequence $(i_1, i_2, \ldots, i_l)$ is called an *integer partition* of k and l is referred to as the *size* of that partition.
- For $1 \leqslant i \leqslant k$ we denote by $P_i(k)$ the set of integer partitions of k of size i, and by $P(k)$ the set of integer partitions of k. In particular note that $P(k) = \coprod_{i=1}^k P_i(k)$.
- Let $f : \mathbb{R}^n \to \mathbb{R}^m$ be k-times differentiable. For any $\pi = (i_1, \ldots, i_l) \in P(k)$ (where possibly $i_p = i_q$ for $p \neq q$) we define $\nabla^\pi f := \operatorname{Sym}\left(\nabla^{i_1}f\otimes\cdots\otimes\nabla^{i_l}f\right)$ where for any tensor T of rank r, $(\operatorname{Sym}T)_{j_1\ldots j_r} := \frac{1}{r!}\sum_{\sigma\in S_r} T_{j_{\sigma(1)}\ldots j_{\sigma(r)}}$.

EXAMPLE 4.11. Examples of integer partitions and derivatives indexed by integer partitions are

- $P_2(4) = \{(3,1), (2,2)\}$,
- $P(4) = \{(4), (3,1), (2,2), (2,1,1), (1,1,1,1)\}$, and
- $\nabla^{(2,1,1)}f = \operatorname{Sym}\left(\nabla^2 f\otimes\nabla f\otimes\nabla f\right) = \operatorname{Sym}\left(\nabla f\otimes\nabla^2 f\otimes\nabla f\right) = \operatorname{Sym}\left(\nabla f\otimes\nabla f\otimes\nabla^2 f\right)$.

REMARK 4.12. Derivatives indexed by integer partitions, denoted by $\nabla^\pi f$, are a convenient shorthand for terms appearing in the Faà di Bruno formula for derivatives of compositions. Their key property which we will use in estimates is that, for any integer partition $\pi = (i_1, \ldots, i_l)$, $|\nabla^\pi f| \leqslant \prod_{j=1}^l \left|\nabla^{i_j} f\right|$. For example $\left|\nabla^{(2,1,1)} f\right| \leqslant \left|\nabla^2 f\right| |\nabla f|^2$.

Having introduced notation for derivatives indexed by integer partitions we now use it to obtain bounds on derivatives of m in Lemma 4.13 below.

LEMMA 4.13 (Bounds on derivatives of m). *The function m from Definition 4.1 is smooth and moreover for every $k \in \mathbb{N}$ there exists $C_k > 0$ such that, for every $A \in \mathfrak{B}$, $\left|\nabla^k m(A)\right| \leqslant C_k |m(A)|^{k+1}$.*

PROOF. First we observe that it suffices to show that, for $\partial_{ij} m := \frac{\partial m(A)}{\partial A_{ij}}$,

$$\partial_{ij} m_{kl} = -m_{ki} m_{jl}, \tag{4.8}$$

To prove that (4.8) holds, note that for any smooth $A : (-1, 1) \to \mathfrak{B}$ (where $\mathfrak{B}$ is as in Definition 4.1), $\frac{\mathrm{d}}{\mathrm{d}t} m(A(t)) = -m(A(t)) \left(\frac{\mathrm{d}}{\mathrm{d}t} A(t)\right) m(A(t))$. Since we can pick A such that $A(0)$ and $\frac{\mathrm{d}}{\mathrm{d}t} A(0)$ are *arbitrarily* specified, it follows that for any $A_0 \in \mathfrak{B}$ and any $V \in \mathbb{R}^{n\times n}$, $\nabla m(A_0) V = -m(A_0) V m(A)$, i.e. indeed $\partial_{ij} m_{kl} = \nabla m_{kl}(e_i \otimes e_j) = -(m(e_i \otimes e_j) m)_{kl} = -m_{ki} m_{jl}$. $\square$

We now use the bounds on m we have just obtained to derive bounds on post-compositions with m.

LEMMA 4.14 (Bounds on derivatives of post-compositions with m). *Let $0 < \delta < 1$ and consider m from Definition 4.1, which is smooth by Lemma 4.13. For every $k \in \mathbb{N}$ there exists $C_{k,\delta} > 0$ such that for every smooth $A : \mathbb{T}^n \to \mathbb{R}^{n\times n}$, if $||A||_\infty < \delta$ then, for every $x \in \mathbb{T}^n$, $\left|\nabla^k (m(A))(x)\right| \leqslant C_{k,\delta} \sum_{\pi\in P(k)} |\nabla^\pi A(x)|$, where $P(k)$ and ∇^π are defined in Notation 4.10.*

PROOF. Note that since $||A||_\infty < \delta < 1$ it follows from Corollary 6.19 that $||m(A)||_\infty < \frac{1}{1-\delta}$. Therefore, by Proposition 6.24 and Lemma 4.13,

$$\begin{aligned}|\nabla^k (m(A))(x)| \leqslant C \sum_{i=1}^k |\nabla^i m(A(x))| \sum_{\pi\in P_i(k)} |\nabla^\pi A(x)| &\leqslant C \sum_{i=1}^k |m(A(x))|^{i+1} \sum_{\pi\in P_i(k)} |\nabla^\pi A(x)| \\ &\leqslant C_{k,\delta} \sum_{\pi\in P(k)} |\nabla^\pi A(x)|.\end{aligned}$$

$\square$

Below we specialize Lemma 4.14 to the only case which matters for us, namely the case of $m\left(JJ_{eq}^{-1}\right)$.

COROLLARY 4.15. *Let δ_0 be as in the small energy regime (c.f. Definition 4.2). For every $k \in \mathbb{N}$ there exists $C_k > 0$ such that if $||Z||_{H^4} \leqslant \delta_0$ then $\left|\nabla^k \left(m\left(JJ_{eq}^{-1}\right)\right)(x)\right| \leqslant C_k \sum_{\pi\in P(k)} |\nabla^\pi J(x)|$, for almost every $x \in \mathbb{T}^3$.*

PROOF. This follow immediately from combining Lemmas 4.4 and 4.14. $\square$

Having obtained good estimates on terms involving m which appear in the nonlinearity we are ready to estimate the interaction terms.

PROPOSITION 4.16 (Estimates of the interactions). *Let δ_0 be as in the small energy regime (c.f. Definition 4.2). For every $k = 0, 1, 2, 3, 4$ there exists $C_{I,k} > 0$ such that if $||Z||_{H^4} \leqslant \delta_0$ then*

$$\left|\int_{\mathbb{T}^3} N(Z)\cdot Z\right| \leqslant C_{I,0} ||Z||_{H^4} ||Z||_{L^2}^2$$

and

$$\sum_{|\alpha|=k} \left|\int_{\mathbb{T}^3} \partial^\alpha N(Z)\cdot\partial^\alpha Z\right| \leqslant C_{I,k} ||Z||_{H^4} \left(\sum_{i=1}^k \left|\left|\nabla^i Z\right|\right|_{L^2}^2 + \left|\left|\nabla^{k+1}(u,\omega)\right|\right|_{L^2}^2\right).$$

PROOF. The nonlinearities are all of one of three types, and so we write $N = N_{\mathrm{I}} + N_{\mathrm{II}} + N_{\mathrm{III}}$ for

$$\begin{aligned}
N_{\mathrm{I}} &:= -\left(\left(u\cdot\nabla\right)u,\ J_{eq}\left(u\cdot\nabla\right)\omega,\ \left(u\cdot\nabla\right)J\right),\\
N_{\mathrm{II}} &:= \Big(0,\ JJ_{eq}^{-1}m\left(JJ_{eq}^{-1}\right)\left(\omega\times J_{eq}\omega_{eq}+\omega_{eq}\times J\omega_{eq}+\omega_{eq}\times J_{eq}\omega+2\kappa\omega\right)\\
&\qquad - m\left(JJ_{eq}^{-1}\right)\left(\omega\times J\omega+\omega\times J\omega_{eq}+\omega\times J_{eq}\omega+\omega_{eq}\times J\omega\right),\left[\Omega,J\right]\Big),\ \text{and}\\
N_{\mathrm{III}} &:= \left(0,\ -JJ_{eq}^{-1}m\left(JJ_{eq}^{-1}\right)\left(\kappa\nabla\times u+\tilde{\alpha}\nabla\left(\nabla\cdot\omega\right)+\tilde{\gamma}\Delta\omega\right),0\right).
\end{aligned}$$

We first consider the case of α nonzero and so for $T \in \{\mathrm{I},\mathrm{II},\mathrm{III}\}$ and $i = 1,2,3,4$ we write $\mathcal{N}_{T,i} := \sum_{|\alpha|=i}\int_{\mathbb{T}^3}\partial^\alpha N_T\left(Z\right)\cdot\partial^\alpha Z$.

Estimating nonlinearities of type I is fairly straightforward. We expand out $\int_{\mathbb{T}^3}\partial^\alpha N_{\mathrm{I}}\left(Z\right)\cdot\partial^\alpha Z$ and use the generalized Hölder inequality, putting two factors in L^2 and putting the remaining factors in L^∞ (thanks to Lemma 4.3). For example, writing for simplicity $N_{\mathrm{I}}\left(Z\right) = \left(u\cdot\nabla\right)Z$ and considering the case where $\partial^\alpha = \partial_{ijkl}$, one of the terms that appears is $\int_{\mathbb{T}^3}\left(\partial_{ijk}u\cdot\nabla\right)\nabla_l Z\cdot\partial_{ijkl}Z$, and it can be estimated in the following way, which is typical of how nonlinear interactions of type I are handled:

$$\left|\int_{\mathbb{T}^3}\left(\partial_{ijk}u\cdot\nabla\right)\nabla_l Z\cdot\partial_{ijkl}Z\right| \leqslant \left|\left|\nabla^3 u\right|\right|_{L^2}\left|\left|\nabla^2 Z\right|\right|_\infty\left|\left|\nabla^4 Z\right|\right|_{L^2} \lesssim \left|\left|Z\right|\right|_{H^4}\left(\left|\left|\nabla^3 Z\right|\right|_{L^2}^2+\left|\left|\nabla^4 Z\right|\right|_{L^2}^2\right).$$

The only subtlety for these nonlinear terms is the fact that when ∂^α hits ∇Z in $\left(u\cdot\nabla\right)Z$, the interaction vanishes due to the incompressibility constraint. Indeed, for any multi-index α,

$$\int_{\mathbb{T}^3}\left(u\cdot\nabla\right)\partial^\alpha Z\cdot\partial^\alpha Z = -\frac{1}{2}\int_{\mathbb{T}^3}\left(\nabla\cdot u\right)\left|\partial^\alpha Z\right|^2 = 0.$$

This cancellation is essential since we have no dissipative control of J and hence we would not be able to control interactions involving $\nabla\partial^\alpha J$ (which is a component of $\nabla\partial^\alpha Z$). Estimating all the nonlinearities of type I in this manner we obtain:

$$\begin{aligned}
&\left|\mathcal{N}_{\mathrm{I},1}\right| \lesssim \left|\left|Z\right|\right|_{H^4}\left|\left|\nabla Z\right|\right|_{L^2}^2, && \left|\mathcal{N}_{\mathrm{I},3}\right| \lesssim \left|\left|Z\right|\right|_{H^4}\left(\left|\left|\nabla^3 Z\right|\right|_{L^2}^2+\left|\left|\nabla^2 Z\right|\right|_{L^2}^2\right),\ \text{and}\\
&\left|\mathcal{N}_{\mathrm{I},2}\right| \lesssim \left|\left|Z\right|\right|_{H^4}\left|\left|\nabla^2 Z\right|\right|_{L^2}^2, && \left|\mathcal{N}_{\mathrm{I},4}\right| \lesssim \left|\left|Z\right|\right|_{H^4}\left(\left|\left|\nabla^4 Z\right|\right|_{L^2}^2+\left|\left|\nabla^3 Z\right|\right|_{L^2}^2\right).
\end{aligned}$$

To estimate nonlinearities of type II we proceed similarly, namely applying the generalized Hölder inequality with two factors in L^2 and the rest in L^∞. In particular we use Lemma 4.4 and Corollary 4.15 to control $m\left(JJ_{eq}^{-1}\right)$ and its derivatives, as well as the second part of Lemma 4.3 for the terms appearing when applying Corollary 4.15 which are cubic or higher-order. As an illustrative example let us write the nonlinearities of type II as $N_{\mathrm{II}}\left(Z\right) = m\left(J\right)b\left(Z,Z\right)$ for some bilinear form b and consider $\int_{\mathbb{T}^3}\partial_{ijk}\left(m\left(JJ_{eq}^{-1}\right)\right)b\left(\partial_l Z,Z\right)\cdot\partial_{ijkl}Z$. This terms appears when $\partial^\alpha = \partial_{ijkl}$ and can be estimated as follows:

$$\begin{aligned}
\left|\int_{\mathbb{T}^3}\partial_{ijk}\left(m\left(JJ_{eq}^{-1}\right)\right)b\left(\partial_l Z,Z\right)\cdot\partial_{ijkl}Z\right| &\lesssim \int_{\mathbb{T}^3}\left(\left|\nabla^3 J\right|+\left|\nabla^2 J\right|\left|\nabla J\right|+\left|\nabla J\right|^3\right)\left|\nabla Z\right|\left|Z\right|\left|\nabla^4 Z\right|\\
&\lesssim \left(\left|\left|\nabla^3 Z\right|\right|_{L^2}+\left|\left|Z\right|\right|_{H^4}\left|\left|\nabla^2 Z\right|\right|_{L^2}+\left|\left|Z\right|\right|_{H^4}^2\left|\left|\nabla Z\right|\right|_{L^2}\right)\left|\left|Z\right|\right|_{H^4}^2\left|\left|\nabla^4 Z\right|\right|_{L^2}\\
&\lesssim \left|\left|Z\right|\right|_{H^4}\left(\left|\left|\nabla^3 Z\right|\right|_{L^2}+\left|\left|\nabla^2 Z\right|\right|_{L^2}+\left|\left|\nabla Z\right|\right|_{L^2}\right)\left|\left|\nabla^4 Z\right|\right|_{L^2}\\
&\lesssim \left|\left|Z\right|\right|_{H^4}\left(\left|\left|\nabla^4 Z\right|\right|_{L^2}^2+\left|\left|\nabla^3 Z\right|\right|_{L^2}^2+\left|\left|\nabla^2 Z\right|\right|_{L^2}^2+\left|\left|\nabla Z\right|\right|_{L^2}^2\right).
\end{aligned}$$

Estimating all terms of type II in this fashion yields, for $i = 1,2,3,4$, $\left|\mathcal{N}_{\mathrm{II},i}\right| \lesssim \left|\left|Z\right|\right|_{H^4}\sum_{j=1}^{i}\left|\left|\nabla^j Z\right|\right|_{L^2}^2$.

Nonlinearities of type III are the most delicate to estimate due to the presence of $\nabla\times u$, $\nabla\left(\nabla\cdot\omega\right)$, and $\Delta\omega$. The presence of these terms causes two difficulties

(1) when ∂^α hits $\Delta\omega$ (or $\nabla\left(\nabla\cdot\omega\right)$) we must integrate by parts since we do not have any control, even through the dissipation, on $\nabla^{|\alpha|+2}\omega$, and
(2) there are precisely two terms in which more than two derivatives of order three or above appear, terms for which we cannot simply use L^2 and L^∞ in the right-hand side of the generalized Hölder inequality. This is easily remedied by more carefully choosing the L^p spaces used, which is done explicitly below.

Let us write the nonlinearity schematically as $N_{\mathrm{III}}(Z) = m\left(JJ_{eq}^{-1}\right) b\left(Z, \nabla^2\omega\right)$ for some bilinear form b. Here is how we handle (1) discussed above: for any multi-index α

$$\begin{aligned}\mathcal{N}_\alpha := \int_{\mathbb{T}^3} m\left(JJ_{eq}^{-1}\right) b\left(Z, \Delta\partial^\alpha\omega\right)\cdot\partial^\alpha\omega &= -\int_{\mathbb{T}^3} \partial_i\left(m\left(JJ_{eq}^{-1}\right)\right) b\left(Z, \partial_i\partial^\alpha\omega\right)\cdot\partial^\alpha\omega \\ &- \int_{\mathbb{T}^3} m\left(JJ_{eq}^{-1}\right) b\left(\partial_i Z, \partial_i\partial^\alpha\omega\right)\cdot\partial^\alpha\omega - \int_{\mathbb{T}^3} m\left(JJ_{eq}^{-1}\right) b\left(Z, \partial_i\partial^\alpha\omega\right)\cdot\partial_i\partial^\alpha\omega\end{aligned}$$

and hence

$$\begin{aligned}|\mathcal{N}_\alpha| &\lesssim \left(||\nabla Z||_\infty ||Z||_\infty + ||\nabla Z||_\infty\right)\left|\left|\nabla^{|\alpha|+1}\omega\right|\right|_{L^2}\left|\left|\nabla^{|\alpha|}\omega\right|\right|_{L^2} + ||Z||_\infty\left|\left|\nabla^{|\alpha|+1}\omega\right|\right|_{L^2} \\ &\lesssim ||Z||_{H^4}\left(\left|\left|\nabla^{|\alpha|+1}\omega\right|\right|_{L^2}^2 + \left|\left|\nabla^{|\alpha|+1}\omega\right|\right|_{L^2}^2\right).\end{aligned}$$

Now we show how to handle (2) discussed above. Both terms under consideration appear when $|\alpha| = 4$, and so we write $\partial^\alpha = \partial_{ijkl}$. Note that we will use Corollary 4.15 to bound $\left|m\left(JJ_{eq}^{-1}\right)\right|$ above by $|\nabla^3 J| + |\nabla^2 J||\nabla J| + |\nabla J|^3$, but below we will only indicate how to deal with the first one amongst these three terms (since the last two can be taken care of by a generalized Hölder inequality using only L^2 and L^∞). We have, using the fact that $H^1\left(\mathbb{T}^3\right) \hookrightarrow L^4\left(\mathbb{T}^3\right)$,

$$\begin{aligned}&\left|\int_{\mathbb{T}^3} m\left(JJ_{eq}^{-1}\right) b\left(\partial_{ijk}Z, \Delta\partial_l\omega\right)\cdot\partial_{ijkl}\omega + \int_{\mathbb{T}^3}\partial_{ijk}\left(m\left(JJ_{eq}^{-1}\right)\right) b\left(Z, \Delta\partial_l\omega\right)\cdot\partial_{ijkl}\omega\right| \\ &\lesssim \int_{\mathbb{T}^3}|\nabla^3 Z||\nabla^3\omega||\nabla^4\omega| + ||Z||_\infty\int_{\mathbb{T}^3}|\nabla^3 J||\nabla^3\omega||\nabla^4\omega| + \dots \\ &\lesssim \int_{\mathbb{T}^3}|\nabla^3 Z||\nabla^3\omega||\nabla^4\omega| \lesssim \left|\left|\nabla^3 Z\right|\right|_{H^1}\left|\left|\nabla^3\omega\right|\right|_{H^1}\left|\left|\nabla^4\omega\right|\right|_{L^2} \\ &\lesssim ||Z||_{H^4}\left(\left|\left|\nabla^3\omega\right|\right|_{L^2} + \left|\left|\nabla^4\omega\right|\right|_{L^2}\right)\left|\left|\nabla^4\omega\right|\right|_{L^2} \lesssim ||Z||_{H^4}\left(\left|\left|\nabla^4\omega\right|\right|_{L^2} + \left|\left|\nabla^3\omega\right|\right|_{L^2}\right).\end{aligned}$$

Estimating all nonlinearities of type III in this fashion yields, for $i = 1, 2, 3, 4$,

$$|\mathcal{N}_{\mathrm{III},i}| \lesssim ||Z||_{H^4}\left(\sum_{j=1}^{i}\left|\left|\nabla^j Z\right|\right|_{L^2}^2 + \left|\left|\nabla^{i+1}(u,\omega)\right|\right|_{L^2}^2\right).$$

Finally we consider the case $\alpha = 0$. Using the fact that $\int_{\mathbb{T}^3}(u\cdot\nabla) Z\cdot Z = 0$ and that $[\Omega, J] : J = 0$ (see Lemma 6.20) we see that

$$\begin{aligned}\int_{\mathbb{T}^3} N(Z)\cdot Z &= -\int_{\mathbb{T}^3} m\left(JJ_{eq}^{-1}\right)\left(\omega\times J\omega + \omega\times J\omega_{eq} + \omega\times J_{eq}\omega + \omega_{eq}\times J\omega\right)\cdot\omega \\ &+ \int_{\mathbb{T}^3} JJ_{eq}^{-1} m\left(JJ_{eq}^{-1}\right)\left(\omega\times J_{eq}\omega_{eq} + \omega_{eq}\times J\omega_{eq} + \omega_{eq}\times J_{eq}\omega + 2\kappa\omega - \kappa\nabla\times u - \tilde{\alpha}\nabla\left(\nabla\cdot\omega\right) - \tilde{\gamma}\Delta\omega\right)\cdot\omega.\end{aligned}$$

It thus follows from Lemmas 4.3 and 4.4 that $\left|\int_{\mathbb{T}^3} N(Z)\cdot Z\right| \lesssim ||Z||_{H^4}||Z||_{L^2}^2$. □

4.4. The chain of energy inequalities. We begin this section by combining the results of Sections 4.2 and 4.3 in order to obtain a chain of energy inequalities.

Proposition 4.17 (Chain of energy inequalities). *There exist $C_0, C_1, C_D > 0$ such that for every $0 < \varepsilon < 1$ there exists $0 < \delta(\varepsilon) < 1$ such that if $\sup_{0\leqslant t\leqslant T}||Z(t)||_{H^4} \leqslant \delta(\varepsilon)$ and Z solves (4.1) then*

$$\frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}t}\left|\left|\sqrt{D}Z\right|\right|_{L^2}^2 + \mathcal{D}(u,\omega) \leqslant \varepsilon||Z||_{L^2}^2 + C_0\left(||\bar{\omega}||_{L^2}^2 + ||a||_{L^2}^2\right)$$

and, for $k = 1, 2, 3, 4$,

$$\frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}t}\left|\left|\nabla^k\left(\sqrt{D}Z\right)\right|\right|_{L^2}^2 + \frac{C_D}{2}\left|\left|\nabla^k(u,\omega)\right|\right|_{H^1}^2 \leqslant \varepsilon\left|\left|\nabla^k Z\right|\right|_{L^2}^2 + C_1\sum_{i=0}^{k-1}\left|\left|\nabla^i Z\right|\right|_{L^2}^2.$$

Proof. Let $\varepsilon > 0$, let C_D and $C_{I,k}$ be as in Lemma 4.9 and Proposition 4.16 respectively, let C_σ be as in Lemma 4.8 for $\sigma := \frac{C_D}{4}$, let $n_k := \#\{\text{multi-index } \alpha : |\alpha| = k\}$, and pick $\delta := \min_{0\leqslant k\leqslant 4}\left\{\delta_0, \frac{\varepsilon}{C_{I,k}n_k}, \frac{C_D}{4C_{I,k}}\right\}$.

First we consider $k=0$. Observe that for $2C_0 := 2(\lambda-\nu) + \left(\frac{\mathfrak{t}}{2\kappa}\right)^2$,

$$B(\bar{\omega}, a) = 2C_0 \int_{\mathbb{T}^3} \bar{\omega}^{\perp} \cdot a \leqslant C_0 \left(||\bar{\omega}||_{L^2}^2 + ||a||_{L^2}^2 \right). \tag{4.9}$$

By Propositions 4.7, 4.16, (4.9), and the fact that $\delta \leqslant \frac{\varepsilon}{C_{I,0}}$ we deduce the energy inequality for $k=0$.

Now we consider $k=1,2,3,4$. For any nonzero multi-index α it follows from Propositions 4.7 and 4.16 and from Lemmas 4.9 and 4.8 that

$$\begin{aligned}\frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}t}\left|\left|\sqrt{D}\left(\partial^\alpha Z\right)\right|\right|_{L^2}^2 + C_D||\partial^\alpha(u,\omega)||_{H^1}^2 &\leqslant \left(\frac{C_D}{4}\left|\left|\partial^{\alpha+1}\omega\right|\right|_{L^2}^2 + C_\sigma\left|\left|\partial^{\alpha-1}a\right|\right|_{L^2}^2\right) \\ &\quad + C_{I,k}||Z||_{H^4}\left(\sum_{i=1}^{k}\left|\left|\nabla^i Z\right|\right|_{L^2}^2 + \left|\left|\nabla^{k+1}(u,\omega)\right|\right|_{L^2}^2\right).\end{aligned}$$

Summing over $|\alpha| = k$ and using that $\delta \leqslant \min\left(\frac{C_D}{4C_{I,k}}, \frac{\varepsilon}{C_{I,k}n_k}\right)$ we observe that, after absorbing $\left|\left|\partial^{\alpha+1}\omega\right|\right|_{L^2}^2$ and $\left|\left|\nabla^{k+1}(u,\omega)\right|\right|_{L^2}^2$ into the dissipation on the left-hand side,

$$\frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}t}\left|\left|\nabla^k\left(\sqrt{D}Z\right)\right|\right|_{L^2}^2 + \frac{C_D}{2}\left|\left|\nabla^k(u,\omega)\right|\right|_{H^1}^2 \leqslant n_k C_\sigma\left|\left|\partial^{k-1}a\right|\right|_{L^2}^2 + \varepsilon\sum_{i=1}^{k}\left|\left|\nabla^i Z\right|\right|_{L^2}^2$$

from which the result follows upon taking $C_1 := \max(1, n_4 C_\sigma)$. □

We now record, in abstract form, a Gronwall-type lemma for chains of differential inequalities.

LEMMA 4.18 (Chain of Gronwall inequalities). *Consider, for $k \geqslant -1$, $E_k : [0,\infty) \to [0,\infty)$. Suppose that there exists $C_{-1}, C > 0$, $0 < \theta \leqslant \theta_0 < \psi$, and $k_{max} \geqslant -1$ such that for every $t \geqslant 0$, $E_{-1}(t) \leqslant C_{-1}e^{\psi t}$ and every $k \geqslant 0$,*

$$\frac{\mathrm{d}}{\mathrm{d}t}E_k(t) \leqslant \theta E_k(t) + C\sum_{i=-1}^{k-1} E_i(t). \tag{4.10}$$

Then, for every $0 \leqslant k \leqslant k_{max}$, there exist $C_k > 0$ such that for every $t \geqslant 0$

$$E_k(t) \leqslant C_k\left(C_{-1} + \sum_{i=0}^{k} E_i(0)\right)e^{\psi t} =: \widetilde{C}_k e^{\psi t}. \tag{4.11}$$

Moreover: if (4.10) holds for every $k \geqslant -1$ then so does (4.11).

PROOF. We induct on k, noting that the base case $k=-1$ holds by assumption. Now suppose that (4.11) holds for every $i=-1,\ldots,k-1$. Then, by (4.10), $\frac{\mathrm{d}}{\mathrm{d}t}\left(E_k(t)e^{-\theta t}\right) \leqslant C\sum_{i=-1}^{k-1}E_i(t)e^{-\theta t}$ and hence, integrating in time and using (4.11), where $\widetilde{C}_{-1} := C_{-1}$,

$$E_k(t) \leqslant E_k(0)e^{\theta t} + C\sum_{i=-1}^{k-1} e^{\theta t}\int_0^t \widetilde{C}_i e^{(\psi-\theta)s}ds \leqslant \left(E_k(0) + \frac{C}{\psi-\theta_0}\sum_{i=-1}^{k-1}\widetilde{C}_i\right)e^{\psi t} \leqslant C_k\left(C_{-1} + \sum_{i=0}^{k}E_i(0)\right)e^{\psi t}.$$

for some $C_k > 0$. □

We conclude this section by applying Lemma 4.18 to the chain of differential inequalities obtained in Proposition 4.17, which yields a bootstrap energy inequality.

PROPOSITION 4.19 (Bootstrap energy inequality). *There exists $0 < \delta_B < 1$ such that if Z solves (4.1) and $\sup_{0\leqslant t\leqslant T}||Z(t)||_{H^4} \leqslant \delta_B$ then for every $\psi > 0$ there exists $C(\psi) > 0$ such that if there exists $C_{ins} > 0$ such that $E_{ins}(t) := ||\bar{\omega}(t)||_{L^2}^2 + ||a(t)||_{L^2}^2$ satisfies $E_{ins}(t) \leqslant C_{ins}e^{\psi t}$ for all $t > 0$ then, for all $t \geqslant 0$,*

$$||Z(t)||_{H^4}^2 \leqslant C(\psi)\left(||Z(0)||_{H^4}^2 + C_{ins}\right)e^{\psi t}.$$

PROOF. Let us define $E_{-1} := E_{\mathrm{ins}}$, $E_k(t) := \left|\left|\nabla^k\left(\sqrt{D}Z\right)\right|\right|_{L^2}^2$ for every $t \geqslant 0$ and every $k \geqslant 0$, and $C := \max(C_0, C_1)$ for C_0 and C_1 as in Proposition 4.17. Observe that $|J_{eq}^{1/2}w|^2 \geqslant \nu/2|w|^2$ for any $w \in \mathbb{R}^3$ and hence $||Z||_{L^2}^2 \leqslant \max(1, 2/\nu)\left|\left|\sqrt{D}Z\right|\right|_{L^2}^2$. Let $\psi > 0$ and note that we may deduce from Proposition 4.17, picking $\varepsilon = \frac{1}{2}\min(1, \psi/2, \psi\nu/2)$, $\delta_B := \delta(\varepsilon)$, and neglecting the dissipation, that for $k = 0, \ldots, 4$ and every $t \geqslant 0$

$$\frac{\mathrm{d}}{\mathrm{d}t}E_k(t) \leqslant \frac{\psi}{2}E_k(t) + C\sum_{i=-1}^{k-1} E_i(t). \tag{4.12}$$

Now suppose that, for every $t \geqslant 0$, $E_{-1}(t) = E_{\text{ins}}(t) \leqslant C_{\text{ins}}e^{\psi t} =: C_{-1}e^{\psi t}$. Using Lemma 4.18 we obtain that for $k = 0, \ldots, 4$ there exists $C_k > 0$ such that $E_k(t) \leqslant C_k\left(C_{-1} + \sum_{i=0}^{k} E_i(0)\right)e^{\psi t}$. Finally, summing over $k = 0, \ldots, 4$ we obtain that

$$\begin{aligned}||Z(t)||_{H^4}^2 \leqslant \max(1, 2/\nu)\sum_{k=0}^{4} E_k(t) &\leqslant \widetilde{C}(\psi)\left(C_{-1} + \left|\left|\sqrt{D}Z(0)\right|\right|_{H^4}^2\right)e^{\psi t}\\ &\leqslant \max(1, \lambda, \nu)\,\widetilde{C}(\psi)\left(C_{\text{ins}} + ||Z(0)||_{H^4}^2\right)e^{\psi t}\end{aligned}$$

for some $\widetilde{C}(\psi) > 0$, so we may simply pick $C(\psi) := \max(1, \lambda, \nu)\,\widetilde{C}(\psi)$. □

5. The bootstrap instability argument

In this section we prove our main result using a Guo-Strauss bootstrapping argument. This technique was introduced by Guo and Strauss in [GS95a], inspired by [GS95b] and [FSV97]. For a cleanly written and very readable form of the bootstrap instability argument we refer to Lemma 1.1 of [GHS07].

For the purpose of the theorem below, we define what we mean by a strong solution of (1.1a)–(1.1d).

Definition 5.1 (Strong solutions). For any $X_0 \in H^2(\mathbb{T}^3)$ and any $T > 0$ we define a *strong solution* of (1.1a)–(1.1d) with initial condition X_0 to be any function $X \in L^\infty([0, T], H^2(\mathbb{T}^3))$ with $\partial_t X \in L^\infty([0, T], L^2(\mathbb{T}^3))$ for which (1.1a)–(1.1d) is satisfied almost everywhere in $(0, T) \times \mathbb{T}^3$ and such that $X(0) = X_0$.

Theorem 5.2 (Bootstrap instability). *Let η_* be as in Proposition 3.10 and assume that μ, κ, $\alpha + \frac{4\beta}{3}$, $\beta + \gamma > 0$. There exists $\theta, \delta > 0$ and $Z_0 \in L^2(\mathbb{T}^3, \mathbb{R}^3 \times \mathbb{R}^3 \times \mathbb{R}^{3\times 3})$ such that for all $0 < \iota < \delta$ if we define $T_I := \frac{1}{\eta_*}\log\frac{\theta}{\iota}$ then there exists a strong solution $X = (u, \omega, J) \in L^\infty([0, T_I], H^4(\mathbb{T}^3))$ of (1.1a)–(1.1d) with pressure $p \in L^\infty([0, T_I], H^4(\mathbb{T}^3))$ and initial condition $X(0) = X_{eq} + \iota Z_0$ such that $||X(T_I) - X_{eq}||_{L^2} > \frac{\theta}{2}$.*

Proof. The crux of the argument is to compare three timescales: the instability timescale T_I, the linear-dominance timescale T_L, and the smallness timescale T_S. We will show that at times living in both the linear-dominance and the smallness timescale (i.e. times anterior to both T_L and T_S) two key estimates hold, namely (5.1) and (5.2). This will allow us, by way of contradiction, to show that the instability timescale is the shortest of the three. It will thus follow that instability occurs while the dynamics are dominated by the linearization and while we are in the small energy regime.

We begin by recalling appropriate notation from previous results. Let $(u_0, \omega_0, a_0) =: Y$ be as in Proposition 3.13 and note that without loss of generality we may assume that $||Y||_{L^2} = 1$. Define $Z_0 := (u_0, \omega_0, J_0)$ where $J_0 = \begin{pmatrix} 0_{2\times 2} & a_0 \\ a_0^T & 0 \end{pmatrix}$. Let δ_0 be as in the small energy regime (c.f. Definition 4.2), let $C_S := C_S\left(\frac{\eta_*}{2}\right)$ as in Proposition 3.12, let C_N be as in Proposition 4.5, let $\psi := 2\eta_*$ such that $C_B := C(\psi)$ and δ_B are as in Proposition 4.17, and let δ_{lwp} be as in Theorem 6.3 with δ_{lwp} being chosen small enough so as to ensure that $L \leqslant \eta_*$.

We can now define the appropriate small scales θ and δ, which in turn will later allow us to precisely define the timescales. Let

$$\theta = \frac{1}{2}\min\left(\delta_0, \delta_B, \frac{1}{C}, ||Z_0||_{L^2}\left(\frac{\delta_{\text{lwp}}}{||Z_0||_{H^4}}\right)^{\eta_*/L}\right),$$

$$\delta = \frac{1}{2}\min\left(1, \delta_{\text{lwp}}, \frac{\theta}{||Z_0||_{H^4}}, \left(C_B\left(||Z_0||_{H^4}^2 + 4\right)\theta\right)^{-1/2}, \frac{1}{2C\theta}\right),$$

and let $0 < \iota < \delta$.

By our local well-posedness result (see Section 6.1 and Corollary 6.6 in particular) there exists $T_E > 0$ and a unique strong solution $Z \in L^\infty([0, T_E]; H^4(\mathbb{T}^3))$ of (4.1) with pressure $p \in L^\infty H^4$ and initial data $Z(0) = \iota Z_0$. Note that our local existence result (Theorem 6.3) tells us moreover that the solution Z may be continued as long as Z remains in an open H^4-ball of radius δ_{lwp}. We may thus without loss of generality assume T_E to be the *maximal* time of existence of the solution in the sense that $T_E := \sup\left\{T > 0 : Z \text{ exists on } [0,T] \text{ and } \sup_{0\leqslant t<T}||Z(t)||_{H^4} < \delta_{\mathrm{lwp}}\right\}$. Expanding out the definition of the notation in (4.1) we see that $X := X_{eq} + Z$ is a strong solution of (1.1a)–(1.1d) with initial condition $X(0) = X_{eq} + \iota Z_0$.

We may now define the timescales. We define

$$T_L := \sup\left\{0 < t < T_E : ||\bar{\omega}(t)||_{L^2} + ||a(t)||_{L^2} \leqslant 2\iota e^{\eta_* t}\right\}, T_I := \frac{1}{\eta_*}\log\frac{\theta}{\iota}, \text{ and}$$
$$T_S := \sup\left\{0 < t < T_E : ||Z(t)||_{H^4} < \theta\right\}.$$

Note that $T_L \geqslant 0$ since $||Z(0)||_{L^2} = \iota$, that $T_S \geqslant 0$ since $||Z(t)||_{H^4} = \iota||Z_0||_{H^4} < \theta$, and that $T_S, T_L \leqslant T_E$

Step 1: Since $\theta \leqslant \delta_B$ we deduce from Proposition 4.19 that if $t \leqslant \min(T_L, T_S)$ then

$$||Z(t)||^2_{H^4} \leqslant C_B \iota^2 \left(||Z_0||^2_{H^4} + 4\right) e^{2\eta_* t}. \tag{5.1}$$

Now we apply the Leray projector to eliminate the pressure and write (4.1) in the reduced form $\partial_t Z = \mathcal{L}Z + \widetilde{N}(Z)$, where $\widetilde{N} := \mathbb{P}D^{-1}N$. More precisely we apply $\mathbb{P}$ and observe that $\bar{\mathbb{P}}\mathcal{B} = \mathcal{B}$ and hence $\mathbb{P}\mathcal{L} = \mathcal{L}$. Indeed this follows from the observation that on one hand, for $k \neq 0$, $\hat{\bar{\mathbb{P}}}(k) = \left(I - \frac{k\otimes k}{|k|^2}\right) \oplus I_3 \oplus I_2 = I - \mathrm{proj}_{V_k}$ and the fact that, since $\hat{\mathcal{B}}_k$ acts on $\mathbb{C}^8/V_k$, it follows that $\mathrm{proj}_{V_k} \circ \hat{\mathcal{B}}_k = \hat{\mathcal{B}}_k$, whilst on the other hand, for $k = 0$, we have that $\hat{P}_L(0) = I_3$ (since constant vector fields are divergence-free) and hence $\hat{\mathbb{P}}(0) = \mathrm{id}$.

We can thus apply the Duhamel formula to obtain

$$Z(t) - e^{t\mathcal{L}}Z(0) = \int_0^t e^{(t-s)\mathcal{L}}\widetilde{N}(Z(s))\,ds$$

which can be estimated, when $t \leqslant \min(T_L, T_S)$, using the fact that $\theta \leqslant \delta_0$, Proposition 3.12, the fact that the Leray projector is bounded on L^2, the inequality $\left|\left|D^{-1}\right|\right| \leqslant \sqrt{\max(1, 2/\nu)}$, and Proposition 4.5 to yield, for $C := \frac{2}{\eta_*}\max(1, 2/\nu)\, C_S C_N C_B \left(||Z_0||^2_{H^4} + 4\right)$,

$$\left|\left|Z(t) - e^{t\mathcal{L}}\iota Z_0\right|\right|_{L^2} \leqslant C\iota^2 e^{2\eta_* t}. \tag{5.2}$$

Step 2: Now we show that $T_I = \min(T_I, T_L, T_S)$, using the key estimates (5.1) and (5.2). First suppose for the sake of contradiction that $T_L = \min(T_I, T_L, T_S)$. By definition of T_L,

$$||\bar{\omega}(T_L)||_{L^2} + ||a(T_L)||_{L^2} = 2\iota e^{\eta_* T_L}. \tag{5.3}$$

Now note that (5.2) applies since $T_L \leqslant T_S$ and thus it follows from Proposition 3.13 and the choice of Z_0 that $||Z(T_L)||_{L^2} \leqslant \left(1 + C\iota e^{\eta_* T_L}\right)\iota e^{\eta_* T_L} < 2\iota e^{\eta_* T_L}$, where we have used that $T_L \leqslant T_I$ and hence $C\iota e^{\eta_* T_L} \leqslant C\iota e^{\eta_* T_I} = C\theta < 1$. This contradicts (5.3) and hence the linear-dominance timescale T_L is not the smallest of the three timescales considered.

Now suppose for the sake of contradiction that $T_S = \min(T_I, T_L, T_S)$. By definition of T_S,

$$||Z(T_S)||_{H^4} = \theta. \tag{5.4}$$

Since $T_S \leqslant T_L$ we may use (5.1) and since $T_S \leqslant T_I$ we have that $e^{2\eta_* T_S} \leqslant e^{2\eta_* T_I} = \theta$. Putting these two facts together tells us that $||Z(T_S)||^2_{H^4} \leqslant C_B\iota^2\left(||Z_0||^2_{H^4} + 4\right)\theta^2 < \theta$ which contradicts (5.4). Therefore the smallness timescale T_S is not the smallest of the three timescales considered. We thus deduce that $T_I = \min(T_I, T_L, T_S)$.

Step 3: Finally we show that $||Z(T_I)||_{L^2} > \frac{\theta}{2}$. Since T_I is smaller than both T_L and T_S (and hence smaller than T_E) we may use (5.1) and (5.2), as well as Proposition 3.13, the choice of Z_0, and the fact that $\iota e^{\eta_* T_I} = \theta$ to see that $||X(T_I) - X_{eq}||_{L^2} = ||Z(T_I)||_{L^2} \geqslant \iota e^{\eta_* T_I} - C\iota^2 e^{2\eta_* T_I} = \theta(1 - C\iota\theta) > \frac{1}{2}$. □

6. Appendix

6.1. Local well-posedness. In this section we prove the local well-posedness of (1.1a)–(1.1d). This is done in two steps: we prove local existence in the small energy regime in Theorem 6.3 and we prove uniqueness within a broader class of solutions in Theorem 6.5. Notably, this uniqueness result makes no smallness assumption and only requires that the unknowns belong to appropriate Sobolev spaces.

A key step on the way to our local existence result is to prove that the nonlinearity is sufficiently regular. We do this below in Lemma 6.1 where we prove that the nonlinearity is analytic.

Lemma 6.1 (Analyticity of the nonlinearity). *Let $0 < \delta \leqslant \delta_0$ for δ_0 as in the small energy regime. For every $s > \frac{3}{2}$, $N : H^{s+2} \cap H^4_{\delta_0} \to H^s$ is analytic (as a mapping from H^{s+2} to H^s). Moreover the Lipschitz constant of N on $H^{s+2} \cap H^4_{\delta_0} \to H^s$ approaches zero as $\delta \downarrow 0$.*

Proof. The two key observations are that (i) we may write $N(Z) = P\left(m\left(JJ_{eq}^{-1}\right), Z, \nabla Z, \nabla^2 Z\right)$ for some polynomial P and that (ii) m is analytic (recall that m is defined in Definition 4.1). Indeed m can be written as a geometric series, namely $m(A) = \sum_{i=0}^{\infty} (-1)^i A^i$ for every $A \in \mathfrak{B}$, where $\mathfrak{B}$ is defined in Definition 4.1.

Using Lemma 6.25, the fact that H^s is a continuous algebra when $s > \frac{3}{2}$, and the fact that polynomials are analytic, it follows that we may write $N = F\left(\mathcal{J}^2 Z\right)$ for some function $F : \operatorname{dom} F \subseteq H^s \to H^s$ which is analytic on its domain (i.e. where it is well-defined), where $\mathcal{J}^2 Z := \left(Z, \nabla Z, \nabla^2 Z\right)$. The last observation we need is that $\mathcal{J}^2\left(H^s \cap H^4_{\delta_0}\right) \subseteq \operatorname{dom} F$. This holds since, if $Z = (u, \omega, J) \in H^{s+2} \cap H^4_{\delta_0}$ for δ_0 as in the small energy regime, then by Lemma 4.4 we know that $J \mapsto m\left(JJ_{eq}^{-1}\right)$ is well-defined, and hence analytic. Since $\mathcal{J}$ is a bounded linear map from H^{s+2} to H^s it is also analytic, and so we may conclude that $N : H^{s+2} \cap H^4_{\delta_0} \to H^s$ is analytic as a map from H^{s+2} to H^s.

Finally, note that the polynomial P above is at least quadratic in $\left(Z, \nabla Z, \nabla^2 Z\right)$ and that therefore $DN(0) = 0$. In particular it follows that the Lipschitz constant of N on balls of vanishingly small radii approaches zero, as claimed. □

Remark 6.2. See [Whi65] for a brief and clean summary of basic results regarding analytic functions between Banach spaces.

With Lemma 6.1 in hand we may now prove our local existence result.

Theorem 6.3 (Local existence and continuous dependence on the data). *There are universal constants $\rho, \delta_{lwp}, C > 0$ such that for any $Z_0 = (u_0, \omega_0, J_0) \in H^4$ with $\nabla \cdot u_0 = 0$, $\fint_{\mathbb{T}^3} u_0 = 0$, and $||Z_0||_{H^4} < \delta_{lwp}$, there exists a time of existence $T_{lwp} > 0$, there exists $Z = (u, \omega, J) \in L^\infty H^4$ with $(u, \omega) \in L^2 H^5$, $\partial_t Z \in L^\infty H^2 \cap L^2 H^3$, and $\partial_t J \in L^\infty H^3$, and there exists $p \in L^\infty H^4 \cap L^2 H^5$ with average zero such that u is divergence-free and has average zero, (u, p, ω, J) solves*

$$\partial_t DZ = \widetilde{\mathcal{L}} Z + \Lambda(p) + N(Z) \ \text{ a.e. in } (0, T_{lwp}) \text{ and } Z(0) = Z_0 \text{ in } H^{4-\frac{1}{4}}, \tag{6.1}$$

and the estimates

$$||Z||_{L^\infty H^4} + ||(u, \omega)||_{L^2 H^5} + ||\partial_t Z||_{L^\infty H^2 \cap L^2 H^3} + ||\partial_t J||_{L^\infty H^3} \leqslant C ||Z_0||_{H^4} \tag{6.2}$$

and

$$||p||_{L^\infty H^4 \cap L^2 H^5} \leqslant C ||u||^2_{L^\infty H^4 \cap L^2 H^5}. \tag{6.3}$$

hold. Moreover we have the lower bound $T_{lwp} \geqslant \frac{1}{\rho} \log \frac{\delta_{lwp}}{||Z_0||_{H^4}}$.

Proof. We proceed via a standard Galerkin scheme and thus omit the fine details of the proof here. A key point is that everything we need to know about the nonlinearity for the purpose of this local well-posedness result is obtained in Lemma 6.1.

We now proceed in five steps. In Step 1 we eliminate the pressure via Leray projection, in Step 2 we prove local well-posedness for a sequence of appropriate approximate problems, in Step 3 we obtain uniform bounds on these approximate solutions, in Step 4 we pass to the limit via a compactness argument, and in Step 5 we reconstruct the pressure.

First we recall some notation from earlier results which is required to define the smallness parameter δ_{lwp}. Let δ_0 be as in the small energy regime, let $\delta = \delta\left(\frac{1}{2}\right)$ be as in Proposition 4.17, and define $C_2 := \max(1, \lambda, \nu) \max(1, 2/\nu)$. Then take $\delta_{\text{lwp}} := \frac{1}{3} \min(\delta_0 / C_2, \delta)$.

Step 1: Leray projection eliminating the pressure.

Recall that we denote the Leray projector by $\mathbb{P}_L$ and that we write $\mathbb{P} = \mathbb{P}_L \oplus I_3 \oplus I_{3\times 3}$. Upon applying $\mathbb{P}$ to (6.1) we thus see that (noting that $\mathbb{P}Z = Z$ since $\nabla \cdot u = 0$ and that $\mathbb{P}$ and $\widetilde{\mathcal{L}}$ commute since they are both Fourier multipliers): $\partial_t DZ = \widetilde{\mathcal{L}}Z + \mathbb{P}N(Z)$.

Step 2: Local well-posedness of a sequence of approximate problems.

Let $V_n := \left\{ Z \in L^2\left(\mathbb{T}^3; \mathbb{R}^3 \times \mathbb{R}^3 \times \mathbb{R}^{3\times 3}\right) \,\middle|\, \hat{Z}(k) = 0 \text{ if } |k| > n \text{ and } \nabla \cdot u = 0 \right\}$, let $\mathcal{U}_n := V_n \cap H^4_{\delta_0/2}$ where H^α_R denotes the open ball around zero of radius R in H^α, and let $\mathbb{P}_n$ be the orthogonal projection onto V_n defined by $\hat{\mathbb{P}}_n(k) = \mathbb{1}(|k| \leqslant n)$.

We approximate the system obtained after Leray projection in Step 1 by

$$\partial_t DZ_n = \widetilde{\mathcal{L}}Z_n + \mathbb{P}_n\mathbb{P}N(Z_n) \text{ and } Z_n(0) = \mathbb{P}_n Z_0. \tag{6.4}$$

In order to use standard finite-dimensional ODE theory we write (6.4) as

$$\partial_t Z_n = F_n(Z_n) \text{ and } Z_n(0) = \mathbb{P}_n Z_0 \tag{6.5}$$

for $F_n = D^{-1}\left(\widetilde{\mathcal{L}} + \mathbb{P}_n\mathbb{P}N\right)$. It follows from Lemma 6.1 that F_n is analytic from $H^4_{\delta_0}$ to H^2, and since $\mathcal{U}_n$ is a subset of $H^4_{\delta_0}$ and $\mathbb{P} \circ \mathbb{P}$ maps onto V_n we deduce that F_n maps $\mathcal{U}_n$ to V_n.

We may now apply standard ODE theory, which tells us that if we pick an initial condition $Z_0 = (u_0, \omega_0, J_0) \in H^4$ which satisfies $\nabla \cdot u_0 =$, $\fint_{\mathbb{T}^3} u_0 = 0$, and $||Z_0||_{H^4} < \delta_{\text{lwp}}$ then there exists a maximal time of existence $T_n > 0$, a unique $Z_n \in C^\infty([0, T_n); \mathcal{U}_n)$ solving (6.5), and the following blow-up criterion holds: for any $T > 0$ if $\sup_{0\leqslant t\leqslant T} ||Z_n(t)||_{H^4} < \frac{\delta_0}{2}$ then $T \leqslant T_n$.

Step 3: Uniform bounds on the approximate solutions.

To obtain uniform bounds it suffices to apply Proposition 4.17 to the approximate solutions Z_n. Since Proposition 4.17 is only applicable in a small energy regime we must first ensure that $||Z_n||_{H^4}$ remains sufficiently small. We defined $\tilde{T}_n$ to this effect below.

Let $\delta_u = \frac{1}{3}\min(\delta_0, \delta)$, and let $\tilde{T}_n = \sup\{t > 0 \mid ||Z_n||_{H^4} \leqslant \delta_u\}$. Note that $\tilde{T}_n \geqslant T_n$ by the blow-up criterion from Step 1. We may now apply a time-integrated version of Proposition 4.17 (with $\varepsilon = \frac{1}{2}$) to obtain

$$\frac{1}{2}\left|\left|\sqrt{D}Z_n(t)\right|\right|^2_{L^2} - \frac{1}{2}\left|\left|\sqrt{D}Z_n(0)\right|\right|^2_{L^2} + \int_0^t \mathcal{D}(u_n, \omega_n)(s)\,ds \leqslant \int_0^t \left(\frac{1}{2} + C_0\right)||Z_n(s)||^2_{L^2} ds \tag{6.6}$$

and, for $k = 1, 2, 3, 4$,

$$\begin{aligned}\frac{1}{2}\left|\left|\nabla^k\left(\sqrt{D}Z_n(t)\right)\right|\right|^2_{L^2} - \frac{1}{2}\left|\left|\nabla^k\left(\sqrt{D}Z_n(0)\right)\right|\right|^2_{L^2} + \int_0^t \frac{C_D}{2}\left|\left|\nabla^k(u_n, \omega_n)(s)\right|\right|^2_{H^1} ds \\ \leqslant \int_0^t \max\left(\frac{1}{2}, C_1\right)\sum_{i=0}^k \left|\left|\nabla^i Z_n(s)\right|\right|^2_{L^2} ds.\end{aligned} \tag{6.7}$$

where C_0, C_1, and C_D are as in Proposition 4.17. Note that Proposition 4.17 as stated applies to solutions of $\partial_t DZ = \widetilde{\mathcal{L}}Z + N(Z) + \Lambda(p)$ whereas Z_n satisfies $\partial_t DZ_n = \widetilde{\mathcal{L}}Z_n + \mathbb{P}_n\mathbb{P}_L N(Z_n)$. Nonetheless, Proposition 4.17 applies to Z_n as well since this theorem relies solely on energy estimates, and in particular, since $\fint_{\mathbb{T}^3} \Lambda(p) \cdot Z = 0$ when $\nabla \cdot u = 0$ and $\fint_{\mathbb{T}^3} \mathbb{P}_n\mathbb{P}N(Z_n) \cdot Z_n = \fint_{\mathbb{T}^3} N(Z_n) \cdot Z_n$ since Z_n belongs to the image of the projection $\mathbb{P}_n \circ \mathbb{P}$, it follows that the estimate obtained for Z in Proposition 4.17 also holds for Z_n.

Summing (6.6) and (6.7) and using the integral form of the Gronwall inequality tells us that, for any $0 < t < \tilde{T}_n$,

$$||Z_n(t)||^2_{H^4} + \int_0^t ||(u_n, \omega_n)||^2_{H^5} \leqslant C_2 e^{\rho t}||Z_0||^2_{H^4} \tag{6.8}$$

where $\rho := 2(1 + C_0 + C_1)\max(1, 2/\nu)$. In particular we deduce from the blow-up criterion that if we denote by T_{lwp} the infimum of T_n over n then $T_{\text{lwp}} \geqslant \frac{1}{\rho}\log\frac{\delta_{\text{lwp}}}{||Z_0||_{H^4}}$. In other words we have a uniform lower bound on the time of existence of the approximate solutions.

Now we obtain bounds on the time derivative $\partial_t Z_n$, which are required for the compactness argument in Step 4. Note first that (6.8) tells us that, for $C_4 = C_2 e^{\rho T_{\text{lwp}}}$,

$$\sup_n \left(||(u_n, \omega_n, J_n)||^2_{L^\infty H^4} + ||(u_n, \omega_n)||^2_{L^2 H^5}\right) \leqslant C_4||Z_0||^2_{H^4} \tag{6.9}$$

where L^pH^s denote $L^p([0,T_{\text{lwp}}];H^s)$. Using Lemma 6.1 and the boundedness of $\widetilde{\mathcal{L}}$, $\mathbb{P}_n$, and $\mathbb{P}$ we deduce from (6.9) that, for some $C_5>0$,

$$\sup_n\left(||\partial_t(u_n,\omega_n,J_n)||^2_{L^\infty H^2}+||\partial_t(u_n,\omega_n)||^2_{L^2H^3}\right)\leqslant C_5||Z_0||^2_{H^4}. \tag{6.10}$$

Finally we improve this bound on $\partial_t Z_n$ by paying closer attention to the structure of the PDE (6.4). Specifically: since $\widetilde{\mathcal{L}}_3$ and N_3 lose fewer derivatives than $\widetilde{\mathcal{L}}$ and N do, we obtain an improved estimate for $\partial_t J_n$:

$$\sup_n||\partial_t J_n||^2_{L^\infty H^3}\leqslant C_4||Z_0||^2_{H^4}. \tag{6.11}$$

Step 4: Passing to the limit by compactness.

By applying Banach-Alaoglu (i.e. the weak-$*$ compactness of bounded sets) to the bounds provided by (6.9), (6.10), and (6.11) we obtain a subsequence of (Z_n), which for simplicity we do not relabel, such that

$$Z_n\overset{*}{\rightharpoonup}Z \text{ in } L^\infty H^4,(u_n,\omega_n)\rightharpoonup(u,\omega)\text{ in } L^2H^5, \tag{6.12}$$

$$\partial_t Z_n\overset{*}{\rightharpoonup}\partial_t Z\text{ in } L^\infty H^2,\partial_t Z_n\rightharpoonup\partial_t Z\text{ in } L^2H^3,\text{ and } \partial_t J_n\overset{*}{\rightharpoonup}\partial_t J\text{ in } L^\infty H^3 \tag{6.13}$$

for some $Z=(u,\omega,J)\in L^\infty H^4$ with $(u,\omega)\in L^2H^5$, $\partial_t Z\in L^\infty H^2\cap L^2H^3$, and $\partial_t J\in L^\infty H^3$. Moreover, it follows from Aubin-Lions-Simon that, passing to another subsequence which we do not relabel,

$$Z_n\to Z\text{ in } C^0H^{4-\frac14} \tag{6.14}$$

and that $Z\in C^0H^{4-\frac14}$.

We now pass to the limit. It follows immediately from (6.12) and (6.13) that

$$\partial_t DZ_n\overset{*}{\rightharpoonup}\partial_t DZ\text{ and }\widetilde{\mathcal{L}}Z_n\overset{*}{\rightharpoonup}\widetilde{\mathcal{L}}Z\text{ in } L^\infty H^2. \tag{6.15}$$

To pass to the limit in the nonlinearity we write

$$\mathbb{P}_n\mathbb{P}N(Z_n)-\mathbb{P}N(Z)=\mathbb{P}_n\mathbb{P}(N(Z_n)-N(Z))+(\mathbb{P}_n-I)\mathbb{P}N(Z):=A+B.$$

Passing to the limit in B is immediate: by weak-$*$ lower semi-continuity of the $L^\infty H^4$ norm we know that $\sup_{0\leqslant t\leqslant T_0}||Z(t)||_{H^4}\leqslant\frac{\delta_0}{2}<\delta_0$ such that $N(Z)$ is a well-defined element of $L^\infty H^2$. In particular, since $||(I-\mathbb{P}_n)f||_{H^s}\to 0$ for all $s\geqslant 0$ and all $f\in H^s$, it follows that

$$||B||_{L^\infty H^2}=||(I-\mathbb{P}_n)\mathbb{P}N(Z)||_{L^\infty H^2}\to 0. \tag{6.16}$$

Passing to the limit in A relies on the analyticity of the nonlinearity obtained in Lemma 6.1: since $Z_n\to Z$ in $C^0H^{4-\frac14}$ and since, as observed above, both the sequence (Z_n) and its limit Z lie in $H^4_{\delta_0/2}$, it follows from Lemma 6.1 (since $2-\frac14>\frac32$) that $N(Z_n)\to N(Z)$ in $C^0H^{4-\frac14}$. So finally:

$$||A||_{L^\infty H^{2-\frac14}}=||\mathbb{P}_n\mathbb{P}(N(Z_n)-N(Z))||_{L^\infty H^{2-\frac14}}\leqslant||N(Z_n)-N(Z)||_{L^\infty H^{2-\frac14}}\to 0. \tag{6.17}$$

We conclude from (6.4), (6.15), (6.16), and (6.17) that Z is a strong solution of $\partial_t DZ=\widetilde{\mathcal{L}}Z+\mathbb{P}N(Z)$. As a consequence we deduce that the conditions $\nabla\cdot u=0$ and $⨍_{\mathbb{T}^3}u=0$ are propagated in time, i.e. they hold for every $0\leqslant t<T_{\text{lwp}}$.

Finally we deduce from (6.9), (6.10), and (6.11) and the weak and weak-$*$ lower semi-continuity of the appropriate norms that, for some $C>0$,

$$||(u,\omega,J)||_{L^\infty H^4}+||(u,\omega)||_{L^2H^5}+||\partial_t(u,\omega,J)||_{L^\infty H^2\cap L^2H^3}+||\partial_t J||_{L^\infty H^3}\leqslant C||Z_0||_{H^4}. \tag{6.18}$$

Step 5: Reconstructing the pressure.

The key observation is that since $\mathbb{P}=\mathbb{P}_L\oplus I_3\oplus I_{3\times3}$ we may reconstruct p via $I-\mathbb{P}_L$, where $I-\mathbb{P}_L=\nabla\Delta^{-1}\nabla\cdot$ as per Lemma 6.26. More precisely: let $p:=\Delta^{-1}(\nabla\cdot N_1(Z))$ and note that p thus defined has average zero. Then, by Lemma 6.26, $\nabla p=(I-\mathbb{P}_L)N_1(Z)$ and hence $\Lambda(p)=-(I-\mathbb{P})N(Z)$ such that (6.1) holds. Finally, since $N_1(Z)=-(u\cdot\nabla)u$ and since H^s is an algebra for $s>3/2$ we have that, for $s=3$ or 4,

$$||p||_{H^s}\lesssim||N_1(Z)||_{H^{s-1}}=||(u\cdot\nabla)u||_{H^{s-1}}\lesssim||u||_{H^{s-1}}||u||_{H^s}.$$

Combining these estimates with (6.18) yields (6.3). □

REMARK 6.4. It may appear somewhat odd that the initial condition $Z(0) = Z_0$ of (6.1) holds in $H^{4-\frac{1}{4}}$ and not in H^4 as one might expect. This is due to the loss of spatial regularity incurred when applying the Aubin-Lions-Simon lemma to obtain strong convergence of the approximate solutions in $C^0 H^{4-\frac{1}{4}}$. In particular, note that the only thing which is special about $\frac{1}{4}$ is that it sits squarely between 0 and $\frac{1}{2}$ and that we use that $\left(4 - \frac{1}{4}\right) - 2 > \frac{3}{2}$ when we leverage Lemma 6.1 to pass to the limit in the nonlinearity in Step 4 of the proof of Theorem 6.3. This means that we can actually show that $Z(0) = Z_0$ in $H^{4-\varepsilon}$ for any $0 < \varepsilon < \frac{1}{2}$, since then $4 - \varepsilon < 4$ such that Aubin-Lions-Simon applies and $(4 - \varepsilon) - 2 > \frac{3}{2}$ such that we may still use Lemma 6.1.

We now state and prove our uniqueness result. Note that the only assumptions made are boundedness of appropriate Sobolev norms of the solutions. No smallness assumptions are made here.

THEOREM 6.5 (Uniqueness). *Suppose that, for $i = 1, 2$, $(u_i, p_i, \omega_i, J_i)$ are strong solutions of*

$$\begin{cases} \partial_t u_i + (u_i \cdot \nabla) u_i = (\nabla \cdot T)(u_i, p_i, \omega_i), \\ \nabla \cdot u_i = 0, \\ J_i (\partial_t \omega_i + (u_i \cdot \nabla) \omega_i) + \omega_i \wedge J_i \omega_i = 2 \operatorname{vec} T(u_i, p_i, \omega_i) + (\nabla \cdot M)(\omega_i) + \tau e_3, \text{ and} \\ \partial_t J_i + (u_i \cdot \nabla) J_i = [\Omega_i, J_i] \end{cases}$$

on some common time interval $(0, T)$ which agree initially, i.e. which agree at time $t = 0$. If J_1 is uniformly positive-definite, $p_i, \partial_t (u_i, \omega_i, J_i) \in L^2_T L^2$, $(u_i, \omega_i, J_i), \nabla (u_i, \omega_i, J_i) \in L^\infty_T L^\infty$, and $\partial_t J_1, \partial_t \omega_2 \in L^\infty_T L^\infty$, then these solutions coincide on $(0, T)$.

PROOF. This follows from simple energy estimates for the equations satisfied by the difference of the two solutions. The difference $(u, p, \omega, J) = (u_1 - u_2, p_1 - p_2, \omega_1 - \omega_2, J_1 - J_2)$ satisfies

$$\begin{cases} (\partial_t + u_1 \cdot \nabla) u = (\nabla \cdot T)(u, p, \omega) + f, & (6.19a) \\ \nabla \cdot u_1 = 0, & (6.19b) \\ (J_1 (\partial_t + u_1 \cdot \nabla) + \omega_1 \wedge J_1) \omega = 2 \operatorname{vec} T(u, p, \omega) + (\nabla \cdot M)(\omega) + g, & (6.19c) \\ (\partial_t + u_1 \cdot \nabla) J_1 = [\Omega_1, J_1], \text{ and} & (6.19d) \\ (\partial_t + u_1 \cdot \nabla) J = [\Omega, J] + h & (6.19e) \end{cases}$$

for

$$\begin{cases} f = -(u \cdot \nabla) u_2 \\ g = -J \partial_t \omega_2 - J_1 (u \cdot \nabla) \omega_2 - J (u_2 \cdot \nabla) \omega_2 - \omega_1 \wedge J \omega_2 - \omega \wedge J_2 \omega_2, \text{ and} \\ h = -(u \cdot \nabla) J_2 + [\Omega, J_2]. \end{cases}$$

We can thus multiply (6.19a), (6.19c), and (6.19e) by u, ω, and J respectively to see that, for every $0 < t < T$,

$$\int_{\mathbb{T}^3} \frac{1}{2}|u|^2 + \frac{1}{2} J_1 \omega \cdot \omega + \frac{1}{2}|J|^2 \bigg|_{s=t} - \int_{\mathbb{T}^3} \frac{1}{2}|u|^2 + \frac{1}{2} J_1 \omega \cdot \omega + \frac{1}{2}|J|^2 \bigg|_{s=0}$$
$$+ \int_0^t \int_{\mathbb{T}^3} \frac{\mu}{2}|\mathbb{D}u|^2 + 2\kappa \left| \frac{1}{2} \nabla \times u - \omega \right|^2 + \alpha |\nabla \cdot \omega|^2 + \frac{\beta}{2} |\mathbb{D}^0 \omega|^2 + 2\gamma |\nabla \times \omega|^2 = \int_0^t \int_{\mathbb{T}^3} f \cdot u + g \cdot \omega + h : J.$$

We can write this energy-dissipation-interaction relation more succintly as $\mathcal{E}(t) - \mathcal{E}(0) + \int_0^t \mathcal{D} = \int_0^t \mathcal{I}$ for $\mathcal{I} = \int_{\mathbb{T}^3} f \cdot u + g \cdot \omega + h : J$. It follows from straightforward application of the Hölder and Cauchy-Schwartz inequalities that the interactions are controlled by the energy, i.e. $|\mathcal{I}| \leqslant C\mathcal{E}$ for some constant $C > 0$. Note that since the two solutions agree initially we have that $\mathcal{E}(0) = 0$. Therefore the integral version of Gronwall's inequality tells us that $\mathcal{E}(t) = 0$ for all $0 < t < T$. Since J_1 is uniformly positive definite we deduce that $(u, \omega, J) = 0$. Finally, since $-\Delta p = \nabla u_1 : \nabla u^T + \nabla u : \nabla {u_2}^T = 0$, we conclude that indeed the two solutions coincide. □

Putting Theorem 6.3 and Theorem 6.5 together yields our local well-posedness result, stated below.

COROLLARY 6.6 (Local well-posedness). *The solution obtained in Theorem 6.3 is unique.*

PROOF. This is immediate since the assumptions of Theorem 6.3 ensure that Theorem 6.5 applies. □

6.2. Auxiliary results. Here we record auxiliary results which are used throughout the main body of the chapter. Whilst these results are typically either elementary lemmas or well-known theorems, they are of interest since they are applicable beyond the scope of this chapter.

Lemma 6.7 (Lower bound on the real part of complex square roots). *Let $x, y \in \mathbb{R}$ with $y \neq 0$ and let $\alpha > 0$. We follow the convention according to which the square root of a complex number with non-trivial imaginary part is chosen to have a strictly positive real part. Then* $\operatorname{Re}\sqrt{x+iy} > \alpha$ *if and only if* $x > \alpha^2 - \frac{y^2}{4\alpha^2}$.

Proof. Let us write $\sqrt{x+iy} = u + iv$ for some $u > 0$ and $v \in \mathbb{R}$, such that $x = u^2 - v^2$ and $y = 2uv$. What we wish to prove can then be written as $u > \alpha$ if and only if $u^2 - v^2 > \alpha^2 - \frac{u^2v^2}{\alpha^2}$. The latter inequality can be rearranged as $u^2 - \alpha^2 > -\frac{v^2}{\alpha^2}\left(u^2 - \alpha^2\right)$. This can be simplified, using the fact that $u + \alpha > 0$, to $(u-\alpha)\left(1 + \frac{v^2}{\alpha^2}\right) > 0$. This is indeed equivalent to $u > \alpha$ so we are done. □

Lemma 6.8 (Similarity of matrices acting on quotient spaces). *Let V be a subspace of $\mathbb{C}^n$ and let A, G, and H be complex n-by-n matrices which act on $\mathbb{C}^n/V$ (c.f. Definition 3.1) such that $GH = HG = \operatorname{proj}_{V^\perp}$. Then (1) $B := GAH$ acts on $\mathbb{C}^n/V$, (2) $A = HBG$, and (3) A and B are similar.*

Proof. First we show that B acts on $\mathbb{C}^n/V$. We know that $\operatorname{im} B \subseteq \operatorname{im} G \subseteq V^\perp$ and that $V = \ker H \subseteq \ker B$, so it is enough to show that $\ker B \subseteq V$. Let $x \in \ker B$. Since $Hx \in V^\perp$, it suffices to show that $Hx \in V$ as then $Hx = 0$, i.e. $x \in \ker H = V$. The key observation is that since $\operatorname{im} A \subseteq V^\perp$ and since G and H are inverses on $V^\perp$, we obtain that $A = HGA$. It follows that $AHx = HGAHx = HBx = 0$, i.e. $Hx \in \ker A = V$, and hence (1) holds.

Now observe that in order to prove that $A = HBG$ it is enough to show that $HGA = A$, which was done above, and that $AHG = A$, which we do now. Pick any $x \in \mathbb{C}^n$ and write $x = x_\parallel + x_\perp$ for $x_\parallel \in V$ and $X_\perp \in V^\perp$. Since $\ker G = \ker A = V$ and since $HG = \operatorname{proj}_{V^\perp}$ it follows that $AHGx = AHGx_\perp = Ax_\perp = Ax$, i.e. indeed $AHG = A$.

Finally we show that A and B are similar by explicitly finding an appropriate change-of-basis matrix. Let P be the orthogonal projection onto V, i.e. $\ker P = V^\perp$ and $P|_V = \operatorname{id}|_V$. Observe that, since $\ker B = V = \operatorname{im} P$ and since $\operatorname{im} B \subseteq V^\perp = \ker P$, we may deduce that $BP = PB = 0$. Therefore

$$(H+P)\,B\,(G+P) = A. \tag{6.20}$$

We will now show that $G+P$ and $H+P$ are invertible and $(G+P)^{-1} = H+P$, from which it follows that (6.20) witnesses (3). Let $x \in \ker(G+P)$ and let us write $x = x_\parallel + x_\perp$ as above. Then $0 = (G+P)\,x = Gx_\perp + x_\parallel$ with $Gx_\perp \in V^\perp$ and $x_\parallel \in V$, and hence we must have $Px_\perp = 0$ and $x_\parallel = 0$. In particular, since $\ker G = V$, we know that $x_\perp$ belongs to both V and $V^\perp$ and hence $x_\perp = 0$, such that $x = 0$. This shows that $G+P$ has trivial kernel and is thus invertible. We may deduce in exactly the same way that $H+P$ is invertible. To conclude we simply compute $(H+P)(G+P) = HG + HP + PG + P^2 = HG + P = I$. □

Lemma 6.9 (Bounds on the real parts of the eigenvalues of a matrix using the spectrum of its symmetric part). *Let S and A be symmetric and antisymmetric real n-by-n matrices respectively. It then holds that* $\min\sigma(S) \leqslant \operatorname{Re}\sigma(S+A) \leqslant \max\sigma(S)$.

Proof. Let us denote by λ_+ and λ_- the maximal and minimal eigenvalues of S, respectively, let us define $M = S + A$, and let $a + ib$, $a, b \in \mathbb{R}$, be an eigenvalue of M with eigenvector $x + iy$, $x, y \in \mathbb{R}^n$. Then, since $M(x+iy) = (a+ib)(x+iy)$ it follows that $Mx = ax - by$ and $My = bx + ay$. In particular $Sx\cdot x + Sy\cdot y = Mx\cdot x + My\cdot y = a\left(|x|^2 + |y|^2\right)$ where $Sx\cdot x + Sy\cdot y \leqslant \lambda_+\left(|x|^2 + |y|^2\right)$, and therefore $a \leqslant \lambda_+$. We may obtain in exactly the same way that $a \geqslant \lambda_-$, and hence indeed $\lambda_- \leqslant \operatorname{Re}\sigma(S+A) \leqslant \lambda_+$. □

Theorem 6.10 (Gershgorin disk theorem). *Let A be a complex n-by-n matrix and let $R_i := \sum_{j\neq i}|A_{ij}|$ for $i = 1, \ldots, n$. Every eigenvalue of A lies in one of the closed disks $\overline{B(A_{ii}, R_i)}$, where $i = 1, \ldots, n$. These disks are called the* Gershgorin disks *of A.*

Proof. Let v be an eigenvector of A with eigenvalue λ. Without loss of generality (otherwise we may divide v by $\pm\|v\|_\infty$): $v_i = 1$ for some index i and $|v_j| \leqslant 1$ for all indices j different from i. Now observe that

$$(Av)_i = \lambda v_i \quad \Leftrightarrow \quad A_{ii}v_i + \sum_{j\neq i} A_{ij}v_j = \lambda v_i \quad \Leftrightarrow \quad \lambda - A_{ii} = \sum_{j\neq i} A_{ij}v_j$$

and thus $|\lambda - A_{ii}| \leqslant \sum_{j \neq i} |A_{ij}||v_j| \leqslant \sum_{j \neq i} |A_{ij}| = R_i$ i.e. indeed λ lies in $\overline{B(A_{ii}, R_i)}$, which is one of the Gershgorin disks of A. $\square$

Corollary 6.11 (Bounds on the imaginary parts of the eigenvalues of a matrix using the Frobenius norm of its antisymmetric part). *Let S and A be symmetric and antisymmetric real n-by-n matrices respectively. Then $|\mathrm{Im}\, \sigma(S+A)| \leqslant \sqrt{n-1} ||A||_2$, where $||A||_2 := \sqrt{A : A}$ is the* Frobenius norm *of A.*

Proof. Since S is symmetric, there exists an orthogonal matrix Q and a diagonal matrix D such that $QSQ^T = D$. Therefore $Q(S+A)Q^T = D + QAQ^T$. In particular, for $\tilde{A} := QAQ^T$, we know that $S+A$ and $D+\tilde{A}$ have the same spectrum. Writing $D = \mathrm{diag}(\lambda_1, \dots, \lambda_n)$ where the λ_i's are the eigenvalues of S, we may apply Theorem 6.10 to deduce that the eigenvalues of $D + \tilde{A}$ lie within closed disks centered at λ_i (since $\tilde{A}$ is antisymmetric and hence all its diagonal entries are equal to zero) and with corresponding radii $R_i = \sum_{j \neq i} |\tilde{A}_{ij}| \leqslant \sqrt{n-1} ||\tilde{A}||_2$. The result then follows from the observation that the eigenvalues λ_i of the symmetric matrix S are real and the fact that $||\tilde{A}||_2^2 = QAQ^T : QAQ^T = Q^T QAQ^T Q : A = ||A||_2^2$. $\square$

Lemma 6.12 (Bounds on matrix exponentials using the symmetric part). *Let M be a real n-by-n matrix, let $S := \frac{1}{2}\left(M + M^T\right)$ denote its symmetric part, and let σ denote the largest eigenvalue of S. Then, for every $t > 0$, $\left|\left|e^{tM}\right|\right|_{\mathcal{L}(l^2, l^2)} \leqslant e^{\sigma t}$.*

Proof. This follows from a simple Gronwall inequality upon noticing that, for any $x \in \mathbb{R}^n$, $Mx \cdot x = Sx \cdot x$. More precisely: pick any $x_0 \in \mathbb{R}^n$ and define $x(t) := e^{tM} x_0$ for every $t \geqslant 0$. Observe that $\frac{\mathrm{d}}{\mathrm{d}t} x(t) = Mx(t)$ and hence $\frac{\mathrm{d}}{\mathrm{d}t} ||x(t)||_2^2 = 2Sx(t) \cdot x(t) \leqslant 2\sigma ||x(t)||_2^2$. Since $x(0) = x_0$, applying Gronwall's inequality yields that, for every $t \geqslant 0$, $\left|\left|e^{tM} x_0\right|\right|_2^2 = ||x(t)||_2^2 \leqslant e^{2\sigma t} ||x_0||_2^2$, from which the result follows. $\square$

Lemma 6.13 (Bounds on matrix exponentials for Jordan canonical forms). *For any matrix norm $|\cdot|$ there exists a constant $C_n > 0$ such that for every complex n-by-n matrix M in Jordan canonical form, if $\eta := \max \mathrm{Re}\, \sigma(M)$ then, for every $t \geqslant 0$, $|e^{tM}| \leqslant C_n (1 + t^n) e^{\eta t}$.*

Proof. Since M is in Jordan canonical form it can be written as $M = J_{a_1}(\lambda_1) \oplus \cdots \oplus J_{a_k}(\lambda_k)$ where the λ_i's are eigenvalues of M and $J_a(\lambda) = \lambda I_a + N_a$ for $(N_a)_{ij} = 1$ if $j = i+1$ and $(N_a)_{ij} = 0$ otherwise, Note that, since N_a is an a-by-a matrix whose only non-zero entries are those immediately above the diagonal, it is nilpotent of order a. In particular, note that since the identity commutes with all matrices, it follows that $e^{J_a(\lambda)} = e^{\lambda} e^{N_a}$, and recall that for any nilpotent matrix N of order q its matrix exponential is given by a finite sum, i.e. $e^N = \sum_{j=0}^{q-1} \frac{1}{j!} N^j$. We can thus compute the matrix exponential of M to be $e^{tM} = e^{\lambda_1 t} e^{tN_{a_1}} \oplus e^{\lambda_k t} e^{tN_{a_k}}$ which can be estimated by $|e^{tM}| \leqslant \sum_{i=1}^{k} e^{(\mathrm{Re}\, \lambda_i)t} \left| \sum_{j=0}^{a_i} \frac{1}{j!} (tN_{a_i})^j \right| \lesssim e^{\eta t}(1 + t^n)$ where have used that polynomials of degree q in a real variable x can be bounded above (up to a constant) by $1 + x^q$, and where the constants up to which the inequalities above hold only depends on n and the choice of the matrix norm. $\square$

Corollary 6.14 (Bounds on matrix exponentials). *Let M be a real n-by-n matrix and let $\eta := \max \mathrm{Re}\, \sigma(M)$. For any matrix norm $|\cdot|$ there exists a constant $C = C(M) > 0$ such that, for every $t \in \mathbb{R}$, it holds that $|e^{tM}| \leqslant C(1 + t^n) e^{\eta t}$.*

Proof. This follows from Lemma 6.13 since every matrix M is similar to a matrix in Jordan canonical form. The constant obtained depends on M since the norm of the matrices used to conjugate M to put it in Jordan canonical form depend on M. $\square$

Proposition 6.15 (Construction of a semigroup via matrix exponentials as Fourier multipliers). *Let $M : \mathbb{Z}^n \to \mathbb{R}^{l \times l}$ be a family of matrices for which there exists $\eta \in \mathbb{R}$ and $C_F > 0$ such that, for every $k \in \mathbb{Z}^n$ and every $t > 0$,*

$$\left|\left|e^{tM(k)}\right|\right|_{\mathcal{L}(l^2,\, l^2)} \leqslant C_F e^{\eta t}. \tag{6.21}$$

For any $t \geqslant 0$ the operator $e^{t\mathcal{L}}$ defined by the multiplier $\left(e^{t\mathcal{L}}\right)^{\wedge}(k) := e^{tM(k)}$ is a bounded operator on $L^2\left(\mathbb{T}^n; \mathbb{R}^l\right)$ such that $\left(e^{t\mathcal{L}}\right)_{t \geqslant 0}$ defines an η-contractive semigroup, i.e.

(1) *$e^{0\mathcal{L}}$ is the identity,*
(2) *for every $t, s \geqslant 0$, $e^{t\mathcal{L}} e^{s\mathcal{L}} = e^{s\mathcal{L}} e^{t\mathcal{L}} = e^{(t+s)\mathcal{L}}$,*

(3) *for every* $f \in L^2(\mathbb{T}^n; \mathbb{R}^l)$, $t \to e^{t\mathcal{L}} f$ *is a continuous map from* $[0, \infty)$ *to* $L^2(\mathbb{T}^n; \mathbb{R}^l)$, *and*
(4) *for every* $r \geqslant 0$, $\left|\left|e^{t\mathcal{L}}\right|\right|_{\mathcal{L}(H^r(\mathbb{T}^n;\mathbb{R}^l); H^r(\mathbb{T}^n;\mathbb{R}^l))} \leqslant C_F e^{\eta t}$.

Moreover, let us write $v = (v_1, \dots, v_p) \in \mathbb{R}^{q_1} \times \cdots \times \mathbb{R}^{q_p}$, *where* $q_1 + \cdots + q_p = l$, *and suppose that there exists* $\alpha_1, \dots, \alpha_p \in \mathbb{N}$ *and* $C_D > 0$ *such that for every* $k \in \mathbb{Z}^n$ *and every* $v \in \mathbb{R}^l$,

$$|M(k)\, v|^2 \leqslant C_D \sum_{i=1}^{p} \langle k \rangle^{2\alpha_i} |v_i|^2. \tag{6.22}$$

Then

(5) *the domain of the semigroup* $\left(e^{t\mathcal{L}}\right)_{t \geqslant 0}$ *contains* $H^{\alpha_1}(\mathbb{T}^n, \mathbb{R}^{q_1}) \times \cdots \times H^{\alpha_p}(\mathbb{T}^n, \mathbb{R}^{q_p})$ *and*
(6) *its generator is the linear differential operator* $\mathcal{L}$ *with symbol* M, *i.e.* $\widehat{\mathcal{L}}(k) := M(k)$.

Proof. The boundedness of $e^{t\mathcal{L}}$ and (4) follow directly from (6.21). (1) and (2) follow from the fact that, for any matrix M, $\left(e^{tM}\right)_{t \geqslant 0}$ is a representation of the semigroup $(\mathbb{R}_{\geqslant 0}, +)$, i.e. $e^{0M} = I$ and $e^{tM} e^{sM} = e^{sM} e^{tM} = e^{(t+s)M}$. To prove that (3) holds it suffices to show that $t \mapsto e^{t\mathcal{L}} f$ is continuous at $t = 0$. This is immediate since

$$\left|\left|e^{t\mathcal{L}} f - f\right|\right|_{L^2} \leqslant \sum_{|k| \leqslant K} \left|\left(e^{tM_k} - I\right) \hat{f}(k)\right|^2 + \left(e^{\eta t} + 1\right)^2 \underbrace{\sum_{|k| > K} |\hat{f}(k)|^2}_{=: R_f(K)}$$

where $R_f(K) \to 0$ as $K \to \infty$ since $f \in L^2$, and hence, since for any fixed K the collection $\left\{t \mapsto e^{tM_k}\right\}_{|k| \leqslant K}$ is as finite collection of continuous maps, we indeed obtain that $e^{t\mathcal{L}} f \to f$ in L^2 as $t \to 0$.

Finally, to prove (5) and (6) we proceed as we did for (3). First we note that, by the mean-value theorem, for every $k \in \mathbb{Z}^n$ and every $t > 0$, $\frac{e^{tM_k} - I}{t} - M_k = \int_0^1 \left(e^{stM_k} - I\right) M_k ds$. Therefore, for any $f \in L^2$ and any $0 < t < \delta$, if we write $\hat{f} = \left(\hat{f}_1, \dots, \hat{f}_p\right) \in \mathbb{R}^{q_1} \times \cdots \times \mathbb{R}^{q_p}$ then

$$\begin{aligned}
\left|\left|\frac{e^{t\mathcal{L}} f - f}{t} - \mathcal{L} f\right|\right|_{L^2}^2 &\leqslant \sum_{k \in \mathbb{Z}^l} \left|\left|\int_0^1 \left(e^{stM_k} - I\right) ds\right|\right|_{\mathcal{L}(l^2,\, l^2)}^2 \left|M_k \hat{f}(k)\right|^2 \\
&\leqslant C(K, f) \sum_{|k| < K} \left|\left|\int_0^1 \left(e^{stM_k} - I\right) ds\right|\right|_{\mathcal{L}(l^2,\, l^2)}^2 + C(\eta, \delta) \underbrace{\sum_{|k| > K} \sum_{i=1}^{p} \langle k \rangle^{2\alpha_i} \left|\hat{f}_i(k)\right|^2}_{=: H_f(K)}
\end{aligned}$$

In particular, if $f \in H^{\alpha_1} \times \cdots \times H^{\alpha_p}$ then $H_f(K) \to 0$ as $K \to \infty$ and thus, since, for any fixed K, $\left\{t \mapsto e^{tM_k}\right\}_{|k| \leqslant K}$ is as finite collection of continuous maps, we may conclude that indeed $\frac{e^{t\mathcal{L}} f - f}{t} \to \mathcal{L} f$ in L^2 as $t \to 0$. ☐

Theorem 6.16 (Rouché). *Let* $\Omega \subseteq \mathbb{C}$ *be a connected open set whose boundary is a simple curve and let* f *and* g *be holomorphic in* Ω. *If* $|f - g| < |f|$ *on* $\partial\Omega$ *then* f *and* g *have the same number of zeros in* Ω.

Proof. See Chapter 4 of [**Ahl78**]. ☐

Theorem 6.17 (Implicit Function Theorem for mixed real-complex functions). *Let* $f : \mathcal{O} \subseteq \mathbb{C} \times \mathbb{R}^m \to \mathbb{C}$, *where* $\mathcal{O}$ *is open, be continuously differentiable in the real sense (i.e. after identifying* $\mathbb{C}$ *with* $\mathbb{R}^2$ *in the canonical way) is continuously differentiable. Let* $(z_0, v_0) \in \mathcal{O}$ *and let us write* $f = f(z, v)$ *for* $z \in \mathbb{C}$ *and* $v \in \mathbb{R}^m$. *If (1)* $f(z_0, v_0) = 0$ *and (2)* $\partial_z f(z_0, v_0) \neq 0$ *then there exist open sets* $\mathcal{U} \subseteq \mathbb{C} \times \mathbb{R}^m$ *and* $W \subseteq \mathbb{R}^m$ *and a function* $g : W \to \mathbb{C}$ *which is continuously differentiable in the real sense such that (1)* $(z_0, v_0) \in \mathcal{U}, v_0 \in W$, *(2)* $g(v_0) = z_0$, *(3)* $(g(v), v) \in \mathcal{U}$ *for every* $v \in W$, *(4)* $f(g(v), v) = 0$ *for every* $v \in W$ *and*

$$\nabla_v g(v_0) = \frac{-\nabla_v f(z_0, v_0)}{\partial_z f(z_0, v_0)}.$$

Moreover, if f *is more regular, in the real sense, then so is* g.

Proof. See Chapter 9 of [**Rud76**]. ☐

LEMMA 6.18 (Coercivity implies invertibility and bounds on the inverse). *Let B be a real n-by-n matrix. If B is coercive, i.e. if there exists $C_0 > 0$ such that for every $x \in \mathbb{R}^n$, $|Bx| \geqslant C_0|x|$, then B is invertible and $\left|\left|B^{-1}\right|\right|_{op} \leqslant \frac{1}{C_0}$.*

PROOF. Observe that since B is coercive, it has trivial kernel, and is hence invertible. To obtain the bound on the operator norm of B^{-1} simply observe that for every $y \in \mathbb{R}^n$, $|y| = |BB^{-1}y| \geqslant C_0|B^{-1}y|$. □

COROLLARY 6.19 (Invertibility and bounds for perturbations of the identity). *Let B be a real n-by-n matrix. If $||B||_{op} < 1$ then $I + B$ is invertible and $\left|\left|(I+B)^{-1}\right|\right|_{op} \leqslant \frac{1}{1-||B||_{op}}$.*

PROOF. The key observation is that $I + B$ is coercive with coercivity constant $1 - ||B||_{\text{op}}$. The result then follows from Lemma 6.18. □

LEMMA 6.20. *Let A and N be real n-by-n matrices such that N is normal, i.e. $NN^T = N^TN$. Then $[A, N] : N = 0$.*

PROOF. This follows from a direct computation: $NA : N = A : N^TN = A : NN^T = AN : N$ and hence $[A, N] : N = AN : N - NA : N = 0$. □

PROPOSITION 6.21 (Korn inequality). *There exists $C_K > 0$ such that for every $u \in H^1\left(\mathbb{T}^3, \mathbb{R}^3\right)$, $||\nabla u||_{L^2} \leqslant C_K \left(||u||_{L^2} + ||\mathbb{D}u||_{L^2}\right)$.*

PROOF. See Lemma IV.7.6 in [BF13]. □

PROPOSITION 6.22 (Korn-Poincaré inequality). *There exists $C_{KP} > 0$ such that for every $u \in H^1\left(\mathbb{T}^3, \mathbb{R}^3\right)$, $||u||_{L^2} \leqslant C_{KP}\left(\left|⨍\, u\right| + ||\mathbb{D}u||_{L^2}\right)$.*

PROOF. This is a consequence of Proposition 6.21 – see for example Lemma IV.7.7 in [BF13] – noting that $\nabla \times u$ has average zero on the torus. □

LEMMA 6.23 (A div-curl identity on the torus). *For any $v \in H^1\left(\mathbb{T}^3, \mathbb{R}^3\right)$, it holds that $||\nabla v||_{L^2}^2 = ||\nabla \cdot v||_{L^2}^2 + ||\nabla \times v||_{L^2}^2$.*

PROOF. The key observation is that for any $w \in \mathbb{R}^3$ and any nonzero $k \in \mathbb{Z}^3$, $w \mapsto \frac{k \times w}{|k|}$ is an isometry on $\text{span}_k^\perp$, and hence $|w|^2 = |\text{proj}_k\, w|^2 + |\text{proj}_{k^\perp}\, w|^2 = \frac{|k \cdot w|^2}{|k|^2} + \frac{|k \times w|^2}{|k|^2}$. Combining this observation with Parseval's identity allows us to conclude:

$$||\nabla v||_{L^2}^2 = \sum_{k \in \mathbb{Z}^3} |k \otimes \hat{v}\,(k)|^2 = \sum_{k \in \mathbb{Z}^3 \setminus \{0\}} |k|^2 |\hat{v}\,(k)|^2 = \sum_{k \in \mathbb{Z}^3} |k \cdot \hat{v}\,(k)|^2 + \sum_{k \in \mathbb{Z}^3} |k \times \hat{v}\,(k)|^2 = ||\nabla \cdot v||_{L^2}^2 + ||\nabla \times v||_{L^2}^2.$$

□

PROPOSITION 6.24 (Estimates from the Faà di Bruno formula). *Let $\mathcal{U} \subseteq \mathbb{R}^n$ and $\mathcal{V} \subseteq \mathbb{R}^p$ be open and let $g : \mathcal{U} \to \mathcal{V}$ and $F : \mathcal{V} \to \mathbb{R}^q$ be k-times differentiable. There exists a constant $C = C\,(n, p, q, k) > 0$ which does not depend on F or g such that, for every $x \in \mathcal{U}$,*

$$\left|\nabla^k\,(F \circ g)\,(x)\right| \leqslant C \sum_{i=1}^{k} \left|\nabla^i F\,(g\,(x))\right| \sum_{\pi \in P_i(k)} |\nabla^\pi g\,(x)|\,.$$

PROOF. This estimate follows immediately from the Faà di Bruno formula, which was first proven in [Arb00] and can be found in a rather clean form in [Har06]. □

LEMMA 6.25 (Post-compositions by analytic functions are analytic). *Suppose that $F : \mathbb{R}^k \to \mathbb{R}^l$ is analytic about zero and let $s > \frac{n}{2}$. There exists $\delta > 0$ such that $F^* : H^s_\delta\left(\mathbb{T}^n; \mathbb{R}^k\right) \to H^s\left(\mathbb{T}^n; \mathbb{R}^l\right)$, defined by $F^*\,(G) = F \circ G$ for every $G \in H^s_\delta$, is analytic.*

PROOF. Let $\delta = \frac{R}{C_s}$ where R is the radius of convergence of F about zero and C_s is the constant from the continuous embedding $H^s \cdot H^s \hookrightarrow H^s$ and suppose that $F\,(x) = \sum_{i=0}^\infty F_i \bullet X^{\otimes i}$ for every $x \in B\,(0, R)$, for some fixed tensorial coefficients F_i. Then indeed, for every $G \in H^s_\delta$, $F^*\,(G) = \sum_{i=0}^\infty F_i \bullet G^{\otimes i}$ with

$$\sum_{i=0}^{\infty} |F_i| \left|\left|G^{\otimes i}\right|\right|_{H^s} \leqslant \sum_{i=0}^{\infty} |F_i|\, C_s^i\, ||G||_{H^s}^i \leqslant \sum_{i=0}^{\infty} |F_i|\, R^i < \infty.$$

□

LEMMA 6.26 (Formula for the Leray projector and its complement). *Let $\mathbb{P}_L$ denote the Leray projector on the torus. Then $\mathbb{P}_L = -\nabla \times \Delta^{-1} \nabla\times$ and $I - \mathbb{P}_L = \nabla\Delta^{-1}\nabla\cdot$.*

PROOF. This is immediate since $\hat{\mathbb{P}}_L(0) = I$ and $\hat{\mathbb{P}}_L(k) = I - \frac{k\otimes k}{|k|^2}$ if $k \neq 0$ and since $k \times k \times \cdot = |k|^2 - k \otimes k$. □

6.3. Derivation of the perturbative energy-dissipation relation. In this section we derive the energy-dissipation relation (2.1), which is satisfied by solutions of (1.1a)–(1.1d). First recall that the Cauchy stress tensor and the couple stress tensor are denoted by T and M respectively. We will write $T_{eq} = -\kappa\Omega_{eq}$ for the equilibrium version of the stress tensor. For simplicity we will also write $D_t := \partial_t + u \cdot \nabla$ for the advective derivative. The conservation of linear momentum (1.1a) can then be written as $D_t u = \nabla \cdot T$ such that multiplying by u yields

$$\frac{\mathrm{d}}{\mathrm{d}t}\int_{\mathbb{T}^3} \frac{1}{2}|u|^2 = \int_{\mathbb{T}^3} D_t\left(\frac{1}{2}|u|^2\right) = \int_{\mathbb{T}^3} D_t u \cdot u = \int_{\mathbb{T}^3} (\nabla \cdot T)\cdot u = -\int_{\mathbb{T}^3} T : \nabla u. \tag{6.23}$$

Similarly, the conservation of angular momentum (1.1c) can be written as

$$JD_t\omega + [\Omega, J]\,\omega = 2\,\mathrm{vec}\,(T - T_{eq}) + \nabla \cdot M$$

and hence multiplying by $\omega - \omega_{eq}$ yields

$$JD_t\omega \cdot (\omega - \omega_{eq}) + [\Omega, J]\,\omega \cdot (\omega - \omega_{eq}) = 2\,\mathrm{vec}\,(T - T_{eq}) \cdot (\omega - \omega_{eq}) + (\nabla \cdot M) \cdot (\omega - \omega_{eq})\,. \tag{6.24}$$

The right-hand side of (6.24) is dealt with in the usual way:

$$\int_{\mathbb{T}^3} 2\,\mathrm{vec}\,(T - T_{eq}) \cdot (\omega - \omega_{eq}) + (\nabla \cdot M) \cdot (\omega - \omega_{eq}) = \int_{\mathbb{T}^3} (T - T_{eq}) : (\Omega - \Omega_{eq}) - M : \nabla(\omega - \omega_{eq})\,. \tag{6.25}$$

Dealing with the left-hand side of (6.24) requires further rearranging. Using the fact that the conservation of micro-inertia (1.1d) can be written as $D_t J = [\Omega, J]$ and adding and subtracting $\frac{1}{2}D_t J\,(\omega - \omega_{eq}) \cdot (\omega - \omega_{eq})$ yields

$$JD_t\omega \cdot (\omega - \omega_{eq}) + [\Omega, J]\,\omega \cdot (\omega - \omega_{eq}) = D_t\left(\frac{1}{2}J\,(\omega - \omega_{eq}) \cdot (\omega - \omega_{eq})\right) + \frac{1}{2}D_t J\,(\omega + \omega_{eq}) \cdot (\omega - \omega_{eq})\,. \tag{6.26}$$

The key observation that allows us to conclude is the identity $[\Omega, J]\,(\omega + v) \cdot (\omega - v) = -[\Omega, J]\,v \cdot v$ for every $v \in \mathbb{R}^3$. Combining this identity with $D_t J = [\Omega, J]$ tells us that

$$\frac{1}{2}D_t J\,(\omega + \omega_{eq}) \cdot (\omega - \omega_{eq}) = -\frac{1}{2}\,(D_t J)\,\omega_{eq} \cdot \omega_{eq} = -D_t\left(\frac{1}{2}J\omega_{eq} \cdot \omega_{eq}\right). \tag{6.27}$$

Finally: combining (6.25), (6.26), and (6.27) yields

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(\int_{\mathbb{T}^3} J\,(\omega - \omega_{eq}) \cdot (\omega - \omega_{eq}) - \frac{1}{2}J\omega_{eq} \cdot \omega_{eq}\right) = \int_{\mathbb{T}^3} (T - T_{eq}) : (\Omega - \Omega_{eq}) - M : \nabla(\omega - \omega_{eq})\,.$$

Adding this equation to (6.23) yields the energy-dissipation relation (2.1).

6.4. The 8-by-8 matrix M in all its glory. In this section we record the matrix M_k in an explicit form. Recall that M_k is introduced in Section 3.2, and is written there in a compact form well-suited to the analysis of its spectrum. However, in order to compute the characteristic polynomial of M, we employed the assistance of a symbolic algebra package, and this thus requires providing an explicit form of the matrix M_k. M_k can be written in block form as

$$M_k = \begin{pmatrix} A & B & 0_{3\times 2} \\ B^T & C & D \\ 0_{2\times 3} & E & F \end{pmatrix}$$

where

$$A = -\left(\mu+\kappa/2\right)\left(|k|^2 I_3 - k\otimes k\right) = \left(\mu+\kappa/2\right)\begin{pmatrix} -k_2^2-k_3^2 & k_1k_2 & k_1k_3 \\ k_1k_2 & -k_1^2-k_3^2 & k_2k_3 \\ k_1k_3 & k_2k_3 & -k_2^3-k_3^2 \end{pmatrix},$$

$$B = \frac{\kappa}{|k|}\left(|k|^2 I_3 - k\otimes k\right)\operatorname{diag}\left(\lambda^{-1/2},\lambda^{-1/2},\nu^{-1/2}\right)$$

$$= \frac{\kappa}{\sqrt{k_1^2+k_2^2+k_3^2}}\begin{pmatrix} \left(k_2^2+k_3^2\right)/\sqrt{\lambda} & -k_1k_2/\sqrt{\lambda} & -k_1k_3/\sqrt{\nu} \\ -k_1k_2/\sqrt{\lambda} & \left(k_1^2+k_3^2\right)/\sqrt{\lambda} & -k_2k_3/\sqrt{\nu} \\ -k_1k_3/\sqrt{\lambda} & -k_2k_3/\sqrt{\lambda} & \left(k_1^2+k_2^2\right)/\sqrt{\nu} \end{pmatrix}$$

$$D = \frac{\tau}{2\kappa}\sqrt{1-\frac{\nu}{\lambda}}\begin{pmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \end{pmatrix},\ E = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix},\ F = \frac{\tau}{2\kappa}\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$$

and

$$C = -\operatorname{diag}\left(\lambda^{-1/2},\lambda^{-1/2},\nu^{-1/2}\right)\left(2\kappa I_3 + \left(\alpha+\beta/3-\gamma\right)k\otimes k + \left(\beta+\gamma\right)|k|^2 I_3\right)\operatorname{diag}\left(\lambda^{-1/2},\lambda^{-1/2},\nu^{-1/2}\right)$$
$$-\left(1-\frac{\nu}{\lambda}\right)\frac{\tau}{2\kappa}\left(e_2\otimes e_1 - e_1\otimes e_2\right)$$

such that

$$C_{11} = -\lambda^{-1}\left(2\kappa + \left(\alpha+4\beta/3\right)k_1^2 + \left(\beta+\gamma\right)\left(k_2^2+k_3^2\right)\right),\quad C_{12} = -\lambda^{-1}\left(\alpha+\beta/3-\gamma\right)k_1k_2 + \frac{\tau}{2\kappa}\left(1-\frac{\nu}{\lambda}\right),$$
$$C_{22} = -\lambda^{-1}\left(2\kappa + \left(\alpha+4\beta/3\right)k_2^2 + \left(\beta+\gamma\right)\left(k_1^2+k_3^2\right)\right),\quad C_{21} = -\lambda^{-1}\left(\alpha+\beta/3-\gamma\right)k_1k_2 - \frac{\tau}{2\kappa}\left(1-\frac{\nu}{\lambda}\right),$$
$$C_{33} = -\nu^{-1}\left(2\kappa + \left(\alpha+4\beta/3\right)k_3^2 + \left(\beta+\gamma\right)\left(k_1^2+k_2^2\right)\right),\quad C_{13} = C_{31} = -\lambda^{-1/2}\nu^{-1/2}\left(\alpha+\beta/3-\gamma\right)k_1k_3,$$
$$C_{23} = C_{32} = -\lambda^{-1/2}\nu^{-1/2}\left(\alpha+\beta/3-\gamma\right)k_2k_3.$$

CHAPTER 3

The viscous surface wave problem with generalized surface energies

Abstract.

We study a three-dimensional incompressible viscous fluid in a horizontally periodic domain with finite depth whose free boundary is the graph of a function. The fluid is subject to gravity and generalized forces arising from a surface energy. The surface energy incorporates both bending and surface tension effects. We prove that for initial conditions sufficiently close to equilibrium the problem is globally well-posed and solutions decay to equilibrium exponentially fast, in an appropriate norm. Our proof is centered around a nonlinear energy method that is coupled to careful estimates of the fully nonlinear surface energy.

Contents

1. Introduction

In this chapter we study the dynamics of a three-dimensional periodic layer of viscous incompressible fluid bounded below by a rigid interface and above by a moving free boundary. The free boundary is advected with the fluid, but the configuration of the free boundary gives rise to surface stresses that act as forcing terms on the fluid. In this introductory section we discuss the origin and nature of the surface stresses and then record the equations of motion.

1.1. Surface energies. We will restrict our attention in this chapter to surface stresses that are generated as generalized forces associated to an energy functional that depends on the configuration of the surface. Here the generalized force is understood in the sense that it is the negative gradient of the energy. The classical example of such a force is surface tension, which is associated to the energy functional given by a constant multiple of the area functional (the constant is known as the coefficient of surface tension). The generalized force is then the mean curvature operator, which is the trace of the second fundamental form. We can account for higher-order geometric effects by considering more general functionals depending on the second fundamental form itself. The classical example of such an energy is the Willmore functional, which is the square of the mean curvature integrated over the surface. Our goal here is to briefly survey the vast literature associated to Willmore-type energies and their relation to interfacial mechanics.

The Willmore energy was popularized in the differential geometry literature by Willmore's initial work on it [**Wil65**] and in his books on Riemannian geometry [**Wil82**, **Wil93**]. Willmore also formulated the so-called Willmore conjecture, which predicted the minimizers of the energy among immersed tori. The Willmore conjecture was proved recently by Marques–Neves [**MN14**]. Critical points of the Willmore energy remain an active topic of study in geometric analysis and PDE: for example, Kuwert–Schätzle [**KS04**] studied removable singularities, Rivière [**Riv08**] developed a theory of weak Willmore immersions, and Bernard–Rivière [**BR14**] proved results about energy quantization and compactness.

Remarkably, energies of Willmore-type arise naturally in many areas of applied mathematics, and so such energies have received much attention outside of differential geometry. Roughly speaking, one can justify the widespread appearance of Willmore-type energies in applications through the lens of dimension reduction in elasticity. In many applications one considers a thin three-dimensional elastic material. When the size of the thin direction is very small relative to the two other directions, then it is natural to seek an effective two-dimensional model, thereby reducing the dimension. A rigorous derivation of Willmore-type energies as $\Gamma-$limits of three-dimensional elastic energies was carried out by Friesecke–James–Müller [**FMJ02**, **FJM02**] for plates and Friesecke–James–Mora–Müller [**FJMM03**] for shells.

One major area of interest in these energies is the study of biological membranes, lipid bi-layers, and vesicles. All of these structures can be thought of as very thin elastic materials, and should thus have some relation to Willmore-type energies. In [**Hel73**] Helfrich introduced such an energy to model the structure of lipid bi-layers, which led to these energies being standard modeling tools in membrane biology. More recent advances have considered coupled models of fluid-membrane dynamics: Du–Li–Liu [**DLL07**] and Du–Liu–Ryham–Wang [**DLRW09**] used phase field models to model fluid dynamics coupled to vesicles, Farshbaf-Shaker–Garcke [**FSG11**] developed thermodynamically consistent higher order phase field models, and Ryham–Klotz–Yao–Cohen [**RKYC16**] used Willmore-type energies to study the energetics of membrane fusion.

The coupling of the full fluid equations to surface stresses generated by Willmore-type energies presents numerous analytical challenges. Cheng–Coutand–Shkoller [**CCS07**] proved a local existence result for a viscous fluid coupled to a nonlinear elastic biofluid shell, and Cheng–Shkoller [**CS10**] proved local existence for a model with a Koiter shell. Local existence results for similar models related to hemodynamics were proved by Muha–Ĉanić [**MC15**, **MC16**]. We refer to the work of Bonito–Nochetto–Pauletti [**BNP11**] and Barett–Garcke–Nürnberg [**BGN17**] and the references contained therein for a discussion of the numerical analysis of such models.

A second major area of interest in energies of this type is the study of thin layers of ice, which can be thought of as thin elastic materials. We refer to the book by Squire–Hosking–Kerr–Langhorne [**SHKL96**] and the references therein for an overview of the physics specific to ice sheets. We refer to the work of Plotnikov–Toland [**PT11**] for a discussion of how the Willmore functional is related to bending energies for thin sheets of ice. The question of how fluids couple to the dynamics of ice sheets has attracted much attention in recent years, though most attention has focused on inviscid fluids. Solitary and traveling wave solutions and effective equations were studied by Milewski–Vanden-Broeck–Wang [MVBW11], Wang–Vanden-Broeck–Milewski [WVBM13], and Trichtchenko–Milewski–Parau–Vanden-Broeck [TMPVB19] in two dimensions and by Milewski–Wang [MW13] and Trichtchenko–Parau–Vandedn-Broeck–Milewski [TPVBM18] in three dimensions. For two-dimensional irrotational two-fluid flows, Liu–Ambrose [LA17] proved well-posedness and Akers–Ambrose–Sulon [AAS17] constructed traveling wave solutions for a two-fluid model. The one-fluid model was studied by Ambrose–Siegel [AS17].

Interestingly, Willmore-type energies also appear in other applications with no clear connection to thin elastic structures. In [Rub17] Rubinstein details how the energy appears in optics in questions related to optimal lens design. Hawking [Haw68] also introduced a Willmore-like energy in his study of gravitational radiation.

1.2. Examples of surface energies. In this chapter we are concerned with periodic slab-like geometries, which in particular means that we restrict our attention to surfaces given as the graph of a function $\eta : \mathbb{T}^2 \to \mathbb{R}$, where $\mathbb{T}^2 = \mathbb{R}^2/\mathbb{Z}^2$ is the usual 2−torus. This has the benefit of significantly simplifying the differential geometry of the surface. The area element, the unit normal, and the shape operator (the matrix in coordinates whose trace is the mean curvature) are then, respectively,

$$\sqrt{1+|\nabla\eta|^2},\quad \frac{(-\nabla\eta,1)}{\sqrt{1+|\nabla\eta|^2}},\ \text{and}\ \frac{1}{\sqrt{1+|\nabla\eta|^2}}\left(I-\frac{\nabla\eta\otimes\nabla\eta}{1+|\nabla\eta|^2}\right)\nabla^2\eta. \tag{1.1}$$

We consider generalized Willmore-type energies that depend both on $\nabla\eta$ and $\nabla^2\eta$, which allows for a combination of surface stresses of surface tension and bending type. We specify the energy functional $\mathcal{W}$ through the use of an energy density $f : \mathbb{R}^2 \times \mathbb{R}^{2\times 2} \to \mathbb{R}$:

$$\mathcal{W}(\eta) = \int_{\mathbb{T}^2} f\left(\nabla\eta, \nabla^2\eta\right).$$

Note in particular that we neglect to allow the energy density to depend on η directly since this is the case for surface energies that only depend on the geometric quantities defined in (1.1). We now consider various examples of energies of this type. Along the way we will record both the first and second variations of the energies.

Willmore energy: We consider the Willmore energy, which arises in the Helfrich model of elasticity for a lipid membrane [Hel73], modeled as a surface Σ:

$$\mathcal{W}_H = \int_\Sigma C_1 + C_2(H-H_0)^2 + C_3 K$$

for some non-negative constants C_1, C_2, C_3 and H_0, where

- $H := \operatorname{tr} s$ is the mean curvature,
- $K := \det s$ the Gaussian curvature,
- $h := \frac{\nabla^2\eta}{A}$ is the scalar extrinsic curvature, or scalar second fundamental form, and
- $s := h^{\#}$ is the shape operator, i.e. for any vector fields X, Y, $g(s(X), Y) = h(X, Y)$, where g is the metric on Σ.

Note that since $\int_\Sigma K$ is a topological invariant (due to Gauss-Bonnet), and that since $\int_\Sigma 1$ yields a lower-order differential operator (see the surface area discussion below), we can simply consider the energy

$$\mathcal{W} = \int_\Sigma \frac{1}{2}H^2.$$

We may rewrite this energy as

$$2\mathcal{W}(\eta) = \int_\Sigma H^2 = \int_{\mathbb{T}^2} H^2 A = \int_{\mathbb{T}^2} \left|g^{-1} : h\right|^2 A = \int_{\mathbb{T}^2} \frac{1}{\sqrt{1+|\nabla\eta|^2}}\left|\left(I-\frac{\nabla\eta\otimes\nabla\eta}{1+|\nabla\eta|^2}\right) : \nabla^2\eta\right|^2,$$

where

- $A := \sqrt{1+|\nabla\eta|^2}$ is the area element, and
- $g^{-1} := I - \frac{\nabla\eta\otimes\nabla\eta}{1+|\nabla\eta|^2}$ is the inverse of the metric tensor.

The first variation is non-trivial to compute, so we skip it here and refer to [Wil93], where a detailed computation shows that [1]

$$\delta\mathcal{W}(\eta) = \Delta_\Sigma H + \frac{1}{2}H\left(H^2 - 4K\right),$$

where

$$\Delta_\Sigma f := -\frac{1}{A}\nabla\cdot\left(Ag^{-1}\cdot\nabla f\right)$$

is the Laplace-Beltrami operator on the surface Σ. The second variation about a flat equilibrium is the same as the linearization of $\delta\mathcal{W}(\eta)$ about a flat equilibrium, and is the bi-Laplacian:

$$\delta_0^2\mathcal{W} = \Delta^2.$$

'Scalar' Willmore energy: Computing the general second variation $\delta_\eta^2\mathcal{W}$ of the Willmore energy presented above is a harrowing experience, and therefore we now discuss a toy model similar to the full Willmore energy but simple enough to yield tractable computations. This is what we call the 'scalar' Willmore energy, namely

$$\mathcal{W}(\eta) = \int_{\mathbb{T}^2}\frac{1}{2}m(\nabla\eta)\left|\Delta\eta\right|^2$$

for some smooth $m:\mathbb{R}^2\to(0,\infty)$ with $m(0)>0$. Simple computations then show that the variations of $\mathcal{W}$ are given by

$$\delta\mathcal{W}(\eta) = \Delta\Big(m(\nabla\eta)\,\Delta\eta\Big) - \nabla\cdot\left(\frac{1}{2}\nabla m(\nabla\eta)\left|\Delta\eta\right|^2\right)$$

and

$$\left(\delta_\eta^2\mathcal{W}\right)\phi = \Delta\Big(m(\nabla\eta)\,\Delta\phi\Big) + \nabla\cdot\Big(\nabla\left(\nabla\eta\cdot\nabla m(\nabla\eta)\right)\cdot\nabla\phi - \frac{1}{2}|\Delta\eta|^2\nabla^2 m(\nabla\eta)\cdot\nabla\phi\Big).$$

In particular, the second variation at the flat equilibrium is

$$\delta_0^2\mathcal{W} = \left(\sqrt{m(0)}\Delta\right)^2.$$

Anisotropic Willmore energy: The last surface energy we discuss that yields a fourth-order differential operator is one which, by contrast with the previous two, does not linearize to the bi-Laplacian. This surface energy is thus a prototypical example of anisotropic bending energies. In particular, we consider the surface energy

$$\mathcal{W}(\eta) := \frac{1}{2}\int_{\mathbb{T}^2}\left|C(\nabla\eta):\nabla^2\eta\right|^2$$

for some $C:\mathbb{R}^2\to\mathrm{Sym}\left(\mathbb{R}^{2\times 2}\right)$ such that $C(0)$ is positive-definite. Then the linearization about the equilibrium of the first variation of $\mathcal{W}$ is

$$\delta_0^2\mathcal{W} = \left(C(0):\nabla^2\right)^2.$$

Note that for $C(w) = \sqrt{m(w)}I$ we recover the 'scalar' Willmore energy and for

$$C(w) = \frac{1}{\left(1+|w|^2\right)^{1/4}}\left(I - \frac{w\otimes w}{1+|w|^2}\right)$$

we recover the Willmore energy discussed above.

Surface area: We now discuss how surface energies related to surface area yield second order differential operators that describe, for example, the forces due to surface tension. Consider the surface energy

$$\int_\Sigma 1 = \int_{\mathbb{T}^2} A,$$

where as above (in the discussion of the Willmore energy) $A = \sqrt{1+|\nabla\eta|^2}$. It is well-known that the first variation of the area functional written above is precisely (minus) the mean curvature, and that it models

[1] Note that our conventions differ slightly from those used by Willmore: we define the mean curvature to be the sum of the principal curvatures, and not half of that sum, and we define the Willmore energy to be half of the square of the mean curvature.

the effect of surface tension seeking to minimize the surface area of the free surface. More precisely, its variations are given by

$$\delta\mathcal{W}(\eta) = -H = -\nabla \cdot \left(\frac{\nabla \eta}{\sqrt{1 + |\nabla \eta|^2}} \right)$$

and

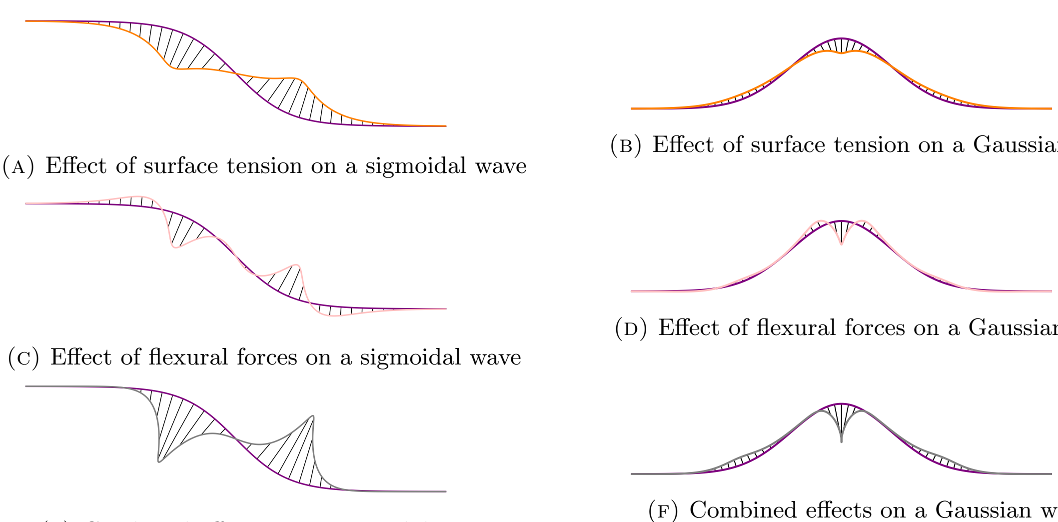
$$\left(\delta_\eta^2 \mathcal{W}\right)\phi = -\nabla \cdot \left(g^{-1} \cdot \frac{\nabla \phi}{A} \right) = -\nabla \cdot \left(\left(I - \frac{\nabla \eta \otimes \nabla \eta}{1 + |\nabla \eta|^2} \right) \cdot \frac{\nabla \phi}{\sqrt{1 + |\nabla \eta|^2}} \right).$$

In particular, its linearization about equilibrium is

$$\delta_0^2 \mathcal{W} = -\Delta.$$

Competing effects of surface tension and flexural forces: Our general form of the surface energy allows for energetic contributions due to bending as well as area, and as such we will allow for surface stresses of flexural and surface tension type. Here we record some examples of what these forces look like in terms of the local geometry of the surface. In particular, we see that there are instances in which the bending and surface tension stresses are in opposition.

- Circular arc: In a circular (one-dimensional) arc surface tension and flexural forces act in opposite directions, the former pushing inward and the latter pushing outward. This is due to the simple observation regarding the scaling of these surface energies:
$$\mathcal{A} = \int_\Sigma 1 \sim R \quad \text{and} \quad \mathcal{W} = \int_\Sigma H^2 \sim \frac{1}{R^2} R = \frac{1}{R}.$$
- Sigmoidal wave: Surface tension and flexural forces acting in opposite directions can also be seen locally in some more complicated geometries, such as that of the sigmoidal wave shown in Figure 1. In particular, these forces act in opposite directions to one another at the front and tail of the wave.
- Gaussian wave: This is another example, shown in Figure 1, of a geometry in which, locally, surface tension and flexural forces may act in opposite directions.



(A) Effect of surface tension on a sigmoidal wave

(B) Effect of surface tension on a Gaussian wave

(C) Effect of flexural forces on a sigmoidal wave

(D) Effect of flexural forces on a Gaussian wave

(E) Combined effects on a sigmoidal wave

(F) Combined effects on a Gaussian wave

FIGURE 1. The purple curve is the profile of a free surface Σ given as the graph of $\eta = \tanh$ on the left and of $\eta(x) = e^{-x^2/2}$ on the right. The black segments show the force $\delta\mathcal{W}(\eta)\,\nu_\Sigma$ exercised on the free surface corresponding to a surface energy $\mathcal{W}$. In (A) and (B), $\mathcal{W} = \int_\Sigma 1$; in (C) and (D), $\mathcal{W} = \int_\Sigma H^2$; and in (E) and (F), $\mathcal{W} = \int_\Sigma \alpha + \beta H^2$ for some $\alpha, \beta > 0$. The other curve (orange in (A) and (B), pink in (C) and (D), and grey in (E) and (F)) illustrates the new profile of the free surface after application of the force $\delta\mathcal{W}(\eta)\,\nu_\Sigma$.

1.3. Fluid equations. We now consider a slab of periodic fluid occupying the moving domain

$$\Omega(t) := \left\{ x = (\bar{x}, x_3) \in \mathbb{T}^2 \times \mathbb{R} \;\middle|\; -b < x_3 < \eta(t, \bar{x}) \right\}$$

for an unknown height function $\eta : [0,\infty) \times \mathbb{T}^2 \to (-b, \infty)$. The lower boundary of $\Omega(t)$ is the rigid unmoving interface

$$\Sigma_b := \left\{ x \in \mathbb{T}^2 \times \mathbb{R} \;\middle|\; x_3 = -b \right\},$$

while the upper boundary is the moving interface

$$\Sigma(t) := \left\{ x \in \mathbb{T}^2 \times \mathbb{R} \;\middle|\; x_3 = \eta(t, \bar{x}) \right\}.$$

We assume that the fluid is subject to a uniform gravitational field of strength $g \in \mathbb{R}$ acting perpendicularly to Σ_b. Note in particular that we do not require $g \geq 0$: more will be said about this below in the latter part of Section 2.2. We assume that the free interface is subject to surface stresses generated by the energy

$$\mathcal{W}(\eta) = \int_{\mathbb{T}^2} f\left(\nabla\eta, \nabla^2\eta\right) \tag{1.2}$$

for a function $f : \mathbb{R}^2 \times \mathbb{R}^{2\times 2} \to \mathbb{R}$ satisfying the assumptions enumerated below in Section 2.2. If ν denotes the unit normal pointing *out* of $\Omega(t)$, then the surface stress is

$$-\delta\mathcal{W}(\eta)(-\nu) = \delta\mathcal{W}(\eta)\,\nu, \tag{1.3}$$

i.e. the magnitude of the stress is $-\delta\mathcal{W}(\eta)$ but the direction is $-\nu$, which indicates that the surface stress acts on the fluid. This form of $\mathcal{W}$ allows us to consider a generalized mixture of bending and surface tension stresses. Due to this general form, we will not attribute the source of the energy (and hence the stress) to any particular model, but as elaborated on above in Section 1.1, such an energy would arise if we viewed the surface as a thin biological membrane or as a thin layer of ice. Our assumptions on f will always require that $\delta\mathcal{W}(\eta)$ is a fourth-order differential operator, typically of quasilinear form.

We will assume that the fluid is incompressible and viscous, which means that we can describe its state by specifying its velocity $v(t,\cdot) : \Omega(t) \to \mathbb{R}^3$ and pressure $q(t,\cdot) : \Omega(t) \to \mathbb{R}$. For simplicity we will assume that the fluid density and viscosity are normalized to unity. The equations of motion are then the free boundary Navier-Stokes equations coupled to surface stresses of the form (1.3) generated by the free energy (1.2). These read

$$\begin{cases} \partial_t v + (v \cdot \nabla) v = -\nabla q + \Delta v & \text{in } \Omega(t), \quad (1.4a) \\ \nabla \cdot v = 0 & \text{in } \Omega(t), \quad (1.4b) \\ (qI - \mathbb{D}v)\,\nu = \Big(\delta\mathcal{W}(\eta) + g\eta\Big)\nu & \text{on } \Sigma(t), \quad (1.4c) \\ \partial_t \eta = (v \cdot \nu)\sqrt{1 + |\nabla\eta|^2} & \text{on } \Sigma(t), \text{ and} \quad (1.4d) \\ v = 0 & \text{on } \Sigma_b, \quad (1.4e) \end{cases}$$

where

$$(\mathbb{D}v)_{ij} = \partial_i v_j + \partial_j v_i \tag{1.5}$$

is the symmetrized gradient and I is the 3×3 identity matrix. By a minor standard abuse of notation in (1.4c) and (1.4d), all quantities involving η and its derivatives at a point (x', x_3) in $\Sigma(t)$ are understood to be determined by their values at $x' \in \mathbb{T}^2$. The first two equations are the usual incompressible Navier-Stokes system, the third is the balance of stresses on the free interface, the fourth is the kinematic transport equation, and the fifth is the no-slip condition at the rigid interface. Note that what we call the pressure q is really the difference between the standard pressure $\bar{q}$ and hydrostatic pressure $-gx_3$, i.e. $q = \bar{q} + gx_3$. Making this substitution in the first and third equations reveals that the gravitational term is originally a bulk force acting in $\Omega(t)$.

Sufficiently regular solutions to (1.4a)–(1.4e) obey the following equations: the energy-dissipation identity

$$\frac{d}{dt}\left(\int_{\Omega(t)} \frac{1}{2}|v|^2 + \int_{\mathbb{T}^2} \frac{g}{2}|\eta|^2 + \mathcal{W}(\eta)\right) + \int_{\Omega(t)} \frac{1}{2}|\mathbb{D}v|^2 = 0, \tag{1.6}$$

and the mass conservation identity

$$\frac{d}{dt}\int_{\mathbb{T}^2} \eta = 0. \tag{1.7}$$

The first term in parentheses in (1.6) is the kinetic energy of the fluid, the second is the total gravitational potential energy stored in the fluid, and the third is the surface energy (1.2). The term outside parentheses is the usual viscous dissipation, which in particular forces the total energy (the sum of the three terms) to be non-increasing in time. The equation (1.7) is understood as the integral form of mass conservation since $b+\int_{\mathbb{T}^2}\eta(\cdot,t)$ is the mass of the fluid body at time $t\geq 0$. We will assume that the parameter b is chosen such that the initial mass of fluid is b, which means that

$$\int_{\mathbb{T}^2}\eta_0=0 \text{ and hence } \int_{\mathbb{T}^2}\eta\,(t,\cdot)=0 \text{ for } t\geq 0. \tag{1.8}$$

From the no-slip condition, Korn's inequality (see Proposition 8.22), and (1.6) we conclude that any equilibrium (time-independent) solutions must satisfy $v=0$. In turn, this, (1.4a)–(1.4e), and (1.8) imply that $p=0$, which reduces to η solving

$$\delta\mathcal{W}\,(\eta)+g\eta=0. \tag{1.9}$$

It's clear that $\eta=0$ is a solution to this, but it does not follow from our assumptions on the energy density f (enumerated below in Section 2.2) that 0 is the only solution to this equation. However, our assumptions do require that 0 is a local minimum of the total surface energy (the sum of $\mathcal{W}$ and the gravitational potential $\mathcal{P}$)

$$\mathcal{W}\,(\eta)+\mathcal{P}\,(\eta):=\mathcal{W}\,(\eta)+\int_{\mathbb{T}^2}\frac{g}{2}|\eta|^2 \tag{1.10}$$

and that the second variation of $\mathcal{W}+\mathcal{P}$ is positive definite at 0, when restricted to functions of zero average. It is a simple matter to check that (1.9) corresponds to the Euler-Lagrange equation $\delta(\mathcal{W}+\mathcal{P})\,(\eta)=0$, which means that 0 is an isolated critical point of $\mathcal{W}+\mathcal{P}$. Then $\eta=0$ is the only solution to (1.9) within an open set containing 0. Thus, there is a locally unique equilibrium corresponding to a flat slab of quiescent fluid. Our main goal in this chapter is to show that this equilibrium solution is asymptotically stable and to characterize the rate of decay to equilibrium.

Much is known about problems of the form (1.4a)–(1.4e) when $\mathcal{W}$ is a multiple $\sigma\geq 0$ of the area function, $g>0$, and the cross-section is either periodic ($\mathbb{T}^2$) or infinite ($\mathbb{R}^2$). The case $\sigma>0$ corresponds to surface tension, and $\sigma=0$ corresponds to no surface tension. Beale [**Bea81**] proved the first local well-posedness results for the infinite cross section without surface tension. Beale [**Bea84**] also proved global existence of solutions near equilibrium for the infinite problem with surface tension. Beale–Nishida[**BN85**] then proved that these global solutions decay at an algebraic rate. The existence of global solutions with and without surface tension was also studied by Tani–Tanaka [**TT95**], but no decay information was obtained. Guo–Tice [**GT13b**] proved that for the infinite problem without surface tension, small data leads to global solutions that decay algebraically. For the periodic problem without surface tension, Hataya [**Hat09**] constructed global solutions decaying at a fixed algebraic rate, and Guo–Tice [**GT13a**] proved that solutions decay almost exponentially, with the decay rate determined by the data. Nishida–Teramoto–Yoshihara [**NTY04**] proved that the periodic problem with surface tension leads to global solutions near equilibrium that decay exponentially. Tan–Wang [**TW14**] established a sort of continuity result, proving that the global solutions with surface tension converge to the global solutions without surface tension as $\sigma\to 0$.

As mentioned in Section 1.1, there are results on the local existence of solutions to models coupling incompressible Navier-Stokes to free boundaries with elastic and bending stresses: [**CCS07, CS10, MC15, MC16**]. However, to the best of our knowledge, there are no global existence or asymptotic stability results on the problem (1.4a)–(1.4e) with $\mathcal{W}$ combining bending and surface tension stresses.

2. Main result

2.1. Reformulation in a fixed domain. In order to solve the problem (1.4a)–(1.4e) we flatten the domain, which has the benefit of allowing us to work with a domain that is no longer time-dependent. More precisely, we move from the Eulerian domain $\Omega\,(t)$ to the fixed equilibrium domain $\Omega:=\mathbb{T}^2\times(-b,0)$ via a map $\Phi:[0,T)\times\mathbb{T}^2\times\mathbb{R}\to\mathbb{T}^2\times\mathbb{R}$ such that for every $0\leqslant t<T$, $\Phi(t,\cdot):\Omega\to\Omega\,(t)$ is a diffeomorphism that maps the lower/upper boundary of Ω to the upper/lower boundary of $\Omega(t)$.

To precisely define this map we need two tools. The first is any smooth cutoff function $\mathcal{X}:\Omega\to\mathbb{R}$ such that $\mathcal{X}=1$ on Σ and $\mathcal{X}=0$ on Σ_b. For instance, we can define $\mathcal{X}\,(x_3)=1+\frac{x_3}{b}$. The second tool is the harmonic extension map ext, the precise definition of which can be found in Section 8.2. For $0\leq t<T$, the

extension allows us to extend $\eta(t,\cdot):\mathbb{T}^2\to\mathbb{R}$ to the function $\operatorname{ext}\eta(t,\cdot):\Omega\to\mathbb{R}$, defined in the bulk. The extension is done to help with regularity issues when taking the trace of Φ onto Σ.

With these tools in hand, we define

$$\Phi(t,\cdot)=\mathrm{id}+\operatorname{ext}\eta(t,\cdot)\chi e_3 \tag{2.1}$$

for the choice of cutoff χ as above. An important observation is that if η is sufficiently small (which is made precise in item (2) of Remark 4.3), then $\Phi(t,\cdot)$ is a diffeomorphism onto $\Omega(t)$. In particular, if we denote by $\Sigma=\mathbb{T}^2\times\{0\}$ the upper boundary of the fixed domain Ω, then $\Phi(t,(\Sigma))=\Sigma(t)$ and $\Phi(t,\cdot)=\mathrm{id}$ on Σ_b: see Figure 2. Any function f defined on the Eulerian domain $\Omega(t)$ thus gives rise to a function $F:=f\circ\Phi$

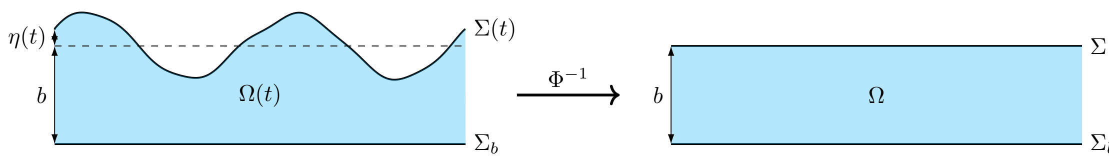


FIGURE 2. A cartoon of the diffeomorphism fixing the domain

defined on the fixed domain Ω. In particular, the manifestations on Ω of the temporal and spatial derivatives of f are given by

$$\begin{cases}\nabla^{\mathcal{G}}F:=\nabla\left(F\circ\Phi^{-1}\right)\circ\Phi \text{ and}\\ \partial_t^{\mathcal{G}}F:=\partial_t\left(F\circ\Phi^{-1}\right)\circ\Phi\end{cases}$$

i.e. $f=F\circ\Phi^{-1}$ and $\nabla f=\left(\nabla^{\mathcal{G}}F\right)\circ\Phi^{-1}$ (and similarly for temporal derivatives). The differential operators $\nabla^{\mathcal{G}}$ and $\partial_t^{\mathcal{G}}$ are called $\mathcal{G}$-differential operators. In more concrete terms, the $\mathcal{G}$-differential operators may be written as

$$\begin{cases}\nabla^{\mathcal{G}}=\mathcal{G}\cdot\nabla \text{ and}\\ \partial_t^{\mathcal{G}}=\partial_t-(\partial_t\Phi)\cdot\nabla^{\mathcal{G}}\end{cases}$$

for $\mathcal{G}:=(\nabla\Phi)^{-T}$. Similarly, we define the $\mathcal{G}$-versions of the symmetrized gradient and of the Laplacian via $\mathbb{D}^{\mathcal{G}}F:=\nabla^{\mathcal{G}}F+\left(\nabla^{\mathcal{G}}F\right)^T$ and $\Delta^{\mathcal{G}}F:=\nabla^{\mathcal{G}}\cdot\left(\nabla^{\mathcal{G}}F\right)$. We may now reformulate (1.4a)–(1.4e) as a system of PDEs on the fixed domain Ω. Indeed, solutions $\mathcal{X}^*=(v,q,\eta)$ on $\Omega(t)$ of (1.4a)–(1.4e) correspond to solutions $\mathcal{X}=(v\circ\Phi,q\circ\Phi,\eta)=:(u,p,\eta)$ on Ω of

$$\begin{cases}\partial_t^{\mathcal{G}}u+\left(u\cdot\nabla^{\mathcal{G}}\right)u=-\nabla^{\mathcal{G}}p+\Delta^{\mathcal{G}}u & \text{in } \Omega, \quad (2.2a)\\ \nabla^{\mathcal{G}}\cdot u=0 & \text{in } \Omega, \quad (2.2b)\\ \left(pI-\mathbb{D}^{\mathcal{G}}u\right)\nu_{\partial\Omega}=\Big(\delta\mathcal{W}(\eta)+g\eta\Big)\nu_{\partial\Omega} & \text{on } \Sigma, \quad (2.2c)\\ \partial_t\eta=u\cdot\nu_{\partial\Omega}^{\mathcal{G}} & \text{on } \Sigma, \text{ and} \quad (2.2d)\\ u=0 & \text{on } \Sigma_b\ . \quad (2.2e)\end{cases}$$

where $\nu_{\partial\Omega}^{\mathcal{G}}$ is defined in (4.1). The rest of this chapter is therefore concerned with the study of this system.

2.2. Assumptions on the surface energy density. We now make precise the assumptions that we impose on the surface energy density $f:\mathbb{R}^2\times\mathbb{R}^{2\times 2}\to\mathbb{R}$ throughout the chapter. We assume the following.

(1) f is smooth, i.e. infinitely differentiable. If we keep track of the regularity needed on f at the lowest level of regularity to close the estimates in this chapter, then we only need $f\in C^{7,1}$. However, no effort has been made to make this regularity optimal in light of the fact that if we sought smooth solutions, then f would have to be smooth as well.

(2) $f(0,0)=0$ and $\nabla f(0,0)=0$. This is an assumption that can be made without loss of generality because we may reduce the general case to this one by adding a null Lagrangian and a constant to the surface energy. Indeed, for an arbitrary $\tilde f$, we may define

$$f(w,M):=\tilde f(w,M)-\tilde f(0,0)+\nabla_w\tilde f(0,0)\cdot w+\nabla_M\tilde f(0,0):M$$

such that indeed, $f(0)=0$, $\nabla f(0)=0$ and for $\eta:\mathbb{T}^2\to\mathbb{R}$ sufficiently regular we have that

$$\begin{aligned}\int_{\mathbb{T}^2} f\left(\nabla\eta,\nabla^2\eta\right) &= \int_{\mathbb{T}^2}\tilde f\left(\nabla\eta,\nabla^2\eta\right)-\int_{\mathbb{T}^2}\tilde f(0)-\int_{\mathbb{T}^2}\nabla_w\tilde f(0,0)\cdot\nabla\eta+\nabla_M\tilde f(0,0):\nabla^2\eta\\ &=\int_{\mathbb{T}^2}\tilde f\left(\nabla\eta,\nabla^2\eta\right)-\int_{\mathbb{T}^2}\tilde f(0),\end{aligned}$$

i.e. the surface energies defined by f and $\tilde f$ only differ by an irrelevant constant. Note that the third integral on the right side of the first equality vanishes by integrating by parts.

(3) The Hessian of f satisfies

$$\nabla^2_{M,M}f(0)\bullet\left(k^{\otimes 4}\right)-\nabla^2_{w,w}f(0)\bullet\left(k^{\otimes 2}\right)+g\gtrsim|k|^4 \tag{2.3}$$

for all $k\in\mathbb{Z}^2\setminus\{0\}$, i.e. $\delta_0^2\mathcal{W}+g$ is strictly elliptic over functions of average zero. See Section 8.8 for a more detailed discussion of the ellipticity of $\delta_0^2\mathcal{W}+g$.

Note in particular that our assumptions on f do not necessarily imply that $\mathcal{W}$ is positive definite. However, the third assumption requires that the total surface energy $\mathcal{W}+\mathcal{P}$ defined in (1.10) is positive definite for sufficiently small perturbations of 0.

The third assumption can also be understood as saying that flexural effects dominate. For example, if we consider

$$\mathcal{W}=\int_\Sigma \alpha+\beta H^2,$$

then $\delta_0^2\mathcal{W}=-\alpha\Delta+\beta\Delta^2$. If $\alpha,g<0$, then upon applying the Fourier transform we see that

$$\left(\delta_0^2\mathcal{W}+g\right)^\wedge(k)=16\pi^4\beta|k|^4+4\pi^2\alpha|k|^2+g\geqslant\left(16\pi^4\beta+4\pi^2\alpha+g\right)|k|^4$$

since $|k|\geqslant 1$ for all $k\in\mathbb{Z}^2\setminus\{0\}$. In particular, even if $\alpha,g<0$, as long as $16\pi^4\beta>-\left(4\pi^2\alpha+g\right)$, then $\delta_0^2\mathcal{W}+g$ is strictly elliptic over functions of average zero. In physically meaningful terms (c.f. Figure 3 for a sketch), this means that sufficiently strong flexural effects dominate over adverse surface tension and gravity effects. In particular, we can allow for $g<0$ in general.

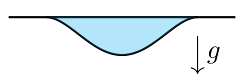


FIGURE 3. Sufficiently strong flexural effects dominate adverse gravitational effects.

2.3. Statement of the main result. In order state the main result, it is convenient to introduce the notion of an admissible initial condition and to introduce the energy and dissipation functionals.

An admissible initial condition is, loosely speaking, a pair (u_0,η_0) such that u_0 is incompressible, the boundary conditions are satisfied, an appropriate compatibility condition holds, and η_0 has average zero. The precise definition of an admissible initial condition may be found in Definition 6.4, and a more detailed discussion of the compatibility condition is included in Remark 6.5.

Now let us introduce the energy and dissipation functionals. Given a triple $\mathcal{X}=(u,p,\eta)$, the associated energy and dissipation functionals are

$$\mathcal{E}(\mathcal{X}):=||u||^2_{H^2(\Omega)}+||\partial_t u||^2_{L^2(\Omega)}+||p||^2_{H^1(\Omega)}+||\eta||^2_{H^{9/2}(\mathbb{T}^2)}+||\partial_t\eta||^2_{H^2(\mathbb{T}^2)}$$

and

$$\mathcal{D}(\mathcal{X}):=||u||^2_{H^3(\Omega)}+||\partial_t u||^2_{H^1(\Omega)}+||p||^2_{H^2(\Omega)}+||\eta||^2_{H^{11/2}(\mathbb{T}^2)}+||\partial_t\eta||^2_{H^{5/2}(\mathbb{T}^2)}+\left|\left|\partial_t^2\eta\right|\right|^2_{H^{1/2}(\mathbb{T}^2)},$$

respectively. We will sometimes abuse notation slightly and write $\mathcal{E}(t):=\mathcal{E}(\mathcal{X}(t))$ and $\mathcal{D}(t):=\mathcal{D}(\mathcal{X}(t))$ when it is clear from context which triple $\mathcal{X}$ is being used. We may now state the main result of this chapter.

THEOREM 2.1. *Assume that f satisfies the conditions enumerated in Section 2.2. Then there exist universal constants $C,\lambda,\epsilon>0$ such that for every admissible initial condition (u_0,η_0) satisfying*

$$||\eta_0||^2_{H^{9/2}(\mathbb{T}^2)}+||u_0||^2_{H^2(\Omega)}+||u_0\cdot(-\nabla\eta_0,1)||^2_{H^2(\Sigma)}\leqslant\epsilon \tag{2.4}$$

there exists a unique solution $\mathcal{X} = (u, p, \eta)$ *of* (2.2a)–(2.2e) *on* $[0, \infty)$ *such that*

$$\sup_{t \geqslant 0} \mathcal{E}(t) e^{\lambda t} + \int_0^\infty \mathcal{D}(t) e^{\lambda t} dt \leqslant C \mathcal{E}(0).$$

Note that requiring the smallness of the third term in (2.4) comes from the compatibility condition (c.f. Section 6 for a more detailed discussion). Theorem 2.1 is proved in Section 6 in the somewhat more precise form of Theorem 6.11. Theorem 2.1 guarantees that η is regular and small enough to transform the solution back to the Eulerian system, which then gives rise to a global decaying solution to (1.4a)–(1.4e), obeying similar estimates.

3. Discussion

In order to prove global well-posedness and decay, we employ a *nonlinear energy method*. We outline this method in Section 3.1, discuss the difficulties that arise in Section 3.2, and provide a strategy of the proof in Section 3.3. We also discuss how the work presented in this chapter fits with respect to previous work considering other types of surface forces, highlighting that the present work may be viewed, in some sense, as 'supercritical.'

3.1. Nonlinear energy method. In this section we provide a high-level overview of the nonlinear energy method employed to prove global well-posedness and decay. The *nonlinear energy method* informs the scheme of a priori estimates that we employ, and begin as follows: we multiply the PDE by the unknown u and integrate by parts with respect to the *nonlinear* differential operators $\nabla^{\mathcal{G}}$. This yields the energy-dissipation relation

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(\int_\Omega \frac{1}{2}|u|^2 J + \mathcal{W}(\eta) + \int_{\mathbb{T}^2} \frac{g}{2}|\eta|^2\right) + \left(\int_\Omega \frac{1}{2}|\mathbb{D}^{\mathcal{G}} u|^2 J\right) = 0$$

where $J := \det \nabla\Phi$ accounts for the local deformation in volume due to the change of coordinates Φ. Close to the equilibrium solution $(u, p, \eta) = 0$, the energy-dissipation relation becomes the same as that which is obtained by a standard energy estimate for the linearization of the PDE about the equilibrium, namely

$$\frac{\mathrm{d}}{\mathrm{d}t}\underbrace{\left(\int_\Omega \frac{1}{2}|u|^2 + \mathcal{Q}_0(\eta) + \int_{\mathbb{T}^2} \frac{g}{2}|\eta|^2\right)}_{E} + \underbrace{\left(\int_\Omega \frac{1}{2}|\mathbb{D}u|^2\right)}_{D} = 0,$$

where $\mathcal{Q}_0$ denotes the quadratic approximation of $\mathcal{W}$ about the equilibrium, and is defined precisely in Section 4.4.1. The good news is that if we restrict our attention to terms involving u, i.e. consider only $E_u := \frac{1}{2}\int_\Omega |u|^2$, then it follows from the no-slip boundary condition $u = 0$ on Σ_b and Korn's inequality, Proposition 8.22, that the dissipation is coercive over the energy, i.e.

$$E_u = \int_\Omega \frac{1}{2}|u|^2 \lesssim \int_\Omega \frac{1}{2}|\mathbb{D}u|^2.$$

If for the moment we ignore the terms in the energy depending on η, then a Gronwall-type argument shows that we should expect exponential decay of E_u:

$$\begin{cases} \frac{\mathrm{d}}{\mathrm{d}t} E_u + D = 0 \\ E_u \leqslant CD \end{cases} \quad \Rightarrow \quad \frac{\mathrm{d}}{\mathrm{d}t}\left(E_u(t) e^{Ct}\right) \leqslant 0 \quad \Rightarrow \quad E_u(t) \leqslant e^{-Ct} E_u(0).$$

Of course, we are not actually able to ignore the η terms in the energy, so we must find a mechanism for controlling these terms with the dissipation functional.

Such a mechanism is found by appealing to the equations (2.2c) and (2.2d), which allow us to estimate η and $\partial_t \eta$. Indeed, in order to obtain this coercivity we may use the elliptic nature of the dynamic boundary condition (2.2c) to transfer control of u (and p) onto additional control of η. However at this stage we can only conclude that $u \in H^1$, which is insufficient to make sense of the trace of the stress tensor in the dynamic boundary condition, and so this mechanism for regularity transfer is not available to us.

To resolve this issue we take derivatives of the problem that are compatible with the no-slip boundary condition (temporal and horizontal spatial derivatives) and apply a version of the energy-dissipation estimate. The extra control that this provides then allows us to use a host of auxiliary estimates that permit the transfer of regularity between u and η. For example, the dynamic boundary condition allows us to gain control of higher-order derivatives of η. Proceeding in this fashion, we can close the estimates by showing that the dissipation is coercive over the *full* energy.

3.2. Difficulties. We now turn to a discussion of the difficulties encountered when employing the nonlinear energy method described above. The central difficulty is that there is a nontrivial interdependence between two essential features of the problem, namely the regularity gain and transfer mechanisms on one hand and the energy-dissipation structure on the other hand. This difficulty is exacerbated by two components of the problem in particular:

- $\mathcal{W}$ is of order two, which is supercritical (in a sense made precise below),
- $\delta\mathcal{W}$ is generally a quasilinear differential operator of order four.

In order to describe the difficulties we encounter, it is helpful to write the problem in a more compact form as $N(\mathcal{X}) = 0$ for $\mathcal{X} = (u, p, \eta)$ the unknown and N the nonlinear differential operator given by

$$N(\mathcal{X}) = N(u, p, \eta) = \begin{pmatrix} \partial_t^{\mathcal{G}} u + \left(u \cdot \nabla^{\mathcal{G}}\right) u + \nabla^{\mathcal{G}} p - \Delta^{\mathcal{G}} u \\ \nabla^{\mathcal{G}} \cdot u \\ \operatorname{tr}_\Sigma \left(pI - \mathbb{D}^{\mathcal{G}} u\right) \nu_\Sigma - \Big(\delta\mathcal{W}(\eta) + g\eta\Big)\nu_\Sigma \\ \partial_t \eta - \operatorname{tr}_\Sigma u \cdot \nu_\Sigma \sqrt{1 + |\nabla\eta|^2} \\ \operatorname{tr}_{\Sigma_b} u \end{pmatrix}.$$

3.2.1. *Structured estimates.* Most terms in N may be viewed as linear operators with multilinear dependence on geometric coefficients under control (such as $\mathcal{G}$ and J). When computing the commutators between N and partial derivatives, the contribution from these kind of terms is relatively benign. A more detailed description of these operators and the corresponding commutators may be found in Section 5.1. However the term $\delta\mathcal{W}(\eta)$, which comes from the fully nonlinear surface energy, cannot be written in this form and as a consequence it gives rise to commutators that are too singular to be controlled in a *structured* manner.

More precisely: the first attempt would be to write the equation $\partial^\alpha(N(\mathcal{X})) = 0$ as a perturbation of $L\mathcal{X} = 0$, where L denotes the linearization of N about the equilibrium. In other words, we would seek to write $\partial^\alpha(N(\mathcal{X})) = L\partial^\alpha\mathcal{X} + C(\partial^\alpha\mathcal{X})$ for some commutators C. Then upon integrating by parts and deriving the corresponding energy-dissipation relation, we would obtain commutators that are too singular to be controlled in a structured manner.

To elucidate what we mean by this, let us consider the following cartoon: consider the following energy-dissipation relation

$$\frac{\mathrm{d}}{\mathrm{d}t} E + D = C$$

where C denotes some commutators. If we can show that

$$|C| \leqslant \sqrt{E} D, \tag{3.1}$$

then for $E \leqslant \frac{1}{4}$ (i.e. in the cartoon version of what we will later call the small energy regime) we have that

$$\frac{\mathrm{d}}{\mathrm{d}t} E + \frac{1}{2} D \leqslant 0.$$

Moreover, if the dissipation D is coercive over the energy E (i.e. $E \leqslant D$), then we can conclude that the energy decays exponentially fast. However, if instead of (3.1) we can only show that

$$|C| \leqslant D^{3/2}, \tag{3.2}$$

then we cannot conclude anything about the boundedness or decay of E. In other words: whilst both (3.1) and (3.2) show that the commutators C can be controlled, only (3.1) shows that the commutators can be controlled in a manner *respectful of the energy-dissipation structure*. In particular, note that unstructured estimates like (3.2) are typically easier to obtain than structured estimates like (3.1) due to the fact that the dissipation is coercive over the energy, and hence $\sqrt{E} D \leqslant D^{3/2}$.

A more specific discussion of why our scheme of a priori estimates would fail due to the term coming from the nonlinear surface energy may be found in Remark 5.3.

3.2.2. *Parabolic criticality.* As hinted at earlier, a particular source of difficulty when attempting to estimate these commutators comes from the fact that energies of order two, like the energies of Willmore-type considered here, are 'supercritical.' This critical phenomenon comes from the fact that the Stokes system embedded into our problem imposes parabolic scaling on u, but when we use the equations of motion to gain dissipative control of spatial and temporal derivatives of η this generally induces non-parabolic scaling for η estimates. This mismatch between the u scaling and the η scaling is precisely the source of the critical threshold. In particular, as will be detailed below, previous work dealing with capillary forces due to surface tension may be viewed as 'subcritical' whilst this work dealing with flexural forces due to bending may be viewed as 'supercritical.'

To better understand this difficulty it is helpful to consider a toy example in which

$$\mathcal{W}(\eta) = \int_{\mathbb{T}^2} \left| |\nabla|^\alpha \eta \right|^2$$

for some $\alpha > 0$. We then observe that if $u \in H^s(\Omega)$ (and so $p \in H^{s-1}(\Omega)$), then we may use the kinematic and dynamic boundary conditions,

$$\begin{cases} \left(\delta_0^2 \mathcal{W}\right)\eta = \operatorname{tr}\left(pI - \mathbb{D}u\right) : (e_3 \otimes e_3) \in H^{s-\frac{3}{2}}\left(\mathbb{T}^2\right) \text{ and} \\ \partial_t \eta = \operatorname{tr} u \cdot e_3 \in H^{s-\frac{1}{2}}\left(\mathbb{T}^2\right), \end{cases}$$

to obtain the following control over η and $\partial_t \eta$:

$$\begin{cases} ||\eta||_{H^{s+2\alpha-3/2}(\mathbb{T}^2)} \lesssim ||u||_{H^s(\Omega)} + ||p||_{H^{s-1}(\Omega)} \text{ and} \\ ||\partial_t \eta||_{H^{s-1/2}(\mathbb{T}^2)} \lesssim ||u||_{H^s(\Omega)}. \end{cases}$$

Therefore the difference in regularity between η and $\partial_t \eta$ is $\left(s + 2\alpha - \frac{3}{2}\right) - \left(s - \frac{1}{2}\right) = 2\alpha - 1$. To summarize schematically, the induced dissipative η scaling is:

$$\partial_t \eta \sim |\nabla|^{2\alpha-1} \eta,$$

where this should be understood in the sense that if we control $\partial_t \eta$ in H^s, then we expect to control η in $H^{s+(2\alpha-1)}$,and vice-versa (i.e. control of η in H^s is expected to correspond to control of $\partial_t \eta$ in $H^{s-(2\alpha-1)}$).

This scaling mismatch complicates the design of a scheme of a priori estimates in which control of time derivatives is leveraged to gain control of spatial derivatives, but temporal differentiation of the equations leads to high-order commutators. In particular:

- For $\alpha < \frac{3}{2}$, temporal derivatives of η are cheap relative to spatial derivatives (by contrast with parabolic scaling). This is what we refer to as the subcritical case. The case of surface tension, which corresponds to $\alpha = 1$, falls into this category.
- For $\alpha = \frac{3}{2}$, η follows parabolic scaling.
- For $\alpha > \frac{3}{2}$, temporal derivatives of η are expensive relative to spatial derivatives (by contrast with parabolic scaling). This is what we refer to as the supercritical case. The case of flexural forces, which corresponds to $\alpha = 2$ and which is considered in this chapter, falls into this category.

Since the Willmore-type energies we consider here are supercritical, we must therefore be very wary of commutators involving time derivatives of η. Again, the precise manner in which this can be an issue for the scheme of a priori estimates presented here is discussed in Remark 5.3.

3.2.3. *Appropriate linearization.* To summarize the difficulties discussed so far: we seek to estimate the commutators in a structured manner, and we have to be particularly careful regarding terms involving time derivatives of η due to the supercriticality of the Willmore-type energies discussed here. To address both of these issues we proceed as follows: instead of linearizing the PDE system directly (whether about the equilibrium or about any $\mathcal{X}$), we find a quadratic approximation of the energy and dissipation, and then derive the associated PDE - which is also linear but *not* the same as the linearization of the nonlinear operator N. In some sense, it is beneficial to perform the linearization in this manner since it is more respectful of the structure of the fully nonlinear surface energy. In a more precise sense, we will see below that performing the linearization in this manner leads to commutators that can be controlled.

We thus view N as a perturbation of some linear operator $L_{\mathcal{X}}$ (i.e. a linear operator whose coefficients depend on $\mathcal{X}$) different from its linearization but such that the energy-dissipation relation associated with $L_{\mathcal{X}}$ has 'good commutators'. Note that we write $L_{\mathcal{X}}$ to emphasize that the coefficients of this linear operators depend on $\mathcal{X}$. We will thus consider the commutators (called this by a slight abuse of notation) $\partial^\alpha \circ N - L_{\mathcal{X}} \circ \partial^\alpha$, where $L_{\mathcal{X}}$ is given by

$$L_{\mathcal{X}}(\mathcal{Y}) := L_{\mathcal{X}}(v, q, \zeta) = \begin{pmatrix} \partial_t^{\mathcal{G}} v + \left(u \cdot \nabla^{\mathcal{G}}\right) v + \nabla^{\mathcal{G}} p - \Delta^{\mathcal{G}} v \\ \nabla^{\mathcal{G}} \cdot v \\ \operatorname{tr}_\Sigma \left(qI - \mathbb{D}^{\mathcal{G}} v\right) \nu_\Sigma - \left(\delta_\eta^2 \mathcal{W}(\zeta) + g\zeta\right) \nu_\Sigma \\ \partial_t \zeta - \operatorname{tr}_\Sigma v \cdot \nu_\Sigma \sqrt{1 + |\nabla \eta|^2} \\ \operatorname{tr}_{\Sigma_b} u. \end{pmatrix}.$$

Note here that $\mathcal{G} = \mathcal{G}(\eta)$ and $\nu_\Sigma = \nu_\Sigma(\eta)$, i.e. these geometric coefficients depend on η (i.e. on $\mathcal{X}$) and not ζ (i.e. not on $\mathcal{Y}$).

This is where the subtle interdependence between the energy-dissipation structure and the regularity gain and transfer structure is most apparent. On one hand the linearization of N about the equilibrium, denoted by L, tells us how much regularity can be gained and therefore tell us which commutators can be controlled, and on the other hand the energy-dissipation structure associated with N tells us which form of control of these commutators is allowed in order to close the estimates. The precise form of $L_{\mathcal{X}}$ is then chosen such that it yields 'good' commutators respectful of both of these features, i.e. commutators upon which we have *structured control*, and which are also tame enough despite the supercriticality of the surface energy. In particular, note that when $\mathcal{X}$ is the equilibrium solution, i.e. $\mathcal{X} = 0$, then $L_0 = L$.

3.2.4. *Failure of coercivity.* We discussed above that surface energies of order $\alpha = 3/2$ are critical, in some sense. Nonetheless, close to that exponent, i.e. whether in the case of surface tension where $\alpha = 1$ or in the case of bending energies where $\alpha = 2$, exponential decay of the energy can be obtained. Whilst this is not addressed directly in this chapter, it is worth pointing out that this is no longer true when $\alpha < 1/2$ or $\alpha > 5/2$.

- When $\alpha < 1/2$ one does not obtain exponential decay of the energy for the linearized problem about equilibrium, but only algebraic decay. We refer to Tice–Zbarsky [TZ18] for details.
- When $\alpha > 5/2$, the scheme of a priori estimates is not sufficient to obtain coercivity of the dissipation over the energy. Recall that in order to show that the dissipation is coercive over the energy, we must differentiate the PDE. Indeed, upon differentiating we obtain enough control on u to make sense of the trace of the stress tensor $pI - \mathbb{D}u$, which in turn allows us to leverage the dynamic boundary condition to turn control of u into higher-order control of η, thus obtaining coercivity. Taking derivatives up to parabolic order two (i.e. taking one temporal and two spatial derivatives) we see that the only appearance of $\partial_t \eta$ in the energy is via the term
$$\mathcal{Q}_0(\partial_t \eta) \asymp ||\partial_t \eta||^2_{H^\alpha(\mathbb{T}^2)},$$
whilst the kinematic boundary tells us that
$$\mathcal{D} \gtrsim ||u||^2_{H^3(\Omega)} \gtrsim ||\operatorname{tr} u||^2_{H^{5/2}(\mathbb{T}^2)} \gtrsim ||\partial_t \eta||^2_{H^{5/2}(\mathbb{T}^2)}.$$
So indeed, for $\alpha > \frac{5}{2}$, $\mathcal{D} \not\gtrsim \mathcal{E}$. Note that this problem cannot be circumvented by applying more time derivatives, as it will always occur for the highest order term.

3.3. Strategy of the proof. In this section we sketch the strategy of the proof. We describe the key moving pieces in Section 3.3.1, then discuss how they interact in Section 3.3.2. This allows us to identify in Section 3.3.3 the 'hard analysis' estimates that have to be made to close the estimates and thus conclude the proof. Throughout this section we also outline the plan of the chapter, pointing to the location of each step of the proof.

3.3.1. *The moving pieces.* The key moving pieces are: 1. L, 2. $L_{\mathcal{X}}$, and 3. the various versions of the energy and the dissipation.

(1) We denote by L the linearization of N about the equilibrium, which is responsible for the regularity gain and transfer mechanisms.
(2) We denote by $L_{\mathcal{X}}$ a linear approximation of N about $\mathcal{X}$, which is responsible for the energy-dissipation structure of the problem. In particular $L_{\mathcal{X}}$ dictates the precise form of the energy-dissipation relation and of the commutators $\partial^\alpha \circ N - L_{\mathcal{X}} \circ \partial^\alpha$.
(3) The various versions of the energy and the dissipation (precisely defined in Section 4.5.2):

- The *equilibrium* versions, denoted by $\overline{\mathcal{E}}$ and $\overline{\mathcal{D}}$, which come from the energy-dissipation relation corresponding to the linearized problem about the equilibrium and consist of functional norms of the unknowns.
- The *improved* versions, denoted by $\mathcal{E}$ and $\mathcal{D}$, which are obtained by bootstrapping from the equilibrium versions, using the regularity gain and transfer mechanisms embedded in L. In other words, if $L\mathcal{X} = 0$ then $\overline{\mathcal{E}}$ controls $\mathcal{E}$ and $\overline{\mathcal{D}}$ controls $\mathcal{D}$.
- The *geometric* versions, denoted by $\widetilde{\mathcal{E}}$ and $\widetilde{\mathcal{D}}$, which come from the energy-dissipation relation corresponding to N and $L_{\mathcal{X}}$ and consist of functional norms of the unknowns involving the $\mathcal{G}$-differential operators and weighted by the geometric coefficient (such as J).

In particular, note that since $L_{\mathcal{X}}$ depends on $\mathcal{X}$, so do $\widetilde{\mathcal{E}}$ and $\widetilde{\mathcal{D}}$, and so we also write them as $\widetilde{\mathcal{E}}(\cdot\,;\mathcal{X})$ and $\widetilde{\mathcal{D}}(\cdot\,;\mathcal{X})$, respectively. Moreover, note that the notation we use is consistent since on one hand, when $\mathcal{X} = 0$ we have that $L_0 = L$, and on the other hand $\widetilde{\mathcal{E}}(\cdot\,;0) = \overline{\mathcal{E}}$ and $\widetilde{\mathcal{D}}(\cdot\,;0) = \overline{\mathcal{D}}$. This is summarized in the diagram below, where 'IBP' denotes integration by parts.

$$\begin{tikzcd} L_{\mathcal{X}}\mathcal{Y} = 0 \arrow[r, "IBP"] \arrow[d, "\mathcal{X}=0"] & \frac{\mathrm{d}}{\mathrm{d}t}\widetilde{\mathcal{E}}(\mathcal{Y}\,;\mathcal{X}) + \widetilde{\mathcal{D}}(\mathcal{Y}\,;\mathcal{X}) = 0 \arrow[d, "\mathcal{X}=0"] \\ L\mathcal{Y} = 0 \arrow[r, "IBP"] & \frac{\mathrm{d}}{\mathrm{d}t}\overline{\mathcal{E}}(\mathcal{Y}) + \overline{\mathcal{D}}(\mathcal{Y}) = 0 \end{tikzcd}$$

The precise derivation of the energy-dissipation relations can be found at the start of Section 5.1.

3.3.2. *How the moving pieces interact.* As discussed earlier, there are two key features of the problem that our proof relies on:

(1) Given the equilibrium versions of the energy and the dissipation, the regularity gain and transfer mechanisms embedded in the linearization L dictate the form of the improved versions. The general form of the auxiliary estimates obtained from those regularity gain and transfer mechanisms can be found at the start of Section 5.2.
(2) The form of $L_{\mathcal{X}}$ dictates the energy-dissipation structure, which thus determines the form of the geometric versions of the energy and dissipation, as well as the form of the commutators $\partial^\alpha \circ N - L_{\mathcal{X}} \circ \partial^\alpha$. The derivation of the energy-dissipation relation and the computation of the commutators can be found at the start of Section 5.1.

The interaction of the moving pieces is also summarized more tersely in Figure 4.

$$\begin{tikzcd} L\mathcal{X} = R \arrow[r, leftrightarrow] \arrow[d] & N(\mathcal{X}) = 0 \arrow[r, "\mathcal{G}-IBP"] \arrow[d, "\partial^\alpha"] & \frac{\mathrm{d}}{\mathrm{d}t}\widetilde{E}^0(\mathcal{X}) + \widetilde{D}^0(\mathcal{X}) = 0 \\ \begin{cases} \mathcal{E} \lesssim \overline{\mathcal{E}} + \mathcal{N}_E \\ \mathcal{D} \lesssim \overline{\mathcal{D}} + \mathcal{N}_D \end{cases} & L_{\mathcal{X}}(\partial^\alpha \mathcal{X}) = C^\alpha \arrow[r, "\mathcal{G}-IBP"] & \frac{\mathrm{d}}{\mathrm{d}t}\widetilde{E}(\partial^\alpha\mathcal{X};\mathcal{X}) + \widetilde{D}(\partial^\alpha\mathcal{X};\mathcal{X}) = \langle C^\alpha, \partial^\alpha\mathcal{X}\rangle_{\mathcal{X}} \end{tikzcd}$$

FIGURE 4. Schematic overview of the strategy of the proof, where $\mathcal{G} - IBP$ refers to integration by part with respect to the $\mathcal{G}$-differential operators (c.f. Section 7.1 for the relevant integration theorems).

3.3.3. *The 'hard analysis' estimates.* In order to close the estimates, we need to show that, in the small energy regime,

- the commutators are small, which is done in the latter part of Section 5.1, and
- all versions of the energy are comparable (and similarly for the dissipation), which is done in the latter part of Section 5.2 (where we essentially show that the equilibrium and improved versions are comparable) and in Section 5.3 (where we essentially show that the equilibrium and geometric versions are comparable).

4. Notation

The purpose of this section is to collect in a single place all of the notational conventions we will use throughout the rest of the chapter.

4.1. Basics. Here we collect notation for variables, derivatives, and tensor manipulations.

4.1.1. *Variables and derivatives.* We use the following notation for space-time variables.

- $T \in (0, \infty]$ denotes a time.
- For any $x = (x_1, x_2, x_3) \in \mathbb{R}^3$, we write $\bar{x} := (x_1, x_2) \in \mathbb{R}^2$ and $\tilde{x} := (\bar{x}, 0) = (x_1, x_2, 0) \in \mathbb{R}^3$.
- Similarly, we employ the following notation for derivatives: $\nabla = (\partial_1, \partial_2, \partial_3)$, $\overline{\nabla} := (\partial_1, \partial_2)$, and $\widetilde{\nabla} := \left(\overline{\nabla}, 0\right) = (\partial_1, \partial_2, 0)$.

4.1.2. *Parabolic order of multi-indices.* For any $\alpha = (\alpha_0, \bar{\alpha}) \in \mathbb{N}^{1+n}$ such that $\partial^\alpha = \partial_t^{\alpha_0} \partial_{\bar{x}}^{\bar{\alpha}}$, we define $|\alpha|_{t,\bar{x}^2} := 2\alpha_0 + \bar{\alpha}$, and call it the *parabolic order* of α.

4.1.3. *Inequalities.* We say a constant C is universal if it only depends on the various parameters of the problem, the dimension, etc., but not on the solution or the data. The notation $\alpha \lesssim \beta$ will be used to mean that there exists a universal constant $C > 0$ such that $\alpha \leq C\beta$.

4.1.4. *Contractions, inner products, and derivatives of tensors.* Throughout the chapter we will use the Einstein summation convention of summing over repeated indices. We will also need the following scalar products:

- $a \cdot b = a_i b_i$ for any $a, b \in \mathbb{R}^n$,
- $A : B = A_{ij} B_{ij}$ for any $A, B \in \mathbb{R}^{n \times n}$,
- $T \bullet S = T_{i_1 \dots i_k} S_{i_1 \dots i_k}$ for any $T, S \in \mathbb{R}^{\overbrace{n \times \cdots \times n}^{k \text{ times}}} = (\mathbb{R}^n)^{\otimes k}$.

When contracting tensors of different ranks we will write

- $(T \bullet S)_{j_1 \dots j_p k_1 \dots k_r} = T_{j_1 \dots j_p i_1 \dots i_q} S_{i_1 \dots i_q k_1 \dots k_r}$ for any $T \in (\mathbb{R}^n)^{\otimes(p+q)}$ and $S \in (\mathbb{R}^n)^{\otimes(q+r)}$, such that $T \bullet S \in (\mathbb{R}^n)^{\otimes(p+r)}$.

For derivatives of tensors we write:

- $\left(\nabla^l S\right)_{i_1 \dots i_k a_1 \dots a_l} = \partial_{a_1} \dots \partial_{a_l} S_{i_1 \dots i_k}$ for any $S : \mathbb{R}^n \to (\mathbb{R}^n)^{\otimes k}$,
- $\left(\left(\nabla^l\right)^T S\right)_{a_1 \dots a_l i_1 \dots i_k} = \partial_{a_1} \dots \partial_{a_l} S_{i_1 \dots i_k}$ for any $S : \mathbb{R}^n \to (\mathbb{R}^n)^{\otimes k}$.

4.2. Sobolev spaces. Here we record our notation for Sobolev spaces.

- For sets of the form $D = \mathbb{T}^2$ or Ω we write $H^s(D)$ to denote the usual L^2-based Sobolev space of order $s \geq 0$, and write $\dot{H}^s(D)$ to denote their homogeneous counterparts. When $D = \mathbb{T}^2$ we extend this to include $s < 0$ using the standard Fourier characterization.
- For sets of the form $D = \mathbb{T}^2$ or Ω, the notation $H^{s+}(D)$ will be employed to mean the following:

$$\begin{cases} \alpha \lesssim ||f||_{H^{s+}(D)} & \text{means that } \forall \epsilon > 0, \exists C > 0 \text{ s.t. } \alpha \leqslant C||f||_{H^{s+\epsilon}(D)} \\ ||f||_{H^{s+}(D)} \lesssim \beta & \text{means that } \exists \epsilon > 0, \exists C > 0 \text{ s.t. } ||f||_{H^{s+\epsilon}(D)} \leqslant C\beta. \end{cases}$$

4.3. Domains and coefficients. Here we record notation related to the Eulerian and fixed domains and the coefficients associated to them.

4.3.1. *Eulerian and flattened domains.* We recall that the Eulerian and fixed or equilibrium domains satisfy the following.

The Eulerian domain

- $\Omega(t) := \left\{x \in \mathbb{T}^2 \times \mathbb{R} \,\middle|\, -b < x_3 < \eta(t, \bar{x})\right\}$
- $\Sigma(t) := \left\{x \in \mathbb{T}^2 \times \mathbb{R} \,\middle|\, x_3 = \eta(t, \bar{x})\right\}$
- $\Sigma_b := \left\{x \in \mathbb{T}^2 \times \mathbb{R} \,\middle|\, x_3 = -b\right\}$
- $\partial\Omega(t) = \Sigma(t) \sqcup \Sigma_b$

The fixed domain

- $\Omega := \mathbb{T}^2 \times (-b, 0)$
- $\Sigma := \mathbb{T}^2 \times \{0\}$
- Σ_b as before
- $\partial\Omega = \Sigma \sqcup \Sigma_b$

4.3.2. *Geometric coefficients.* Recall that the flattening map Φ defined by (2.1) allows us to map Ω to $\Omega(t)$. Associated to the flattening map are the following essential geometric coefficients.

- $J := \det \nabla\Phi = 1 + \partial_3 (\mathcal{X} \operatorname{ext} \eta)$
- $\mathcal{G} := (\nabla\Phi)^{-T} = (I + e_3 \otimes \nabla (\mathcal{X} \operatorname{ext} \eta))^{-T} = I - \frac{\nabla(\mathcal{X} \operatorname{ext} \eta) \otimes e_3}{1 + \partial_3(\mathcal{X} \operatorname{ext} \eta)}$

See Lemma 7.1 for the computations of J and $\mathcal{G}$.

4.3.3. *Differential operators with variable coefficients.* Given any matrix field $M : \Omega \to \mathbb{R}^{3\times 3}$ and any vector field $v : \Omega \to \mathbb{R}^3$, we define

- $\nabla^M := M \cdot \nabla$, i.e. $\partial_i^M = M_{ij}\partial_j$
- $\mathbb{D}^M v := 2\ \mathrm{Sym}\left(\nabla^M v\right) = \nabla^M v + \left(\nabla^M v\right)^T$

When $M = \mathcal{G}$, these operators arise naturally as "Φ-conjugates" of the usual differential operators ∇ and $\mathbb{D}$. More precisely, upon changing variables via Φ we have that $\nabla^{\mathcal{G}} f = \nabla\left(f \circ \Phi^{-1}\right) \circ \Phi$ (and similarly for the symmetrized gradient). Note that, as illustrated in Figure 5, horizontal slices in the fixed domain correspond to curved hypersurfaces in the Eulerian domain. In particular, horizontal derivatives in the fixed domain correspond to derivatives tangential to these hypersurfaces in the Eulerian domain.

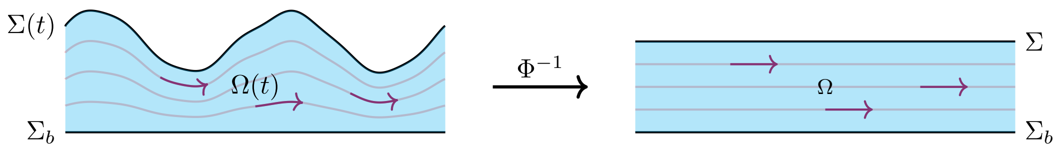


FIGURE 5. Horizontal slices in the fixed domain are mapped to curved hypersurfaces in the Eulerian domain by the diffeomorphism flattening the domain.

Since Φ is time-dependent, we also define $\mathcal{G}$-versions of time derivatives:

- $\partial_t^{\mathcal{G}} := \partial_t - (\partial_t \Phi) \cdot \nabla^{\mathcal{G}} = \partial_t - \frac{1}{J}\chi\,\mathrm{ext}\,(\partial_t \eta)\,\partial_3$
- $D_t^{v,\mathcal{G}} := \partial_t^{\mathcal{G}} + v \cdot \nabla^{\mathcal{G}}$

(c.f. Lemma 7.1 for the computation of $\partial_t \Phi$). Once again, these differential operators arise naturally when changing variables since $\partial_t^{\mathcal{G}} f = \partial_t\left(f \circ \Phi^{-1}\right) \circ \Phi$. Similarly, $D_t^{v,\mathcal{G}}$ arises naturally in the context of the $\mathcal{G}$-Reynolds transport theorem (c.f. Proposition 7.3). Finally, when integrating by parts, since $\nabla^{\mathcal{G}} \neq \nabla$ we will pick up a normal $\nu_{\partial\Omega}^{\mathcal{G}} \neq \nu_{\partial\Omega}$, where $\nu_{\partial\Omega}$ denotes the outer unit normal to $\partial\Omega$, defined as

$$\nu_{\partial\Omega}^{\mathcal{G}} := \underbrace{(\mathcal{G}J)}_{\mathrm{cof}\nabla\Phi} \cdot \nu_{\partial\Omega} = \begin{cases} -\widetilde{\nabla}\eta + e_3 & \text{on } \Sigma \\ -e_3 = \nu_{\partial\Omega} & \text{on } \Sigma_b \end{cases} \tag{4.1}$$

(see Proposition 7.2 for the statement of the $\mathcal{G}$-divergence theorem and Lemma 7.1 for the computation of $\nu_{\partial\Omega}^{\mathcal{G}}$).

4.4. Terms related to the surface energy. Here we record notation related to the surface energy.

4.4.1. *Functionals and operators associated with the surface energy.* We consider some surface energy density $f : \mathbb{R}^2 \times \mathbb{R}^{2\times 2} \to \mathbb{R}$, and define the following for any sufficiently regular $\eta, \phi, \psi, \phi_i : \mathbb{T}^2 \to \mathbb{R}$, where $i = 1, \ldots, k$.

- Jet: $\mathcal{J}\eta := \left(\nabla\eta, \nabla^2\eta\right)$, i.e. $\mathcal{J} = \left(\nabla, \nabla^2\right)$ such that $\mathcal{J}^*(w, M) = -\nabla \cdot w + \nabla^2 : M$.
- Surface energy: $\mathcal{W}(\eta) := \int_{\mathbb{T}^2} f(\mathcal{J}\eta)$.
- Directional derivatives: $\delta_\phi \mathcal{W}(\eta) := \frac{\mathrm{d}}{\mathrm{d}t}\mathcal{W}(\eta + t\phi)\,|_{t=0}$.
- Derivative: $D\mathcal{W}$ defined via $\langle D\mathcal{W}(\eta), \phi\rangle := \delta_\phi \mathcal{W}(\eta)$.
- Second derivative: $D^2\mathcal{W}$ defined via $\left\langle D^2\mathcal{W}(\eta), (\phi, \psi)\right\rangle := \delta_\phi\delta_\psi\mathcal{W}(\eta) = \delta_\psi\delta_\phi\mathcal{W}(\eta)$.
- Higher-order derivatives: for $k \in \mathbb{N}$, $D^k\mathcal{W}$ defined via

$$\left\langle D^k\mathcal{W}(\eta), (\phi_1, \phi_2, \ldots, \phi_k)\right\rangle := \delta_{\phi_1}\delta_{\phi_2}\ldots\delta_{\phi_k}\mathcal{W}(\eta).$$

- First variation: $\delta\mathcal{W}(\eta) := \mathcal{J}^*(\nabla f(\mathcal{J}\eta))$ such that

$$\langle D\mathcal{W}(\eta), \phi\rangle = \int_{\mathbb{T}^2} \delta\mathcal{W}(\eta)\,\phi = \int_{\mathbb{T}^2} \nabla f(\mathcal{J}\eta) \cdot \mathcal{J}\phi.$$

- Second variation: $\left(\delta_\eta^2\mathcal{W}\right)\phi := \mathcal{J}^*\left(\nabla^2 f(\mathcal{J}\eta) \cdot \mathcal{J}\phi\right)$ such that

$$\left\langle D^2\mathcal{W}(\eta), (\phi, \psi)\right\rangle = \int_{\mathbb{T}^2} \left(\left(\delta_\eta^2\mathcal{W}\right)\phi\right)\psi = \int_{\mathbb{T}^2} \nabla^2 f(\mathcal{J}\eta) \bullet (\mathcal{J}\phi \otimes \mathcal{J}\psi).$$

- Higher-order variations: for $k \in \mathbb{N}$,

$$\left(\delta_\eta^k \mathcal{W}\right)(\phi_1, \phi_2, \dots, \phi_{k-1}) := \mathcal{J}^* \left(\nabla^k f(\mathcal{J}\eta) \bullet (\mathcal{J}\phi_1 \otimes \mathcal{J}\phi_2 \otimes \cdots \otimes \mathcal{J}\phi_{k-1})\right)$$

  such that

$$\begin{aligned} \left\langle D^k \mathcal{W}(\eta), (\phi_1, \phi_2, \dots, \phi_{k-1}, \phi_k)\right\rangle &= \int_{\mathbb{T}^2} \Big(\left(\delta_\eta^k \mathcal{W}\right)(\phi_1, \phi_2, \dots, \phi_{k-1}) \Big) \phi_k \\ &= \int_{\mathbb{T}^2} \nabla^k f(\mathcal{J}\eta) \bullet (\mathcal{J}\phi_1 \otimes \mathcal{J}\phi_2 \otimes \cdots \otimes \mathcal{J}\phi_{k-1} \otimes \mathcal{J}\phi_k). \end{aligned}$$

- Quadratic approximation:

$$\mathcal{Q}_\eta(\phi) := \frac{1}{2}\int_{\mathbb{T}^2} \nabla^2 f(\mathcal{J}\eta) \bullet (\mathcal{J}\phi \otimes \mathcal{J}\phi) = \frac{1}{2}\int_{\mathbb{T}^2} \left(\left(\delta_\eta^2 \mathcal{W}\phi\right)\right)\phi = \frac{1}{2}\left\langle D^2 \mathcal{W}(\eta), (\phi, \phi)\right\rangle.$$

- Derivatives of the quadratic approximation: for any $\alpha \in \mathbb{N}^2$,

$$(\partial^\alpha \mathcal{Q}_\eta)(\phi) := \frac{1}{2}\int_{\mathbb{T}^2} \partial^\alpha \Big(\nabla^2 f(\mathcal{J}\eta)\Big) \bullet (\mathcal{J}\phi \otimes \mathcal{J}\phi)$$

  and in particular $\mathcal{Q}_{\dot\eta} := \partial_t \mathcal{Q}_\eta$.

4.4.2. *Constants associated to f.* At several points in our analysis we will need to refer to special constants related to the surface energy density f. We define these now.

Definition 4.1 (Universal constants). We define the following.

- Define

$$C_1 := ||\mathcal{J}||_{\mathcal{L}\left(H^{9/2}(\mathbb{T}^2); L^\infty(\mathbb{T}^2)\right)}$$

  and observe that C_1 is a finite universal constant since it only depends on the Sobolev embedding $H^s\left(\mathbb{T}^2\right) \hookrightarrow L^\infty\left(\mathbb{T}^2\right)$ for all $s > 1$.
- Define, for all $k \in \mathbb{N}$,

$$C_f^{(k)} := \left|\left|\nabla^k f\right|\right|_{L^\infty\left(\overline{B(0,C_1)}\right)}.$$

  Crucially, note that if we are in the small energy regime (see Definition 4.2), where in particular $\mathcal{E} \leqslant 1$, then

$$\left|\left|\nabla^k f(\mathcal{J}\eta)\right|\right|_{L^\infty(\mathbb{T}^2)} \leqslant C_f^{(k)} < \infty$$

  since for all $\eta : \mathbb{T}^2 \to \mathbb{R}$ sufficiently regular

$$||\mathcal{J}\eta||_{L^\infty(\mathbb{T}^2)} \leqslant C_1 ||\eta||_{H^{9/2}(\mathbb{T}^2)} \leqslant C_1\sqrt{\mathcal{E}} \leqslant C_1.$$

  Note that this will be helpful to recall when we are performing the a priori estimates since the term $\left|\left|\nabla^k f(\mathcal{J}\eta)\right|\right|_{L^\infty(\mathbb{T}^2)}$ frequently appears (for various values of k).

4.5. Quantities associated with the unknowns. Here we collect notation associated with the unknowns.

4.5.1. *Unknown variables.* We will use the following notation to refer to unknowns in the fluid equations.

- Velocities are $u, v : [0,T) \times \Omega \to \mathbb{R}^3$.
- Pressures are $p, q : [0,T) \times \Omega \to \mathbb{R}$.
- Stress tensors are $S^{\mathcal{G}}, T^{\mathcal{G}} : [0,T) \times \Omega \to \operatorname{Sym}\left(R^{3\times 3}\right)$ defined by $S^{\mathcal{G}} := pI - \mathbb{D}^{\mathcal{G}} u$ and $T^{\mathcal{G}} := qI - \mathbb{D}^{\mathcal{G}} v$.
- Surface elevations are $\eta, \zeta : [0,T) \times \mathbb{T}^2 \to (-b, \infty)$.

4.5.2. *The different versions of the energy and dissipation.* We will need various forms of the energy and dissipation functionals. We record the definitions of these now.

Geometric versions: For $\mathcal{X}_0 = (u, p, \eta)$ and $\mathcal{Y} = (v, q, \zeta)$, we define

$$\begin{cases} \widetilde{E}^0(\mathcal{X}_0) := \dfrac{1}{2}\displaystyle\int_\Omega |u|^2 J(\eta) + \mathcal{W}(\eta) + \dfrac{g}{2}\int_{\mathbb{T}^2} |\eta|^2, \\ \widetilde{E}(\mathcal{Y}; \mathcal{X}_0) := \dfrac{1}{2}\displaystyle\int_\Omega |v|^2 J(\eta) + \mathcal{Q}_\eta(\zeta) + \dfrac{g}{2}\int_{\mathbb{T}^2} |\zeta|^2, \\ \widetilde{D}^0(\mathcal{X}_0) := \dfrac{1}{2}\displaystyle\int_\Omega \left|\mathbb{D}^{\mathcal{G}(\eta)} u\right|^2 J(\eta), \text{ and} \\ \widetilde{D}(\mathcal{Y}; \mathcal{X}_0) := \dfrac{1}{2}\displaystyle\int_\Omega \left|\mathbb{D}^{\mathcal{G}(\eta)} v\right|^2 J(\eta) \end{cases}$$

where we have written $J(\eta)$ and $\mathcal{G}(\eta)$ instead of writing, as we do elsewhere, J and $\mathcal{G}$ respectively in order to emphasize the dependence on η of these geometric coefficients. We also define

$$\begin{cases} \widetilde{\mathcal{E}}(\mathcal{Y}; \mathcal{X}_0) := \widetilde{E}^0(\mathcal{X}_0) + \widetilde{E}(\partial_t \mathcal{Y}; \mathcal{X}_0) + \widetilde{E}\left(\overline{\nabla}\mathcal{Y}; \mathcal{X}_0\right) + \widetilde{E}\left(\overline{\nabla}^2\mathcal{Y}; \mathcal{X}_0\right) \text{ and} & (4.3a) \\ \widetilde{\mathcal{D}}(\mathcal{Y}; \mathcal{X}_0) := \widetilde{D}^0(\mathcal{X}_0) + \widetilde{D}(\partial_t \mathcal{Y}; \mathcal{X}_0) + \widetilde{D}\left(\overline{\nabla}\mathcal{Y}; \mathcal{X}_0\right) + \widetilde{D}\left(\overline{\nabla}^2\mathcal{Y}; \mathcal{X}_0\right) & (4.3b) \end{cases}$$

i.e. sum up to derivatives of parabolic order two, where we write $F\left(\overline{\nabla}\mathcal{Y}\right)$ to mean $\sum_i F\left(\overline{\nabla}_i\mathcal{Y}\right)$ and $F\left(\overline{\nabla}^2\mathcal{Y}\right)$ to mean $\sum_{i,j} F\left(\overline{\nabla}_{ij}\mathcal{Y}\right)$.

Note that $\widetilde{E}^0(\mathcal{X}_0)$ and $\widetilde{D}^0(\mathcal{X}_0)$ are functions whose domain is the space where $\mathcal{X}_0$ lives, but $\widetilde{E}(\mathcal{Y}; \mathcal{X}_0)$ and $\widetilde{D}(\mathcal{Y}; \mathcal{X}_0)$ are approximations of theses functions about $\mathcal{X}_0$, taking values $\mathcal{Y}$ in the *tangent space* to the space where $\mathcal{X}_0$ lives, hence why they are quadratic in $\mathcal{Y}$.

Equilibrium versions: For $\mathcal{X}_{eq} = (u_{eq}, p_{eq}, \eta_{eq}) = (0, 0, 0)$, i.e. the equilibrium configuration, and $\mathcal{Y} = (v, q, \zeta)$, we define

$$\begin{cases} \overline{E}(\mathcal{Y}) := \widetilde{E}(\mathcal{Y}; \mathcal{X}_{eq}) = \dfrac{1}{2}\displaystyle\int_\Omega |v|^2 + \mathcal{Q}_0(\zeta) + \dfrac{g}{2}\int_{\mathbb{T}^2} |\zeta|^2 \\ \quad = \dfrac{1}{2}\displaystyle\int_\Omega |v|^2 + \dfrac{1}{2}\int_{\mathbb{T}^2} \left(\left(\delta_0^2\mathcal{W} + g\right)\zeta\right)\zeta \text{ and} \\ \overline{D}(\mathcal{Y}) := \widetilde{D}(\mathcal{Y}; \mathcal{X}_{eq}) = \dfrac{1}{2}\displaystyle\int_\Omega |\mathbb{D}v|^2. \end{cases}$$

Note that, using the uniform ellipticity of $\delta_0^2\mathcal{W} + g$ stated in Section 2.2, we obtain that

$$\begin{cases} \overline{E}(\mathcal{Y}) \asymp ||v||^2_{L^2(\Omega)} + ||\zeta||^2_{H^2(\mathbb{T}^2)}, \\ \overline{D}(\mathcal{Y}) \asymp ||\mathbb{D}v||^2_{L^2(\Omega)}. \end{cases}$$

Then we define, once again summing up to derivatives of parabolic order two:

$$\begin{cases} \overline{\mathcal{E}}(\mathcal{Y}) := \overline{E}(\mathcal{Y}) + \overline{E}(\partial_t\mathcal{Y}) + \overline{E}\left(\overline{\nabla}\mathcal{Y}\right) + \overline{E}\left(\overline{\nabla}^2\mathcal{Y}\right) & (4.5a) \\ \overline{\mathcal{D}}(\mathcal{Y}) := \overline{D}(\mathcal{Y}) + \overline{D}(\partial_t\mathcal{Y}) + \overline{D}\left(\overline{\nabla}\mathcal{Y}\right) + \overline{D}\left(\overline{\nabla}^2\mathcal{Y}\right). & (4.5b) \end{cases}$$

Improved versions: For $\mathcal{Y} = (v, q, \zeta)$, we define

$$\begin{cases} \mathcal{E}(\mathcal{Y}) := ||u||^2_{H^2(\Omega)} + ||\partial_t u||^2_{L^2(\Omega)} + ||p||^2_{H^1(\Omega)} + ||\eta||^2_{H^{9/2}(\mathbb{T}^2)} + ||\partial_t\eta||^2_{H^2(\mathbb{T}^2)} \text{ and} & (4.6a) \\ \mathcal{D}(\mathcal{Y}) := ||u||^2_{H^3(\Omega)} + ||\partial_t u||^2_{H^1(\Omega)} + ||p||^2_{H^2(\Omega)} & \\ \qquad + ||\eta||^2_{H^{11/2}(\mathbb{T}^2)} + ||\partial_t\eta||^2_{H^{5/2}(\mathbb{T}^2)} + \left|\left|\partial_t^2\eta\right|\right|^2_{H^{1/2}(\mathbb{T}^2)}. & (4.6b) \end{cases}$$

Note that defined this way, coercivity is immediate, i.e. we have that $\mathcal{E} \lesssim \mathcal{D}$.

4.5.3. *Small energy regime.* We now define the 'small energy regime' that is used throughout the chapter.

DEFINITION 4.2 (Small energy regime). Let $C_0 > 0$ be defined by

$$C_0 := ||\text{ext}||_{\mathcal{L}\left(H^{3/2}(\mathbb{T}^2);\, L^\infty(\Omega)\right)} \left(\frac{1}{b} + \left|\left|\sqrt{-\Delta}\right|\right|_{\mathcal{L}\left(H^{5/2}(\mathbb{T}^2);\, H^{3/2}(\mathbb{T}^2)\right)}\right)$$

and fix some $0 < \delta_0 < \min\left(\frac{1}{C_0^2}, 1\right)$. We say that we are in the 'small energy regime' if and only if there exists a solution $\mathcal{X} = (u, p, \eta)$ on $[0, T)$ such that

$$\sup_{t\in[0,T)} \mathcal{E}\left(\mathcal{X}\right) \leqslant \delta_0 \quad \text{and} \quad \sup_{t\in[0,T)} \mathcal{D}\left(\mathcal{X}\right) < \infty.$$

The following remarks will be important later.

REMARK 4.3.

(1) $C_0 < \infty$ since

$$\begin{cases} \sqrt{-\Delta} \in \mathcal{L}\left(H^{5/2}\left(\mathbb{T}^2\right); H^{3/2}\left(T^2\right)\right), \\ \text{ext} \in \mathcal{L}\left(H^{3/2}\left(\mathbb{T}^2\right); H^2\left(\Omega\right)\right), \text{ and} \\ H^2\left(\Omega\right) \hookrightarrow L^\infty\left(\Omega\right). \end{cases}$$

(2) If $\mathcal{X}$ is a solution such that $\mathcal{E}\left(\mathcal{X}\right) \leqslant \delta_0$, then in particular, by definition of χ (c.f. Section 2.1), and by Lemma 8.7,

$$\begin{aligned} ||\partial_3\left(\chi\, \text{ext}\,\eta\right)||_{L^\infty(\Omega)} &= \left|\left|\frac{\text{ext}\,\eta}{b} + \chi\,\text{ext}\,\sqrt{-\Delta}\eta\right|\right|_{L^\infty(\Omega)} \\ &\leqslant ||\text{ext}||_{\mathcal{L}\left(H^{3/2}(\mathbb{T}^2);L^\infty(\Omega)\right)}\left(\frac{1}{b}||\eta||_{H^{3/2}(\mathbb{T}^2)} + \left|\left|\sqrt{-\Delta}\right|\right|_{\mathcal{L}\left(H^{5/2}(\mathbb{T}^2);H^{3/2}(\mathbb{T}^2)\right)}||\eta||_{H^{5/2}(\mathbb{T}^2)}\right) \\ &\leqslant C_0||\eta||_{H^{5/2}(\mathbb{T}^2)} \quad \leqslant C_0\sqrt{\mathcal{E}} \quad \leqslant C_0\sqrt{\delta_0} \quad < 1 \end{aligned}$$

and therefore $\inf \text{ext}\,\eta \geqslant -bC_0\sqrt{\delta_0} > -b$ such that Φ is well-defined, and

$$\inf J = 1 + \inf\left(\frac{\text{ext}\,\eta}{b} + \chi\,\text{ext}\,\sqrt{(-\Delta\eta)}\right) \geqslant 1 - C_0\sqrt{\delta_0} > 0$$

such that Φ is diffeomorphism.

(3) We require $\delta < 1$ in order to simplify the a priori estimates by not having to track powers of the energy. Indeed, for $\mathcal{E} \leqslant \delta_0 < 1$, $\mathcal{E}^{\alpha_1} + \cdots + \mathcal{E}^{\alpha_n} \lesssim \mathcal{E}^{\min \alpha_i}$.

5. A priori estimates

5.1. Energy-dissipation estimates. In this section we record the energy-dissipation relations arising from the original problem (known as the zeroth-order energy-dissipation relation) in Proposition 5.1 and from the differentiated problem (known as the higher-order energy-dissipation relation) in Proposition 5.2. We then sketch the computation of the commutators, relegating the full details to the appendix, and we estimate these commutators in Lemma 5.4.

We start by recording, immediately below, the energy-dissipation relation arising from the original problem. Note that in the notation of Section 3.2 this is the energy-dissipation relation corresponding to the system of PDEs $N\left(\mathcal{X}\right) = 0$.

PROPOSITION 5.1 (Zeroth-order energy-dissipation relation). *If (u, p, η) solves (2.2a)–(2.2e), then*

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(\int_\Omega \frac{1}{2}|u|^2 J + \mathcal{W}\left(\eta\right) + \int_{\mathbb{T}^2} \frac{g}{2}|\eta|^2\right) + \left(\int_\Omega \frac{1}{2}|\mathbb{D}^{\mathcal{G}} u|^2 J.\right) = 0$$

In other words, for $\widetilde{E}^0$ and $\widetilde{D}^0$ as defined in Section 4.5.2 and $\mathcal{X}_0 = (u, p, \eta)$, we have that

$$\frac{\mathrm{d}}{\mathrm{d}t}\widetilde{E}^0\left(\mathcal{X}_0\right) + \widetilde{D}^0\left(\mathcal{X}_0\right) = 0.$$

Proof. We take the dot product of (2.2a) with u, multiply by J to account for the geometry, and integrate over Ω. This results in:

$$
\begin{aligned}
0 &= \int_\Omega \left(D_t^{u,\mathcal{G}} u\right) \cdot uJ + \int_\Omega \left(\nabla^{\mathcal{G}} \cdot S^{\mathcal{G}}\right) \cdot uJ \\
&\overset{(1)}{=} \int_\Omega D_t^{u,\mathcal{G}} \left(\frac{1}{2}|u|^2\right) J - \int_\Omega S^{\mathcal{G}} : \left(\nabla^{\mathcal{G}} u\right) J + \int_\Omega \left(\left(S^{\mathcal{G}}\right)^T \cdot u\right) \cdot \nu^{\mathcal{G}}_{\partial\Omega} \\
&\overset{(2)}{=} \frac{\mathrm{d}}{\mathrm{d}t}\left(\int_\Omega \frac{1}{2}|u|^2 J\right) + \int_\Omega \frac{1}{2}\left|\mathbb{D}^{\mathcal{G}} u\right|^2 J + \int_{\partial\Omega} \left(S^{\mathcal{G}} \cdot \nu^{\mathcal{G}}_{\partial\Omega}\right) \cdot u \\
&= \frac{\mathrm{d}}{\mathrm{d}t}\left(\int_\Omega \frac{1}{2}|u|^2 J\right) + \int_\Omega \frac{1}{2}\left|\mathbb{D}^{\mathcal{G}} u\right|^2 J + \int_{\mathbb{T}^2} \left(\delta\mathcal{W}(\eta) + g\eta\right)\left(u \cdot \nu^{\mathcal{G}}_{\partial\Omega}\right) \\
&= \frac{\mathrm{d}}{\mathrm{d}t}\left(\int_\Omega \frac{1}{2}|u|^2 J\right) + \int_\Omega \frac{1}{2}\left|\mathbb{D}^{\mathcal{G}} u\right|^2 J + \int_{\mathbb{T}^2} \left(\delta\mathcal{W}(\eta) + g\eta\right)\partial_t \eta \\
&= \frac{\mathrm{d}}{\mathrm{d}t}\left(\int_\Omega \frac{1}{2}|u|^2 J\right) + \int_\Omega \frac{1}{2}\left|\mathbb{D}^{\mathcal{G}} u\right|^2 J + \frac{\mathrm{d}}{\mathrm{d}t}\left(\mathcal{W}(\eta) + \int_{\mathbb{T}^2} \frac{g}{2}|\eta|^2\right) \\
&= \frac{\mathrm{d}}{\mathrm{d}t}\left(\int_\Omega \frac{1}{2}|u|^2 J + \mathcal{W}(\eta) + \int_{\mathbb{T}^2} \frac{g}{2}|\eta|^2\right) + \int_\Omega \frac{1}{2}\left|\mathbb{D}^{\mathcal{G}} u\right|^2 J.
\end{aligned}
$$

Here in (1) we have used the $\mathcal{G}$-divergence theorem (Proposition 7.2) and the fact that $\nabla \cdot \left(M^T \cdot v\right) = M : \nabla v + (\nabla \cdot M) \cdot v$. In (2) we have used the $\mathcal{G}$-Reynolds transport theorem (Proposition 7.3) and the fact that $\nabla^{\mathcal{G}} \cdot u = 0$. □

Having recorded the energy-dissipation relation associated with the original problem above in Proposition 5.1, we now record the energy-dissipation relation associated with the differentiated problem below in Proposition 5.2. Note that in the notation of Section 3.2 and for $C = \left(C^1, C^2, C^3, C^4\right)$ this is the energy-dissipation relation corresponding to the system of PDEs $L_{\mathcal{X}_0}(\mathcal{Y}) = C$.

Proposition 5.2 (Higher-order energy-dissipation relation). *If $\mathcal{X}_0 = (u, p, \eta)$ and $\mathcal{Y} = (v, q, \zeta)$ solve*

$$
\begin{cases}
D_t^{u,\mathcal{G}} v + \nabla^{\mathcal{G}} \cdot T^{\mathcal{G}} = C^1 & \text{in } \Omega, \qquad (5.1a)\\
\nabla^{\mathcal{G}} \cdot v = C^2 & \text{in } \Omega, \qquad (5.1b)\\
\left(\left(\delta^2_\eta \mathcal{W}\right)\zeta + g\zeta\right)\nu^{\mathcal{G}}_{\partial\Omega} - T^{\mathcal{G}} \cdot \nu^{\mathcal{G}}_{\partial\Omega} = C^3 & \text{on } \Sigma, \qquad (5.1c)\\
\partial_t \zeta - v \cdot \nu^{\mathcal{G}}_{\partial\Omega} = C^4 & \text{on } \Sigma, \text{ and} \qquad (5.1d)\\
v = 0 & \text{on } \Sigma_b, \qquad (5.1e)
\end{cases}
$$

where recall that $T^{\mathcal{G}} := qI - \mathbb{D}^{\mathcal{G}} v$ (c.f. Section 4.5.1) and where $\mathcal{G} = \mathcal{G}(\eta)$, then

$$
\begin{aligned}
&\frac{\mathrm{d}}{\mathrm{d}t}\left(\int_\Omega \frac{1}{2}|v|^2 J(\eta) + \mathcal{Q}_\eta(\zeta) + \int_{\mathbb{T}^2} \frac{g}{2}|\zeta|^2\right) + \left(\int_\Omega \frac{1}{2}\left|\mathbb{D}^{\mathcal{G}(\eta)} v\right|^2 J(\eta)\right) = \\
&= \mathcal{Q}_{\dot\eta}(\zeta) + \int_\Omega \left(C^1 \cdot v\right) J(\eta) + \int_\Omega C^2 q J(\eta) + \int_{\mathbb{T}^2} C^3 \cdot v + \int_{\mathbb{T}^2} C^4 \left(\delta^2_\eta \mathcal{W} + g\right)\zeta =: \langle C, \mathcal{Y}\rangle_{\mathcal{X}_0}
\end{aligned}
$$

for $C = \left(C^1, C^2, C^3, C^4\right)$. In other words, for $\widetilde{E}$ and $\widetilde{D}$ as defined in Section 4.5.2,

$$
\frac{\mathrm{d}}{\mathrm{d}t}\widetilde{E}(\mathcal{Y}; \mathcal{X}_0) + \widetilde{D}(\mathcal{Y}; \mathcal{X}_0) = \langle C, \mathcal{Y}\rangle_{\mathcal{X}_0},
$$

where we have written $J(\eta)$ and $\mathcal{G}(\eta)$ instead of writing, as we do elsewhere, J and $\mathcal{G}$ respectively in order to emphasize the dependence on η of these geometric coefficients.

PROOF. Taking the dot product of (5.1a) with uJ and integrating over Ω yields

$$
\begin{aligned}
\overbrace{\int_\Omega \left(C^1\cdot v\right)J}^{\mathrm{I}} &= \int_\Omega \left(D_t^{\mathcal{G}} v\right)\cdot vJ + \int_\Omega \left(\nabla^{\mathcal{G}}\cdot T^{\mathcal{G}}\right)\cdot vJ \\
&= \int_\Omega D_t^{\mathcal{G}}\left(\frac{1}{2}|v|^2\right)J - \int_\Omega \left(T^{\mathcal{G}}:\nabla^{\mathcal{G}} v\right)J + \int_{\partial\Omega}\left(T^{\mathcal{G}}\cdot v\right)\cdot \nu_{\partial\Omega}^{\mathcal{G}} \\
&= \underbrace{\frac{\mathrm{d}}{\mathrm{d}t}\left(\int_\Omega \frac{1}{2}|v|^2 J\right)}_{\mathrm{II}} - \underbrace{\int_\Omega qC^2 J}_{\mathrm{III}} + \underbrace{\int_\Omega \frac{1}{2}|\mathbb{D}^{\mathcal{G}} v|^2 J}_{\mathrm{IV}} + \underbrace{\int_{\mathbb{T}^2}\left(T^{\mathcal{G}}\cdot \nu_{\partial\Omega}^{\mathcal{G}}\right)\cdot v}_{(\star)}
\end{aligned}
$$

where

$$
\begin{aligned}
(\star) &= \int_{\mathbb{T}^2}\left(\left(\delta_\eta^2\mathcal{W}+g\right)\zeta\right)\left(v\cdot\nu_{\partial\Omega}^{\mathcal{G}}\right) - \int_{\mathbb{T}^2} C^3\cdot v \\
&= \int_{\mathbb{T}^2}\left(\left(\delta_\eta^2\mathcal{W}+g\right)\zeta\right)\partial_t\zeta - \int_{\mathbb{T}^2}\left(\left(\delta_\eta^2\mathcal{W}+g\right)\zeta\right)C^4 - \int_{\mathbb{T}^2} C^3\cdot v \\
&= \left(\underbrace{\frac{\mathrm{d}}{\mathrm{d}t}\left(\mathcal{Q}_\eta(\zeta)\right)}_{\mathrm{V}} - \underbrace{\mathcal{Q}_{\dot\eta}(\zeta)}_{\mathrm{VI}}\right) + \underbrace{\frac{\mathrm{d}}{\mathrm{d}t}\left(\int_{\mathbb{T}^2}\frac{g}{2}\zeta^2\right)}_{\mathrm{VII}} - \underbrace{\int_{\mathbb{T}^2}\left(\left(\delta_\eta^2\mathcal{W}+g\right)\zeta\right)C^4}_{\mathrm{VIII}} - \underbrace{\int_{\mathbb{T}^2} C^3\cdot v}_{\mathrm{IX}}.
\end{aligned}
$$

So finally

$$
\begin{aligned}
& \mathrm{I} = \mathrm{II} - \mathrm{III} + \mathrm{IV} + \mathrm{V} - \mathrm{VI} + \mathrm{VII} - \mathrm{VIII} - \mathrm{IX} \\
\iff\quad & (\mathrm{II}+\mathrm{V}+\mathrm{VII}) + \mathrm{IV} = \mathrm{VI} + \mathrm{I} + \mathrm{III} + \mathrm{IX} + \mathrm{VIII} \\
\iff\quad & \frac{\mathrm{d}}{\mathrm{d}t}\left(\int_\Omega \frac{1}{2}|v|^2 J + \mathcal{Q}_\eta(\zeta) + \int_{\mathbb{T}^2}\frac{g}{2}|\zeta|^2\right) + \left(\int_\Omega \frac{1}{2}|\mathbb{D}^{\mathcal{G}} v|^2 J\right) = \\
& = \mathcal{Q}_{\dot\eta}(\zeta) + \int_\Omega \left(C^1\cdot v\right)J + \int_\Omega C^2 qJ + \int_{\mathbb{T}^2} C^3\cdot v + \int_{\mathbb{T}^2} C^4\left(\delta_\eta^2\mathcal{W}+g\right)\zeta.
\end{aligned}
$$

□

Using the notation from the sketch in Section 3.2, we can rephrase Proposition 5.2 as follows: if $\mathcal{X}_0$ and $\mathcal{Y}$ solve $L_{\mathcal{X}_0}(\mathcal{Y}) = C$, then $\frac{\mathrm{d}}{\mathrm{d}t}\widetilde{\mathcal{E}}(\mathcal{Y};\mathcal{X}_0) + D(\mathcal{Y};\mathcal{X}_0) = \langle C,\mathcal{Y}\rangle_{\mathcal{X}_0}$. We thus seek to compute $C^\alpha = L_{\mathcal{X}}(\partial^\alpha\mathcal{X}) - \partial^\alpha(N(\mathcal{X}))$.

As discussed in Section 3.2, the 'commutator' C^α is not quite equal to $[N,\partial^\alpha]$ because of the 'fully nonlinear' term coming from the surface energy. In particular, the terms in N are of two types: almost all terms can be written as non-constant coefficient linear operators which have a *multilinear dependence on their coefficients*, and one term (coming from the surface energy) is 'fully nonlinear' and cannot be written in that form. For terms of the first type, we have genuine commutators, and these are easy to compute: if $L = \hat{L}(\pi_a,\dots,\pi_k)$, then

$$[\partial^\alpha, L] = \sum_{\substack{\beta+\sum_{i=1}^k\gamma_i=\alpha \\ \beta<\alpha}} \hat{L}\left(\partial^{\gamma_1}\pi_1,\dots,\partial^{\gamma_k}\pi_k\right)\circ\partial^\beta.$$

See Proposition 8.10 for the full computations. For the term of the second type, we do not compute

$$\left[\nu_{\partial\Omega}^{\mathcal{G}}\delta\mathcal{W},\partial^\alpha\right] = \left(\nu_{\partial\Omega}^{\mathcal{G}}\delta\mathcal{W}\right)\circ\partial^\alpha - \partial^\alpha\circ\left(\nu_{\partial\Omega}^{\mathcal{G}}\delta\mathcal{W}\right)$$

but instead compute

$$\mathcal{C}^{\mathcal{W},\alpha}(\eta) := \left(\left(\nu_{\partial\Omega}^{\mathcal{G}}\delta_\eta^2\mathcal{W}\right)\circ\partial^\alpha - \partial^\alpha\circ\left(\nu_{\partial\Omega}^{\mathcal{G}}\delta\mathcal{W}\right)\right)(\eta).$$

REMARK 5.3. Using $\delta_\eta^2\mathcal{W}$, as opposed to $\delta_0^2\mathcal{W}$ in the differentiated version of the PDE is natural since it is precisely this operator which appears when differentiating $\delta\mathcal{W}$, i.e. since

$$\partial^\alpha\left(\delta\mathcal{W}(\eta)\right) = \left(\delta_\eta^2\mathcal{W}\right)\left(\partial^\alpha\eta\right).$$

Using $\delta_0^2\mathcal{W}$ instead of $\delta_\eta^2\mathcal{W}$ would also make it difficult to close the estimates since it would yield (due to commutators arising when differentiating the PDE in time) interactions of the form

$$\int_{\mathbb{T}^2} \underbrace{\left(\delta_\eta^2\mathcal{W} - \delta_0^2\mathcal{W}\right)}_{(\star)} (\partial_t\eta)\,(\mathrm{tr}\,\partial_t u)$$

where typically, i.e. unless the surface energy density f has a special structure, $(\star)$ involves fourth-order derivatives. For example, in the case of the 'scalar' Willmore energy, i.e.

$$\mathcal{W}(\eta) := \int_{\mathbb{T}^2} m(\nabla\eta)\,|\Delta\eta|^2$$

for some smooth $m : \mathbb{R}^2 \to (0,\infty)$ with $m(0) > 0$, we have that

$$\left(\delta_\eta^2\mathcal{W} - \delta_0^2\mathcal{W}\right)\phi = \left(m(\nabla\eta) - m(0)\right)\Delta^2\phi + 2\nabla\left(m(\nabla\eta)\right)\cdot\nabla\Delta\phi + \Delta\left(m(\nabla\eta)\right)\Delta\phi.$$

In general

$$\left(\delta_\eta^2\mathcal{W} - \delta_0^2\mathcal{W}\right)\phi = \mathcal{J}^*\left(\left(\nabla^2 f(\mathcal{J}\eta) - \nabla^2 f(0)\right)\bullet\mathcal{J}\phi\right)$$

which (again, unless f has some special structure) typically involves fourth-order derivatives of ϕ. Such interactions would be troublesome because they would thus take the form

$$\int_{\mathbb{T}^2} \left(\nabla^4\partial_t\eta\right)(\mathrm{tr}\,\partial_t u)\,(\text{l.o.t.})$$

for some lower order terms that could be controlled via the energy. Terms like this cannot be controlled in our scheme of a priori estimates because we have insufficient control of $\partial_t\eta$ and $\partial_t u$, since we only know that

$$\mathcal{D} \gtrsim \left|\left|\nabla^4\partial_t\eta\right|\right|^2_{H^{-3/2}(\mathbb{T}^2)} + \left|\left|\mathrm{tr}\,\partial_t u\right|\right|^2_{H^{1/2}(\mathbb{T}^2)}.$$

The detailed computations of $\mathcal{C}^{\mathcal{W},\alpha}(\eta)$ are in Lemma 7.5. Putting it all together, we obtain that:

$$\begin{aligned}
\langle C^\alpha, \partial^\alpha\mathcal{X}\rangle_{\mathcal{X}} = \mathcal{Q}_{\dot\eta}(\partial^\alpha\eta) &- \int_\Omega \left(\left[\partial_t\Phi\cdot\nabla^{\mathcal{G}}, \partial^\alpha\right]u\right)\cdot(\partial^\alpha u)\,J + \int_\Omega \left(\left[u\cdot\nabla^{\mathcal{G}}, \partial^\alpha\right]u\right)\cdot(\partial^\alpha u)\,J \\
&- \int_\Omega \left(\left[\left(\nabla^{\mathcal{G}}\cdot\mathcal{G}^T\right)\cdot\nabla, \partial^\alpha\right]u\right)\cdot(\partial^\alpha u)\,J - \int_\Omega \left(\left[\left(\mathcal{G}^T\cdot\mathcal{G}\right):\nabla^2, \partial^\alpha\right]u\right)\cdot(\partial^\alpha u)\,J \\
&+ \int_\Omega \left(\left[\nabla^{\mathcal{G}}, \partial^\alpha\right]p\right)\cdot(\partial^\alpha u)\,J + \int_\Omega \left(\left[\nabla^{\mathcal{G}}\cdot, \partial^\alpha\right]u\right)(\partial^\alpha p)\,J \\
&+ \int_{\mathbb{T}^2} \left(\left[\nu^{\mathcal{G}}_{\partial\Omega}\cdot\mathbb{D}^{\mathcal{G}}, \partial^\alpha\right]u\right)\cdot\partial^\alpha u - \int_{\mathbb{T}^2} \left(\left[\nu^{\mathcal{G}}_{\partial\Omega}, \partial^\alpha\right]p\right)\cdot\partial^\alpha u \\
&+ g\int_{\mathbb{T}^2} \left(\left[\nu^{\mathcal{G}}_{\partial\Omega}, \partial^\alpha\right]\eta\right)\cdot\partial^\alpha u + \int_{\mathbb{T}^2} \mathcal{C}^{\mathcal{W},\alpha}(\eta)\cdot\partial^\alpha u \\
&- \int_{\mathbb{T}^2} \left(\left[\nu^{\mathcal{G}}_{\partial\Omega}\cdot, \partial^\alpha\right]u\right)\Big(\left(\delta_\eta^2\mathcal{W} + g\right)(\partial^\alpha\eta)\Big) \\
&=: \mathrm{I} + \mathrm{II} + \mathrm{III} + \mathrm{IV} + \mathrm{V} + \mathrm{VI} + \mathrm{VII} + \mathrm{VIII} + \mathrm{IX} + \mathrm{X} + \mathrm{XI} + \mathrm{XII}.
\end{aligned} \tag{5.2}$$

The following lemma shows how these terms may be estimated.

Lemma 5.4. *If the small energy assumptions hold (see Definition 4.2), then there are functionals $\mathcal{C}^1, \mathcal{C}^2$ such that*

$$\sum_{|\alpha|_{t,\overline{x}^2}\le 2} \langle C^\alpha, \partial^\alpha\mathcal{X}\rangle_{\mathcal{X}} = \mathcal{C}^1 + \frac{\mathrm{d}}{\mathrm{d}t}\mathcal{C}^2$$

with

$$|\mathcal{C}^1| \lesssim \sqrt{\mathcal{E}}\mathcal{D} \text{ and } |\mathcal{C}^2| \lesssim \sqrt{\mathcal{E}}\mathcal{E}.$$

Proof. We begin with a sketch of the general argument. Most of the commutators appearing in I − XII in (5.2) are multilinear in terms of quantities that we control (such as the unknowns u, p, η and geometric coefficients J, $\mathcal{G}$, Φ, $\nu^{\mathcal{G}}_{\partial\Omega}$). To handle such commutators, we use the Hölder and Sobolev inequalities. See Proposition 8.11 for how we control terms of the form

$$\left|\int f_1 \dots f_k\right|$$

when we control the f_i's in some H^{s_i} spaces.

In some cases, we may need to use a couple of other tools to be able to place functions in Sobolev spaces of sufficiently high regularity. We may need to 'borrow' regularity, i.e. use that $H^{s+\alpha}(\mathbb{R}^n) \cdot H^{s+\beta}(\mathbb{R}^n) \hookrightarrow H^s(\mathbb{R}^n)$: see Propositions 8.12, 8.13, and 8.14. We also need to use post-composition results, i.e. use that $C^{k,\alpha}(H^s(\mathbb{R}^n)) \hookrightarrow H^s(\mathbb{R}^n)$: see Proposition 8.15.

For a few commutators, namely XI and XII, we will need to use the smallness and boundedness of variations of the surface energy, i.e. Lemmas 7.9, 7.10, and 7.11.

Estimates of these forms ultimately contribute to $\mathcal{C}^1$. We now turn to the question of how $\mathcal{C}^2$ arises. The term VII involves an appearance of $\partial_t p$, which is not controlled in either the energy or dissipation, though it is defined through the local existence theory in a manner that allows us to integrate by parts in time:

$$\int_\Omega (\partial_t p)\, w = \frac{\mathrm{d}}{\mathrm{d}t}\left(\int_\Omega pw\right) - \int_\Omega p\,(\partial_t w)\,.$$

Note that the non-time-differentiated term can be controlled by $\sqrt{\mathcal{E}}\mathcal{D}$ like any of the other commutators contributing to $\mathcal{C}^1$, but the time-differentiated term must be controlled at a lower regularity level by $\mathcal{E}^{3/2}$. In particular, the term of the form $\int_\Omega pw$ arising from commutator VII is the only contribution to $\mathcal{C}^2$.

We now provide detailed proofs for the estimates of four terms that are particularly delicate. For example, three of them are 'critical' in the sense that they lead to a full factor of $\mathcal{D}$ appearing, suggesting that they are precisely at the limit of what the improved energy and dissipation allow us to control. Moreover, these four terms are representative of various difficulties encountered. We thus detail how to control:

(1) the commutator I when $\partial^\alpha = \partial_t$ since it highlights how to handle terms of the form $|\int f_1 \dots f_k|$,
(2) the commutator VII when $\partial^\alpha = \partial_t$ since this is precisely the term that requires integration by parts in time in order to be brought under control,
(3) the commutator XI when $\partial^\alpha = \overline{\nabla}^2$ since it requires intermediate results about the smallness of $\delta\mathcal{W}$, $\delta^2_\eta\mathcal{W}$, and $\delta^3_\eta\mathcal{W}$, and since it highlights how post-composition and product estimates in Sobolev spaces are used, and
(4) the commutator XII when $\partial^\alpha = \partial_t$, for the same reasons.

Estimating the remaining commutators follows a similar procedure and thus we omit those estimates.

(1) *A typical estimate on the surface.* We detail how to control the commutator I when $\partial^\alpha = \partial_t$. The commutator is

$$\mathcal{Q}_{\dot\eta}(\partial_t\eta) = \frac{1}{2}\int_{\mathbb{T}^2} \nabla^3 f(\mathcal{J}\eta)\cdot\left((\mathcal{J}\partial_t\eta)^{\otimes 3},\right)$$

and it can be controlled as follows:

$$\begin{aligned}|\mathcal{Q}_{\dot\eta}(\partial_t\eta)| &\lesssim \left|\left|\nabla^3 f(\mathcal{J}\eta)\right|\right|_{L^\infty(\mathbb{T}^2)} ||\mathcal{J}\partial_t\eta||_{L^2(\mathbb{T}^2)} ||\mathcal{J}\partial_t\eta||^2_{L^4(\mathbb{T}^2)}\\ &\lesssim C^{(3)}_f ||\partial_t\eta||_{H^2(\mathbb{T}^2)} ||\mathcal{J}\partial_t\eta||^2_{H^{1/2}(\mathbb{T}^2)}\\ &\lesssim \sqrt{\mathcal{E}}||\partial_t\eta||^2_{H^{5/2}(\mathbb{T}^2)} \quad \lesssim \sqrt{\mathcal{E}}\mathcal{D}.\end{aligned}$$

Recall that $C^{(3)}_f$ is defined in Definition 4.1.

(2) *Integration by parts in time.* We detail how to control the commutator VII when $\partial^\alpha = \partial_t$. The commutator is

$$\int_\Omega (\partial_t\mathcal{G}) : (\nabla u)\,(\partial_t p)\, J.$$

Schematically, we have:

$$\int_\Omega (\partial_t p)\, w = \frac{\mathrm{d}}{\mathrm{d}t}\left(\int_\Omega pw\right) - \int_\Omega p\,(\partial_t w)$$

where we may *only* use the energy (and not the dissipation) to control $\int_\Omega pw$ since it is time-differentiated, and where we may proceed as usual, i.e. using both the energy and the dissipation, but not using the dissipation more than twice, to control $\int_\Omega p\,(\partial_t w)$. The first term is

$$\int_\Omega (\partial_t \mathcal{G} : \nabla u)\, pJ,$$

and it can be estimated in the following way:

$$\begin{aligned}|\ldots| &\lesssim ||\partial_t \mathcal{G}||_{L^3(\Omega)} ||\nabla u||_{L^3(\Omega)} ||p||_{L^3(\Omega)} ||J||_{L^\infty(\Omega)}\\ &\lesssim ||\partial_t \mathcal{G}||_{H^{1/2}(\Omega)} ||\nabla u||_{H^{1/2}(\Omega)} ||p||_{H^{1/2}(\Omega)} ||J||_{H^{3/2+}(\Omega)}\\ &\lesssim \left(1+\sqrt{\mathcal{E}}\right)\mathcal{E}^{3/2} \quad \lesssim \mathcal{E}^{3/2}.\end{aligned}$$

The second term is

$$\int_\Omega \left(\partial_t^2 \mathcal{G} : \nabla u\right) Jp + \int_\Omega (\partial_t \mathcal{G} : \nabla \partial_t u)\, Jp + \int_\Omega (\partial_t \mathcal{G} : \nabla u)\,(\partial_t J)\, p,$$

and can be estimated in the following way:

$$\begin{aligned}|\ldots| &\lesssim \left|\left|\partial_t^2 \mathcal{G}\right|\right|_{L^2(\Omega)} ||\nabla u||_{L^6(\Omega)} ||J||_{L^\infty(\Omega)} ||p||_{L^6(\Omega)}\\ &\quad + ||\partial_t \mathcal{G}||_{L^6(\Omega)} ||\nabla \partial_t u||_{L^2(\Omega)} ||J||_{L^\infty(\Omega)} ||p||_{L^6(\Omega)}\\ &\quad + ||\partial_t \mathcal{G}||_{L^3(\Omega)} ||\nabla u||_{L^3(\Omega)} ||\partial_t J||_{L^\infty(\Omega)} ||p||_{L^3(\Omega)}\\ &\lesssim \left|\left|\partial_t^2 \mathcal{G}\right|\right|_{L^2(\Omega)} ||\nabla u||_{H^1(\Omega)} ||J||_{H^{3/2+}(\Omega)} ||p||_{H^1(\Omega)}\\ &\quad + ||\partial_t \mathcal{G}||_{H^1(\Omega)} ||\nabla \partial_t u||_{L^2(\Omega)} ||J||_{H^{3/2+}(\Omega)} ||p||_{H^1(\Omega)}\\ &\quad + ||\partial_t \mathcal{G}||_{H^{1/2}(\Omega)} ||\nabla u||_{H^{1/2}(\Omega)} ||\partial_t J||_{H^{3/2+}(\Omega)} ||p||_{H^{1/2}(\Omega)}\\ &\lesssim \sqrt{\mathcal{D}}\left(1+\sqrt{\mathcal{E}}\right)\mathcal{E} + \left(1+\sqrt{\mathcal{E}}\right)\mathcal{E}^{3/2} + \mathcal{E}^2 \quad \lesssim (1+\mathcal{E})\,\mathcal{E}\sqrt{\mathcal{D}}.\end{aligned}$$

(3) *Another typical estimate on the surface.* We detail how to control the commutator XI when $\partial^\alpha = \overline{\nabla}^2$. The commutator is

$$\begin{aligned}\int_{\mathbb{T}^2} \left(\nabla^3 \eta\right) \delta\mathcal{W}(\eta)\left(\operatorname{tr}\nabla^2 u\right) + \int_{\mathbb{T}^2}\left(\nabla^2\eta\right)\left(\left(\delta_\eta^2 \mathcal{W}\right)(\nabla\eta)\right)\left(\operatorname{tr}\nabla^2 u\right)\\ + \int_{\mathbb{T}^2} \nu_{\partial\Omega}^{\mathcal{G}}\left(\left(\delta_\eta^3\mathcal{W}\right)(\nabla\eta,\nabla\eta)\right)\left(\operatorname{tr}\nabla^2 u\right)\\ =: \mathrm{XI}_1 + \mathrm{XI}_2 + \mathrm{XI}_3.\end{aligned}$$

The first two terms can be estimated in the following way:

$$\begin{aligned}|\mathrm{XI}_1 + \mathrm{XI}_2| &\lesssim \left|\left|\nabla^3\eta\right|\right|_{L^\infty(\mathbb{T}^2)} ||\delta\mathcal{W}(\eta)||_{L^2(\mathbb{T}^2)} \left|\left|\operatorname{tr}\nabla^2 u\right|\right|_{L^2(\mathbb{T}^2)}\\ &\quad + \left|\left|\nabla^2\eta\right|\right|_{L^\infty(\mathbb{T}^2)} \left|\left|\left(\delta_\eta^2\mathcal{W}\right)(\nabla\eta)\right|\right|_{L^2(\mathbb{T}^2)} \left|\left|\nabla^2 u\right|\right|_{L^2(\mathbb{T}^2)}\\ &\lesssim \left|\left|\nabla^3\eta\right|\right|_{H^{1+}(\mathbb{T}^2)} ||\delta\mathcal{W}(\eta)||_{L^2(\mathbb{T}^2)} \left|\left|\nabla^2 u\right|\right|_{H^{1/2}(\Omega)}\\ &\quad + \left|\left|\nabla^2\eta\right|\right|_{H^{1+}(\mathbb{T}^2)} \left|\left|\left(\delta_\eta^2\mathcal{W}\right)(\nabla\eta)\right|\right|_{L^2(\mathbb{T}^2)} \left|\left|\nabla^2 u\right|\right|_{H^{1/2}(\Omega)}\\ &\lesssim \sqrt{\mathcal{E}}\sqrt{\mathcal{E}}\sqrt{\mathcal{D}} + \sqrt{\mathcal{E}}\sqrt{\mathcal{D}}\sqrt{\mathcal{D}} \quad \lesssim \mathcal{E}\sqrt{\mathcal{D}} + \sqrt{\mathcal{E}}\mathcal{D},\end{aligned}$$

where we have used that $||\delta\mathcal{W}(\eta)||_{H^{1/2}(\mathbb{T}^2)} \lesssim \sqrt{\mathcal{E}}$, and that

$$\begin{aligned}\left|\left|\left(\delta_\eta^2\mathcal{W}\right)(\nabla\eta)\right|\right|_{L^2(\mathbb{T}^2)} &= \left|\left|\mathcal{J}^*\left(\nabla^2 f(\mathcal{J}\eta)\bullet\mathcal{J}\nabla\eta\right)\right|\right|_{L^2(\mathbb{T}^2)}\\ &\lesssim \left|\left|\nabla^2 f(\mathcal{J}\eta)\bullet\mathcal{J}\nabla\eta\right|\right|_{H^2(\mathbb{T}^2)}\\ &\lesssim \left|\left|\nabla^2 f(\mathcal{J}\eta)\right|\right|_{H^2(\mathbb{T}^2)} ||\mathcal{J}\nabla\eta||_{H^2(\mathbb{T}^2)}\\ &\lesssim \left(C_f^{(2)} + C_f^{(5)}\left(||\mathcal{J}\eta||_{H^2(\mathbb{T}^2)} + ||\mathcal{J}\eta||_{H^2(\mathbb{T}^2)}^2\right)\right)||\eta||_{H^5(\mathbb{T}^2)}\\ &\lesssim \left(1+||\eta||_{H^4(\mathbb{T}^2)} + ||\eta||_{H^4(\mathbb{T}^2)}^2\right)\sqrt{\mathcal{D}} \quad \lesssim \sqrt{\mathcal{D}},\end{aligned}$$

recalling that $C_f^{(3)}$ is defined in Definition 4.1. The last term requires a bit more precaution:

$$\begin{aligned}|\mathrm{XI}_3| &\lesssim \left|\left|\nu^{\mathcal{G}}_{\partial\Omega}\left(\mathrm{tr}\,\nabla^2 u\right)\right|\right|_{H^{1/2}(\mathbb{T}^2)}\left|\left|\left(\delta^3_\eta\mathcal{W}\right)(\nabla\eta,\nabla\eta)\right|\right|_{H^{-1/2}(\mathbb{T}^2)}\\ &\lesssim \left|\left|\nu^{\mathcal{G}}_{\partial\Omega}\right|\right|_{H^{\frac{3}{2}+}(\mathbb{T}^2)}\left|\left|\mathrm{tr}\,\nabla^2 u\right|\right|_{H^{1/2}(\mathbb{T}^2)}\left|\left|\nabla^3 f\left(\mathcal{J}\eta\right)\bullet\left(\mathcal{J}\eta\otimes\mathcal{J}\eta\right)\right|\right|_{H^{3/2}(\mathbb{T}^2)}\\ &\lesssim \left(1+\sqrt{\mathcal{E}}\right)\sqrt{\mathcal{D}}\left|\left|\nabla^3 f\left(\mathcal{J}\eta\right)\right|\right|_{H^{3/2}(\mathbb{T}^2)}\left|\left|\mathcal{J}\nabla\eta\right|\right|^2_{H^{3/2}(\mathbb{T}^2)}\\ &\lesssim \left(1+\sqrt{\mathcal{E}}\right)\sqrt{\mathcal{D}}\sqrt{\mathcal{E}}\mathcal{E} \quad \lesssim \left(1+\sqrt{\mathcal{E}}\right)\mathcal{E}^{3/2}\sqrt{\mathcal{D}} \quad \lesssim \mathcal{E}\sqrt{\mathcal{D}},\end{aligned}$$

where we have used that

$$\begin{aligned}\left|\left|\nabla^3 f\left(\mathcal{J}\eta\right)\right|\right|_{H^{3/2}(\mathbb{T}^2)} &\lesssim C_f^{(3)} + C_f^{(5)}\left(||\mathcal{J}\eta||_{H^{3/2}(\mathbb{T}^2)} + ||\mathcal{J}\eta||^2_{H^{3/2}(\mathbb{T}^2)}\right)\\ &\lesssim 1+\sqrt{\mathcal{E}}+\mathcal{E}^{3/2} \quad \lesssim 1.\end{aligned}$$

(4) *One last typical estimate on the surface.* We detail how to control the commutator XII when $\partial^\alpha = \partial_t$. The commutator is

$$\int_{\mathbb{T}^2}(\nabla\partial_t\eta)\,(\mathrm{tr}\,u)\left(\left(\delta^2_\eta\mathcal{W}+g\right)(\partial_t\eta)\right),$$

and it can be estimated in the following way

$$\begin{aligned}|\ldots| &\lesssim ||(\nabla\partial_t\eta)\,(\mathrm{tr}\,u)||_{H^{3/2}(\mathbb{T}^2)}\left|\left|\left(\delta^2_\eta\mathcal{W}+g\right)(\partial_t\eta)\right|\right|_{H^{-3/2}(\mathbb{T}^2)}\\ &\lesssim ||\nabla\partial_t\eta||_{H^{3/2}(\mathbb{T}^2)}||\mathrm{tr}\,u||_{H^{3/2}(\mathbb{T}^2)}\left(\left|\left|\delta^2_\eta\mathcal{W}\right|\right|_{\mathcal{L}\left(H^{5/2};\,H^{-3/2}\right)}+1\right)||\partial_t\eta||_{H^{5/2}(\mathbb{T}^2)}\\ &\lesssim \sqrt{\mathcal{D}}\sqrt{\mathcal{E}}\sqrt{\mathcal{D}} \quad \lesssim \sqrt{\mathcal{E}}\mathcal{D}.\end{aligned}$$

□

5.2. Regularity gain. In this section we record the auxiliary estimates arising from the linearized problem (about the equilibrium) in Proposition 5.5, we compute the nonlinear remainders obtained when writing the full nonlinear problem as a perturbation of its linearization, and finally we estimate these nonlinear remainders in Lemma 5.7.

We begin by recording our auxiliary estimates in a general form.

PROPOSITION 5.5 (Generic form of the auxiliary estimates). *Let $R = \left(R^1, R^2, R^3, R^4\right)$ be given and suppose that (u,p,η) solves*

$$\begin{cases}\partial_t u - \Delta u + \nabla p = R^1 & \text{in } \Omega,\\ \nabla\cdot u = R^2 & \text{in } \Omega,\\ \left(\delta^2_0\mathcal{W}+g\right)\eta e_3 + \mathbb{D}u\cdot e_3 - pe_3 = R^3 & \text{on } \Sigma,\\ \partial_t\eta - u\cdot e_3 = R^4 & \text{on } \Sigma, \text{ and}\\ u = 0 & \text{on } \Sigma_b.\end{cases}$$

Then [2]

$$\begin{aligned}&||u||_{H^2(\Omega)} + ||\partial_t u||_{L^2(\Omega)} + ||p||_{H^1(\Omega)} + ||\eta||_{H^{9/2}(\mathbb{T}^2)} + ||\partial_t\eta||_{H^2(\mathbb{T}^2)}\\ &\quad\lesssim ||\partial_t u||_{L^2(\Omega)} + ||\eta||_{H^4(\mathbb{T}^2)} + ||\partial_t\eta||_{H^2(\mathbb{T}^2)} + ||u||_{L^2(\Omega)}\\ &\quad + \left|\left|R^1\right|\right|_{L^2(\Omega)} + \left|\left|R^2\right|\right|_{H^1(\Omega)} + \left|\left|R^3\right|\right|_{H^{1/2}(\mathbb{T}^2)} + \left|\left|R^4\right|\right|_{H^{3/2}(\mathbb{T}^2)}\end{aligned} \tag{5.3}$$

and [3]

$$\begin{aligned}&||u||_{H^3(\Omega)} + ||\partial_t u||_{H^1(\Omega)} + ||p||_{H^2(\Omega)} + ||\eta||_{H^{11/2}(\mathbb{T}^2)} + ||\partial_t\eta||_{H^{5/2}(\mathbb{T}^2)} + \left|\left|\partial_t^2\eta\right|\right|_{H^{1/2}(\mathbb{T}^2)}\\ &\quad\lesssim ||\mathbb{D}u||_{L^2(\Omega)} + ||\mathbb{D}\partial_t u||_{L^2(\Omega)} + \left|\left|\mathbb{D}\overline{\nabla}^2 u\right|\right|_{L^2(\Omega)}\\ &\quad + \left|\left|R^1\right|\right|_{H^1(\Omega)} + \left|\left|R^2\right|\right|_{H^2(\Omega)} + \left|\left|R^3\right|\right|_{H^{3/2}(\mathbb{T}^2)} + \left|\left|R^4\right|\right|_{H^{5/2}(\mathbb{T}^2)} + \left|\left|\partial_t R^4\right|\right|_{H^{1/2}(\mathbb{T}^2)}\end{aligned} \tag{5.4}$$

---

[2]Note that the terms $||\overline{\nabla}u||_{L^2}$ and $\left|\left|\overline{\nabla}^2 u\right|\right|_{L^2}$ are present in $\overline{\mathcal{E}}$ but are absent from the right-hand side of the estimate.

[3]Note that the term $||\mathbb{D}\overline{\nabla}u||_{L^2}$ is present in $\overline{\mathcal{D}}$ but are absent from the right-hand side of the estimate.

i.e.

$$\begin{cases} \mathcal{E} \lesssim \overline{\mathcal{E}} + \mathcal{N}_E \text{ and} \\ \mathcal{D} \lesssim \overline{\mathcal{D}} + \mathcal{N}_D \end{cases}$$

for

$$\begin{cases} \mathcal{N}_E := \left|\left|R^1\right|\right|^2_{L^2(\Omega)} + \left|\left|R^2\right|\right|^2_{H^1(\Omega)} + \left|\left|R^3\right|\right|^2_{H^{1/2}(\mathbb{T}^2)} + \left|\left|R^4\right|\right|^2_{H^{3/2}(\mathbb{T}^2)} \\ \mathcal{N}_D := \left|\left|R^1\right|\right|^2_{H^1(\Omega)} + \left|\left|R^2\right|\right|^2_{H^2(\Omega)} + \left|\left|R^3\right|\right|^2_{H^{3/2}(\mathbb{T}^2)} + \left|\left|R^4\right|\right|^2_{H^{5/2}(\mathbb{T}^2)} + \left|\left|\partial_t R^4\right|\right|^2_{H^{1/2}(\mathbb{T}^2)}. \end{cases}$$

PROOF. We begin with the estimates related to the energy. We divide the argument into several steps.

(1) We initiate our scheme of estimates in the usual way for parabolic problems: treat temporal derivatives as forcing terms in the stationary equations in order to recover control of the spatial derivatives from control of the temporal derivatives. In particular, note that (u, p, η) solves a Stokes problem with mixed boundary conditions where $\partial_t u$ and $\partial_t \eta$ are treated as forcing terms, i.e.

$$\begin{cases} -\Delta u + \nabla p = -\partial_t u + R^1 & \text{in } \Omega, \\ \nabla \cdot u = R^2 & \text{in } \Omega, \\ u \cdot e_3 = \partial_t \eta - R^4 & \text{on } \Sigma, \\ (\mathbb{D} u e_3)_{tan} = \left(R^3\right)_{tan} & \text{on } \Sigma, \text{ and} \\ u = 0 & \text{on } \Sigma_b, \end{cases}$$

where for any vector field $w : \Sigma \to \mathbb{R}^3$ we denote by w_{tan} its tangential part, i.e. $w_{tan} = (I - e_3 \otimes e_3)\, w$. Therefore, by using elliptic regularity estimates for the Stokes problem (i.e. the auxiliary estimate 8.17) we obtain that

$$\begin{aligned} ||u||_{H^2(\Omega)} + ||\nabla p||_{L^2(\Omega)} &\lesssim \left|\left|-\partial_t u + R^1\right|\right|_{L^2(\Omega)} + \left|\left|R^2\right|\right|_{H^1(\Omega)} + \left|\left|\partial_t \eta - R^4\right|\right|_{H^{3/2}(\mathbb{T}^2)} + \left|\left|\left(R^3\right)_{tan}\right|\right|_{H^{1/2}(\mathbb{T}^2)} \\ &\leqslant ||\partial_t u||_{L^2(\Omega)} + ||\partial_t \eta||_{H^{3/2}(\mathbb{T}^2)} + \left|\left|R^1\right|\right|_{L^2(\Omega)} + \left|\left|R^2\right|\right|_{H^1(\Omega)} + \left|\left|\left(R^3\right)_{tan}\right|\right|_{H^{1/2}(\mathbb{T}^2)} + \left|\left|R^4\right|\right|_{H^{3/2}(\mathbb{T}^2)}. \end{aligned}$$

(2) Ultimately, we wish to control the full H^1 norm of p via the improved energy, but so far we only control the gradient of p. In order to proceed further we therefore use the normal component of the dynamic boundary condition to obtain control of the trace of p on the top boundary. Indeed, since

$$p = \mathbb{D}u : (e_3 \otimes e_3) + \left(\delta_0^2 \mathcal{W} + g\right) \eta - R^3 \cdot e_3 \qquad \text{on } \mathbb{T}^2 \, (\sim \Sigma)$$

it follows that

$$\begin{aligned} ||\mathrm{tr}_\Sigma\, p||_{L^2(\mathbb{T}^2)} &\lesssim ||\mathrm{tr}_\Sigma\, \mathbb{D}u||_{L^2(\mathbb{T}^2)} + ||\eta||_{H^4(\mathbb{T}^2)} + \left|\left|R^3 \cdot e_3\right|\right|_{L^2(\mathbb{T}^2)} \\ &\lesssim ||u||_{H^{3/2}(\Omega)} + ||\eta||_{H^4(\mathbb{T}^2)} + \left|\left|R^3 \cdot e_3\right|\right|_{L^2(\mathbb{T}^2)}. \end{aligned}$$

(3) We can now, as intended, recover control of the full H^1 norm of p by using a Poincaré-type inequality (i.e. auxiliary estimate 8.1):

$$||p||_{H^1(\Omega)} \lesssim ||\mathrm{tr}_\Sigma\, p||_{L^2(\mathbb{T}^2)} + ||\nabla p||_{L^2(\Omega)}.$$

(4) Now that we have enough control on the stress tensor to obtain estimates for its trace onto the boundary, we can use the normal component of the dynamic boundary condition to obtain control of higher-order spatial derivatives of η. Indeed, since

$$\left(\delta_0^2 \mathcal{W} + g\right) \eta = p - \mathbb{D}u : (e_3 \otimes e_3) + R^3 \cdot e_3$$

it follows from the elliptic regularity of $\delta_0^2 \mathcal{W} + g$ (i.e. the auxiliary estimate 8.20) that

$$\begin{aligned} ||\eta||_{H^{9/2}(\mathbb{T}^2)} &\lesssim ||\mathrm{tr}_\Sigma\, p||_{H^{1/2}(\mathbb{T}^2)} + ||\mathrm{tr}_\Sigma\, \mathbb{D}u||_{H^{1/2}(\mathbb{T}^2)} + \left|\left|R^3 \cdot e_3\right|\right|_{H^{1/2}(\mathbb{T}^2)} \\ &\lesssim ||p||_{H^1(\Omega)} + ||u||_{H^2(\Omega)} + \left|\left|R^3 \cdot e_3\right|\right|_{H^{1/2}(\mathbb{T}^2)}. \end{aligned}$$

Assembling the above estimates, we see that

$$\begin{aligned} ||u||_{H^2(\Omega)} + ||p||_{H^1(\Omega)} + ||\eta||_{H^{9/2}(\mathbb{T}^2)} &\lesssim ||\partial_t u||_{L^2(\Omega)} + ||\partial_t \eta||_{H^{3/2}(\mathbb{T}^2)} + ||\eta||_{H^4(\mathbb{T}^2)} \\ &\quad + \left|\left|R^1\right|\right|_{L^2(\Omega)} + \left|\left|R^2\right|\right|_{H^1(\Omega)} + \left|\left|R^3\right|\right|_{H^{1/2}(\mathbb{T}^2)} + \left|\left|R^4\right|\right|_{H^{3/2}(\mathbb{T}^2)}. \end{aligned}$$

Then (5.3) follows immediately from this.

We now turn our attention to estimates related to the dissipation. Again, we divide the argument into steps.

(1) We begin by trading control of the symmetrized gradient for control of full H^1 norms. This is possible due to the no-slip boundary conditions and a Korn-type inequality (i.e. auxiliary estimate 8.22):

$$\begin{cases} \|u\|_{H^1(\Omega)} & \lesssim \|\mathbb{D}u\|_{L^2(\Omega)}, \\ \|\partial_t u\|_{H^1(\Omega)} & \lesssim \|\mathbb{D}\partial_t u\|_{L^2(\Omega)}, \text{ and} \\ \left\|\overline{\nabla}^2 u\right\|_{H^1(\Omega)} & \lesssim \left\|\mathbb{D}\overline{\nabla}^2 u\right\|_{L^2(\Omega)}. \end{cases}$$

(2) Next we use the fact that the horizontal derivatives of the trace of u are equal to the trace of the horizontal derivatives, i.e. $\overline{\nabla} \circ \operatorname{tr}_\Sigma = \operatorname{tr}_\Sigma \circ \overline{\nabla}$. From this and standard trace estimates we obtain:

$$\begin{aligned} \|\operatorname{tr}_\Sigma u\|_{H^{5/2}(\mathbb{T}^2)} &\lesssim \|\operatorname{tr}_\Sigma u\|_{H^{1/2}(\mathbb{T}^2)} + \left\|\overline{\nabla}^2 (\operatorname{tr}_\Sigma u)\right\|_{H^{1/2}(\mathbb{T}^2)} \\ &\lesssim \|u\|_{H^1(\Omega)} + \left\|\operatorname{tr}_\Sigma \overline{\nabla}^2 u\right\|_{H^{1/2}(\Omega)} \\ &\lesssim \|u\|_{H^1(\Omega)} + \left\|\overline{\nabla}^2 u\right\|_{H^1(\Omega)}. \end{aligned}$$

(3) We can now recover control of all the derivatives of u by using the trace of u as datum in a Stokes problem with Dirichlet boundary conditions. Indeed, since

$$\begin{cases} -\Delta u + \nabla p = -\partial_t u + R^1 & \text{in } \Omega, \\ \nabla \cdot u = R^2 & \text{in } \Omega, \\ u = u & \text{on } \Sigma, \\ u = 0 & \text{on } \Sigma_b \end{cases}$$

it follows from elliptic regularity estimates for the Stokes problem (i.e. the auxiliary estimate 8.16) that

$$\begin{aligned} \|u\|_{H^3(\Omega)} + \|\nabla p\|_{H^1(\Omega)} &\lesssim \left\|-\partial_t u + R^1\right\|_{H^1(\Omega)} + \left\|R^2\right\|_{H^2(\Omega)} + \|\operatorname{tr}_\Sigma u\|_{H^{5/2}(\mathbb{T}^2)} \\ &\leqslant \|\partial_t u\|_{H^1(\Omega)} + \|\operatorname{tr}_\Sigma u\|_{H^{5/2}(\mathbb{T}^2)} + \left\|R^1\right\|_{H^1(\Omega)} + \left\|R^2\right\|_{H^2(\Omega)}. \end{aligned}$$

(4) Next we observe that

$$\left(\delta_0^2 \mathcal{W} + g\right)\left(\overline{\nabla}\eta\right) = \overline{\nabla} p - \mathbb{D}\overline{\nabla} u : (e_3 \otimes e_3) + \overline{\nabla} R^3 \qquad \text{on } \mathbb{T}^2 \, (\sim \Sigma)$$

and therefore elliptic estimates for the operator $\delta_0^2 \mathcal{W} + g$ (i.e. the auxiliary estimate 8.20) provide the bounds

$$\begin{aligned} \left\|\overline{\nabla}\eta\right\|_{H^{9/2}(\mathbb{T}^2)} &\lesssim \left\|\operatorname{tr}_\Sigma \overline{\nabla} p\right\|_{H^{1/2}(\mathbb{T}^2)} + \left\|\operatorname{tr}_\Sigma \mathbb{D}\overline{\nabla} u\right\|_{H^{1/2}(\mathbb{T}^2)} + \left\|\overline{\nabla} R^3\right\|_{H^{1/2}(\mathbb{T}^2)} \\ &\lesssim \|\nabla p\|_{H^1(\Omega)} + \|u\|_{H^3(\Omega)} + \left\|\overline{\nabla} R^3\right\|_{H^{1/2}(\mathbb{T}^2)}. \end{aligned}$$

Moreover, since $\int_{\mathbb{T}^2} \eta = 0$, we have that $\|\eta\|_{H^{11/2}} \lesssim \left\|\overline{\nabla}\eta\right\|_{H^{9/2}}$ (by auxiliary estimate 8.2), and so, finally, we have

$$\|\eta\|_{H^{11/2}(\mathbb{T}^2)} \lesssim \|\nabla p\|_{H^1(\Omega)} + \|u\|_{H^3(\Omega)} + \left\|\overline{\nabla} R^3\right\|_{H^{1/2}(\mathbb{T}^2)}.$$

(5) We now parlay the η estimates into full H^2 control of the pressure by arguing as we did for the energy, obtaining control of the trace of the pressure. Since

$$p = \mathbb{D}u : (e_3 \otimes e_3) + \left(\delta_0^2 \mathcal{W} + g\right)\eta - R^3 \cdot e_3 \qquad \text{on } \mathbb{T}^2 \, (\sim \Sigma),$$

it follows that

$$\begin{aligned} \|\operatorname{tr}_\Sigma p\|_{L^2(\mathbb{T}^2)} &\lesssim \|\operatorname{tr}_\Sigma \mathbb{D}u\|_{L^2(\mathbb{T}^2)} + \|\eta\|_{H^4(\mathbb{T}^2)} + \left\|R^3\right\|_{L^2(\mathbb{T}^2)} \\ &\lesssim \|u\|_{H^{3/2}(\Omega)} + \|\eta\|_{H^4(\mathbb{T}^2)} + \left\|R^3\right\|_{L^2(\mathbb{T}^2)}. \end{aligned}$$

We then use a Poincare-type inequality (i.e. auxiliary estimate 8.1) to bound

$$\|p\|_{H^2(\Omega)} \lesssim \|\operatorname{tr}_\Sigma p\|_{L^2(\mathbb{T}^2)} + \|\nabla p\|_{H^1(\Omega)}.$$

(6) Finally, we use the kinematic boundary condition and its time-differentiated version to obtain control of $\partial_t \eta$ and $\partial_t^2 \eta$. Indeed, the kinematic boundary condition tells us that

$$\partial_t \eta = u \cdot e_3 + R^4 \qquad \text{on } \mathbb{T}^2 \,(\sim \Sigma)\,,$$

and therefore

$$||\partial_t \eta||_{H^{5/2}(\mathbb{T}^2)} \lesssim ||\mathrm{tr}_\Sigma\, u||_{H^{5/2}(\mathbb{T}^2)} + \left|\left|R^4\right|\right|_{H^{5/2}(\mathbb{T}^2)} \lesssim ||u||_{H^3(\Omega)} + \left|\left|R^4\right|\right|_{H^{5/2}(\mathbb{T}^2)}.$$

The time-differentiated kinematic boundary condition tells us that

$$\partial_t^2 \eta = (\partial_t u) \cdot e_3 + \partial_t R^4 \qquad \text{on } \mathbb{T}^2 \,(\sim \Sigma)$$

and therefore

$$\left|\left|\partial_t^2 \eta\right|\right|_{H^{1/2}(\mathbb{T}^2)} \lesssim ||\partial_t u||_{H^{1/2}(\Omega)} + \left|\left|\partial_t R^4\right|\right|_{H^{1/2}(\mathbb{T}^2)}.$$

Combining these estimates then shows that

$$\begin{aligned} &||u||_{H^3(\Omega)} + ||\partial_t u||_{H^1(\Omega)} + ||p||_{H^2(\Omega)} + ||\eta||_{H^{11/2}(\mathbb{T}^2)} + ||\partial_t \eta||_{H^{5/2}(\mathbb{T}^2)} + \left|\left|\partial_t^2 \eta\right|\right|_{H^{1/2}(\mathbb{T}^2)} \\ &\quad \lesssim ||\mathbb{D} u||_{L^2(\Omega)} + ||\mathbb{D}\partial_t u||_{L^2(\Omega)} + \left|\left|\mathbb{D}\overline{\nabla}^2 u\right|\right|_{L^2(\Omega)} \\ &\quad + \left|\left|R^1\right|\right|_{H^1(\Omega)} + \left|\left|R^2\right|\right|_{H^2(\Omega)} + \underbrace{\left|\left|R^3\right|\right|_{L^2(\mathbb{T}^2)} + \left|\left|\nabla R^3\right|\right|_{H^{1/2}(T^2)}}_{\lesssim ||R^3||_{H^{3/2}(\mathbb{T}^2)}} + \left|\left|R^4\right|\right|_{H^{5/2}(\mathbb{T}^2)} + \left|\left|\partial_t R^4\right|\right|_{H^{1/2}(\mathbb{T}^2)}, \end{aligned}$$

and then (5.4) follows immediately.

□

Proposition 5.5 tells us in which norm we need to be able to control the nonlinear remainders. In the notation used in the sketch in Section 3.2, these remainders R are given by $R = (L - N)(\mathcal{X})$. Here N corresponds to the system (2.2a)–(2.2e), while L corresponds to the system

$$\begin{cases} \partial_t u + \nabla \cdot S = 0 & \text{in } \Omega, \\ \nabla \cdot u = 0 & \text{in } \Omega, \\ \left(\delta_0^2 \mathcal{W} + g\right) \eta e_3 - S \cdot e_3 = 0 & \text{on } \Sigma, \\ \partial_t \eta - u \cdot e_3 = 0 & \text{on } \Sigma, \text{ and} \\ u = 0 & \text{on } \Sigma_b. \end{cases}$$

It follows that the remainders are given by

$$\begin{cases} R^1 = \left(D_t^{u,\mathcal{G}} u - \partial_t u\right) + \left(\nabla^{\mathcal{G}} \cdot S^{\mathcal{G}} - \nabla \cdot S\right), & (5.6a) \\ R^2 = \nabla^{\mathcal{G}} \cdot u - \nabla \cdot u, & (5.6b) \\ R^3 = \left(\delta \mathcal{W}(\eta)\, \nu_{\partial\Omega}^{\mathcal{G}} - \left(\delta_0^2 \mathcal{W}\right) \eta e_3\right) + g\eta \left(\nu_{\partial\Omega}^{\mathcal{G}} - e_3\right) - \left(S^{\mathcal{G}} \cdot \nu_{\partial\Omega}^{\mathcal{G}} - S \cdot e_3\right), \text{ and} & (5.6c) \\ R^4 = u \cdot \left(\nu_{\partial\Omega}^{\mathcal{G}} - e_3\right). & (5.6d) \end{cases}$$

Before recording our estimates for these terms we discuss how to Taylor expand the surface energy terms.

Remark 5.6. An important subtetly in performing the estimates in this section arises from the fact that the surface energy density may be fully nonlinear. This plays a role in two terms in particular: $\delta \mathcal{W}(\eta)$ and $\left(\delta \mathcal{W} - \delta_0^2 \mathcal{W}\right)(\eta)$. We write these terms in a manner more amenable to estimates by performing a Taylor expansion of ∇f, i.e.

- For $\delta \mathcal{W}$:

$$\begin{aligned} \delta \mathcal{W}(\eta) &= \mathcal{J}^* \left(\nabla f (\mathcal{J}\eta)\right) = \mathcal{J}^* \left(\nabla f(\mathcal{J}\eta) - \nabla f(0)\right) \\ &= \mathcal{J}^* \left(\int_0^1 \nabla^2 f(t\mathcal{J}\eta)\, dt \bullet \mathcal{J}\eta\right) = \mathcal{J}^* \left(\mathfrak{h}(\mathcal{J}\eta) \bullet \mathcal{J}\eta\right), \end{aligned}$$

where

$$\mathfrak{h}(z) := \int_0^1 \nabla^2 f(tz)\, dt$$

for $z = (w, M) \in \mathbb{R}^n \times \mathbb{R}^{n\times n}$. Note that we may also write

$$\delta\mathcal{W}(\eta) = \mathcal{J}^* \left(\mathcal{R}_0 [\nabla f, 0] (\mathcal{J}\eta)\right),$$

where $\mathcal{R}_0$ is defined in Proposition 8.24. This is a useful way of writing $\delta\mathcal{W}(\eta)$ since it provides us with a unified way of estimating a certain number of terms showing up in the remainders.

- For $\left(\delta\mathcal{W} - \delta_0^2\mathcal{W}\right)(\eta)$:

$$\begin{aligned}\delta\mathcal{W}(\eta) - \delta_0^2\mathcal{W}(\eta) &= \mathcal{J}^*\Big(\nabla f(\mathcal{J}\eta) - \nabla^2 f(0) \bullet \mathcal{J}\eta\Big)\\ &= \mathcal{J}^*\Big(\nabla f(\mathcal{J}\eta) - \mathcal{P}_1 [\nabla f, 0](\mathcal{J}\eta)\Big)\\ &= \mathcal{J}^*\Big(\mathcal{R}_1 [\nabla f, 0](\mathcal{J}\eta)\Big)\\ &= \mathcal{J}^*\left(\left(\frac{1}{2}\int_0^1 (1-t)\nabla^3 f(t\mathcal{J}\eta)\,dt\right) \bullet (\mathcal{J}\eta \otimes \mathcal{J}\eta)\right)\\ &= \mathcal{J}^*\Big(\mathfrak{q}(\mathcal{J}\eta) \bullet (\mathcal{J}\eta \otimes \mathcal{J}\eta),\Big)\end{aligned}$$

where

$$\mathfrak{q}(z) := \frac{1}{2}\int_0^1 (1-t)\nabla^3 f(tz)\,dt$$

for $z = (w, M) \in \mathbb{R}^n \times \mathbb{R}^{n\times n}$ and where $\mathcal{P}_1$ and $\mathcal{R}_1$ are defined in Proposition 8.24.

Summarizing, we have:

$$\begin{cases}\delta\mathcal{W}(\eta) = \mathcal{J}^*\left(\mathcal{R}_0[\nabla f, 0](\mathcal{J}\eta)\right) = \mathcal{J}^*\left(\mathfrak{h}(\mathcal{J}\eta) \bullet \mathcal{J}\eta\right)\\ \left(\delta\mathcal{W} - \delta_0^2\mathcal{W}\right)(\eta) = \mathcal{J}^*\left(\mathcal{R}_1[\nabla f, 0](\mathcal{J}\eta)\right) = \mathcal{J}^*\left(\mathfrak{q}(\mathcal{J}\eta) \bullet (\mathcal{J}\eta \otimes \mathcal{J}\eta)\right)\end{cases} \tag{5.7}$$

where $\mathcal{R}_0$ and $\mathcal{R}_1$ are defined in Proposition 8.24 and where

$$\mathfrak{h}(z) := r_0[\nabla f, 0](\mathcal{J}\eta) = \int_0^1 \nabla^2 f(tz)\,dt \quad \text{and} \quad \mathfrak{q}(z) := r_1[\nabla f, 0](\mathcal{J}\eta) = \frac{1}{2}\int_0^1 (1-t)\nabla^3 f(tz)\,dt$$

for $z = (w, M) \in \mathbb{R}^n \times \mathbb{R}^{n\times n}$ and for r_0 and r_1 defined in Proposition 8.24.

Our next result records estimates for the remainder terms.

Lemma 5.7. *Let $\mathcal{N}_E$ and $\mathcal{N}_D$ be as defined in Proposition 5.5, and R^1, R^2, R^3, R^4 be as defined by (5.6a)–(5.6d). If the small energy assumptions hold (see Definition 4.2), then*

$$\mathcal{N}_E \lesssim \mathcal{E}^2 \text{ and } \mathcal{N}_D \lesssim \mathcal{E}\mathcal{D}.$$

Proof. First let us sketch the argument. As in the proof of Proposition 5.5, most terms are easily handled via the standard combination of Hölder and Sobolev inequalities (c.f. Proposition 8.11) since they are multilinear, but some terms arising from the fully nonlinear surface energy have to be handled differently. Essentially, to control those, we make use of the fact that we are in a small energy regime and use Taylor expansions (c.f. Proposition 8.24 for the notation used) to bring it back to the multilinear (i.e. polynomial) case. More precisely, the troublesome terms are $\delta\mathcal{W} - \delta_0^2\mathcal{W}$ and $\delta\mathcal{W}$, which we handle by employing (5.7).

Let us now estimate each remainder in detail. R^2 and R^4 are easy to deal with since

$$R^2 = (\mathcal{G} - I) : \nabla u \text{ and } R^4 = -(\operatorname{tr} u) \cdot \nabla\eta$$

and therefore we can use standard product estimates in Sobolev spaces (c.f. Propositions 8.12, 8.13, and 8.14).

R^1 is similar and only requires expanding out further before being estimated in the same way as R^2 and R^4 above:

$$\begin{aligned}R^1 = &-(\partial_t\Phi) \cdot \mathcal{G} \cdot (\nabla u)^T + u \cdot \mathcal{G} \cdot (\nabla u)^T + (\mathcal{G} - I) \cdot \nabla p\\ &- \nabla\left(\operatorname{Sym}\left((\nabla u) \cdot (\mathcal{G} - I)^T\right)\right) : (\mathcal{G} - I)\\ &- (\nabla\mathbb{D}u) : (\mathcal{G} - I) - \nabla \cdot \left(\operatorname{Sym}\left((\nabla u) \cdot (\mathcal{G} - I)^T\right)\right).\end{aligned}$$

R^3 requires more care, since it can be expanded out to be

$$\begin{aligned} R^3 = &- (\delta\mathcal{W}(\eta) + g\eta)(\nabla\eta) + \left(\delta\mathcal{W}(\eta) - \delta_0^2\mathcal{W}(\eta)\right) e_3 + p\nabla\eta \\ &- \operatorname{Sym}\left((\nabla u)\,\mathcal{G}^T\right)\cdot(\nabla\eta) + \operatorname{Sym}\left((\nabla u)(\mathcal{G}-I)\right)\cdot e_3, \end{aligned}$$

where we have used that R^3 is defined on Σ and $\nu^{\mathcal{G}}_{\partial\Omega}|_\Sigma = -\widetilde{\nabla}\eta + e_3$. Most terms in the expansion of R^3 can be handled by standard product estimates, but as sketched above, two terms require particular care, namely the ones involving $\delta\mathcal{W}(\eta)$ and $\delta\mathcal{W}(\eta) - \delta_0^2\mathcal{W}(\eta)$.

According to (5.7), the key estimates required to control $\delta\mathcal{W}$ and $\delta\mathcal{W} - \delta_0^2\mathcal{W}$ in H^s are therefore the control of $\mathfrak{h}(\mathcal{J}\eta)$ and $\mathfrak{q}(\mathcal{J}\eta)$ in H^s. The details of this estimate rely on post-composition estimates in Sobolev spaces, and are recorded in the appendix in Lemma 7.12 and Corollary 7.13. From this we obtain that for any $s \geqslant 2$,

$$\begin{cases} ||\mathfrak{h}(\mathcal{J}\eta)||_{H^s(\mathbb{T}^2)} \lesssim C_f^{(2)} + C_f^{(\lceil s\rceil+2)}\left(||\eta||_{H^{s+2}(\mathbb{T}^2)} + ||\eta||^{\lceil s\rceil}_{H^{s+2}(\mathbb{T}^2)}\right) \text{ and} \\ ||\mathfrak{q}(\mathcal{J}\eta)||_{H^s(\mathbb{T}^2)} \lesssim C_f^{(3)} + C_f^{(\lceil s\rceil+3)}\left(||\eta||_{H^{s+2}(\mathbb{T}^2)} + ||\eta||^{\lceil s\rceil}_{H^{s+2}(\mathbb{T}^2)}\right), \end{cases}$$

where here we recall that the constants $C_f^{(k)}$ are defined in Definition 4.1.

We may now proceed with the estimates. Since, as detailed above, most terms in the remainder are easy to control, we only highlight those which are more delicate and representative of the difficulties encountered. More precisely, we estimate in detail:

(1) the term involving $\delta\mathcal{W}(\eta) - \delta_0^2\mathcal{W}(\eta)$ in $\mathcal{N}_E$,
(2) the term involving $\delta\mathcal{W}(\eta) + g\eta$ in $\mathcal{N}_D$, and
(3) the term involving $\delta\mathcal{W}(\eta) - \delta_0^2\mathcal{W}(\eta)$ in $\mathcal{N}_D$.

These estimates are obtained as follows.

(1) We seek to control $\left|\left|\left(\delta\mathcal{W}(\eta) - \delta_0^2\mathcal{W}(\eta)\right)e_3\right|\right|_{H^{1/2}(\mathbb{T}^2)}$:

$$\begin{aligned} \left|\left|\left(\delta\mathcal{W}(\eta) - \delta_0^2\mathcal{W}(\eta)\right)e_3\right|\right|_{H^{1/2}(\mathbb{T}^2)} &= \left|\left|\mathcal{J}^*\Big(\mathfrak{q}(\mathcal{J}\eta)\bullet(\mathcal{J}\eta\otimes\mathcal{J}\eta)\Big)\right|\right|_{H^{1/2}(\mathbb{T}^2)} \lesssim ||\mathfrak{q}(\mathcal{J}\eta)\bullet(\mathcal{J}\eta\otimes\mathcal{J}\eta)||_{H^{5/2}(\mathbb{T}^2)} \\ &\lesssim ||\mathfrak{q}(\mathcal{J}\eta)||_{H^{5/2}(\mathbb{T}^2)}||\mathcal{J}\eta||^2_{H^{5/2}(\mathbb{T}^2)} \lesssim 1\cdot\sqrt{\mathcal{E}}\sqrt{\mathcal{E}} = \mathcal{E}. \end{aligned}$$

(2) We seek to control $||(\delta\mathcal{W}(\eta) + g\eta)\nabla\eta||_{H^{3/2}(\mathbb{T}^2)}$ and thus the key term to control is $||\delta\mathcal{W}(\eta)||_{H^{3/2}(\mathbb{T}^2)}$. Since $\delta\mathcal{W}(\eta)$ is a differential operator of order 4, and since $\mathcal{D} \gtrsim ||\eta||_{H^{11/2}(\mathbb{T}^2)} \gtrsim \left|\left|\nabla^4\eta\right|\right|_{H^{3/2}(\mathbb{T}^2)}$, we cannot get away with writing $\delta\mathcal{W}(\eta) = \mathcal{J}^*(\nabla f(\mathcal{J}\eta))$ and estimating $||\nabla f(\mathcal{J}\eta)||_{H^{7/2}(\mathbb{T}^2)}$. Instead, we use Lemma 7.8 to obtain

$$\begin{aligned} ||\delta\mathcal{W}(\eta)||_{H^{3/2}(\mathbb{T}^2)} \leqslant\ & \left|\left|\nabla^2_{M,M}f(\mathcal{J}\eta)\bullet\nabla^4\eta\right|\right|_{H^{3/2}(\mathbb{T}^2)} + \left|\left|\nabla^2_{w,w}f(\mathcal{J}\eta)\bullet\nabla^2\eta\right|\right|_{H^{3/2}(\mathbb{T}^2)} \\ &+ \left|\left|\nabla^3_{M,M,M}f(\mathcal{J}\eta)\bullet\left(\nabla^3\eta\otimes\nabla^3\eta\right)\right|\right|_{H^{3/2}(\mathbb{T}^2)} + \left|\left|\nabla^3_{M,M,w}f(\mathcal{J}\eta)\bullet\left(\nabla^3\eta\otimes\nabla^2\eta\right)\right|\right|_{H^{3/2}(\mathbb{T}^2)} \\ &+ \left|\left|\nabla^3_{w,M,w}f(\mathcal{J}\eta)\bullet\left(\nabla^2\eta\otimes\nabla^2\eta\right)\right|\right|_{H^{3/2}(\mathbb{T}^2)} \\ \lesssim\ & \left|\left|\nabla^2 f(\mathcal{J}\eta)\right|\right|_{H^{3/2}(\mathbb{T}^2)}\left|\left|\nabla^4\eta\right|\right|_{H^{3/2}(\mathbb{T}^2)} + \left|\left|\nabla^2 f(\mathcal{J}\eta)\right|\right|_{H^{3/2}(\mathbb{T}^2)}\left|\left|\nabla^2\eta\right|\right|_{H^{3/2}(\mathbb{T}^2)} \\ &+ \left|\left|\nabla^3 f(\mathcal{J}\eta)\right|\right|_{H^{3/2}(\mathbb{T}^2)}\left|\left|\nabla^3\eta\right|\right|^2_{H^{3/2}(\mathbb{T}^2)} + \left|\left|\nabla^3 f(\mathcal{J}\eta)\right|\right|_{H^{3/2}(\mathbb{T}^2)}\left|\left|\nabla^3\eta\right|\right|_{H^{3/2}(\mathbb{T}^2)}\left|\left|\nabla^2\eta\right|\right|_{H^{3/2}(\mathbb{T}^2)} \\ &+ \left|\left|\nabla^3 f(\mathcal{J}\eta)\right|\right|_{H^{3/2}(\mathbb{T}^2)}\left|\left|\nabla^2\eta\right|\right|^2_{H^{3/2}(\mathbb{T}^2)} \\ \lesssim\ & ||\eta||_{H^{11/2}(\mathbb{T}^2)} + ||\eta||_{H^{7/2}(\mathbb{T}^2)} + ||\eta||^2_{H^{9/2}(\mathbb{T}^2)} + ||\eta||_{H^{9/2}(\mathbb{T}^2)}||\eta||_{H^{7/2}(\mathbb{T}^2)} + ||\eta||^2_{H^{7/2}(\mathbb{T}^2)} \\ \lesssim\ & \sqrt{\mathcal{D}} + \sqrt{\mathcal{E}} + 3\mathcal{E} \quad \lesssim \sqrt{\mathcal{D}}, \end{aligned}$$

where we have used that for $k = 2, 3$,

$$\begin{aligned} \left|\left|\nabla^k f(\mathcal{J}\eta)\right|\right|_{H^{3/2}(\mathbb{T}^2)} \leqslant \left|\left|\nabla^k f(\mathcal{J}\eta)\right|\right|_{H^2(\mathbb{T}^2)} &\lesssim C_f^{(k)} + C_f^{(k+3)}\left(||\eta||_{H^4(\mathbb{T}^2)} + ||\eta||^2_{H^4(\mathbb{T}^2)}\right) \\ &\lesssim 1 + \sqrt{\mathcal{E}} + \mathcal{E} \quad \lesssim 1 \end{aligned}$$

for the constants $C_f^{(k)}$ as defined in Definition 4.1. So finally:

$$||(\delta\mathcal{W}(\eta)+g\eta)\nabla\eta||_{H^{3/2}(\mathbb{T}^2)} \lesssim ||\delta\mathcal{W}(\eta)+g\eta||_{H^{3/2}(\mathbb{T}^2)}||\nabla\eta||_{H^{3/2}(\mathbb{T}^2)} \lesssim \left(\sqrt{\mathcal{D}}+||\eta||^2_{H^{3/2}(\mathbb{T})}\right)||\eta||_{H^{5/2}(\mathbb{T}^2)}$$
$$\lesssim \left(\sqrt{\mathcal{D}}+\sqrt{\mathcal{E}}\right)\sqrt{\mathcal{E}} \quad \lesssim \sqrt{\mathcal{E}}\sqrt{\mathcal{D}}.$$

(3) We seek to control $\left|\left|\left(\delta\mathcal{W}(\eta)-\left(\delta_0^2\mathcal{W}\right)\eta\right)\right|\right|_{H^{3/2}(\mathbb{T}^2)}$. Observe that (using Lemma 7.8 again)

$$\begin{aligned}\delta\mathcal{W}(\eta)-\left(\delta_0^2\mathcal{W}\right)\eta &= \left(\nabla^2_{M,M}f(\mathcal{J}\eta)-\nabla^2_{M,M}f(0)\right)\bullet\nabla^4\eta-\left(\nabla^2_{w,w}f(\mathcal{J}\eta)-\nabla^2_{w,w}f(0)\right)\bullet\nabla^2\eta\\ &+\nabla^3_{M,M,M}f(\mathcal{J}\eta)\bullet\left(\nabla^3\eta\otimes\nabla^3\eta\right)+\nabla^3_{M,M,w}f(\mathcal{J}\eta)\bullet\left(\nabla^3\eta\otimes\nabla^2\eta\right)\\ &+\nabla^3_{w,M,w}f(\mathcal{J}\eta)\bullet\left(\nabla^2\eta\otimes\nabla^2\eta\right).\end{aligned}$$

In particular, for

$$F(z):=\int_0^1\nabla\nabla^2_{M,M}f(tz)\,dt \text{ and } G(z):=\int_0^1\nabla\nabla^2_{p,p}f(tz)\,dt,$$

where $z=(w,M)\in\mathbb{R}^n\times\mathbb{R}^{n\times n}$, we have (by the Fundamental Theorem of Calculus)

$$\begin{aligned}\delta\mathcal{W}(\eta)-\left(\delta_0^2\mathcal{W}\right)\eta &= (F(\mathcal{J}\eta)\bullet\mathcal{J}\eta)\bullet\nabla^4\eta+(G(\mathcal{J}\eta)\bullet\mathcal{J}\eta)\bullet\nabla^2\eta\\ &+\nabla^3_{M,M,M}f(\mathcal{J}\eta)\bullet\left(\nabla^3\eta\otimes\nabla^3\eta\right)+\nabla^3_{M,M,w}f(\mathcal{J}\eta)\bullet\left(\nabla^3\eta\otimes\nabla^2\eta\right)\\ &+\nabla^3_{w,M,w}f(\mathcal{J}\eta)\bullet\left(\nabla^2\eta\otimes\nabla^2\eta\right).\end{aligned}$$

Crucially, all terms have a part which is *quadratic* in η. By an argument similar to that of Lemma 7.12 we have, in the small energy regime, the estimates

$$||F(\mathcal{J}\eta)||_{H^s(\mathbb{T}^2)}\lesssim 1 \text{ and } ||G(\mathcal{J}\eta)||_{H^s(\mathbb{T}^2)}\lesssim 1$$

for any $s\in\left[2,\frac{5}{2}\right]$. So finally, we can perform the estimate:

$$\begin{aligned}\left|\left|\left(\delta\mathcal{W}(\eta)-\left(\delta_0^2\mathcal{W}\right)\eta\right)\right|\right|_{H^{3/2}(\mathbb{T}^2)} &\lesssim ||F(\mathcal{J}\eta)||_{H^{3/2}(\mathbb{T}^2)}||\mathcal{J}\eta||_{H^{3/2}(\mathbb{T}^2)}\left|\left|\nabla^4\eta\right|\right|_{H^{3/2}(\mathbb{T}^2)}\\ &+||G(\mathcal{J}\eta)||_{H^{3/2}(\mathbb{T}^2)}||\mathcal{J}\eta||_{H^{3/2}(\mathbb{T}^2)}\left|\left|\nabla^2\eta\right|\right|_{H^{3/2}(\mathbb{T}^2)}+\text{l.o.t.}\\ &\lesssim ||F(\mathcal{J}\eta)||_{H^2(\mathbb{T}^2)}||\eta||_{H^{7/2}(\mathbb{T}^2)}||\eta||_{H^{11/2}(\mathbb{T}^2)}+||G(\mathcal{J}\eta)||_{H^2(\mathbb{T}^2)}||\eta||^2_{H^{7/2}(\mathbb{T}^2)}+\text{l.o.t.}\\ &\lesssim \sqrt{\mathcal{E}}\sqrt{\mathcal{D}}+\mathcal{E}+\text{l.o.t.} \quad \lesssim \sqrt{\mathcal{E}}\sqrt{\mathcal{D}}\end{aligned}$$

where we have omitted the details for the lower order terms involving $\nabla^3 f$ (denoted l.o.t. above) since they follow exacty as in the second item above.

□

5.3. Geometric corrections. In this section we compute the geometric corrections to the energy and dissipation (i.e. the difference between their geometric and equilibrium versions) in Remark 5.8, and we estimate these corrections in Lemma 5.9.

REMARK 5.8. The geometric corrections are

$$\mathcal{G}_E(\mathcal{X})=\widetilde{\mathcal{E}}(\mathcal{X};\mathcal{X})-\overline{\mathcal{E}}(\mathcal{X}) \text{ and } \mathcal{G}_D(\mathcal{X})=\widetilde{\mathcal{D}}(\mathcal{X};\mathcal{X})-\overline{\mathcal{D}}(\mathcal{X})$$

(c.f. equations 4.3a, 4.3b, 4.5a, and 4.5b for the definitions of the geometric and equilibrium versions of the energy and dissipation). For $\mathcal{X}=(u,p,\eta)$ we can compute the geometric corrections to be

$$\mathcal{G}_E(\mathcal{X})=\sum_{|\alpha|_{t,\overline{x}^2}\leqslant 2}\left(\frac{1}{2}\int_\Omega|\partial^\alpha u|^2(J-1)+\frac{1}{2}\int_{\mathbb{T}^2}\left(\int_0^1 g_\alpha(t)\nabla^3 f(t\mathcal{J}\eta)\,\mathrm{d}t\right)\bullet\left(\mathcal{J}\eta\otimes\mathcal{J}\partial^\alpha\eta\otimes\mathcal{J}\partial^\alpha\eta\right)\right)$$

and

$$\mathcal{G}_D(\mathcal{X})=\sum_{|\alpha|_{t,\overline{x}^2}\leqslant 2}\left(\frac{1}{2}\int_\Omega|\mathbb{D}^{\mathcal{G}-I}\partial^\alpha u|^2 J-\int_\Omega\left(\mathbb{D}^{\mathcal{G}-I}\partial^\alpha u:\mathbb{D}\partial^\alpha u\right)J+\frac{1}{2}\int_\Omega|\mathbb{D}\partial^\alpha u|^2(J-1)\right)$$

(see Section 7.3 for the details of the computation of $\mathcal{G}_E$ and $\mathcal{G}_D$ and the definition of g_α). All we need to know about g_α in order to estimate the geometric corrections is that $|g_\alpha|\leqslant 1$ on $[0,1]$.

Note that $\nabla^3 f$ appears in the geometric corrections to the energy. This is as expected since $\mathcal{G}_E(\mathcal{X}) = \widetilde{\mathcal{E}}(\mathcal{X};\mathcal{X}) - \overline{\mathcal{E}}\mathcal{X} \sim \widetilde{\mathcal{E}}(\mathcal{X};\mathcal{X}) - \widetilde{\mathcal{E}}(\mathcal{X},0)$ where $\widetilde{\mathcal{E}}$ depends on $\nabla^2 f$. Therefore, upon Taylor expanding about the equilibrium solution $\mathcal{X}=0$ we pick up a term involving $\nabla^3 f$.

We now estimate the geometric corrections.

LEMMA 5.9. *In the small energy regime (see Definition 4.2) we have the estimates*

$$|\mathcal{G}_E| \lesssim \sqrt{\mathcal{E}}\mathcal{E} \text{ and } |\mathcal{G}_D| \lesssim \sqrt{\mathcal{E}}\mathcal{D},$$

where $\mathcal{G}_E$ and $\mathcal{G}_D$ are defined in Remark 5.8.

PROOF. First we check that the term involving g_α is small. Observe that since $|g_\alpha(t)| \le 1$ when $t \in [0,1]$, it follows that

$$\sup_{|z|\leqslant R}\left|\int_0^1 g_\alpha(t)\,\nabla^3 f(tz)\,\mathrm{d}t\right| \leqslant \left|\left|\nabla^3 f\right|\right|_{L^\infty\left(\overline{B(0,R)}\right)}$$

and hence

$$\left|\left|\int_0^1 g_\alpha(t)\,\nabla^3 f(t\mathcal{J}\eta)\,\mathrm{d}t\right|\right|_\infty \le \left|\left|\nabla^3 f\right|\right|_{L^\infty\left(\overline{B\left(0,||\mathcal{J}\eta||_\infty\right)}\right)} \leqslant C_f^{(3)},$$

where the constant $C_f^{(3)}$ is defined in Definition 4.1. In particular, in a small energy regime, $||\mathcal{J}\eta||_\infty \lesssim \sqrt{\mathcal{E}}$, and hence

$$\left|\left|\int_0^1 g_\alpha(t)\,\nabla^3 f(t\mathcal{J}\eta)\,\mathrm{d}t\right|\right|_\infty \lesssim C_f^{(3)} \lesssim 1.$$

Note that due to the fashion in which we perform the estimates, it is sufficient to handle the case $\partial^\alpha = \partial_t, \overline{\nabla}^2$. Recall that the control we have over the geometric coefficients $\mathcal{G}$ and J is recorded in Lemma 7.1.

We now estimate the corrections to the energy.

$\boxed{\partial_t}$ The geometric correction is

$$\frac{1}{2}\int_\Omega |\partial_t u|^2 (J-1) + \frac{1}{2}\int_{\mathbb{T}^2}\left(\int_0^1 \nabla^3 f(t\mathcal{J}\eta)\,\mathrm{d}t\right) \bullet (\mathcal{J}\eta \otimes \mathcal{J}\partial_t\eta \otimes \mathcal{J}\partial_t\eta),$$

and it can be estimated in the following way:

$$\begin{aligned}
|\dots| &\lesssim ||\partial_t u||^2_{L^2(\Omega)}||J-1||_{L^\infty(\Omega)} + \left|\left|\int_0^1 g_\alpha(t)\,\nabla^3 f(t\mathcal{J}\eta)\,\mathrm{d}t\right|\right|_{L^\infty(\mathbb{T}^2)} ||\mathcal{J}\eta||_{L^\infty(\mathbb{T}^2)}||\mathcal{J}\partial_t\eta||^2_{L^2(\mathbb{T}^2)} \\
&\lesssim ||\partial_t u||^2_{L^2(\Omega)}||J-1||_{H^{3/2+}(\Omega)} + ||\mathcal{J}\eta||_{H^{1+}(\mathbb{T}^2)}||\mathcal{J}\partial_t\eta||^2_{L^2(\mathbb{T}^2)} \\
&\lesssim \mathcal{E}\sqrt{\mathcal{E}} + \sqrt{\mathcal{E}}\mathcal{E} \quad \lesssim \mathcal{E}^{3/2}.
\end{aligned}$$

$\boxed{\overline{\nabla}^2}$ Note that the control of η in the energy is similar to parabolic scaling, but with a little bit more spatial regularity. Consequently we handle this term as we did the previous one involving ∂_t and obtain (omitting the details)

$$\left|\frac{1}{2}\int_\Omega |\nabla^2 u|^2 (J-1) + \frac{1}{2}\int_{\mathbb{T}^2}\left(\int_0^1 \nabla^3 f(t\mathcal{J}\eta)\,\mathrm{d}t\right) \bullet \left(\mathcal{J}\eta \otimes \mathcal{J}\nabla^2\eta \otimes \mathcal{J}\nabla^2\eta\right)\right| \lesssim \mathcal{E}^{3/2}.$$

Next we estimate the dissipative corrections. Note that $|\mathbb{D}^M v| = |2\,\mathrm{Sym}\left(\nabla^M v\right)| \lesssim |M||\nabla v|$.

$\boxed{\partial_t}$ The geometric correction is

$$\frac{1}{2}\int_\Omega |\mathbb{D}^{\mathcal{G}-I}\partial_t u|^2 J - \int_\Omega \left(\mathbb{D}^{\mathcal{G}-I}\partial_t u : \mathbb{D}\partial_t u\right) J + \frac{1}{2}\int_\Omega |\mathbb{D}\partial_t u|^2 (J-1),$$

and it can be estimated in the following way:

$$\begin{aligned}|\ldots| &\lesssim ||\mathcal{G}-I||^2_{L^\infty(\Omega)}||\nabla\partial_t u||^2_{L^2(\Omega)}||J||_{L^\infty(\Omega)} + ||\mathcal{G}-I||_{L^\infty(\Omega)}||\nabla\partial_t u||^2_{L^2(\Omega)}||J||_{L^\infty(\Omega)} \\ &\quad + ||\nabla\partial_t u||^2_{L^2(\Omega)}||J-1||_{L^\infty(\Omega)} \\ &\lesssim ||\mathcal{G}-I||^2_{H^{3/2+}(\Omega)}||\nabla\partial_t u||^2_{L^2(\Omega)}||J||_{H^{3/2+}(\Omega)} + ||\mathcal{G}-I||_{H^{3/2+}(\Omega)}||\nabla\partial_t u||^2_{L^2(\Omega)}||J||_{H^{3/2+}(\Omega)} \\ &\quad + ||\nabla\partial_t u||^2_{L^2(\Omega)}||J-1||_{H^{3/2+}(\Omega)} \\ &\lesssim \mathcal{E}\mathcal{D}\left(1+\sqrt{\mathcal{E}}\right) + \sqrt{\mathcal{E}}\mathcal{D}\left(1+\sqrt{\mathcal{E}}\right) + \mathcal{D}\sqrt{\mathcal{E}} \quad \lesssim \sqrt{\mathcal{E}}\mathcal{D}.\end{aligned}$$

$\boxed{\overline{\nabla}^2}$ Since the control we have on u follows parabolic scaling precisely, upon replacing $\partial_t u$ by $\nabla^2 u$ we can proceed in exactly the same way we did above. We therefore obtain that

$$\left|\frac{1}{2}\int_\Omega |\mathbb{D}^{\mathcal{G}-I}\nabla^2 u|^2 J - \int_\Omega \left(\mathbb{D}^{\mathcal{G}-I}\nabla^2 u : \mathbb{D}\nabla^2 u\right) J + \frac{1}{2}\int_\Omega |\mathbb{D}\nabla^2 u|^2 (J-1)\right| \lesssim \sqrt{\mathcal{E}}\mathcal{D}.$$

□

5.4. Synthesis. In this section we piece together the various elements of the a priori estimates into our main ‘a priori’ theorem.

THEOREM 5.10 (A priori estimates). *There exist $\delta, \lambda, C_{ap} > 0$ such that if there exists a solution $\mathcal{X} = (u, p, \eta)$ on $[0, T)$ with initial condition $\mathcal{X}_0 = (u_0, p_0, \eta_0)$ satisfying*

$$\sup_{t\in[0,T)} \mathcal{E}(\mathcal{X}) \leqslant \delta \quad \textit{and} \quad \int_0^T \mathcal{D}(\mathcal{X}) < \infty$$

(and so in particular, for $\delta \leqslant 1$, we are in the small energy regime as defined in 4.2), then

$$\sup_{t\in[0,T)} \mathcal{E}(\mathcal{X}) e^{\lambda t} + \int_0^T \mathcal{D}(\mathcal{X}) e^{\lambda s}\, ds \leqslant C_{ap}\overline{\mathcal{E}}(\mathcal{X}_0).$$

PROOF. In order to define δ, λ, and C_{ap}, we must keep track of the constants in Lemmas 5.4, 5.7, 5.9 and Proposition 5.5. In particular, we take $C_{C,E}, C_{C,D}, C_{N,E}, C_{N,D}, C_{G,E}, C_{G,D}, C_{A,E}, C_{A,D} > 0$ such that

$$\begin{cases} |\mathcal{C}^1| \leqslant C_{C,E}\sqrt{\mathcal{E}}\mathcal{D}, \\ |\mathcal{C}^2| \leqslant C_{C,D}\sqrt{\mathcal{E}}\mathcal{E}, \end{cases} \quad \begin{cases} \mathcal{N}_E \leqslant C_{N,E}\mathcal{E}^2, \\ \mathcal{N}_D \leqslant C_{N,D}\mathcal{E}\mathcal{D}, \end{cases} \quad \begin{cases} |\mathcal{G}_E| \leqslant C_{G,E,}\sqrt{\mathcal{E}}\mathcal{E}, \\ |\mathcal{G}_D| \leqslant C_{G,D}\sqrt{\mathcal{E}}\mathcal{D}, \end{cases} \quad \begin{cases} \mathcal{E} \leqslant C_{A,E}\left(\overline{\mathcal{E}} + \mathcal{N}_E\right), \text{ and} \\ \mathcal{D} \leqslant C_{A,D}\left(\overline{\mathcal{D}} + \mathcal{N}_D\right). \end{cases}$$

Moreover we assume without loss of generality that $C_{C,E}, C_{C,D}, C_{N,E}, C_{N,D}, C_{G,E}, C_{G,D}, C_{A,E}, C_{A,D} \geqslant 1$. Now pick

$$\begin{aligned}\delta = \min\Bigg(&\delta_0, \frac{1}{2C_{A,E}C_{N,E}}, \left(\frac{1}{2C_{G,E}C_{A,E}}\right)^2, \left(\frac{1}{8C_{C,E}C_{A,E}}\right)^2, \\ &\frac{1}{2C_{A,D}C_{N,D}}, \left(\frac{1}{2C_{G,D}C_{A,D}}\right)^2, \left(\frac{1}{8C_{C,D}C_{A,D}}\right)^2\Bigg),\end{aligned}$$

$$\frac{1}{2\lambda} = 8C_{A,D}\left(1 + C_{C,E}\sqrt{\delta_0}\right)\left(1 + 2C_{G,E}C_{A,E}\sqrt{\delta_0}\right), \text{ and}$$

$$C := \max\left(8C_{A,E}, 16C_{A,D}\right) 4C_{A,E}\left(1 + 2C_{A,E}C_{G,E}\sqrt{\delta_0}\right)\left(1 + \left(1 + 2C_{A,E}\right) C_{G,E}\sqrt{\delta_0}\right) > 0.$$

We divide the remainder of the proof into several steps.

Step 1: We show that in the δ-small energy regime, i.e. when $\sup \mathcal{E} \leqslant \delta$ and $\int \mathcal{D} < \infty$, all versions of the energy, and all versions of the dissipation, are equivalent. The key observation is that the difference between various versions of the energy and the dissipation can be controlled (by Lemmas 5.4, 5.7, and 5.9) by quantities of the form $\mathcal{E}^\alpha\mathcal{E}$ and $\mathcal{E}^\alpha\mathcal{D}$ respectively, for some $\alpha > 0$. In particular, by picking δ small and imposing that $\mathcal{E} \leqslant \delta$ we may ensure that $\mathcal{E}^\alpha$ be small enough to perform absorption arguments. More precisely, we show that

$$\mathcal{E} \asymp \overline{\mathcal{E}} \asymp \widetilde{\mathcal{E}} \text{ and } \mathcal{D} \asymp \overline{\mathcal{D}} \asymp \widetilde{\mathcal{D}},$$

and in particular we show that

$$\begin{cases} \overline{\mathcal{E}} \leqslant \mathcal{E}, & (5.8a) \\ \mathcal{E} \leqslant C^E_{imp,eq}\overline{\mathcal{E}}, & (5.8b) \\ \widetilde{\mathcal{E}} \leqslant C^E_{geo,eq}\overline{\mathcal{E}}, & (5.8c) \\ \mathcal{E} \leqslant C^E_{imp,geo}\widetilde{\mathcal{E}}, & (5.8d) \end{cases}$$

and

$$\begin{cases} \overline{\mathcal{D}} \leqslant \mathcal{D}, & (5.9a) \\ \mathcal{D} \leqslant C^D_{imp,eq}\overline{\mathcal{D}}, & (5.9b) \\ \widetilde{\mathcal{D}} \leqslant C^D_{geo,eq}\overline{\mathcal{D}}, & (5.9c) \\ \mathcal{D} \leqslant C^D_{imp,geo}\widetilde{\mathcal{D}}, & (5.9d) \end{cases}$$

where

$$\begin{cases} C^E_{imp,eq} = 2C_{A,E}, \\ C^E_{geo,eq} = 1 + 2C_{G,E}C_{A,E}\sqrt{\delta_0}, \\ C^E_{imp,geo} = 4C_{A,E}, \end{cases} \qquad \begin{cases} C^D_{imp,eq} = 2C_{A,D}, \\ C^D_{geo,eq} = 1 + 2C_{G,D}C_{A,D}\sqrt{\delta_0}, \text{ and} \\ C^D_{imp,geo} = 4C_{A,D}. \end{cases}$$

To start, note that (5.8a) follows immediately from the definition of $\overline{\mathcal{E}}$ and $\mathcal{E}$. To obtain (5.8b), we apply Proposition 5.5, Lemma 5.7, and note that since $\mathcal{E} \leqslant \delta \leqslant \frac{1}{2C_{A,E}C_{N,E}}$ it follows that $\frac{C_{A,E}}{1-C_{A,E}C_{N,E}\mathcal{E}} \leqslant 2C_{A,E} = C^E_{imp,eq}$. Thus

$$\mathcal{E} \leqslant C_{A,E}\left(\overline{\mathcal{E}} + \mathcal{N}_E\right) \leqslant C_{A,E}\left(\overline{\mathcal{E}} + C_{N,E}\mathcal{E}^2\right)$$
$$\Rightarrow \quad \mathcal{E} \leqslant \frac{C_{A,E}}{1 - C_{A,E}C_{N,E}\mathcal{E}}\overline{\mathcal{E}} \leqslant C^E_{imp,eq}\overline{\mathcal{E}}.$$

To obtain (5.8c), we use Remark 5.8, Lemma 5.9, and (5.8b) to see that

$$\widetilde{\mathcal{E}} = \overline{\mathcal{E}} + \mathcal{G}_E \leqslant \overline{\mathcal{E}} + C_{G,E}\sqrt{\mathcal{E}}\mathcal{E} \leqslant \left(1 + 2C_{G,E}C_{A,E}\sqrt{\delta_0}\right)\overline{\mathcal{E}} = C^E_{geo,eq}\overline{\mathcal{E}}.$$

To obtain (5.8d) we apply (5.8b), Remark 5.8, and Lemma 5.9 to see that

$$\mathcal{E} \leqslant C^E_{imp,eq}\overline{\mathcal{E}} = C^E_{imp,eq}\left(\widetilde{\mathcal{E}} - \mathcal{G}_E\right) \leqslant C^E_{imp,eq}\left(\widetilde{\mathcal{E}} + C_{G,E}\sqrt{\mathcal{E}}\mathcal{E}\right)$$
$$\Rightarrow \quad \mathcal{E} \leqslant \frac{C^E_{imp,eq}}{1 - C_{G,E}C^E_{imp,eq}\sqrt{\mathcal{E}}}\widetilde{\mathcal{E}} = \frac{2C_{A,E}}{1 - 2C_{G,E}C_{A,E}\sqrt{\mathcal{E}}}\widetilde{\mathcal{E}} \overset{(\star)}{\leqslant} 4C_{A,E}\widetilde{\mathcal{E}} = C^E_{imp,geo}\widetilde{\mathcal{E}},$$

where $(\star)$ holds since $\mathcal{E} \leqslant \delta \leqslant \left(\frac{1}{2C_{G,E}C_{A,E}}\right)^2$. The bound (5.9a) follows immediately from the definition of $\overline{\mathcal{D}}$ and $\mathcal{D}$. To obtain (5.9b), we apply Proposition 5.5 and Lemma 5.7 to see that

$$\mathcal{D} \leqslant C_{A,D}\left(\overline{\mathcal{D}} + \mathcal{N}_D\right) \leqslant C_{A,D}\left(\overline{\mathcal{D}} + C_{N,D}\mathcal{E}\mathcal{D}\right)$$
$$\Rightarrow \quad \mathcal{D} \leqslant \frac{C_{A,D}}{1 - C_{A,D}C_{N,D}\mathcal{E}}\overline{\mathcal{D}} \overset{(\star)}{\leqslant} 2C_{A,D}\overline{\mathcal{D}} = C^D_{imp,eq}\overline{\mathcal{D}},$$

where $(\star)$ holds since $\mathcal{E} \leqslant \delta \leqslant \frac{1}{2C_{A,D}C_{N,D}}$. To obtain (5.9c), we use Remark 5.8, Lemma 5.9, and (5.9b) to see that

$$\widetilde{\mathcal{D}} = \overline{\mathcal{D}} + \mathcal{G}_D \leqslant \overline{\mathcal{D}} + C_{G,D}\sqrt{\mathcal{E}}\mathcal{D} \leqslant \left(1 + 2C_{G,D}C_{A,D}\sqrt{\delta_0}\right)\overline{\mathcal{D}} = C^D_{geo,eq}\overline{\mathcal{D}}.$$

To obtain (5.9d) we apply (5.9b), Remark 5.8, and Lemma 5.9 to see that

$$\mathcal{D} \leqslant C^D_{imp,eq}\overline{\mathcal{D}} = C^D_{imp,eq}\left(\widetilde{\mathcal{D}} - \mathcal{G}_D\right) \leqslant C^D_{imp,eq}\left(\widetilde{\mathcal{D}} + C_{G,D}\sqrt{\mathcal{E}}\mathcal{D}\right)$$
$$\Rightarrow \quad \mathcal{D} \leqslant \frac{C^D_{imp,eq}}{1 - C_{G,D}C^D_{imp,eq}\sqrt{\mathcal{E}}}\widetilde{\mathcal{D}} = \frac{2C_{A,D}}{1 - 2C_{G,D}C_{A,D}\sqrt{\mathcal{E}}}\widetilde{\mathcal{D}} \overset{(\star)}{\leqslant} 4C_{A,D}\widetilde{\mathcal{D}} = C^D_{imp,geo}\widetilde{\mathcal{D}},$$

where $(\star)$ holds since $\mathcal{E} \leqslant \delta \leqslant \left(\frac{1}{2C_{G,D}C_{A,D}}\right)^2$.

Step 2: We apply the generic energy-dissipation relations computed in Propositions 5.1 and 5.2 to the case where $\mathcal{Y} = \partial^\alpha \mathcal{X}$, and then sum over $|\alpha|_{t,\overline{x}^2} \leqslant 2$ to obtain the energy-dissipation relation:

$$\frac{\mathrm{d}}{\mathrm{d}t}\widetilde{\mathcal{E}} + \widetilde{\mathcal{D}} = \mathcal{C}^1 + \frac{\mathrm{d}}{\mathrm{d}t}\mathcal{C}^2 \quad \Leftrightarrow \quad \frac{\mathrm{d}}{\mathrm{d}t}\left(\widetilde{\mathcal{E}} - \mathcal{C}^2\right) + \left(\widetilde{\mathcal{D}} - \mathcal{C}^1\right) = 0.$$

Step 3: Recall that $\mathcal{D} \geqslant \mathcal{E}$, i.e. the dissipation is coercive over the energy. We now use Steps 1 and 2 with this coercivity, as well as Lemma 5.4, to obtain a Gronwall-type inequality:

$$\begin{aligned}
\widetilde{\mathcal{D}} - \mathcal{C}^1 &\geqslant \frac{1}{4C_{A,D}}\mathcal{D} - \mathcal{C}^1 \geqslant \frac{1}{4C_{A,D}}\mathcal{D} - C_{C,D}\sqrt{\mathcal{E}}\mathcal{D} && \text{by (5.9d) and Lemma 5.4}\\
&= \left(\frac{1}{4C_{A,D}} - C_{C,D}\sqrt{\mathcal{E}}\right)\mathcal{D} \geqslant \frac{1}{8C_{A,D}}\mathcal{D} \geqslant \frac{1}{8C_{A,D}}\mathcal{E} && \text{by } \mathcal{E} \leqslant \delta \leqslant \left(\frac{1}{8C_{A,D}C_{C,D}}\right)^2 \text{ and coercivity}\\
&\geqslant \frac{1}{8C_{A,D}\left(1 + C_{C,E}\sqrt{\delta_0}\right)}\left(\mathcal{E} - \mathcal{C}^2\right) && \text{by Lemma 5.4 and since } \mathcal{E} \leqslant \delta \leqslant \delta_0\\
&\geqslant \frac{1}{8C_{A,D}\left(1 + C_{C,E}\sqrt{\delta_0}\right)}\left(\frac{1}{1 + 2C_{G,E}C_{A,E}\sqrt{\delta_0}}\widetilde{\mathcal{E}} - \mathcal{C}^2\right) && \text{by (5.8c) and (5.8a)}\\
&\geqslant \frac{1}{8C_{A,D}\left(1 + C_{C,E}\sqrt{\delta_0}\right)\left(1 + 2C_{G,E}C_{A,E}\sqrt{\delta_0}\right)}\left(\widetilde{\mathcal{E}} - \mathcal{C}^2\right) = 2\lambda\left(\widetilde{\mathcal{E}} - \mathcal{C}^2\right),
\end{aligned}$$

and therefore

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(\widetilde{\mathcal{E}} - \mathcal{C}^2\right) + \lambda\left(\widetilde{\mathcal{E}} - \mathcal{C}^2\right) + \frac{1}{16C_{A,D}}\mathcal{D} \leqslant 0.$$

Upon integrating in time, we obtain that for all $t \in [0, T)$,

$$\left(\widetilde{\mathcal{E}} - \mathcal{C}^2\right)(\mathcal{X})\, e^{\lambda t} + \int_0^t \frac{1}{16C_{A,D}}\mathcal{D}(\mathcal{X})\, e^{\lambda s} ds \leqslant \left(\widetilde{\mathcal{E}} - \mathcal{C}^2\right)(\mathcal{X}_0).$$

Now observe that using (5.8d), Lemma 5.4, and the fact that $\mathcal{E} \leqslant \delta \leqslant \left(\frac{1}{8C_{A,E}C_{C,E}}\right)^2$, we obtain that

$$\widetilde{\mathcal{E}} - \mathcal{C}^2 \geqslant \frac{1}{4C_{A,E}}\mathcal{E} - C_{C,E}\sqrt{\mathcal{E}}\mathcal{E} = \left(\frac{1}{4C_{A,E}} - C_{C,E}\sqrt{\mathcal{E}}\right)\mathcal{E} \geqslant \frac{1}{8C_{A,E}}\mathcal{E},$$

whilst using (5.8c), (5.8a), and Lemma 5.4, we obtain that

$$\widetilde{\mathcal{E}} - \mathcal{C}^2 \leqslant \left(1 + 2C_{G,E}C_{A,E}\sqrt{\delta_0}\right)\mathcal{E} + C_{C,E}\sqrt{\delta_0}\mathcal{E} = \left(1 + (1 + 2C_{A,E})\, C_{G,E}\sqrt{\delta_0}\right)\mathcal{E}.$$

Therefore, for all $t \in [0, T)$,

$$\begin{aligned}
\frac{1}{8C_{A,E}}\mathcal{E}(\mathcal{X})\, e^{\lambda t} + \int_0^t \frac{1}{16C_{A,D}}\mathcal{D}(\mathcal{X})\, e^{\lambda s} ds &\leqslant \left(1 + (1 + 2C_{A,E})\, C_{G,E}\sqrt{\delta_0}\right)\mathcal{E}(\mathcal{X}_0)\\
&\leqslant 4C_{A,E}\left(1 + 2C_{A,E}C_{G,E}\sqrt{\delta_0}\right)\left(1 + (1 + 2C_{A,E})\, C_{G,E}\sqrt{\delta_0}\right)\overline{\mathcal{E}}(\mathcal{X}_0),
\end{aligned}$$

so indeed we have that

$$\sup_{t\in[0,T)} \mathcal{E}(\mathcal{X})\, e^{\lambda t} + \int_0^T \mathcal{D}(\mathcal{X})\, e^{\lambda s}\, ds \leqslant C_{ap}\overline{\mathcal{E}}(\mathcal{X}_0).$$

□

6. Global well-posedness and decay

In this section we prove the main result of the chapter, namely Theorem 6.11. Before proving this global existence and decay result, we first consider the issue of local well-posedness.

6.1. Local well-posedness. The local existence theory can be rigorously developed by modifying the techniques used to prove the a priori estimates (see for instance [**CCS07**, **CS10**, **GT13c**, **Wu14**, **WTK14**, **Zhe17**, **ZT17**]), so for the sake of brevity we will only sketch what can be obtained in this manner.

In order to discuss the local well-posedness theory, we will need the following notation.

Definition 6.1 (Norm measuring the size of the initial condition). We define the following.

- For $\mathcal{Z} = (u_0, \eta_0)$ we write
$$\mathcal{I}(\mathcal{Z}) := \|\eta_0\|^2_{H^{9/2}(\mathbb{T}^2)} + \|u_0\|^2_{H^2(\Omega)} + \left\|u_0 \cdot \nu^{\mathcal{G}}_{\partial\Omega_0}\right\|^2_{H^2(\Sigma)},$$
where we recall from Section 4.3.3 that, on Σ,
$$\nu_{\partial\Omega_0}|_\Sigma = \frac{(-\overline{\nabla}\eta_0, 1)}{\sqrt{1+|\overline{\nabla}\eta|^2}} \quad \text{and} \quad \nu^{\mathcal{G}}_{\partial\Omega_0}|_\Sigma = \sqrt{1+|\overline{\nabla}\eta|^2}\,\nu_{\partial\Omega_0}|_\Sigma = \left(-\overline{\nabla}\eta_0, 1\right).$$
- For $\mathcal{X} = (u, p, \eta)$ we abuse notations slightly and also write $\mathcal{I}(\mathcal{X}) := \mathcal{I}(u, \eta)$.

It is most natural to specify the initial data u_0 and η_0, but in our analysis we also need $\mathcal{E}(0)$, which means we must construct $\partial_t u|_{t=0}$, $p|_{t=0}$, and $\partial_t \eta|_{t=0}$. We sketch how this construction proceeds in the following remark.

Remark 6.2. In this remark we sketch how to construct p_0, $\partial_t u_0$ and $\partial_t \eta_0$ from u_0 and η_0. Recall that the PDE is (2.2a)–(2.2e).

Constructing p_0: Taking the $\mathcal{G}$-divergence of (2.2a) and using (2.2b) yields
$$-\Delta^{\mathcal{G}} p = \nabla^{\mathcal{G}} u : \left(\nabla^{\mathcal{G}} u\right)^T.$$
Dotting (2.2c) with $\nu^{\mathcal{G}}_{\partial\Omega}$ and dividing by $\left(1+|\overline{\nabla}\eta|^2\right)$ then yields
$$p = \left(\mathbb{D}^{\mathcal{G}} u\right)_{33} + \delta\mathcal{W}(\eta) + g.$$
Finally, taking the trace of (2.2a)$\cdot e_3$ onto Σ_b yields
$$\partial_3^{\mathcal{G}} p = \Delta^{\mathcal{G}} u_3.$$
So p solves
$$\begin{cases} -\Delta^{\mathcal{G}} p = \nabla^{\mathcal{G}} u : \left(\nabla^{\mathcal{G}} u\right)^T & \text{in } \Omega, \\ p = \left(\mathbb{D}^{\mathcal{G}} u\right)_{33} + \delta\mathcal{W}(\eta) + g & \text{on } \Sigma, \text{ and} \\ \partial_3^{\mathcal{G}} p = \Delta^{\mathcal{G}} u_3 & \text{on } \Sigma_b. \end{cases}$$
In particular, in the small energy regime where $\mathcal{G} \sim I$, standard elliptic estimates coupled with product estimates in Sobolev spaces (to handle the nonlinear but small remainders) allows us to recover p_0 from u_0 and η_0 using this PDE.

Constructing $\partial_t u_0$ and $\partial_t \eta_0$: We use (2.2a) and (2.2d) to define
$$\begin{cases} \partial_t u_0 := -\left(u_0 \cdot \nabla^{\mathcal{G}}\right) u_0 - \nabla^{\mathcal{G}} p_0 + \Delta^{\mathcal{G}} u_0 \text{ and} \\ \partial_t \eta_0 := u_0 \cdot \nu^{\mathcal{G}}_{\partial\Omega_0}. \end{cases}$$

Following the procedure outlined in Remark 6.2 leads to the following result, which not only constructs the data, but provides an estimate in the small energy regime.

Proposition 6.3 (Constructing the initial conditions). *There exist $\beta, C_{IC} > 0$ such that for every $T > 0$ for which $\mathcal{X} = (u, p, \eta)$ is a solution on $[0, T]$, if $\mathcal{I}(\mathcal{X}(0)) \leqslant \beta$ then $\overline{\mathcal{E}}(\mathcal{X}(0)) \leqslant C_{IC}\mathcal{I}(\mathcal{X}(0))$.*

Next we define the notion of admissible data.

Definition 6.4 (Admissible initial condition). We say that $(u_0, \eta_0) \in H^2\left(\Omega; \mathbb{R}^3\right) \times H^{9/2}\left(\mathbb{T}^2; \mathbb{R}\right)$ is an *admissible initial condition* if it satisfies

- $\nabla \cdot u_0 = 0$,
- $\operatorname{tr}_{\Sigma_b} u_0 = 0$,
- $\left(I - \nu_{\partial\Omega_0} \otimes \nu_{\partial\Omega_0}\right)\left(\operatorname{tr}_\Sigma \mathbb{D} u_0 \cdot \nu_{\partial\Omega_0}\right) = 0$,
- $\operatorname{tr}_\Sigma u_0 \cdot \nu^{\mathcal{G}}_{\partial\Omega_0} \in H^2(\Sigma)$,

- $\int_{\mathbb{T}^2} \eta_0 = 0$, and
- $\mathcal{I}(u,\eta) \leqslant \beta$ for β as in Proposition 6.3.

A few remarks are in order.

Remark 6.5.

(1) The first three items are nothing more than incompressibility and parts of the boundary conditions.
(2) The fifth condition, namely requiring that $\int_{\mathbb{T}^2} \eta = 0$, is related to (1.8).
(3) The fourth condition, namely requiring that $\operatorname{tr} u_0 \cdot \nu^{\mathcal{G}}_{\partial\Omega_0}$ be in H^2, is a *compatibility condition.* Indeed, knowing that u belong to H^2 and η belongs to $H^{9/2}$ only allows us to conclude that $\operatorname{tr} u_0 \cdot \nu_{\partial\Omega_0} = \operatorname{tr} u_0 \cdot \left(-\overline{\nabla}\eta_0, 1\right)$ belongs to $H^{3/2}$. This gap in regularity means the procedure sketched in Remark 6.2 cannot close without assuming that this additional compatibility condition holds a priori. Note that we will prove that this condition persists in time, so there is no trouble iteratively applying the local theory.
(4) The sixth condition is there to ensure that the nonlinear PDEs used in the sketch from Remark 6.2 are sufficiently close to their linear counterpart (corresponding to $\eta = 0$ and $\mathcal{G} = I$) such that the appropriate estimates can be made to produce the result from Proposition 6.3.

We now state the local existence result.

Theorem 6.6 (Local well-posedness). *There exist $T, \kappa_0, C_{lwp} > 0$ such that for every $\kappa \in (0, \kappa_0]$, if $(u_0, \eta_0) \in H^2\left(\Omega; \mathbb{R}^3\right) \times H^{9/2}\left(\mathbb{T}^2; \mathbb{R}\right)$ is an admissible initial condition (c.f. Definition 6.4) satisfying*

$$\mathcal{I}(u_0, \eta_0) \leqslant \kappa$$

(c.f. Definition 6.1 for the definition of $\mathcal{I}$), then there exists a unique solution $\mathcal{X} = (u, p, \eta)$ of (2.2a)–(2.2e) *on $[0, T]$ that satisfies*

$$\sup_{0\leqslant t\leqslant T} \mathcal{E}(\mathcal{X}(t)) + \int_0^T \mathcal{D}(\mathcal{X}(t))\, dt + \left|\left|\partial_t^2 u\right|\right|^2_{\mathcal{V}_T^*} \leqslant C_{lwp}\kappa,$$

where

$$\mathcal{V}_T := \left\{ u \in L^2\left([0,T]; H^1(\Omega)\right) \,\middle|\, \operatorname{tr}_{\Sigma_b} u(t) = 0 \text{ and } \nabla^{\mathcal{G}(t)} \cdot u(t) = 0 \text{ for a.e. } t \in [0,T] \right\}.$$

Remark 6.7. The local existence theorem is sufficient to justify our a priori estimates.

Note that in light of Remark 6.5 (and item (2) therein, in particular) the admissibility of initial conditions is propagated by the flow (provided the solution remains small enough).

Proposition 6.8 (Propagation of admissibility for initial conditions). *Suppose that (u, p, η) is a solution on $[0, T]$ such that (u_0, η_0) is an admissible initial condition. For every $t \in [0, T]$, if $\mathcal{I}(u(t), \eta(t)) \leqslant \beta$, then $(u(t), \eta(t))$ is an admissible initial condition (c.f. Definition 6.4).*

6.2. Proof of the main result. Before stating and proving the main result, i.e. the global well-posedness and decay result, we state and prove two preliminary lemmas. The first lemma, Lemma 6.9, is an eventual global well-posedness result that shows that if small solutions exist past a critical time, then they exist globally in time. The second lemma, Lemma 6.10, is a result about the existence of solution on arbitrarily large finite time intervals, provided the initial data is small enough. Combining these two lemmas will then allow us to prove global well-posedness in Theorem 6.11.

We now prove our first lemma. It says that past a critical time T_{crit}, the exponential decay from the a priori estimates is sufficiently strong to ensure that we remain in a regime where the energy is small enough for the local well-posedness to hold at every time thereafter. This is *eventual* well-posedness since it tells us that there exists a critical time past which the solution is globally well-defined.

Lemma 6.9 (Eventual global well-posedness). *Let δ, λ, and C_{ap} be as in Theorem 5.10. Let κ_0, C_{lwp}, and T be as in Theorem 6.6 and assume without loss of generality that $C_{lwp} \geqslant 1$. Let C_{IC} be as in Proposition 6.3, and let $T_{crit} > 0$ be such that $e^{\lambda\left(T_{crit} - \frac{T}{2}\right)} \geqslant C_{ap} C_{IC}$.*

If $\mathcal{X} = (u, p, \eta)$ is a solution on $[0, \tau]$ for some $\tau \geqslant T_{crit}$, with (u_0, η_0) an admissible initial condition that also satisfies the smallness conditions

$$\begin{cases} \mathcal{I}(\mathcal{X}_0) \leqslant \min\left\{\kappa_0, \frac{\delta}{C_{lwp}}\right\} \\ \sup\limits_{0\leqslant t\leqslant\tau} \mathcal{E}(\mathcal{X}) \leqslant \delta \quad \text{and} \quad \int_0^\tau \mathcal{D}(\mathcal{X}) < \infty, \end{cases} \tag{6.1}$$

then the solution can be uniquely extended to a solution on $[0, \infty)$ satisfying

$$\sup_{t\geqslant 0} \mathcal{E}(\mathcal{X}(t)) \leqslant \delta \quad \text{and} \quad \int_0^\infty \mathcal{D}(\mathcal{X}(t))\, dt < \infty.$$

Proof. Let $\tau \geqslant T_{crit}$ and let $\mathcal{X} = (u, p, \eta)$ be a solution on $[0, \tau]$ starting from admissible data (u_0, η_0) and satisfying (6.1). Define $T_{max} > 0$ to be

$$T_{max} := \sup\left\{ T \geqslant 0 \;\middle|\; \text{solution } \mathcal{X} \text{ exists on } [0, T] \text{ and satisfies } \sup_{0\leqslant t\leqslant T} \mathcal{E}(\mathcal{X}) \leqslant \delta \text{ and } \int_0^T \mathcal{D}(\mathcal{X}) < \infty \right\}.$$

First note that $T_{max} \geqslant \tau \geqslant T_{crit}$. Now suppose, by way of contradiction, that $T_{max} < \infty$. Let $\tilde{T} := T_{max} - \frac{T}{2} > 0$. By Theorem 5.10, Proposition 6.3, and the definition of T_{crit}, which is smaller than T_{max}, we have

$$\mathcal{E}\left(\mathcal{X}\left(\tilde{T}\right)\right) \leqslant C_{ap} e^{-\lambda\left(T_{max} - \frac{T}{2}\right)} \overline{\mathcal{E}}(\mathcal{X}(0)) \leqslant C_{ap} C_{IC} e^{-\lambda\left(T_{crit} - \frac{T}{2}\right)} \mathcal{I}(\mathcal{X}(0)) \leqslant \mathcal{I}(\mathcal{X}(0)).$$

Therefore, since $\mathcal{I}(\mathcal{X}(0)) \leqslant \min\left\{\kappa_0, \frac{\delta}{C_{lwp}}\right\}$, we may employ Proposition 6.8 and Theorem 6.6 to obtain a unique extension of the solution on $\left[0, T_{max} + \frac{T}{2}\right]$ satisfying

$$\sup_{0\leqslant t\leqslant T_{max}+\frac{T}{2}} \mathcal{E}(\mathcal{X}) + \int_0^{T_{max}+\frac{T}{2}} \mathcal{D}(\mathcal{X}) \leqslant C_{lwp} \frac{\delta}{C_{lwp}} \leqslant \delta.$$

We can thus use Theorem 5.10, Proposition 6.3, and the definition of T_{crit} once more, this time on $\left[0, T_{max} + \frac{T}{2}\right]$, to obtain that

$$\mathcal{E}\left(\mathcal{X}\left(T_{max} + \frac{T}{2}\right)\right) \leqslant C_{ap} e^{-\lambda\left(T_{max} + \frac{T}{2}\right)} \overline{\mathcal{E}}(\mathcal{X}(0)) \leqslant C_{ap} C_{IC} e^{-\lambda T_{crit}} \mathcal{I}(\mathcal{X}(0)) \leqslant \mathcal{I}(\mathcal{X}(0)) \leqslant \delta$$

which contradicts the definition of T_{max}. So indeed $T_{max} = \infty$. □

We now prove our second key lemma.

Lemma 6.10 (Arbitrary finite-time well-posedness). *For every $\tau > 0$ there exists $\gamma > 0$ such that if (u_0, η_0) is an admissible initial condition with*

$$\mathcal{I}(u_0, \eta_0) \leqslant \gamma,$$

then there exists a unique solution $\mathcal{X} = (u, p, \eta)$ on $[0, \tau]$ satisfying

$$\sup_{0\leqslant t\leqslant\tau} \mathcal{E}(\mathcal{X}(t)) + \int_0^\tau \mathcal{D}(\mathcal{X}(t))\, dt \leqslant \delta$$

for δ as in Theorem 5.10.

Proof. Let $\tau > 0$, let T be as in Theorem 6.6, and pick $N \in \mathbb{N}$ such that $NT \geqslant \tau$. Let C_{lwp} be as in Theorem 6.6, and note that without loss of generality we may assume that $C_{lwp} > 1$. Let β be as in Proposition 6.3 and let $\gamma := \frac{\beta}{C_{lwp}^N} > 0$.

Let (u_0, η_0) be an admissible initial condition satisfying $\mathcal{I}(u_0, \eta_0) \leqslant \gamma$. Then we apply the local well-posedness result, i.e. Theorem 6.6, N times (using Proposition 6.8 to ensure the 'initial conditions' are admissible at every step). More precisely, at step 1 we use Theorem 6.6 to obtain a unique solution $\mathcal{X} = (u, p, \eta)$ on $[0, T]$ satisfying

$$\sup_{0\leqslant t\leqslant T} \mathcal{E}(\mathcal{X}) + \int_0^T \mathcal{D}(\mathcal{X}) \leqslant C_{lwp}\gamma.$$

Since $\gamma \leqslant \frac{\beta}{C_{lwp}^N} \leqslant \frac{\beta}{C_{lwp}}$, it follows that (u_T, η_T) is an admissible initial condition. Then, at step n for $n = 2, \ldots, N$, suppose that we have solution on $[0, (n-1)T]$ satisfying

$$\sup_{0 \leqslant t \leqslant (n-1)T} \mathcal{E}(\mathcal{X}) + \int_0^{(n-1)T} \mathcal{D}(\mathcal{X}) \leqslant C_{lwp}^{n-1} \gamma$$

such that $\left(u_{(n-1)T}, \eta_{(n-1)T}\right)$ is an admissible initial condition. We may then apply Theorem 6.6 to extend the solution uniquely to $[0, nT]$ such that it satisfies

$$\sup_{0 \leqslant t \leqslant nT} \mathcal{E}(\mathcal{X}) + \int_0^{nT} \mathcal{D}(\mathcal{X}) \leqslant C_{lwp} \left(C_{lwp}^{n-1} \gamma\right) = C_{lwp}^n \gamma.$$

In particular, since $\gamma \leqslant \frac{\beta}{C_{lwp}^N} \leqslant \frac{\beta}{C_{lwp}^n}$, it follows from Proposition 6.8 that (u_{nT}, η_{nT}) is also an admissible initial condition. Finally, after step N, we have a solution on $[0, NT] \supseteq [0, \tau]$ satisfying

$$\sup_{0 \leqslant t \leqslant NT} \mathcal{E}(\mathcal{X}) + \int_0^{NT} \mathcal{D}(\mathcal{X}) \leqslant C_{lwp}^N \gamma \leqslant \delta.$$

$\square$

With the key lemmas in hand, we can now prove our main result.

THEOREM 6.11 (Global well-posedness and decay). *There exists $\epsilon > 0$ such that for every admissible initial condition (u_0, η_0) satisfying*

$$\mathcal{I}(u_0, \eta_0) \leqslant \epsilon$$

there exists a unique solution $\mathcal{X} = (u, p, \eta)$ on $[0, \infty)$ such that

$$\sup_{t \geqslant 0} \mathcal{E}(\mathcal{X}(t)) e^{\lambda t} + \int_0^\infty \mathcal{D}(\mathcal{X}(t)) e^{\lambda t} dt \leqslant C \overline{\mathcal{E}}(\mathcal{X}(0)),$$

where $C = C_{ap} > 0$ and $\lambda > 0$ are as in Theorem 5.10. Recall that admissible initial conditions are defined in Definition 6.4.

PROOF. Let δ be as in the a priori estimates (i.e. Theorem 5.10), let κ_0 and C_{lwp} be as in the local well-posedness result (i.e. Theorem 6.6), let T_{crit} be as in the eventual global well-posedness result (i.e. Lemma 6.9), and let $\gamma = \gamma(T_{crit})$ be as in the arbitrary finite time existence result (i.e. Lemma 6.10). Pick $\epsilon = \min\left(\gamma, \kappa_0, \frac{\delta}{C_{lwp}}\right)$. Now let (u_0, η_0) be an admissible initial condition satisfying $\mathcal{I}(u_0, \eta_0) \leqslant \epsilon$. By the arbitrary finite time existence result (i.e. Lemma 6.10) and the choice of ϵ, there exists a unique solution $\mathcal{X} = (u, p, \eta)$ on $[0, T_{crit}]$ satisfying

$$\sup_{0 \leqslant t \leqslant T_{crit}} \mathcal{E}(\mathcal{X}(t)) + \int_0^{T_{crit}} \mathcal{D}(\mathcal{X}(t)) \leqslant \delta$$

and therefore by the eventual global well-posedness result (i.e. Lemma 6.9) and the choice of ϵ there exists a unique extension of this solution to $[0, \infty)$ satisfying

$$\sup_{t \geqslant 0} \mathcal{E}(\mathcal{X}(t)) \leqslant \delta \quad \text{and} \quad \int_0^\infty \mathcal{D}(\mathcal{X}(t)) < \infty.$$

Finally we establish the exponential decay of the energy of this unique global solution. The a priori estimates (i.e. Theorem 5.10) tell us that for every $T > 0$

$$\sup_{0 \leqslant t \leqslant T} \mathcal{E}(\mathcal{X}(t)) e^{\lambda t} + \int_0^T \mathcal{D}(\mathcal{X}(t)) e^{\lambda t} dt \leqslant C \overline{\mathcal{E}}(\mathcal{X}(0))$$

and so indeed, taking the supremum over $T > 0$ yields the global decay estimate

$$\sup_{t \geqslant 0} \mathcal{E}(\mathcal{X}(t)) e^{\lambda t} + \int_0^\infty \mathcal{D}(\mathcal{X}(t)) e^{\lambda t} dt \leqslant C \overline{\mathcal{E}}(\mathcal{X}(0)).$$

$\square$

7. Appendix: intermediate results

In this first part of the appendix we record various intermediate results of particular interest to the problem discussed in this chapter. We record computations and estimates for the geometric coefficients Φ, $\mathcal{G}$, J and $\nu^{\mathcal{G}}_{\partial\Omega}$, as well as computations and estimates for the variations of the surface energy. We also record details of the computations of various commutators.

7.1. Geometric coefficients and differential operators. In this section we record estimates for the geometric coefficients Φ, $\mathcal{G}$, J and $\nu^{\mathcal{G}}_{\partial\Omega}$ (as defined in Section 4.3) in Lemma 7.1, and we record the $\mathcal{G}$-divergence and $\mathcal{G}$-transport theorems in Propositions 7.2 and 7.3 respectively.

Lemma 7.1 (Estimates for the geometric coefficients). *Recall the notational conventions of Section 4.3. Suppose that we are in the small energy regime (see Definition 4.2). On the upper surface we have the bounds*

$$||\mathrm{tr}\,(\mathcal{G}-I)||_{H^{7/2}(\mathbb{T}^2)} + \left|\left|\nu^{\mathcal{G}}_{\partial\Omega} - e_3\right|\right|_{H^{7/2}(\mathbb{T}^2)} \lesssim \sqrt{\mathcal{E}}.$$

In the bulk we have the bounds

$$||\Phi - id||_{H^5(\Omega)} + ||\partial_t\Phi||_{H^{5/2}(\Omega)} + ||J-1||_{H^4(\Omega)} + ||\partial_t J||_{H^{3/2}(\Omega)} + ||\mathcal{G}-I||_{H^4(\Omega)} + ||\partial_t\mathcal{G}||_{H^{3/2}(\Omega)} \lesssim \sqrt{\mathcal{E}}$$

and

$$||\Phi - id||_{H^6(\Omega)} + ||\partial_t\Phi||_{H^3(\Omega)} + \left|\left|\partial_t^2\Phi\right|\right|_{H^1(\Omega)} + ||J-1||_{H^5(\Omega)} + ||\partial_t J||_{H^2(\Omega)} + ||\partial_t\mathcal{G}||_{H^2(\Omega)} + \left|\left|\partial_t^2\mathcal{G}\right|\right|_{H^0(\Omega)} \lesssim \sqrt{\mathcal{D}}.$$

Proof. We begin with estimating $\Phi = \mathrm{id} + \chi\,\mathrm{ext}\,\eta\, e_3$ and its time derivatives: $\partial_t\Phi = \chi\,\mathrm{ext}\,\partial_t\eta\, e_3$ and $\partial_t^2\Phi = \chi\,\mathrm{ext}\,\partial_t^2\eta\, e_3$. We estimate $\Phi - \mathrm{id}$ using Proposition 8.13, Corollary 8.9, and the definitions of $\mathcal{E}$ and $\mathcal{D}$ (c.f. equations (4.6a) and (4.6b), respectively):

$$||\Phi - \mathrm{id}||_{H^5(\Omega)} = ||\chi\,\mathrm{ext}\,\eta\, e_3||_{H^5(\Omega)} \lesssim ||\chi||_{H^{13/2+}(\Omega)}||\mathrm{ext}\,\eta||_{H^5(\Omega)} \lesssim ||\eta||_{H^{9/2}(\mathbb{T}^2)} \leqslant \sqrt{\mathcal{E}}$$

and

$$||\Phi - \mathrm{id}||_{H^6(\Omega)} \lesssim ||\chi||_{H^{15/2+}(\Omega)}||\eta||_{H^{11/2}(\mathbb{T}^2)} \lesssim \sqrt{\mathcal{D}}.$$

We proceed similarly to estimate the time derivatives of Φ:

$$\begin{cases} ||\partial_t\Phi||_{H^{5/2}(\Omega)} \lesssim ||\partial_t\eta||_{H^2(\mathbb{T}^2)} \leqslant \sqrt{\mathcal{E}}, \\ ||\partial_t\Phi||_{H^3(\Omega)} \lesssim ||\partial_t\eta||_{H^{5/2}(\mathbb{T}^2)} \leqslant \sqrt{\mathcal{D}}, \text{ and} \\ \left|\left|\partial_t^2\Phi\right|\right|_{H^1(\Omega)} \lesssim \left|\left|\partial_t^2\eta\right|\right|_{H^{1/2}(\mathbb{T}^2)} \leqslant \sqrt{\mathcal{D}}. \end{cases}$$

Now we compute J, noting first that $\nabla\Phi = I + e_3 \otimes \nabla\left(\chi\,\mathrm{ext}\,\eta\right)$. Therefore, by Lemma 8.23, by definition of χ (c.f. Section 2.1), and by Lemma 8.7

$$J = \det\nabla\Phi = 1 + \partial_3\left(\chi\,\mathrm{ext}\,\eta\right) = 1 + \frac{\mathrm{ext}\,\eta}{b} + \chi\,\mathrm{ext}\,\sqrt{-\Delta}\eta \text{ and } \partial_t J = \partial_3\left(\chi\,\mathrm{ext}\,\partial_t\eta\right).$$

We may now estimate J and its time derivatives

$$||J-1||_{H^4(\Omega)} = ||\partial_3\left(\chi\,\mathrm{ext}\,\eta\right)||_{H^4(\Omega)} \leqslant ||\chi\,\mathrm{ext}\,\eta||_{H^5(\Omega)} \lesssim ||\chi||_{H^{13/2+}(\Omega)}||\mathrm{ext}\,\eta||_{H^5(\Omega)} \lesssim ||\eta||_{H^{9/2}(\mathbb{T}^2)} \leqslant \sqrt{\mathcal{E}}$$

and similarly

$$\begin{cases} ||J-1||_{H^5(\Omega)} \lesssim ||\eta||_{H^{11/2}(\mathbb{T}^2)} \leqslant \sqrt{\mathcal{D}}, \\ ||\partial_t J||_{H^{3/2}(\Omega)} \lesssim ||\partial_t\eta||_{H^2(\mathbb{T}^2)} \leqslant \sqrt{\mathcal{E}}, \text{ and} \\ ||\partial_t J||_{H^2(\Omega)} \lesssim ||\partial_t\eta||_{H^{5/2}(\mathbb{T}^2)} \leqslant \sqrt{\mathcal{D}}. \end{cases}$$

Now we compute $\mathcal{G}$. Recall that $\mathcal{G} := \left(\nabla\Phi\right)^{-T}$ with $\nabla\Phi = I + e_3 \otimes \nabla\left(\chi\,\mathrm{ext}\,\eta\right)$. Therefore, by Lemma 8.23,

$$\mathcal{G} = I - \frac{\nabla\left(\chi\,\mathrm{ext}\,\eta\right)\otimes e_3}{1 + \partial_3\left(\chi\,\mathrm{ext}\,\eta\right)}$$

i.e. $\mathcal{G} - I = g\left(\nabla\left(\mathcal{X}\operatorname{ext}\eta\right)\right)$ for $g\left(w\right) := \frac{-w\otimes e_3}{1+w\cdot e_3}$ for every $w \in \mathbb{R}^3$ such that $w_3 \neq 1$. We may now estimate $\mathcal{G} - I$ using Proposition 8.15 to obtain

$$\begin{aligned}||\mathcal{G} - I||_{H^4(\Omega)} &\lesssim ||g\left(\nabla\left(\mathcal{X}\operatorname{ext}\eta\right)\right)||_{L^2(\Omega)} \\ &\quad + \underbrace{||g||_{C^{4,1}\left(\overline{B\left(||\nabla(\mathcal{X}\operatorname{ext}\eta)||_\infty\right)}\right)}}_{=:(\star)}\left(||\nabla\left(\mathcal{X}\operatorname{ext}\eta\right)||_{H^4(\Omega)} + ||\nabla\left(\mathcal{X}\operatorname{ext}\eta\right)||^4_{H^4(\Omega)}\right).\end{aligned}$$

Crucially, since we are in the small energy regime (c.f. Definition 4.2),

$$||\partial_3\left(\mathcal{X}\operatorname{ext}\eta\right)||_\infty \leqslant C_0\delta_0 < 1$$

such that $(\star) < \infty$ (since g is and well-defined and hence smooth on the compact set $\overline{B\left(||\nabla\left(\mathcal{X}\operatorname{ext}\eta\right)||_\infty\right)}$) and

$$\begin{aligned}||g\left(\nabla\left(\mathcal{X}\operatorname{ext}\eta\right)\right)||^2_{L^2(\Omega)} &= \int_\Omega \frac{\left|\nabla\left(\mathcal{X}\operatorname{ext}\eta\right)\otimes e_3\right|^2}{\left(1+\partial_3\left(\mathcal{X}\operatorname{ext}\eta\right)\right)^2} \leqslant \int_\Omega \frac{\left|\nabla\left(\mathcal{X}\operatorname{ext}\eta\right)\right|^2}{\left(1-||\partial_3\left(\mathcal{X}\operatorname{ext}\eta\right)||_\infty\right)^2} \\ &\leqslant \frac{1}{\left(1-C_0\delta_0\right)^2}\int_\Omega \left(\nabla\left(\mathcal{X}\operatorname{ext}\eta\right)\right)^2.\end{aligned}$$

Therefore, employing Proposition 8.13, Corollary 8.9, and the definition of $\mathcal{E}$ (c.f. equation (4.6a)), we obtain

$$\begin{aligned}||\mathcal{G} - I||_{H^4(\Omega)} &\lesssim ||\nabla\left(\mathcal{X}\operatorname{ext}\eta\right)||_{L^2(\Omega)} + ||\nabla\left(\mathcal{X}\operatorname{ext}\eta\right)||_{H^4(\Omega)} + ||\nabla\left(\mathcal{X}\operatorname{ext}\eta\right)||^4_{H^4(\Omega)} \\ &\lesssim ||\mathcal{X}||_{H^{13/2+}(\Omega)}||\operatorname{ext}\eta||_{H^5(\Omega)} + \left(||\mathcal{X}||_{H^{13/2+}(\Omega)}||\operatorname{ext}\eta||_{H^5(\Omega)}\right)^4 \\ &\lesssim ||\eta||_{H^{9/2}(\mathbb{T}^2)} + ||\eta||^4_{H^{9/2}(\mathbb{T}^2)} \quad \lesssim \sqrt{\mathcal{E}} + \mathcal{E}^2 \quad \lesssim \sqrt{\mathcal{E}} \quad \text{since } \mathcal{E} \leqslant \delta_0 < 1.\end{aligned}$$

We now compute the time derivatives of $\mathcal{G}$:

$$\begin{cases}\partial_t\mathcal{G} &= \partial_t\left(g\left(\nabla\left(\mathcal{X}\operatorname{ext}\eta\right)\right)\right) = \left(\nabla g\right)\left(\nabla\left(\mathcal{X}\operatorname{ext}\eta\right)\right)\cdot\nabla\left(\mathcal{X}\operatorname{ext}\partial_t\eta\right), \\ \partial_t^2\mathcal{G} &= \partial_t\left(\left(\nabla g\right)\left(\nabla\left(\mathcal{X}\operatorname{ext}\eta\right)\right)\cdot\nabla\left(\mathcal{X}\operatorname{ext}\partial_t\eta\right)\right), \text{ and} \\ &= \left(\nabla^2 g\right)\left(\nabla\left(\mathcal{X}\operatorname{ext}\eta\right)\right) : \left(\nabla\left(\mathcal{X}\operatorname{ext}\partial_t\eta\right)^{\otimes 2}\right) + \left(\nabla g\right)\left(\nabla\left(\mathcal{X}\operatorname{ext}\eta\right)\right)\cdot\nabla\left(\mathcal{X}\operatorname{ext}\partial_t^2\eta\right)\end{cases}$$

such that we may now estimate them, using Proposition 8.13, Proposition 8.15, Corollary 8.9, equations (4.6a) and (4.6b), and the fact that we are in the small energy regime. Doing so, we obtain

$$\begin{aligned}||\partial_t\mathcal{G}||_{H^{3/2}(\Omega)} &\lesssim ||\left(\nabla g\right)\left(\nabla\left(\mathcal{X}\operatorname{ext}\eta\right)\right)||_{H^{7/2}(\Omega)}||\nabla\left(\mathcal{X}\operatorname{ext}\partial_t\eta\right)||_{H^{3/2}(\Omega)} \\ &\lesssim \Bigg(||\nabla g||_{L^\infty\left(\overline{B\left(||\nabla(\mathcal{X}\operatorname{ext}\eta)||_\infty\right)}\right)} \\ &\quad + ||\nabla g||_{C^{3,1}\left(\overline{B\left(||\nabla(\mathcal{X}\operatorname{ext}\eta)||_\infty\right)}\right)}\left(||\nabla\left(\mathcal{X}\operatorname{ext}\eta\right)||_{H^{7/2}(\Omega)} + ||\nabla\left(\mathcal{X}\operatorname{ext}\eta\right)||^4_{H^{7/2}(\Omega)}\right)\Bigg)||\partial_t\eta||_{H^2(\mathbb{T}^2)} \\ &\lesssim \sqrt{\mathcal{E}},\end{aligned}$$

and

$$\begin{aligned}||\partial_t\mathcal{G}||_{H^2(\Omega)} &\lesssim ||\left(\nabla g\right)\left(\nabla\left(\mathcal{X}\operatorname{ext}\eta\right)\right)||_{H^4(\Omega)}||\nabla\left(\mathcal{X}\operatorname{ext}\partial_t\eta\right)||_{H^2(\Omega)} \\ &\lesssim \Bigg(||\nabla g||_{L^\infty\left(\overline{B\left(||\nabla(\mathcal{X}\operatorname{ext}\eta)||_\infty\right)}\right)} \\ &\quad + ||\nabla g||_{C^{4,1}\left(\overline{B\left(||\nabla(\mathcal{X}\operatorname{ext}\eta)||_\infty\right)}\right)}\left(||\nabla\left(\mathcal{X}\operatorname{ext}\eta\right)||_{H^4(\Omega)} + ||\nabla\left(\mathcal{X}\operatorname{ext}\eta\right)||^4_{H^4(\Omega)}\right)\Bigg)||\partial_t\eta||_{H^{5/2}(\mathbb{T}^2)} \\ &\lesssim \sqrt{\mathcal{D}}.\end{aligned}$$

Similarly, using Hölder's inequality, the Sobolev embedding $H^{3/4}(\Omega) \hookrightarrow L^4(\Omega)$, Proposition 8.13, Corollary 8.9, equations (4.6a) and (4.6b), and the fact that we are in the small energy regime, we obtain

$$\begin{aligned} \left|\left|\partial_t^2 \mathcal{G}\right|\right|_{L^2(\Omega)} &\lesssim \left|\left|\left(\nabla^2 g\right)\left(\nabla\left(\chi \operatorname{ext} \eta\right)\right)\right|\right|_{L^\infty(\Omega)} \left|\left|\nabla\left(\chi \operatorname{ext} \partial_t \eta\right)\right|\right|^2_{H^{3/4}(\Omega)} \\ &\quad + \left|\left|\left(\nabla g\right)\left(\nabla\left(\chi \operatorname{ext} \eta\right)\right)\right|\right|_{L^\infty(\Omega)} \left|\left|\nabla\left(\chi \operatorname{ext} \partial_t^2 \eta\right)\right|\right|^2_{L^2(\Omega)} \\ &\lesssim \left|\left|\partial_t \eta\right|\right|^2_{H^{5/4}(\mathbb{T}^2)} + \left|\left|\partial_t^2 \eta\right|\right|_{H^{1/2}(\mathbb{T}^2)} \lesssim \sqrt{\mathcal{E}} + \sqrt{\mathcal{D}} \lesssim \sqrt{\mathcal{D}}. \end{aligned}$$

Finally we estimate $\nu^{\mathcal{G}}_{\partial\Omega}$ on Σ. First we compute:

$$\begin{aligned} \nu^{\mathcal{G}}_{\partial\Omega}|_\Sigma &= J\left(\operatorname{tr}_\Sigma \mathcal{G}\right) \cdot \nu^{\mathcal{G}}_{\partial\Omega}|_\Sigma = J \operatorname{tr}_\Sigma\left(I - \frac{\nabla\left(\chi \operatorname{ext} \eta\right) \otimes e_3}{1 + \partial_3\left(\chi \operatorname{ext} \eta\right)}\right) \cdot e_3 \\ &= J \operatorname{tr}_\Sigma\left(I - \frac{\operatorname{ext} \widetilde{\nabla} \eta \otimes e_3 + \partial_3\left(\chi \operatorname{ext} \eta\right) e_3 \otimes e_3}{1 + \partial_3\left(\chi \operatorname{ext} \eta\right)}\right) \cdot e_3 = J\left(e_3 - \frac{\widetilde{\nabla}\eta + (J-1)\, e_3}{J}\right) = -\widetilde{\nabla}\eta + e_3. \end{aligned}$$

Therefore

$$\left|\left|\nu^{\mathcal{G}}_{\partial\Omega} - e_3\right|\right|_{H^{7/2}(\mathbb{T}^2)} = \left|\left|\nabla \eta\right|\right|_{H^{7/2}(\mathbb{T}^2)} \lesssim \sqrt{\mathcal{E}}.$$

□

We now record versions of the divergence and transport theorem adapted to the differential operators appearing in the PDE after performing the time-dependent change of variables which fixes the domain. In particular, we prove the $\mathcal{G}$-divergence theorem in Proposition 7.2 and we prove the $\mathcal{G}$-transport theorem in Proposition 7.3. The key differences between these theorems and the standard divergence and transport theorems are that:

- standard operators involving ∇ are replaced by their counterparts involving $\nabla^{\mathcal{G}}$, and
- bulk integrands, i.e. integrands over Ω, are multiplied by J

(see Sections 4.3.2 and 4.3.3 for the definitions of J and $\nabla^{\mathcal{G}}$ respectively).

Proposition 7.2 ($\mathcal{G}$-divergence theorem). *For any $v : [0,T] \times \Omega \to \mathbb{R}^3$ sufficiently regular and integrable*

$$\int_\Omega \left(\nabla^{\mathcal{G}} \cdot v\right) J = \int_{\partial\Omega} v \cdot \nu^{\mathcal{G}}_{\partial\Omega}.$$

Proof. This result follows from the divergence theorem and the Piola identity. Indeed, we compute:

$$\begin{aligned} \int_\Omega \left(\nabla^{\mathcal{G}} \cdot v\right) J = \int_\Omega \mathcal{G} : (\nabla v)\, J &= \int_\Omega \nabla \cdot \left(\mathcal{G}^T v J\right) - \int_\Omega \underbrace{\left(\nabla \cdot (\mathcal{G} J)\right)}_{\overset{(\star)}{=} 0} v \\ &= \int_{\partial\Omega} \left(\mathcal{G}^T \cdot v\right) \cdot \nu_{\partial\Omega} J = \int_{\partial\Omega} v \cdot \underbrace{\left(\mathcal{G} \cdot \nu_{\partial\Omega}\right) J}_{= \nu^{\mathcal{G}}_{\partial\Omega}}, \end{aligned}$$

where in $(\star)$ we have used the Piola identity which says that cofactor matrices of gradients are divergence-free: $\nabla \cdot (\mathcal{G} J) = \nabla \cdot (\operatorname{cof} \nabla \Phi) = 0$. □

Next we record a version of the transport theorem.

Proposition 7.3 ($\mathcal{G}$-transport theorem). *For any $f : [0,T] \times \Omega \to \mathbb{R}$ sufficiently regular and integrable*

$$\frac{\mathrm{d}}{\mathrm{dt}}\left(\int_\Omega f J\right) = \int_\Omega \left(D_t^{u,\mathcal{G}} f\right) J,$$

where the differential operator $D_t^{u,\mathcal{G}}$ is as defined in Section 4.3.3.

The proof of the $\mathcal{G}$-transport theorem relies on two small computations, recorded in the following lemma.

Lemma 7.4. *We have that $\partial_t J = \left(\nabla^{\mathcal{G}} \cdot \partial_t \Phi\right) J$, and $u \cdot \nu^{\mathcal{G}}_{\partial\Omega} = \partial_t \Phi \cdot \nu^{\mathcal{G}}_{\partial\Omega}$ on $\partial\Omega$.*

Proof of the $\mathcal{G}$-transport theorem. Using Lemma 7.4, this is a direct computation:

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(\int_\Omega fJ\right) = \int_\Omega (\partial_t f)\, J + \int_\Omega f\,(\partial_t J)\,,$$

where

$$\int_\Omega f\,(\partial_t J) = \int_\Omega f\left(\nabla^{\mathcal{G}}\cdot\partial_t\Phi\right) J = \int_\Omega \nabla^{\mathcal{G}}\cdot(f\partial_t\Phi)\, J - \int_\Omega \left(\partial_t\Phi\cdot\nabla^{\mathcal{G}} f\right) J.$$

To compute $\int_\Omega \nabla^{\mathcal{G}}\cdot(f\partial_t\Phi)\,J$ we use the $\mathcal{G}$-divergence theorem and the fact that $u\cdot\nu^{\mathcal{G}}_{\partial\Omega} = \partial_t\Phi\cdot\nu^{\mathcal{G}}_{\partial\Omega}$ on $\partial\Omega$:

$$\int_\Omega \nabla^{\mathcal{G}}\cdot(f\partial_t\Phi)\,J = \int_{\partial\Omega} f\left(\partial_t\Phi\cdot\nu^{\mathcal{G}}_{\partial\Omega}\right) J = \int_{\partial\Omega} f\left(u\cdot\nu^{\mathcal{G}}_{\partial\Omega}\right) J = \int_\Omega \nabla^{\mathcal{G}}\cdot(fu)\,J.$$

So, finally

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(\int_\Omega fJ\right) = \int_\Omega \underbrace{\left(\partial_t f - \partial_t\Phi\cdot\nabla^{\mathcal{G}} f\right)}_{=\partial_t^{\mathcal{G}} f} J + \nabla^{\mathcal{G}}\cdot(fu)\,J = \int_\Omega \left(D_t^{u,\mathcal{G}} f\right) J$$

since $\nabla^{\mathcal{G}}\cdot u = 0$. □

Proof of Lemma 7.4. Computing $\partial_t J$ and $\partial_t\Phi\cdot\nu^{\mathcal{G}}_{\partial\Omega}$ is nothing more than unpacking the relevant notation (c.f. Section 2.1 for the definition of Φ and Section 4.3 for other associated quantities). Indeed,

$$\partial_t J = \partial_t \det\nabla\Phi = \det\nabla\Phi\,\mathrm{tr}\left(\nabla\Phi^{-1}\cdot\partial_t\nabla\Phi\right) = \underbrace{\det(\nabla\Phi)}_{J}\underbrace{(\nabla\Phi)^{-T}}_{\mathcal{G}} : \nabla\partial_t\Phi = \left(\nabla^{\mathcal{G}}\cdot\partial_t\Phi\right) J,$$

which proves the first identity. For the second note that on $\partial\Omega$,

$$\partial_t\Phi\cdot\nu^{\mathcal{G}}_{\partial\Omega} = \partial_t\left(\mathrm{id} + \chi\,\mathrm{ext}\,\eta\, e_3\right)\cdot\nu^{\mathcal{G}}_{\partial\Omega} = \chi\,\mathrm{ext}\,\partial_t\eta\, e_3\cdot\nu^{\mathcal{G}}_{\partial\Omega} = \begin{cases}\partial_t\eta & \text{on } \Sigma_b\\ 0 & \text{on } \Sigma\end{cases}$$

which means that $\partial_t\Phi\cdot\nu^{\mathcal{G}}_{\partial\Omega} = u\cdot\nu^{\mathcal{G}}_{\partial\Omega}$. □

7.2. Commutators associated with the surface energy. Recall from Section 5.1 that

$$\mathcal{C}^{\mathcal{W},\alpha}(\eta) := \Big(\left(\nu^{\mathcal{G}}_{\partial\Omega}\delta^2_\eta\mathcal{W}\right)\circ\partial^\alpha - \partial^\alpha\circ\left(\nu^{\mathcal{G}}_{\partial\Omega}\delta\mathcal{W}\right)\Big)(\eta)\,.$$

We compute these commutators in the lemma below (for $|\alpha| = 1, 2$), using Remark 8.2.

Lemma 7.5 (Computing the commutators $\mathcal{C}^{\mathcal{W},\alpha}$). *For $|\alpha| = 1$ we have that*

$$\mathcal{C}^{\mathcal{W},\alpha}(\eta) = \left(\partial^\alpha\nu^{\mathcal{G}}_{\partial\Omega}\right)(\delta\mathcal{W})(\eta)\,.$$

Also, for $|\alpha| = |\beta| = 1$ we have that

$$\begin{aligned}\mathcal{C}^{\mathcal{W},\alpha+\beta}(\eta) &= \left(\partial^{\alpha+\beta}\nu^{\mathcal{G}}_{\partial\Omega}\right)(\delta\mathcal{W})(\eta) + \left(\partial^\alpha\nu^{\mathcal{G}}_{\partial\Omega}\right)\left(\delta^2_\eta\mathcal{W}\right)\left(\partial^\beta\eta\right)\\ &\quad+ \left(\partial^\beta\nu^{\mathcal{G}}_{\partial\Omega}\right)\left(\delta^2_\eta\mathcal{W}\right)(\partial^\alpha\eta) + \nu^{\mathcal{G}}_{\partial\Omega}\left(\delta^3_\eta\mathcal{W}\right)\left(\partial^\alpha\eta,\partial^\beta\eta\right).\end{aligned}$$

Proof. Using Remark 8.2, both of these results follow from direct computations: for $|\alpha| = |\beta| = 1$,

$$\partial^\alpha\Big(\nu^{\mathcal{G}}_{\partial\Omega}\delta\mathcal{W}(\eta)\Big) = \left(\partial^\alpha\nu^{\mathcal{G}}_{\partial\Omega}\right)(\delta\mathcal{W})(\eta) + \nu^{\mathcal{G}}_{\partial\Omega}\left(\delta^2_\eta\mathcal{W}\right)(\partial^\alpha\eta)$$

and

$$\begin{aligned}\partial^{\alpha+\beta}\Big(\nu^{\mathcal{G}}_{\partial\Omega}(\delta\mathcal{W})(\eta)\Big) &= \partial^\beta\Big(\left(\partial^\alpha\nu^{\mathcal{G}}_{\partial\Omega}\right)(\delta\mathcal{W})(\eta) + \nu^{\mathcal{G}}_{\partial\Omega}\left(\delta^2_\eta\mathcal{W}\right)(\partial^\alpha\eta)\Big)\\ &= \left(\partial^{\alpha+\beta}\nu^{\mathcal{G}}_{\partial\Omega}\right)(\delta\mathcal{W})(\eta) + \left(\partial^\alpha\nu^{\mathcal{G}}_{\partial\Omega}\right)\left(\delta^2_\eta\mathcal{W}\right)\left(\partial^\beta\eta\right)\\ &\quad+ \left(\partial^\beta\nu^{\mathcal{G}}_{\partial\Omega}\right)\left(\delta^2_\eta\mathcal{W}\right)(\partial^\alpha\eta) + \nu^{\mathcal{G}}_{\partial\Omega}\left(\delta^3_\eta\mathcal{W}\right)\left(\partial^\alpha\eta,\partial^\beta\eta\right) + \nu^{\mathcal{G}}_{\partial\Omega}\left(\delta^2_\eta\mathcal{W}\right)\left(\partial^{\alpha+\beta}\eta\right).\end{aligned}$$

□

7.3. Form of the geometric corrections. Recall from Section 5.3 that the geometric corrections are

$$\mathcal{G}_E(\mathcal{X}) = \widetilde{\mathcal{E}}(\mathcal{X};\mathcal{X}) - \overline{\mathcal{E}}(\mathcal{X}) \text{ and } \mathcal{G}_D(\mathcal{X}) = \widetilde{\mathcal{D}}(\mathcal{X};\mathcal{X}) - \overline{\mathcal{D}}(\mathcal{X}).$$

In this section we show that they can be computed to be

$$\begin{cases} \mathcal{G}_E(\mathcal{X}) &= \sum\limits_{|\alpha|_{t,\overline{x}^2}\leq 1,2} \left(\frac{1}{2}\int_\Omega |\partial^\alpha u|^2 (J-1) + \frac{1}{2}\int_{\mathbb{T}^2}\left(\int_0^1 g_\alpha(t)\nabla^3 f(t\mathcal{J}\eta)\,\mathrm{d}t\right)\bullet(\mathcal{J}\eta\otimes J\partial^\alpha\eta\otimes J\partial^\alpha\eta)\right) \text{ and} \\ \mathcal{G}_D(\mathcal{X}) &= \sum\limits_{|\alpha|_{t,\overline{x}^2}\leq 1} \left(\frac{1}{2}\int_\Omega |\mathbb{D}^{\mathcal{G}-I}\partial^\alpha u|^2 J - \int_\Omega \left(\mathbb{D}^{\mathcal{G}-I}\partial^\alpha u : \mathbb{D}\partial^\alpha u\right) J + \frac{1}{2}\int_\Omega |\mathbb{D}\partial^\alpha u|^2 (J-1)\right) \end{cases} \tag{7.1}$$

where

$$g_\alpha(t) := \begin{cases} \frac{1}{3}(1-t)^2 & \text{for } \alpha = 0 \\ 1 & \text{for } \alpha \neq 0. \end{cases} \tag{7.2}$$

We first compute the geometric correction to the energy:

$$\begin{aligned} \mathcal{G}_E(\mathcal{X}) &= \widetilde{\mathcal{E}}(\mathcal{X};\mathcal{X}) - \overline{\mathcal{E}}(\mathcal{X}) = \left(\widetilde{\mathcal{E}}^0(\mathcal{X}) - \overline{\mathcal{E}}^0(\mathcal{X})\right) + \sum_{|\alpha|_{t,\overline{x}^2}=1,2}\left(\widetilde{\mathcal{E}}(\partial^\alpha\mathcal{X};\mathcal{X}) - \overline{\mathcal{E}}(\partial^\alpha\mathcal{X})\right) \\ &= \frac{1}{2}\int_\Omega |u|^2(J-1) + (\mathcal{W}-\mathcal{Q}_0)(\eta) + \sum_{|\alpha|_{t,\overline{x}^2}=1,2}\frac{1}{2}\int_\Omega |\partial^\alpha u|^2(J-1) + (\mathcal{Q}_\eta - \mathcal{Q}_0)(\partial^\alpha\eta). \end{aligned}$$

Now we can compute $\mathcal{W}-\mathcal{Q}_0$ using Taylor's theorem (using the same notation, namely $\mathcal{P}_2$ and $\mathcal{R}_2$ as in Proposition 8.24), recalling that $f(0) = 0$ and $\nabla f(0) = 0$,

$$\begin{aligned} (\mathcal{W}-\mathcal{Q}_0)(\eta) &= \int_{\mathbb{T}^2} f(\mathcal{J}\eta) - \frac{1}{2}\nabla^2 f(0)\bullet(\mathcal{J}\eta\otimes\mathcal{J}\eta) = \int_{\mathbb{T}^2}(f - \mathcal{P}[f,0])(\mathcal{J}\eta) \\ &= \int_{\mathbb{T}^2}\mathcal{R}[f,0](\mathcal{J}\eta) = \frac{1}{6}\int_{\mathbb{T}^2}\left(\int_0^1 (1-t)^2\nabla^3 f(t\mathcal{J}\eta)\,dt\right)\bullet(\mathcal{J}\eta)^{\otimes 3}. \end{aligned}$$

Similarly we can compute $(\mathcal{Q}_\eta - \mathcal{Q}_0)(\zeta)$ for $\zeta \in \{\partial^\alpha u\}_{|\alpha|_{t,\overline{x}^2}=1,2}$ using the fundamental theorem of calculus:

$$(\mathcal{Q}_\eta - \mathcal{Q}_0)(\zeta) = \frac{1}{2}\int_{\mathbb{T}^2}\left(\nabla^2 f(\mathcal{J}\eta) - \nabla^2 f(0)\right)\bullet(J\zeta\otimes J\zeta) = \frac{1}{2}\int_{\mathbb{T}^2}\left(\int_0^1\nabla^3 f(t\mathcal{J}\eta)\,dt\right)\bullet(\mathcal{J}\eta\otimes J\zeta\otimes J\zeta),$$

which means equations (7.1) hold for g_α given by (7.2).

We now compute the geometric correction to the dissipation. Note that $M \mapsto \mathbb{D}^M v$ is linear, so in particular $|\mathbb{D}^{\mathcal{G}}u|^2 = |\mathbb{D}^{\mathcal{G}-I}u - \mathbb{D}u|^2 = |\mathbb{D}^{\mathcal{G}-I}u|^2 - 2\mathbb{D}^{\mathcal{G}-I} : \mathbb{D}u + |\mathbb{D}u|^2$. Therefore,

$$\begin{aligned} \mathcal{G}_D(\mathcal{X}) &= \widetilde{\mathcal{D}}(\mathcal{X};\mathcal{X}) - \overline{\mathcal{D}}(\mathcal{X}) = \left(\widetilde{\mathcal{D}}^0(\mathcal{X}) - \overline{\mathcal{D}}^0(\mathcal{X})\right) + \sum_{|\alpha|_{t,\overline{x}^2}=1,2}\left(\widetilde{\mathcal{D}}(\partial^\alpha\mathcal{X};\mathcal{X}) - \overline{\mathcal{D}}(\partial^\alpha\mathcal{X})\right) \\ &= \sum_{|\alpha|_{t,\overline{x}^2}=1,2}\left(\frac{1}{2}\int_\Omega\left|\mathbb{D}^{\mathcal{G}-I}u\right|^2 J - \int_\Omega\left(\mathbb{D}^{\mathcal{G}-I}u : \mathbb{D}u\right)J + \frac{1}{2}\int_\Omega|\mathbb{D}u|^2 J - 1\right). \end{aligned}$$

7.4. More commutators. In this section we record the commutators arising when differentiating the problem. We record them in a form readily amenable to estimates by writing them as commutators between partial derivatives and linear operators with multilinear dependence on parameters which we control, namely Φ, $\mathcal{G}$, J, and $\nu^{\mathcal{G}}_{\partial\Omega}$.

Lemma 7.6 (Computation of the commutators in multilinear form). *Suppose that (u, p, η) solves (2.2a)–(2.2e). Then, for each $\partial^\alpha \in \left\{\partial_t, \overline{\nabla}, \overline{\nabla}^2\right\}$, $(\partial^\alpha u, \partial^\alpha p, \partial^\alpha\eta)$ satisfies*

$$\begin{cases} D_t^{u,\mathcal{G}}v + \nabla^{\mathcal{G}}\cdot T^{\mathcal{G}} = C^{1,\alpha} & \text{in } \Omega, \\ \nabla^{\mathcal{G}}\cdot v = C^{2,\alpha} & \text{in } \Omega, \\ \left(\left(\delta^2_\eta\mathcal{W}\right)\zeta + g\zeta\right)\nu^{\mathcal{G}}_{\partial\Omega} - T^{\mathcal{G}}\cdot\nu^{\mathcal{G}}_{\partial\Omega} = C^{3,\alpha} & \text{on } \Sigma, \\ \partial_t\zeta - v\cdot\nu^{\mathcal{G}}_{\partial\Omega} = C^{4,\alpha} & \text{on } \Sigma, \text{ and} \\ v = 0 & \text{on } \Sigma_b \end{cases}$$

where

$$\begin{cases} C^{1,\alpha} = \Big(-\left[\partial^\alpha, \partial_t \Phi \cdot \nabla^{\mathcal{G}}\right] + \left[\partial^\alpha, u \cdot \nabla^{\mathcal{G}}\right] \Big), \\ \quad -\Big(\left[\partial^\alpha, \left(\nabla^{\mathcal{G}} \cdot \mathcal{G}^T\right) \cdot \nabla\right] + \left[\partial^\alpha, \left(\mathcal{G}^T \cdot \mathcal{G}\right) : \nabla^2\right] \Big) + \left[\nabla^{\mathcal{G}}, \partial^\alpha\right] p \\ C^{2,\alpha} = \left[\nabla^{\mathcal{G}} \cdot, \partial^\alpha\right] u, \\ C^{3,\alpha} = \Big(\left[\nu^{\mathcal{G}}_{\partial\Omega} \cdot \mathbb{D}^{\mathcal{G}}, \partial^\alpha\right] u - \left[\nu^{\mathcal{G}}_{\partial\Omega}, \partial^\alpha\right] p \Big) + g \left[\nu^{\mathcal{G}}_{\partial\Omega}, \partial^\alpha\right] \eta + \mathcal{C}^{\mathcal{W},\alpha}(\eta), \text{ and} \\ C^{4,\alpha} = -\left[\nu^{\mathcal{G}}_{\partial\Omega} \cdot, \partial^\alpha\right] u. \end{cases}$$

Proof. Upon applying ∂^α to (2.2a), we find that

$$\begin{aligned} C^{1,\alpha} &= \left[D_t^{u,\mathcal{G}} - \Delta^{\mathcal{G}}, \partial^\alpha\right] u + \left[\nabla^{\mathcal{G}}, \partial^\alpha\right] p \\ &= \left[\partial^\alpha, \left(\partial_t - \partial_t \Phi \cdot \nabla^{\mathcal{G}}\right) + u \cdot \nabla^{\mathcal{G}}\right] - \left[\partial^\alpha, \left(\nabla^{\mathcal{G}} \cdot\right) \circ \left(\nabla^{\mathcal{G}}\right)\right] \\ &= -\left[\partial^\alpha, \partial_t \Phi \cdot \nabla^{\mathcal{G}}\right] + \left[\partial^\alpha, u \cdot \nabla^{\mathcal{G}}\right] - \left[\partial^\alpha, \left(\nabla^{\mathcal{G}} \cdot\right) \circ \left(\nabla^{\mathcal{G}}\right)\right] \\ &= -\left[\partial^\alpha, \partial_t \Phi \cdot \nabla^{\mathcal{G}}\right] + \left[\partial^\alpha, u \cdot \nabla^{\mathcal{G}}\right] - \left[\partial^\alpha, \left(\nabla^{\mathcal{G}} \cdot \mathcal{G}^T\right) \cdot \nabla\right] - \left[\partial^\alpha, \left(\mathcal{G}^T \cdot \mathcal{G}\right) : \nabla^2\right]. \end{aligned}$$

The other commutators are computed by similarly differentiating (2.2b)–(2.2e) □

Remark 7.7 (Explicit form of the commutators). Since the commutators above are written in terms of linear operators with multilinear dependence on parameters, we may use Proposition 8.10 to expand them into pieces that may be estimated using the strategy described in Proposition 8.11. Indeed: (where for the sake of readability we suppress the conditions $\beta + \sum \gamma_i = \alpha$, $\beta < \alpha$, from Proposition 8.10, in the summations below)

$$\begin{cases} \left[\partial^\alpha, v \cdot \nabla^{\mathcal{G}}\right] = \sum \Big(\left(\partial^{\gamma_1} v\right) \cdot \nabla^{\partial^{\gamma_2} \mathcal{G}} \Big) \circ \partial^\beta, & \text{where } v = -\partial_t \Phi, u, \nabla^{\mathcal{G}} \cdot \mathcal{G}^T, \\ \left[\partial^\alpha, M : \nabla^2\right] = \sum \Big(\left(\partial^\gamma M\right) : \nabla^2 \Big) \circ \partial^\beta, & \text{where } M = \mathcal{G}^T \cdot \mathcal{G}, \\ \left[\partial^\alpha, \nabla^{\mathcal{G}}\right] = \sum \nabla^{\partial^\gamma \mathcal{G}} \circ \partial^\beta, & \\ \left[\partial^\alpha, \nu^{\mathcal{G}}_{\partial\Omega} \cdot \mathbb{D}^{\mathcal{G}}\right] = \sum \Big(\partial^{\gamma_1} \left(\nu^{\mathcal{G}}_{\partial\Omega}\right) \cdot \mathbb{D}^{\nabla^{\partial^{\gamma_2} \mathcal{G}}} \Big) \circ \partial^\beta, \text{ and} & \\ \left[\partial^\alpha, \nu^{\mathcal{G}}_{\partial\Omega}\right] = \sum \partial^\gamma \left(\nu^{\mathcal{G}}_{\partial\Omega}\right) \circ \partial^\beta. & \end{cases}$$

7.5. Computing the variations of the surface energy. We record in this section a more explicit expression for the first variation of the surface energy. This is useful when performing some critical estimates where more compact expressions for the first variation are not sufficient to close the estimates.

Lemma 7.8. *Let* $\mathcal{W}(\eta) := \int_{\mathbb{T}^2} f(\mathcal{J}\eta)$ *where we write* $f = f(w, M)$. *Then the first variation of the surface energy can be written as*

$$\begin{aligned} \delta \mathcal{W}(\eta) = {} & \nabla^2_{M,M} f\left(\nabla\eta, \nabla^2\eta\right) \bullet \nabla^4 \eta - \nabla^2_{w,w} f\left(\nabla\eta, \nabla^2\eta\right) \bullet \nabla^2 \eta + \nabla^3_{M,M,M} f\left(\nabla\eta, \nabla^2\eta\right) \bullet \left(\nabla^3\eta \otimes \nabla^3\eta\right) \\ & + 2\nabla^3_{M,M,w} f\left(\nabla\eta, \nabla^2\eta\right) \bullet \left(\nabla^3\eta \otimes \nabla^2\eta\right) s + \nabla^3_{w,M,w} f\left(\nabla\eta, \nabla^2\eta\right) \bullet \left(\nabla^2\eta \otimes \nabla^2\eta\right). \end{aligned}$$

The second variation at the equilibrium is given by

$$\left(\delta_0^2 \mathcal{W}\right) \phi = \nabla^2_{M,M} f(0,0) \bullet \nabla^4 \phi - \nabla^2_{w,w} f(0,0) \bullet \nabla^2 \phi.$$

7.6. Estimates of the variations of the surface energy. In this section we obtain estimates on the variations of the surface energy, obtaining estimates on $\delta\mathcal{W}$ (Lemma 7.9), $\delta^2_\eta \mathcal{W}$ (Lemma 7.10), and $\delta^3_\eta \mathcal{W}$ (Lemma 7.11), as well as estimates on auxiliary functions derived from f by Taylor expanding f about the equilibrium, i.e. about 0 (Lemma 7.12 and Corollary 7.13).

Lemma 7.9 (Smallness of the first variation). *The following hold.*

(1) *For all* $s > -1$ *there exists* $C > 0$ *such that for every* $\eta : \mathbb{T}^2 \to \mathbb{R}$ *sufficiently regular*

$$\|\delta\mathcal{W}(\eta)\|_{H^s(\mathbb{T}^2)} \leqslant C \|\mathfrak{h}(\mathcal{J}\eta)\|_{H^{s+2}(\mathbb{T}^2)} \|\eta\|_{H^{s+4}(\mathbb{T}^2)}$$

for $\mathfrak{h}(z) := \int_0^1 \nabla^2 f(tz)\,dt$*, where* $z = (w, M) \in \mathbb{R}^n \times \mathbb{R}^{n\times n}$.

(2) *In the small energy regime, for all* $s \in \left[0, \frac{1}{2}\right]$ *and for every* $\eta : \mathbb{T}^2 \to \mathbb{R}$ *sufficiently regular,*

$$||\delta\mathcal{W}(\eta)||_{H^s(\mathbb{T}^2)} \lesssim \sqrt{\mathcal{E}}.$$

Proof. The key observation is that we can rewrite $\delta\mathcal{W}$ in a more amenable way using the fundamental theorem of calculus. So let $s > -1$ and observe that

$$\begin{aligned} ||\delta\mathcal{W}(\eta)||_{H^s(\mathbb{T}^2)} &= ||\mathcal{J}^*(\nabla f(\mathcal{J}\eta))||_{H^s(\mathbb{T}^2)} & &= \left|\left|\mathcal{J}^*\left(\int_0^1 \nabla^2 f(t\mathcal{J}\eta)\,dt \bullet \mathcal{J}\eta\right)\right|\right|_{H^s(\mathbb{T}^2)} \\ &= ||\mathcal{J}^*(\mathfrak{h}(\mathcal{J}\eta)\bullet\mathcal{J}\eta)||_{H^s(\mathbb{T}^2)} & &\lesssim ||\mathfrak{h}(\mathcal{J}\eta)\bullet\mathcal{J}\eta||_{H^{s+2}(\mathbb{T}^2)} \\ &\lesssim ||\mathfrak{h}(\mathcal{J}\eta)||_{H^{s+2}(\mathbb{T}^2)}||\eta||_{H^{s+4}(\mathbb{T}^2)} \end{aligned}$$

where in the last step we have used that $s + 2 > 1$ since $s > -1$.

Next note that in the small energy regime we may use Corollary 7.13 to obtain, for any $s \in \left[0, \frac{1}{2}\right]$,

$$||\delta\mathcal{W}(\eta)||_{H^s(\mathbb{T}^2)} \lesssim ||\mathfrak{h}(\mathcal{J}\eta)||_{H^{s+2}(\mathbb{T}^2)}||\eta||_{H^{s+4}(\mathbb{T}^2)} \lesssim ||\eta||_{H^{9/2}(\mathbb{T}^2)} \lesssim \sqrt{\mathcal{E}}.$$

□

Next we consider the second variation.

Lemma 7.10 (Boundedness of the second variation of the surface energy). *Let* $s_0 > 3$ *and recall the constants* $C_f^{(k)}$ *defined in Definition 4.1. If* $\eta \in H^{s_0}(\mathbb{T}^2)$*, then for every* $s \in [2, s_0 - 1)$ *and every* $s \in (3, s_0]$*, there exists a constant* $0 < C = C\left(||\eta||_{H^{s_0}(\mathbb{T}^2)}, C_f^{(\lfloor s_0\rfloor+1)}\right)$ *such that*

$$\delta_\eta^2\mathcal{W} \in \mathcal{L}\left(H^s\left(\mathbb{T}^2\right);\ H^{s-4}\left(\mathbb{T}^2\right)\right) \quad \textit{with} \quad \left|\left|\delta_\eta^2\mathcal{W}\right|\right|_{\mathcal{L}(H^s(\mathbb{T}^2);\,H^{s-4}(\mathbb{T}^2))} \lesssim C,$$

i.e. past a certain regularity threshold for η*, we obtain that* $\delta_\eta^2\mathcal{W}$ *is a differential operator of order 4, as expected.*

Proof. Let $\eta \in H^{s_0}(\mathbb{T}^2)$ and let $\phi \in H^s(\mathbb{T}^2)$ for some $s \in [2, s_0 - 1)$. If $s \in [2, s_0 - 1)$, then we may use Propositions 8.12 and 8.15 to see that

$$\begin{aligned} \left|\left|\left(\delta_\eta^2\mathcal{W}\right)\phi\right|\right|_{H^{s-4}(\mathbb{T}^2)} &= \left|\left|J^*\left(\nabla^2 f(\mathcal{J}\eta)\bullet J\phi\right)\right|\right|_{H^{s-4}(\mathbb{T}^2)} \lesssim \left|\left|\nabla^2 f(\mathcal{J}\eta)\bullet J\phi\right|\right|_{H^{s-2}(\mathbb{T}^2)} \\ &\lesssim \left|\left|\nabla^2 f(\mathcal{J}\eta)\right|\right|_{H^{s_0-2}(\mathbb{T}^2)}||J\phi||_{H^{s-2}(\mathbb{T}^2)} \\ &\lesssim \left(C_f^{(2)} + C_f^{(\lfloor s_0\rfloor)}\left(||\mathcal{J}\eta||_{H^{s_0-2}(\mathbb{T}^2)} + ||\mathcal{J}\eta||_{H^{s_0-2}(\mathbb{T}^2)}^{\lceil s_0\rceil-2}\right)\right)||\phi||_{H^s(\mathbb{T}^2)} \\ &\lesssim \underbrace{\left(C_f^{(\lfloor s_0\rfloor+1)}\left(1 + ||\eta||_{H^{s_0}(\mathbb{T}^2)} + ||\eta||_{H^{s_0}(\mathbb{T}^2)}^{\lfloor s_0\rfloor-2}\right)\right)}_{=:C}||\phi||_{H^s(\mathbb{T}^2)}. \end{aligned}$$

If $s \in (3, s_0]$, we proceed with the same estimates as above, but replacing s_0 with s. In particular, the key difference is that now, since $s > 3$, $H^{s-2}(\mathbb{T}^2)$ is an algebra. □

Next we consider the third variation.

Lemma 7.11 (Boundedness of the third variation of the surface energy). *Let* $s_0 > 4$ *and recall that the constants* $C_f^{(k)}$ *are given in Definition 4.1. If* $\eta \in H^{s_0}(\mathbb{T}^2)$*, then for very* $s \in (3, s_0 - 1)$ *and every* $p, q \geqslant 0$ *such that* $p + q > s + 3$ *there exists a constant* $0 < C = C\left(||\eta||_{H^{s_0}(\mathbb{T}^2)}, C_f^{(\lfloor s_0\rfloor+2)}\right)$ *such that*

$$\delta_\eta^3\mathcal{W} \in \mathcal{L}_2\left(H^p \times H^q;\ H^{s-4}\right) \quad \textit{with} \quad \left|\left|\delta_\eta^3\mathcal{W}\right|\right|_{\mathcal{L}_2(H^p\times H^q;\,H^{s-4})} \lesssim C$$

where for any normed vector spaces V, W, X*,* $\mathcal{L}_2(V \times W; X)$ *denotes the set of continuous bilinear forms on* $V \times W$ *mapping into* X.

Proof. Let $\eta \in H^{s_0}\left(\mathbb{T}^2\right)$ and let $\phi \in H^p\left(\mathbb{T}^2\right)$ and $\psi \in H^q\left(\mathbb{T}^2\right)$ for some $s, p, q \geqslant 0$ such that $s \in (2, s_0 - 1)$ and $p + q > s + 3$. Then, using Propositions 8.12 and 8.15 we obtain that

$$\begin{aligned}
\left|\left|\left(\delta_\eta^3 \mathcal{W}\right)(J\phi, J\psi)\right|\right|_{H^{s-4}(\mathbb{T}^2)} &= \left|\left|J^*\left(\nabla^3 f\left(\mathcal{J}\eta\right) \bullet (J\phi \otimes J\psi)\right)\right|\right|_{H^{s-4}(\mathbb{T}^2)} \\
&\lesssim \left|\left|\nabla^3 f\left(\mathcal{J}\eta\right) \bullet (J\phi \otimes J\psi)\right|\right|_{H^{s-2}(\mathbb{T}^2)} \quad \lesssim \left|\left|\nabla^3 f\left(\mathcal{J}\eta\right)\right|\right|_{H^{s_0-2}(\mathbb{T}^2)} ||J\phi \otimes J\psi||_{H^{s-2}(\mathbb{T}^2)} \\
&\lesssim \left(C_f^{(3)} + C_f^{(\lfloor s_0 \rfloor + 1)}\left(||\mathcal{J}\eta||_{H^{s_0-2}(\mathbb{T}^2)} + ||\mathcal{J}\eta||_{H^{s_0-2}(\mathbb{T}^2)}^{\lceil s_0 \rceil - 2}\right)\right) ||\phi||_{H^{p-2}(\mathbb{T}^2)} ||\psi||_{H^{q-2}(\mathbb{T}^2)} \\
&\lesssim \underbrace{\left(C_f^{(\lfloor s_0 \rfloor + 2)}\left(1 + ||\eta||_{H^{s_0}(\mathbb{T}^2)} + ||\eta||_{H^{s_0}(\mathbb{T}^2)}^{\lfloor s_0 \rfloor - 2}\right)\right)}_{=:C} ||\phi||_{H^p(\mathbb{T}^2)} ||\psi||_{H^q(\mathbb{T}^2)}.
\end{aligned}$$

□

Next we control terms related to Taylor expansions of the surface energy.

Lemma 7.12 (Estimates for the auxiliary functions from the Taylor expansions of the variations of the surface energy). *For any $s \geqslant 2$, $f : \mathbb{R}^2 \times \mathbb{R}^{2\times 2} \to \mathbb{R}$, and $\eta : \mathbb{T}^2 \to \mathbb{R}$ we have that*

$$||r_k[f, 0](\mathcal{J}\eta)||_{H^s(\mathbb{T}^2)} \lesssim C_f^{(k+1)} + C_f^{(\lfloor s \rfloor + k + 2)}\left(||\eta||_{H^{s+2}(\mathbb{T}^2)} + ||\eta||_{H^{s+2}(\mathbb{T}^2)}^{\lceil s \rceil}\right),$$

where r_k is defined in Proposition 8.24 and $C_f^{(k)}$ is defined in Definition 4.1. Moreover, in the small energy regime (see Definition 4.2), if $s \in \left[2, \frac{5}{2}\right]$, then

$$||r_k[f, 0](\mathcal{J}\eta)||_{H^s(\mathbb{T}^2)} \lesssim 1.$$

Proof. The result then follows from post-composition estimates in Sobolev spaces (see Proposition 8.15) and from the observation that

$$\left(\partial^\alpha\left(r_k[f, 0]\right)\right)(z) = \int_0^1 (1-t)^k t^{|\alpha|} \partial^\alpha \nabla^{k+1} f(tz)\, dt$$

such that, for any $R \geqslant 0$ and any $l \in \mathbb{N}$,

$$||r_k[f, 0]||_{C^{l,1}\left(\overline{B(0,R)}\right)} \leqslant \left|\left|\nabla^{k+1} f\right|\right|_{C^{l,1}\left(\overline{B(0,R)}\right)}.$$

Therefore, since $s \geqslant 2$, we obtain from Proposition 8.15 that

$$\begin{aligned}
||r_k[f, 0](\mathcal{J}\eta)||_{H^s(\mathbb{T}^2)} &\lesssim ||r_k[f, 0]||_{L^\infty\left(\overline{B\left(0, ||\mathcal{J}\eta||_{L^\infty(\mathbb{T}^2)}\right)}\right)} \\
&\quad + ||r_k[f, 0]||_{C^{\lfloor s \rfloor, 1}\left(\overline{B\left(0, ||\mathcal{J}\eta||_{L^\infty(\mathbb{T}^2)}\right)}\right)} \left(||\mathcal{J}\eta||_{H^s(\mathbb{T}^2)} + ||\mathcal{J}\eta||_{H^s(\mathbb{T}^2)}^{\lceil s \rceil}\right) \\
&\lesssim C_f^{(k+1)} + C_f^{(\lfloor s \rfloor + k + 2)}\left(||\eta||_{H^{s+2}(\mathbb{T}^2)} + ||\eta||_{H^{s+2}(\mathbb{T}^2)}^{\lceil s \rceil}\right).
\end{aligned}$$

In particular, in the small energy regime where $||\eta||_{H^{9/2}(\mathbb{T}^2)} \lesssim \sqrt{\mathcal{E}} \leqslant \sqrt{\delta_0}$, if $s \in \left[2, \frac{9}{2}\right]$ then

$$||\eta||_{H^{3+}(\mathbb{T}^2)} \lesssim 1 \quad \text{and} \quad ||\eta||_{H^{s+2}(\mathbb{T}^2)} \lesssim 1,$$

and hence

$$||r_k[f, 0](\mathcal{J}\eta)||_{H^s(\mathbb{T}^2)} \lesssim \left|\left|\nabla^{k+1} f\right|\right|_{L^\infty\left(\overline{B(0,C)}\right)} + \left|\left|\nabla^{k+1} f\right|\right|_{C^{k,1}\left(\overline{B(0,C)}\right)} \lesssim 1.$$

□

Lemma 7.12 has the following immediate corollary.

Corollary 7.13. *If for $z = (w, M) \in \mathbb{R}^n \times \mathbb{R}^{n\times n}$ we set*

$$\mathfrak{h}(z) := \int_0^1 \nabla^2 f(tz)\, dt = r_0[\nabla f, 0](z) \;\text{ and }\; \mathfrak{q}(z) := \frac{1}{2}\int_0^1 (1-t)\nabla^3 f(tz)\, dt = r_1[\nabla f, 0](z),$$

then for any $s \geqslant 2$ we have the bounds

$$\begin{cases} ||\mathfrak{h}(\mathcal{J}\eta)||_{H^s(\mathbb{T}^2)} \lesssim C_f^{(2)} + C_f^{(\lfloor s \rfloor + 3)}\left(||\eta||_{H^{s+2}(\mathbb{T}^2)} + ||\eta||_{H^{s+2}(\mathbb{T}^2)}^{\lceil s \rceil}\right) \text{ and} \\ ||\mathfrak{q}(\mathcal{J}\eta)||_{H^s(\mathbb{T}^2)} \lesssim C_f^{(3)} + C_f^{(\lfloor s \rfloor + 4)}\left(||\eta||_{H^{s+2}(\mathbb{T}^2)} + ||\eta||_{H^{s+2}(\mathbb{T}^2)}^{\lceil s \rceil},\right) \end{cases}$$

where the constants $C_f^{(k)}$ are defined in Definition 4.1. In particular, in the small energy regime of Definition 4.2, if $s \in \left[2, \frac{5}{2}\right]$ then

$$\left|\left|\mathfrak{h}\left(\mathcal{J}\eta\right)\right|\right|_{H^s(\mathbb{T}^2)} \lesssim 1 \quad \text{and} \quad \left|\left|\mathfrak{q}\left(\mathcal{J}\eta\right)\right|\right|_{H^s(\mathbb{T}^2)} \lesssim 1.$$

8. Appendix: generic tools

In this second part of the appendix we record generic tools, i.e. results that are employed throughout this chapter but whose applicability is not reduced to the problem in this chapter. In particular, these results are either well-known or slight modifications of standard results. They are therefore recorded here so that they may be precisely stated as reference for when they are invoked elsewhere in this chapter.

8.1. Variations/derivatives of the surface energy. In this section we record various expressions for variations of, and functionals associated with, the surface energy. Recall from Section 4.4.1 that the surface energy associated with a surface given as the graph of η is

$$\mathcal{W}\left(\eta\right) = \int_{\mathbb{T}^2} f\left(\mathcal{J}\eta\right)$$

where the jet $\mathcal{J}\eta$ is given by $\mathcal{J}\eta = \left(\nabla\eta, \nabla^2\eta\right)$. Similarly, the definitions of δ_ϕ, $\delta\mathcal{W}, \delta_\eta^2\mathcal{W}, \dots,$ $D\mathcal{W}, D^2\mathcal{W}$, etc are in Section 4.4.1. We begin by giving the form of variations of $\mathcal{W}$.

Lemma 8.1 (Various representations of the variations/derivatives of the surface energy). *For any sufficiently regular functions $\eta, \phi, \psi, \phi_i : \mathbb{T}^2 \to \mathbb{R}$, $i = 1, \dots, k$, the following hold:*

(1)
$$\left\langle D\mathcal{W}\left(\eta\right), \phi\right\rangle = \int_{\mathbb{T}^2} \delta\mathcal{W}\left(\eta\right)\phi = \int_{\mathbb{T}^2} \nabla f\left(\mathcal{J}\eta\right)\cdot\mathcal{J}\phi,$$

(2)
$$\left\langle D^2\mathcal{W}\left(\eta\right), \left(\phi, \psi\right)\right\rangle = \int_{\mathbb{T}^2} \left(\left(\delta_\eta^2\mathcal{W}\right)\phi\right)\psi = \int_{\mathbb{T}^2} \nabla^2 f\left(\mathcal{J}\eta\right) \bullet \left(\mathcal{J}\phi \otimes \mathcal{J}\psi\right),$$

(3)
$$\begin{aligned}\left\langle D^k\mathcal{W}\left(\eta\right), \left(\phi_1, \phi_2, \dots, \phi_{k-1}, \phi_k\right)\right\rangle &= \int_{\mathbb{T}^2} \Big(\left(\delta_\eta^k\mathcal{W}\right)\left(\phi_1, \phi_2, \dots, \phi_{k-1}\right)\Big)\phi_k \\ &= \int_{\mathbb{T}^2} \nabla^k f\left(\mathcal{J}\eta\right) \bullet \left(\mathcal{J}\phi_1 \otimes \mathcal{J}\phi_2 \otimes \cdots \otimes \mathcal{J}\phi_{k-1} \otimes \mathcal{J}\phi_k\right).\end{aligned}$$

Remark 8.2. We record here formulae for partial derivatives of the first and second variation. For α, β multi-indices such that $|\alpha| = |\beta| = 1$, we have

$$\partial^\alpha\Big(\delta\mathcal{W}\left(\eta\right)\Big) = \delta_\eta^2\mathcal{W}\left(\partial^\alpha\eta\right) \quad \text{and} \quad \partial^{\alpha+\beta}\Big(\delta\mathcal{W}\left(\eta\right)\Big) = \delta_\eta^3\mathcal{W}\left(\partial^\alpha\eta, \partial^\beta\eta\right) + \delta_\eta^2\mathcal{W}\left(\partial^{\alpha+\beta}\eta\right).$$

We now record a lemma that comes in handy when computing second variations.

Lemma 8.3 (Computing the second variation). *For any $\eta, \phi : \mathbb{T}^2 \to \mathbb{R}$ sufficiently regular,*

$$\delta_\phi\left(\delta\mathcal{W}\left(\eta\right)\right) = \left(\delta_\eta^2\mathcal{W}\right)\phi.$$

Proof. For any $\psi \in C_c^\infty\left(\mathbb{T}^2\right)$, we compute

$$\begin{aligned}\int_{\mathbb{T}^2} \delta_\phi\left(\delta\mathcal{W}\left(\eta\right)\right)\psi = \delta_\phi \int_{\mathbb{T}^2} \delta\mathcal{W}\left(\eta\right)\psi = \delta_\phi\left\langle D\mathcal{W}\left(\eta\right), \psi\right\rangle = \delta_\phi\delta_\psi\mathcal{W}\left(\eta\right) \\ = \left\langle D^2\mathcal{W}\left(\eta\right), \left(\phi, \psi\right)\right\rangle = \int_{\mathbb{T}^2} \left(\left(\delta_\eta^2\mathcal{W}\right)\phi\right)\psi.\end{aligned}$$

□

We now record a computation telling us how the quadratic approximation to the surface energy behaves when differentiated in time, which comes in handy when estimating the commutators.

Proposition 8.4. *For any $\eta, \zeta : \mathbb{T}^2 \to \mathbb{R}$ sufficiently regular,*

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(\mathcal{Q}_\eta\left(\zeta\right)\right) = \mathcal{Q}_{\dot\eta}\left(\zeta\right) + \left\langle D^2\mathcal{W}\left(\eta\right), \left(\zeta, \partial_t\zeta\right)\right\rangle = \mathcal{Q}_{\dot\eta}\left(\zeta\right) + \int_{\mathcal{U}} \Big(\left(\delta_\eta^2\mathcal{W}\right)\zeta\Big)\partial_t\zeta.$$

PROOF. This result is nothing more than the product rule transcribed into our notation. This is apparent when rewriting the formula above as:

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{1}{2}\int_{\mathbb{T}^2}\nabla^2 f\left(\mathcal{J}\eta\right)\otimes\left(\mathcal{J}\zeta\otimes\mathcal{J}\zeta\right)\right)=\frac{1}{2}\int_{\mathbb{T}^2}\partial_t\left(\nabla^2 f\left(\mathcal{J}\eta\right)\right)\otimes\left(\mathcal{J}\zeta\otimes\mathcal{J}\zeta\right)+\int_{\mathbb{T}^2}\nabla^2 f\left(\mathcal{J}\eta\right)\otimes\left(\mathcal{J}\zeta\otimes\mathcal{J}\partial_t\zeta\right).$$

□

8.2. Harmonic extension. In this section we record the standard definition and estimates of the harmonic extension of a function from $\mathbb{T}^2$ to $\mathbb{T}^2\times(-\infty,0)$. Although the extension is defined in this large set, we will typically only need in on $\mathbb{T}^2\times(-b,0)$.

DEFINITION 8.5 (Harmonic extension). We define the following.

(1) For any $f\in L^1(\mathbb{T}^2)$, define $\operatorname{ext} f:\mathbb{T}^2\times(-\infty,0)\to\mathbb{R}$ by, for every $x\in\mathbb{T}^2\times(-\infty,0)$,
$$\left(\operatorname{ext} f\right)(x):=\sum_{\bar{k}\in\mathbb{Z}^2}\left(\hat{f}\left(\bar{k}\right)e^{2\pi|\bar{k}|x_3}\right)e^{2\pi i\bar{k}\cdot\bar{x}},$$
where $\hat{\cdot}$ denotes the Fourier transform and where recall that $x=(\bar{x},x_3)$.

(2) For any $f:[0,T)\times\mathbb{T}^2\to\mathbb{R}$, define $\operatorname{ext} f:[0,T)\times\mathbb{T}^2\times(-\infty,0)\to\mathbb{R}$ by $(\operatorname{ext} f)(t,\cdot):=\operatorname{ext}\left(f(t,\cdot)\right)$.

REMARK 8.6. Recall that $\operatorname{ext} f$ as defined above is called the *harmonic* extension of f because it solves

$$\begin{cases}-\Delta\operatorname{ext} f=0 & \text{in } \mathbb{T}^2\times(-b,0)\,,\\ \operatorname{ext} f=f & \text{on } \{x_3=0\}\,.\end{cases}$$

Next we record some identities related to the harmonic extension.

LEMMA 8.7 (Identities for the derivatives of the harmonic extension). *For any $f:\mathbb{T}^2\to\mathbb{R}$ sufficiently regular,*

$$\partial_3\operatorname{ext} f=\operatorname{ext}\sqrt{-\Delta}f\quad\textit{and}\quad\overline{\nabla}\operatorname{ext} f=\operatorname{ext}\overline{\nabla}f,$$

where $\left(\sqrt{-\Delta}f\right)^{\wedge}\left(\bar{k}\right):=2\pi|\bar{k}|$ for all $\bar{k}\in\mathbb{Z}^2$.

PROOF. These results follow directly from short computations on the Fourier side. □

Next we record some useful estimates, starting with L^2 ones.

LEMMA 8.8 (L^2 bound on the harmonic extension). *For any $f:\mathbb{T}^2\to\mathbb{R}$ sufficiently regular,*

$$||\operatorname{ext} f||_{L^2(\Omega)}\leqslant\frac{1}{2\sqrt{\pi}}||f||_{\dot{H}^{-1/2}(\mathbb{T}^2)}$$

where $\Omega=\mathbb{T}^2\times(-b,0)$.

PROOF. To obtain this inequality we proceed as follows: employ Parseval's identity on the horizontal slices, then apply Tonelli's theorem so that we may integrate exactly along the vertical direction, and finally note that $1-e^{-4\pi b|\cdot|}\leqslant 1$. □

The L^2 bounds coupled with the identities for the derivatives of the harmonic extension lead to H^s bounds.

COROLLARY 8.9 (H^s bounds on the harmonic extension). *Recall that $\Omega=\mathbb{T}^2\times(-b,0)$. For any $s\geqslant 0$, there exists $C_s>0$ such that for any $f\in H^{s-1/2}\left(\mathbb{T}^2\right)$,*

$$||\operatorname{ext} f||_{H^s(\Omega)}\lesssim||f||_{H^{s-1/2}(\mathbb{T}^2)},$$

PROOF. This result follows from Lemmas 8.7 and 8.8 when s is an integer and a standard interpolation argument otherwise. □

8.3. Commutators with linear operators with multilinear dependence on their parameters. In this section we record how to compute commutators between partial derivatives and linear operators with multilinear dependence on their parameters.

PROPOSITION 8.10. *Suppose that L is a linear differential operator acting on functions $\eta : [0,T)\times\mathbb{T}^2 \to \mathbb{R}$ that can be written as $L = \hat{L}(\pi_1,\dots,\pi_k)$ for some parameters $\pi_1,\dots,\pi_k : [0,T)\times\mathbb{T}^2 \to \mathbb{R}$, where $\hat{L}$ is multilinear. Then, for any multi-index $\alpha = (\alpha_0,\bar{\alpha}) \in \mathbb{N}^3$ such that $\partial^\alpha = \partial_t^{\alpha_0}\partial_{\bar{x}}^{\bar{\alpha}}$, we have*

$$[\partial^\alpha, L] = \sum_{\substack{\beta+\sum_{i=1}^k \gamma_i=\alpha \\ \beta<\alpha}} \hat{L}\left(\partial^{\gamma_1}\pi_1,\dots,\partial^{\gamma_k}\pi_k\right)\circ\partial^\beta.$$

PROOF. For any $\eta : \mathbb{T}^2 \to \mathbb{R}$, we compute directly, using the multilinearity of $\hat{L}$ in $(\star)$ below:

$$[\partial^\alpha, L]\,\eta = \partial^\alpha\Big(\hat{L}\left(\pi_1,\dots,\pi_k\right)\eta\Big) - L\left(\partial^\alpha\eta\right) \overset{(\star)}{=} \left(\sum_{\beta+\sum_{i=1}^k \gamma_i=\alpha} \hat{L}\left(\partial^{\gamma_1}\pi_1,\dots,\partial^{\gamma_k}\pi_k\right)\partial^\beta\eta\right) - L\left(\partial^\alpha\eta\right)$$

$$= \left(\sum_{\substack{\beta+\sum_{i=1}^k \gamma_i=\alpha \\ \beta\neq\alpha}} \hat{L}\left(\partial^{\gamma_1}\pi_1,\dots,\partial^{\gamma_k}\pi_k\right)\circ\partial^\beta\right)\eta$$

Since $\beta \neq \alpha$ above is equivalent to $\beta < \alpha$ in this context (since necessarily $\beta \leqslant \alpha$), we obtain the result desired. ☐

8.4. General recipe for controlling interactions with Sobolev norms. We record here a general recipe for controlling interactions with Sobolev norms by combining the Hölder inequality and appropriate Sobolev embeddings.

PROPOSITION 8.11. *Let $n,k \in \mathbb{N}$ and let $s_1,\dots,s_k \geqslant 0$ be such that either*

$$\text{(i)}\ \sum_{i=1}^k \min\left(s_i,\frac{n}{2}\right) > n\left(\frac{k}{2}-1\right)\ \text{ or (ii) }\ \sum_{i=1}^k \min\left(s_i,\frac{n}{2}\right) \geqslant n\left(\frac{k}{2}-1\right)\ \text{ and } s_i \neq \frac{n}{2} \text{ for all } i$$

holds. Then there exists $C > 0$ such that for every $f_1 \in H^{s_1}(\mathbb{T}^n),\dots,f_k \in H^{s_k}(\mathbb{T}^n)$,

$$\left|\int_{\mathbb{T}^n} f_1\dots f_k\right| \leqslant C||f_1||_{H^{s_1}(\mathbb{T}^n)}\dots||f_k||_{H^{s_k}(\mathbb{T}^n)}.$$

8.5. Product estimates in Sobolev spaces. In this section we record for which regularity indices s,t,u it holds that $H^s\cdot H^t \hookrightarrow H^u$. Using Fourier analysis, these results boil down to:

(1) The following pointwise bound on the Fourier side:

$$\langle\cdot\rangle^s\left|(fg)^s\right| \lesssim \langle\cdot\rangle^s|\hat{f}| * |\hat{g}| + |\hat{f}| * \langle\cdot\rangle^s|\hat{g}|$$

for $f,g : \mathbb{T}^n \to \mathbb{R}$, which follows from the elementary observation that $\langle k\rangle^2 \lesssim \langle k-l\rangle^2 + \langle l\rangle^2$ for all $k,l \in \mathbb{Z}^n$.

(2) Young's inequality for convolutions.

(3) Using Hölder's inequality on the Fourier side to show that

$$\left|\left|\langle\cdot\rangle^s\hat{f}\right|\right|_{l^p(\mathbb{Z}^n)} \lesssim ||f||_{H^{s+\alpha}(\mathbb{T}^n)}$$

for the appropriate values s, p and α.

PROPOSITION 8.12 (H^s is a Banach algebra when $s > s_*$). *Let $D = \mathbb{T}^2, \Omega$ and correspondingly let $s_* = 1, \frac{3}{2}$. If $s > s_*$, then*

$$H^s(D)\cdot H^s(D) \hookrightarrow H^s(D)$$

i.e. for every $s > s_*$ there exists $C > 0$ such that for every $f,g \in H^s(D)$, the product fg belongs to $H^s(D)$ and satisfies the estimate

$$||fg||_{H^s(D)} \leqslant C||f||_{H^s(D)}||g||_{H^s(D)}.$$

Proposition 8.13 ($H^{s+\alpha}$ is a continuous multiplier on H^s when $\alpha > s_*$). *Let $D = \mathbb{T}^2, \Omega$ and correspondingly let $s_* = 1, \frac{3}{2}$. For every $s \geqslant 0$, if $\alpha > s_*$, then*

$$H^{s+\alpha}(D) \cdot H^s(D) \hookrightarrow H^s(D)$$

i.e. for every such s and α there exists $C > 0$ such that for every $f \in H^{s+\alpha}(D)$ and $g \in H^s(D)$, the product fg belongs to $H^s(D)$ and satisfies the estimate

$$||fg||_{H^s(D)} \leqslant C||f||_{H^{s+\alpha}(D)}||g||_{H^s(D)}.$$

Proposition 8.14 (Borrowing regularity from both factors). *Let $D = \mathbb{T}^2, \Omega$ and correspondingly let $s_* = 1, \frac{3}{2}$. For every $s \geqslant 0$ and $\alpha, \beta > 0$, if $s + (\alpha + \beta) > s_*$, then*

$$H^{s+\alpha}(D) \cdot H^{s+\beta}(D) \hookrightarrow H^s(D)$$

i.e. for every such s, α, and β there exists $C > 0$ such that for every $f \in H^{s+\alpha}(D)$ and $g \in H^{s+\beta}(D)$, the product fg belongs to $H^s(D)$ and satisfies the estimate

$$||fg||_{H^s(D)} \leqslant C||f||_{H^{s+\alpha}(D)}||g||_{H^{s+\beta}(D)}$$

8.6. Post-composition estimates in Sobolev spaces. We record here conditions on s for H^s to be closed under post-composition by a sufficiently smooth function (also known as a Nemytskii operator, or as a superposition operator).

These post-composition estimates boils down to estimates of the multilinear terms involving derivatives of various orders which appear in the Faà di Bruno formula (i.e. the chain rule for higher-order derivatives). The key observation is that these terms can be written as derivatives of polynomials. Coupling this observation with the fact that H^s is an algebra for sufficiently large s (c.f. Proposition 8.12) thus yields the post-composition estimates.

Proposition 8.15. *Let $D = \mathbb{T}^2, \Omega$ and correspondingly let $s_* = 1, \frac{3}{2}$. Let $k \in \mathbb{N}$ and $\alpha \in [0, 1)$. If $k > s_*$ then for every $g \in H^{k+\alpha}(D; \mathbb{R})$ and for every $F \in C^{k,1}_{loc}(\mathbb{R}; \mathbb{R})$, $F \circ g \in H^{k+\alpha}(D; \mathbb{R})$ with*

$$\begin{aligned} ||F \circ g||_{H^{k+\alpha}(D)} &\lesssim ||F \circ g||_{L^2(D)} + ||F||_{C^{k,1}\left(\overline{B\left(0,||g||_\infty\right)}\right)} \left(||g||_{H^{k+\alpha}(D)} + ||g||^{k+\lceil\alpha\rceil}_{H^{k+\alpha}(D)} \right) \\ &\lesssim ||F||_{L^\infty\left(\overline{B\left(0,||g||_\infty\right)}\right)} + ||F||_{C^{k,1}\left(\overline{B\left(0,||g||_\infty\right)}\right)} \left(||g||_{H^{k+\alpha}(D)} + ||g||^{k+\lceil\alpha\rceil}_{H^{k+\alpha}(D)} \right) \end{aligned}$$

where $B(R) = (-R, R)$ and $\lceil x \rceil$ denotes the smallest integer greater than or equal to x.

8.7. Elliptic estimates for the Stokes problem. In this section we record estimates for the Stokes problem. We begin with the case of Dirichlet conditions.

Proposition 8.16 (Estimates for the Stokes problem with Dirichlet boundary condition). *Let $s \geq 0$, let $f \in H^s(\Omega)$, $g \in H^{s+1}(\Omega)$, and $h \in H^{s+3/2}(\partial\Omega)$ satisfy $\int_\Omega f = \int_{\partial\Omega} h \cdot \nu$, and let (u, p) solve*

$$\begin{cases} -\Delta u + \nabla p = f & \text{in } \Omega, \\ \nabla \cdot u = g & \text{in } \Omega, \text{ and} \\ u = h & \text{on } \partial\Omega. \end{cases}$$

Then

$$||u||_{H^{s+2}(\Omega)} + ||\nabla p||_{H^s(\Omega)} \lesssim ||f||_{H^s(\Omega)} + ||g||_{H^{s+1}(\Omega)} + ||h||_{H^{s+3/2}(\partial\Omega)}.$$

Next we consider the Stokes problem with different boundary conditions.

Proposition 8.17 (Estimates for the Stokes problem with mixed Dirichlet-Neumann boundary condition). *Let $s \geq 0$, let $f \in H^s(\Omega)$, $g \in H^{s+1}(\Omega)$, $h_1 \in H^{s+3/2}(\Sigma)$, and $h_2 \in H^{s+1/2}(\Sigma)$, and let (u, p) solve*

$$\begin{cases} -\Delta u + \nabla p = f & \text{in } \Omega, \\ \nabla \cdot u = g & \text{in } \Omega, \\ u \cdot e_3 = h_1 & \text{on } \Sigma, \\ (\mathbb{D}u \cdot e_3)_{tan} = h_2 & \text{on } \Sigma, \text{ and} \\ u = 0 & \text{on } \Sigma_b \end{cases}$$

where $v_{tan} := (I - e_3 \otimes e_3)\, v$*, i.e.* v_{tan} *is the* tangential *part of* v*. Then*

$$||u||_{H^{s+2}(\Omega)} + ||\nabla p||_{H^s(\Omega)} \lesssim ||f||_{H^s(\Omega)} + ||g||_{H^{s+1}(\Omega)} + ||h_1||_{H^{s+3/2}(\Sigma)} + ||h_2||_{H^{s+1/2}(\Sigma)}.$$

8.8. Dynamic boundary conditions. We now turn our attention to estimates related to the dynamic boundary condition (2.2c). We begin with a definition.

DEFINITION 8.18. Let L be a linear differential operator acting on functions $\eta : \mathbb{T}^n \to \mathbb{R}$ and let $k \in \mathbb{N}$. We say that L is a *strictly elliptic k-th order differential operator on functions of average zero* if there exists $C > 0$ such that $\hat{L} \geqslant C|\cdot|^k$.

Next we record elliptic estimates for such operators.

PROPOSITION 8.19. *Let* L *be a strictly elliptic k-th order differential operator on the n-torus. Then there exists* $C > 0$ *such that for every* $s \in \mathbb{R}$ *and every* $f \in H^s(\mathbb{T}^n)$*, if* η *solves* $L\eta = f$ *on* $\mathbb{T}^n$*, then*

$$||\eta||_{\dot{H}^{s+k}(\mathbb{T}^n)} \leqslant C||f||_{\dot{H}^s(\mathbb{T}^n)}.$$

PROOF. This result follows immediately from the assumption on L of strict ellipticity over functions of average zero:

$$||\eta||^2_{\dot{H}^{s+k}(\mathbb{T}^n)} = \left|\left||\cdot|^{2s}\hat{\eta}\right|\right|^2_{l^2(\mathbb{Z}^n)} \leqslant \frac{1}{C^2}\left|\left||\cdot|^{2(s-k)}\hat{L}\hat{\eta}\right|\right|^2_{l^2(\mathbb{Z}^n)} = \frac{1}{C^2}||f||^2_{\dot{H}^{s-k}(\mathbb{T}^n)}.$$

□

A byproduct of Proposition 8.19 is the following estimate, tailored to the dynamic boundary condition.

COROLLARY 8.20 (Estimates for the dynamic boundary condition). *Let* $g > 0$ *and* $f : \mathbb{R}^2 \times \mathbb{R}^{2\times 2} \to \mathbb{R}$*, write* $f = f(w, M)$ *for* $(w, M) \in \mathbb{R}^2 \times \mathbb{R}^{2\times 2}$*, and suppose that* (2.3) *holds. Then for every* $s \geqslant 0$ *there exists* $\tilde{C} > 0$ *such that for every* $f \in H^s(\mathbb{T}^2)$*, if* η *satisfies* $\int_{\mathbb{T}^2} \eta = 0$ *and solves* $\left(\delta_0^2 \mathcal{W} + g\right)\eta = f$ *on* $\mathbb{T}^2$*, then*

$$||\eta||_{H^{s+4}(\mathbb{T}^n)} \leqslant \tilde{C}||f||_{\dot{H}^s(\mathbb{T}^n)}.$$

PROOF. The assumption (2.3) tells us precisely that $\delta_0^2\mathcal{W} + g$ is a strictly elliptic fourth-order operator over functions of average zero. Proposition 8.19 thus yields the desired result, since on $\mathbb{Z}^n \setminus \{0\}$, $|\cdot|^s \asymp \langle\cdot\rangle^s$, and hence for functions of average zero $||\cdot||_{\dot{H}^s(\mathbb{T}^n)} \asymp ||\cdot||_{H^s(\mathbb{T}^n)}$. □

Next we consider Poincaré-type inequalities.

PROPOSITION 8.21 (Poincaré-type inequalities). *The following hold.*

(1) *There exists* $C^P > 0$ *such that for every* $\phi \in H^1(\Omega)$*,*

$$||\phi||_{H^1(\Omega)} \leq C^P \left(||\mathrm{tr}\,\phi||_{L^2(\Sigma)} + ||\nabla\phi||_{L^2(\Omega)},\right) \tag{8.1}$$

(2) *For every* $s \geq 0$*, there exists* $C^P_s > 0$ *such that for every* $\eta \in H^{s+1}(\mathbb{T}^n)$ *satisfying* $\int_{\mathbb{T}^n} \eta = 0$ *we have that*

$$||\eta||_{H^{s+1}} \leq C^P_s ||\nabla\eta||_{H^s}. \tag{8.2}$$

Korn's inequality, which we record now, is a sort of Poincaré-type inequality for the symmetrized gradient. See Lemma 2.7 in [Bea81] for a proof.

PROPOSITION 8.22 (Korn inequality). *There exist* $C_K > 0$ *such that for every* $\phi \in H^1(\Omega)$*, if* $\phi = 0$ *on* Σ_b*, then*

$$||\phi||_{H^1(\Omega)} \leq C_K||\mathbb{D}\phi||_{L^2(\Omega)}.$$

8.9. Linear algebra. In this section we record some simple facts from linear algebra.

LEMMA 8.23 (Determinant of a rank 1 perturbation of the identity). *Let* $a, b \in \mathbb{R}^n$ *and let* $M = I + a\otimes b$*. Then* $\det M = 1 + a \cdot b$*. Moreover, if* $a \cdot b \neq -1$ *then* M *is invertible and* $M^{-1} = I - \frac{a\otimes b}{1+a\cdot b}$*.*

8.10. Taylor's theorem. We record Taylor's theorem here in order to fix notation.

PROPOSITION 8.24 (Taylor's theorem with integral remainder). *For any* $f \in C^{k+1}\left(\mathbb{R}^d;\mathbb{R}\right)$ *and any* $z_0 \in \mathbb{R}^d$,

$$f = \mathcal{P}_k\left[f, z_0\right] + r_k\left[f, z_0\right] \bullet (\cdot - z_0)^{\otimes(k+1)} = \mathcal{P}_k\left[f, z_0\right] + \mathcal{R}_k\left[f, z_0\right],$$

where, for any $z \in \mathbb{R}^d$,

$$\mathcal{P}_k\left[f, z_0\right](z) := \sum_{l=0}^{k} \frac{1}{l!} \nabla^l f\left(z_0\right) \bullet \left(z - z_0\right)^{\otimes l},$$

$\mathcal{R}_k\left[f, z_0\right] := r_k\left[f, z_0\right] \bullet (\cdot - z_0)^{\otimes(k+1)}$, *and*

$$r_k\left[f, z_0\right](z) := \frac{1}{(k+1)!} \int_0^1 (1-t)^k \nabla^{k+1} f\left((1-t)\, z_0 + tz\right) \mathrm{d}t$$

EXAMPLE 8.25. For example, when $k = 2$ we have

$$f(z) = \underbrace{f(0) + \nabla f(0) \cdot z + \frac{1}{2} \nabla^2 f(0) \cdot (z \otimes z)}_{\mathcal{P}_2[f,0](z)} + \underbrace{\frac{1}{6} \left(\int_0^1 (1-t)^2\, \nabla^3 f(tz)\, \mathrm{d}t \right) \bullet (z \otimes z \otimes z)}_{\mathcal{R}_2[f,0](z)}.$$

Bibliography

[AAS17] B. F. Akers, D. M. Ambrose, and D. W. Sulon. Periodic traveling interfacial hydroelastic waves with or without mass. *Z. Angew. Math. Phys.*, 68(6):Art. 141, 27, 2017.

[Ahl78] L. V. Ahlfors. *Complex analysis.* McGraw-Hill Book Co., New York, third edition, 1978. An introduction to the theory of analytic functions of one complex variable, International Series in Pure and Applied Mathematics.

[AK71] S. Allen and K. Kline. Lubrication Theory for Micropolar Fluids. *Journal of Applied Mechanics*, 38(3):646–&, 1971.

[Arb00] L. F. A. Arbogast. Du calculs des dérivations. *Levrault, Strasbourg*, 1800.

[AS74] G. Ahmadi and M. Shahinpoor. Universal stability of magneto-micropolar fluid motions. *Internat. J. Engrg. Sci.*, 12:657–663, 1974.

[AS17] D. M. Ambrose and M. Siegel. Well-posedness of two-dimensional hydroelastic waves. *Proc. Roy. Soc. Edinburgh Sect. A*, 147(3):529–570, 2017.

[BBR+08] O. A. Beg, R. Bhargava, S. Rawat, K. Halim, and H. S. Takhar. Computational modeling of biomagnetic micropolar blood flow and heat transfer in a two-dimensional non-Darcian porous medium. *Meccanica*, 43(4):391–410, AUG 2008.

[Bea81] J. T. Beale. The initial value problem for the Navier-Stokes equations with a free surface. *Comm. Pure Appl. Math.*, 34(3):359–392, 1981.

[Bea84] J. T. Beale. Large-time regularity of viscous surface waves. *Arch. Rational Mech. Anal.*, 84(4):307–352, 1983/84.

[BF13] F. Boyer and P. Fabrie. *Mathematical tools for the study of the incompressible Navier-Stokes equations and related models*, volume 183 of *Applied Mathematical Sciences*. Springer, New York, 2013.

[BGN17] J. W. Barrett, H. Garcke, and R. Nürnberg. Finite element approximation for the dynamics of asymmetric fluidic biomembranes. *Math. Comp.*, 86(305):1037–1069, 2017.

[BŁ96] G. Bayada and G. Łukaszewicz. On micropolar fluids in the theory of lubrication. Rigorous derivation of an analogue of the Reynolds equation. *Internat. J. Engrg. Sci.*, 34(13):1477–1490, 1996.

[BN85] J. T. Beale and T. Nishida. Large-time behavior of viscous surface waves. In *Recent topics in nonlinear PDE, II (Sendai, 1984)*, volume 128 of *North-Holland Math. Stud.*, pages 1–14. North-Holland, Amsterdam, 1985.

[BNP11] A. Bonito, R. H. Nochetto, and M. S. Pauletti. Dynamics of biomembranes: effect of the bulk fluid. *Math. Model. Nat. Phenom.*, 6(5):25–43, 2011.

[BR14] Y. Bernard and T. Rivière. Energy quantization for Willmore surfaces and applications. *Ann. of Math. (2)*, 180(1):87–136, 2014.

[BSAV17] D. Banerjee, A. Souslov, A. G. Abanov, and V. Vitelli. Odd viscosity in chiral active fluids. *Nature Communications*, 8, Nov 17 2017.

[CCS07] C. H. A. Cheng, D. Coutand, and S. Shkoller. Navier-Stokes equations interacting with a nonlinear elastic biofluid shell. *SIAM J. Math. Anal.*, 39(3):742–800, 2007.

[CM12] Q. Chen and C. Miao. Global well-posedness for the micropolar fluid system in critical Besov spaces. *J. Differential Equations*, 252(3):2698–2724, 2012.

[CS10] C. H. A. Cheng and S. Shkoller. The interaction of the 3D Navier-Stokes equations with a moving nonlinear Koiter elastic shell. *SIAM J. Math. Anal.*, 42(3):1094–1155, 2010.

[DC09] B.-Q. Dong and Z.-M. Chen. Asymptotic profiles of solutions to the 2D viscous incompressible micropolar fluid flows. *Discrete Contin. Dyn. Syst.*, 23(3):765–784, 2009.

[dGM62] S. R. de Groot and P. Mazur. *Non-equilibrium thermodynamics.* Series in Physics. North-Holland Publishing Co., Amsterdam; Interscience Publishers Inc., New York, 1962.

[DLL07] Q. Du, M. Li, and C. Liu. Analysis of a phase field Navier-Stokes vesicle-fluid interaction model. *Discrete Contin. Dyn. Syst. Ser. B*, 8(3):539–556, 2007.

[DLRW09] Q. Du, C. Liu, R. Ryham, and X. Wang. Energetic variational approaches in modeling vesicle and fluid interactions. *Phys. D*, 238(9-10):923–930, 2009.

[DZ10] B.-Q. Dong and Z. Zhang. Global regularity of the 2D micropolar fluid flows with zero angular viscosity. *J. Differential Equations*, 249(1):200–213, 2010.

[Eri66] A. C. Eringen. Theory of Micropolar Fluids. *Journal of Mathematics and Mechanics*, 16(1):1–&, 1966.

[Eri99] A. C. Eringen. *Microcontinuum field theories. I. Foundations and solids.* Springer-Verlag, New York, 1999.

[Eri01] A. C. Eringen. *Microcontinuum field theories. II. Fluent media.* Springer-Verlag, New York, 2001.

[FJM02] G. Friesecke, R. D. James, and S. Müller. A theorem on geometric rigidity and the derivation of nonlinear plate theory from three-dimensional elasticity. *Comm. Pure Appl. Math.*, 55(11):1461–1506, 2002.

[FJMM03] G. Friesecke, R. D. James, M. G. Mora, and S. Müller. Derivation of nonlinear bending theory for shells from three-dimensional nonlinear elasticity by Gamma-convergence. *C. R. Math. Acad. Sci. Paris*, 336(8):697–702, 2003.

[FMJ02] G. Friesecke, S. Müller, and R. D. James. Rigorous derivation of nonlinear plate theory and geometric rigidity. *C. R. Math. Acad. Sci. Paris*, 334(2):173–178, 2002.

[FSG11] M. H. Farshbaf-Shaker and H. Garcke. Thermodynamically consistent higher order phase field Navier-Stokes models with applications to biomembranes. *Discrete Contin. Dyn. Syst. Ser. S*, 4(2):371–389, 2011.

[FSV97] S. Friedlander, W. Strauss, and M. Vishik. Nonlinear instability in an ideal fluid. *Ann. Inst. H. Poincaré Anal. Non Linéaire*, 14(2):187–209, 1997.

[FVR07] L. C. F. Ferreira and E. J. Villamizar-Roa. Micropolar fluid system in a space of distributions and large time behavior. *J. Math. Anal. Appl.*, 332(2):1425–1445, 2007.

[GBRT13] F. Gay-Balmaz, T. S. Ratiu, and C. Tronci. Equivalent theories of liquid crystal dynamics. *Arch. Ration. Mech. Anal.*, 210(3):773–811, 2013.

[GHS07] Y. Guo, C. Hallstrom, and D. Spirn. Dynamics near unstable, interfacial fluids. *Comm. Math. Phys.*, 270(3):635–689, 2007.

[GR77] G. P. Galdi and S. Rionero. A note on the existence and uniqueness of solutions of the micropolar fluid equations. *Internat. J. Engrg. Sci.*, 15(2):105–108, 1977.

[GS95a] Y. Guo and W. A. Strauss. Instability of periodic BGK equilibria. *Comm. Pure Appl. Math.*, 48(8):861–894, 1995.

[GS95b] Y. Guo and W. A. Strauss. Nonlinear instability of double-humped equilibria. *Ann. Inst. H. Poincaré Anal. Non Linéaire*, 12(3):339–352, 1995.

[GT13a] Y. Guo and I. Tice. Almost exponential decay of periodic viscous surface waves without surface tension. *Arch. Ration. Mech. Anal.*, 207(2):459–531, 2013.

[GT13b] Y. Guo and I. Tice. Decay of viscous surface waves without surface tension in horizontally infinite domains. *Anal. PDE*, 6(6):1429–1533, 2013.

[GT13c] Y. Guo and I. Tice. Local well-posedness of the viscous surface wave problem without surface tension. *Anal. PDE*, 6(2):287–369, 2013.

[Gur81] M. E. Gurtin. *An introduction to continuum mechanics*, volume 158 of *Mathematics in Science and Engineering*. Academic Press, Inc. [Harcourt Brace Jovanovich, Publishers], New York-London, 1981.

[Har06] M. Hardy. Combinatorics of partial derivatives. *Electron. J. Combin.*, 13(1):Research Paper 1, 13, 2006.

[Hat09] Y. Hataya. Decaying solution of a Navier-Stokes flow without surface tension. *J. Math. Kyoto Univ.*, 49(4):691–717, 2009.

[Haw68] S. W. Hawking. Gravitational Radiation in an Expanding Universe. *Journal of Mathematical Physics*, 9(4):598–604, 1968.

[Hel73] W. Helfrich. Elastic Properties of Lipid Bilayers: Theory and Possible Experiments. *Zeitschrift für Naturforschung C*, 28(11-12):693–03, 1973.

[KLŁ19] P. Kalita, J. A. Langa, and G. Łukaszewicz. Micropolar meets Newtonian. The Rayleigh-Bénard problem. *Phys. D*, 392:57–80, 2019.

[KS04] E. Kuwert and R. Schätzle. Removability of point singularities of Willmore surfaces. *Ann. of Math. (2)*, 160(1):315–357, 2004.

[LA17] S. Liu and D. M. Ambrose. Well-posedness of two-dimensional hydroelastic waves with mass. *J. Differential Equations*, 262(9):4656–4699, 2017.

[LR04] D. Lhuillier and A. Rey. Nematic liquid crystals and ordered micropolar fluids. *Journal of Non-Newtonian Fluid Mechanics*, 120(1-3):169–174, JUL 1 2004. 3rd International Workshop on Nonequilibrium Thermodynamics and Complex Fluids, Princeton, NJ, AUG 14-17, 2003.

[Lu89] G. Ł ukaszewicz. On the existence, uniqueness and asymptotic properties for solutions of flows of asymmetric fluids. *Rend. Accad. Naz. Sci. XL Mem. Mat. (5)*, 13(1):105–120, 1989.

[Łuk90] G. Łukaszewicz. On nonstationary flows of incompressible asymmetric fluids. *Math. Methods Appl. Sci.*, 13(3):219–232, 1990.

[Łuk99] G. Łukaszewicz. *Micropolar fluids*. Modeling and Simulation in Science, Engineering and Technology. Birkhäuser Boston, Inc., Boston, MA, 1999. Theory and applications.

[Łuk01] G. Łukaszewicz. Long time behavior of 2D micropolar fluid flows. *Math. Comput. Modelling*, 34(5-6):487–509, 2001.

[LZ16] Q. Liu and P. Zhang. Optimal time decay of the compressible micropolar fluids. *J. Differential Equations*, 260(10):7634–7661, 2016.

[Mau85] R. P. Maurya. Peripheral-Layer Viscosity and Microstructural Effects on the Capillary-Tissue Fluid Exchange. *Journal of Mathematical Analysis and Applications*, 110(1):59–73, 1985.

[MC15] B. Muha and S. Ĉanić. Fluid-structure interaction between an incompressible, viscous 3D fluid and an elastic shell with nonlinear Koiter membrane energy. *Interfaces Free Bound.*, 17(4):465–495, 2015.

[MC16] B. Muha and S. Ĉanić. Existence of a weak solution to a fluid-elastic structure interaction problem with the Navier slip boundary condition. *J. Differential Equations*, 260(12):8550–8589, 2016.

[Mil60] D. G. Miller. Thermodynamics of irreversible processes. the experimental verification of the Onsager reciprocal relations. *Chemical Reviews*, 60(1):15–37, 1960.

[MK08] K. S. Mekheimer and M. A. E. Kot. The micropolar fluid model for blood flow through a tapered artery with a stenosis. *Acta Mechanica Sinica*, 24(6):637–644, DEC 2008.

[MN14] F. C. Marques and A. Neves. Min-max theory and the Willmore conjecture. *Ann. of Math. (2)*, 179(2):683–782, 2014.

[MRP16] A. Mielke, D. R. M. Renger, and M. A. Peletier. A generalization of onsagers reciprocity relations to gradient flows with nonlinear mobility. *Journal of Non-Equilibrium Thermodynamics*, 41(2):141–149, 2016.

[MVBW11] P. A. Milewski, J.-M. Vanden-Broeck, and Z. Wang. Hydroelastic solitary waves in deep water. *J. Fluid Mech.*, 679:628–640, 2011.

[MW13] P. A. Milewski and Z. Wang. Three dimensional flexural-gravity waves. *Stud. Appl. Math.*, 131(2):135–148, 2013.

[NS12] E. R. Nicodemus and S. C. Sharma. Performance Characteristics of Micropolar Lubricated Membrane-Compensated Worn Hybrid Journal Bearings. *Tribology Transactions*, 55(1):59–70, 2012.

[NST16] R. H. Nochetto, A. J. Salgado, and I. Tomas. The equations of ferrohydrodynamics: modeling and numerical methods. *Math. Models Methods Appl. Sci.*, 26(13):2393–2449, 2016.

[NTY04] T. Nishida, Y. Teramoto, and H. Yoshihara. Global in time behavior of viscous surface waves: horizontally periodic motion. *J. Math. Kyoto Univ.*, 44(2):271–323, 2004.

[Ons31a] L. Onsager. Reciprocal relations in irreversible processes. i. *Phys. Rev.*, 37:405–426, Feb 1931.

[Ons31b] L. Onsager. Reciprocal relations in irreversible processes. ii. *Phys. Rev.*, 38:2265–2279, Dec 1931.

[PT11] P. I. Plotnikov and J. F. Toland. Modelling nonlinear hydroelastic waves. *Philos. Trans. R. Soc. Lond. Ser. A Math. Phys. Eng. Sci.*, 369(1947):2942–2956, 2011.

[Ram85] H. Ramkissoon. Flow of a micropolar fluid past a Newtonian fluid sphere. *Z. Angew. Math. Mech.*, 65(12):635–637, 1985.

[Riv08] T. Rivière. Analysis aspects of Willmore surfaces. *Invent. Math.*, 174(1):1–45, 2008.

[RKYC16] R. J. Ryham, T. S. Klotz, L. Yao, and F. S. Cohen. Calculating Transition Energy Barriers and Characterizing Activation States for Steps of Fusion. *Biophysical Journal*, 110(5):1110–1124, 2016.

[RM97] M. A. Rojas-Medar. Magneto-micropolar fluid motion: existence and uniqueness of strong solution. *Math. Nachr.*, 188:301–319, 1997.

[RTTa] A. Remond-Tiedrez and I. Tice. Anisotropic micropolar fluids subject to a uniform microtorque: the unstable case. Submitted. `[arxiv.org/abs/1910.12942]`.

[RTTb] A. Remond-Tiedrez and I. Tice. The viscous surface wave problem with generalized surface energies. To appear in SIAM J. Math. Anal. `[arxiv.org/abs/1806.07660]`.

[Rub17] J. Rubinstein. The mathematical theory of multifocal lenses. *Chin. Ann. Math. Ser. B*, 38(2):647–660, 2017.

[Rud76] W. Rudin. *Principles of mathematical analysis*. McGraw-Hill Book Co., New York-Auckland-Düsseldorf, third edition, 1976. International Series in Pure and Applied Mathematics.

[SHKL96] V. Squire, R. Hosking, A. Kerr, and P. Langhorne. *Moving Loads on Ice Plates*. Solid Mechanics and Its Applications. Springer Netherlands, 1996.

[Smi71] G. F. Smith. On isotropic functions of symmetric tensors, skew-symmetric tensors and vectors. *Internat. J. Engrg. Sci.*, 9:899–916, 1971.

[SSP82] P. Sinha, C. Singh, and K. PrasaD. Lubrication of Human Joints - A Microcontinuum Approach. *WEAR*, 80(2):159–181, 1982.

[Tar06] A. Tarasińska. Global attractor for heat convection problem in a micropolar fluid. *Math. Methods Appl. Sci.*, 29(11):1215–1236, 2006.

[TMPVB19] O. Trichtchenko, P. Milewski, E. Părău, and J.-M. Vanden-Broeck. Stability of periodic traveling flexural-gravity waves in two dimensions. *Stud. Appl. Math.*, 142(1):65–90, 2019.

[TPVBM18] O. Trichtchenko, E. I. Părău, J.-M. Vanden-Broeck, and P. Milewski. Solitary flexural-gravity waves in three dimensions. *Philos. Trans. Roy. Soc. A*, 376(2129):20170345, 14, 2018.

[TT95] A. Tani and N. Tanaka. Large-time existence of surface waves in incompressible viscous fluids with or without surface tension. *Arch. Rational Mech. Anal.*, 130(4):303–314, 1995.

[TW14] Z. Tan and Y. Wang. Zero surface tension limit of viscous surface waves. *Comm. Math. Phys.*, 328(2):733–807, 2014.

[TZ18] I. Tice and S. Zbarsky. Decay of solutions to the linearized free surface Navier-Stokes equations with fractional boundary operators. *ArXiv e-prints*, June 2018.

[Wan70a] C. C. Wang. A new representation theorem for isotropic functions: An answer to Professor G. F. Smith's criticism of my papers on representations for isotropic functions. I. Scalar-valued isotropic functions. *Arch. Rational Mech. Anal.*, 36:166–197, 1970.

[Wan70b] C. C. Wang. A new representation theorem for isotropic functions: An answer to Professor G. F. Smith's criticism of my papers on representations for isotropic functions. II. Vector-valued isotropic functions, symmetric ten tensor-valued isotropic functions, and skew-symmetric tensor-valued isotropic functions. *Arch. Rational Mech. Anal.*, 36:198–223, 1970.

[Whi65] E. F. Whittlesey. Analytic functions in Banach spaces. *Proc. Amer. Math. Soc.*, 16:1077–1083, 1965.

[Wil65] T. J. Willmore. Note on embedded surfaces. *An. Sti. Univ. "Al. I. Cuza" Iasi Sect. I a Mat. (N.S.)*, 11B:493–496, 1965.

[Wil82] T. J. Willmore. *Total curvature in Riemannian geometry*. Ellis Horwood Series: Mathematics and its Applications. Ellis Horwood Ltd., Chichester; Halsted Press [John Wiley & Sons, Inc.], New York, 1982.

[Wil93] T. J. Willmore. *Riemannian geometry*. Oxford Science Publications. The Clarendon Press, Oxford University Press, New York, 1993.

[WTK14] Y. Wang, I. Tice, and C. Kim. The viscous surface-internal wave problem: global well-posedness and decay. *Arch. Ration. Mech. Anal.*, 212(1):1–92, 2014.

[Wu14] L. Wu. Well-posedness and decay of the viscous surface wave. *SIAM J. Math. Anal.*, 46(3):2084–2135, 2014.

[WVBM13] Z. Wang, J.-M. Vanden-Broeck, and P. A. Milewski. Two-dimensional flexural-gravity waves of finite amplitude in deep water. *IMA J. Appl. Math.*, 78(4):750–761, 2013.

[Yua10] B. Yuan. On regularity criteria for weak solutions to the micropolar fluid equations in Lorentz space. *Proc. Amer. Math. Soc.*, 138(6):2025–2036, 2010.

[Zhe17] Y. Zheng. Local well-posedness for the Bénard convection without surface tension. *Commun. Math. Sci.*, 15(4):903–956, 2017.

[ZT17] Y. Zheng and I. Tice. Local well posedness of the near-equilibrium contact line problem in 2-dimensional Stokes flow. *SIAM J. Math. Anal.*, 49(2):899–953, 2017.