Lagrangian Averaging, Nonlinear Waves, and Shock Regularization

Thesis by
Harish S. Bhat

In Partial Fulfillment of the Requirements
for the Degree of
Doctor of Philosophy

California Institute of Technology
Pasadena, California

2005

(Defended May 6, 2005)

Acknowledgements

This work would not have been initiated, pursued, or completed were it not for the steady encouragement of my adviser, Jerry Marsden. I admire greatly his patience, wisdom, conviction, and insight, and I thank him for taking me on as a student in early 2001. Of course, I should also thank him for his ideas, for being a superb teacher, and for leading by example. If there is one thing that has puzzled me at Caltech, it is trying to understand how he does it all.

My other main pillars of academic support have been Matt West and Razvan Fetecau. I met Matt and Razvan in the Spring of 2002 when they were still graduate students at Caltech. Without Matt, I would be writing an endless stream of personal notes rather than pursuing problems of interest to the community. Without Razvan, I would be continually floating incorrect and wild theories. Thanks are due to both of them for blocking (or, at least attempting to block) my natural efforts to make a fool of myself. I am also indebted to them for their time and support when I was at Stanford University from Dec. 2003 until Sep. 2004.

In CDS, I thank my seven fellow batchmates Melvin Flores, Melvin Leok, Antonis Papachristodoulou, Xin Liu, Stephen Prajna, Domitilla Del Vecchio, and Shreesh Mysore. Our first year here was a truly shared experience and let me say with great optimism that I look forward to our reunions! I thank Jair Koiller for being an awesome instructor, for letting me give a couple lectures, and for writing letters all the way from Brazil. I thank Wang-Sang Koon for the experience of working with him, and also for forgiving my delayed TA efforts. I thank Nawaf Bou-Rabee for taking on many of my TA duties during the final phases of writing this thesis. I thank the members of my candidacy committee, Anthony Leonard and Tapio Schneider. And finally I thank Wendy McKay and Spencer Mortensen, whose efforts to assist me in various tasks were matched only by their kindness.

I thank the numerous members of the Indian student community, who have all altered my life in a positive way. I thank those who participated with me in *Chanakya*, and all those who acted for me in *Raman Leela*. I thank the Caltech Y and all those I got to

know during Y-Hikes and other activities. A proper listing of the names of all persons referred to above could easily take two pages. You all are the reasons my experience here was positive in every possible way.

Rather abstractly, let me simply thank Caltech, the Institute itself, for being everything that it is. I thank all those unseen forces and invisible hands who brought me here, let me grow here, and are now seeing it fit to let me go. If there was some practical way to stay here forever as a graduate student, I am sorry I did not discover it.

I thank those in the academic community who took time to answer my questions while asking me questions of their own. I thank Ehsan Afshari and Ali Hajimiri for calling me out of the blue and inviting me into their research world—let's see what the future holds! I thank my new friends at Columbia University and the University of Washington, who gave me excellent feedback on parts of this thesis.

I thank all my friends from the Bay Area who have managed to keep up with me over the past decade.

There are four people left.

Let me first thank my father and my mother for motivating me to pursue my interest in the mathematical sciences, and for giving me more opportunities than I would have known to ask for. I know I will never be able to repay you.

Next I thank Vinay, my younger brother. Your questions, stories, commentary, and always-open ear have been a constant source of support. May this thesis inspire you to act on your own passions!

And lastly, I thank Lee-May, my source of inspiration, my best friend, and my fianceé. What can I say in this space that you do not already know?

Lagrangian Averaging, Nonlinear Waves, and Shock Regularization

by

Harish S. Bhat

In Partial Fulfillment of the
Requirements for the Degree of
Doctor of Philosophy

Abstract

In this thesis, we explore various models for the flow of a compressible fluid as well as model equations for shock formation, one of the main features of compressible fluid flows.

We begin by reviewing the variational structure of compressible fluid mechanics. We derive the barotropic compressible Euler equations from a variational principle in the material frame. The particle relabeling symmetry of fluid mechanics is explained, and the material-frame Lagrangian is shown to be invariant under this symmetry. We then show how the barotropic compressible Euler equations arise from a variational principle in the spatial frame. Writing the resulting equations of motion requires certain Lie-algebraic calculations that we carry out in detail for expository purposes.

Next, we extend the derivation of the Lagrangian averaged Euler (LAE-α) equations to the case of barotropic compressible flows. The aim of Lagrangian averaging is to regularize the compressible Euler equations by adding dispersion instead of artificial viscosity. Along the way, the derivation of the isotropic and anisotropic LAE-α equations is simplified and clarified. The derivation in this paper involves averaging over a tube of trajectories η^ϵ centered around a given Lagrangian flow η. With this tube framework, the LAE-α equations are derived by following a simple procedure: start with a given action, expand via Taylor series in terms of small-scale fluid fluctuations ξ, truncate, average, and then model those terms that are nonlinear functions of ξ. Closure of the equations is provided through the use of *flow rules*, which prescribe the evolution of the fluctuations along the mean flow.

We then analyze a simple one-dimensional Lagrangian averaged model, a subcase of the general models derived above. We prove the existence of a large family of traveling waves; these solutions correspond to homoclinic orbits in the phase plane. Computing the dispersion relation for this model, we find it is nonlinear, implying that the equation is dispersive. Since the amount of dispersion in the model is controlled by α, we expect that the zero-α limit will be highly oscillatory. We carry out numerical experiments that show that the model possesses smooth, bounded solutions that display interesting pattern formation. Finally, we relate the mathematical features of this model to other models for solitary waves in compressible fluids, and to other solitary wave phenomena that occurs in air.

Finally, we examine a Hamiltonian partial differential equation (PDE) that regularizes the inviscid Burgers equation without the addition of standard viscosity. Here α is a small parameter that controls the amount of regularization we have added to the inviscid Burgers equation. We show the existence of a large family of traveling front solutions that connect two different states. These solutions correspond to heteroclinic orbits in the phase plane. Using the method of characteristics, we also analyze the initial-value problem and prove well-posedness for a certain class of initial data. We prove that in the zero-α limit, without any standard viscosity, solutions of the PDE converge strongly to weak solutions of the inviscid Burgers equation. We provide numerical evidence that this strong limit satisfies an entropy inequality for the inviscid Burgers equation. We demonstrate a non-local Hamiltonian structure for the PDE.

Contents

Chapter 1

Variational Compressible Fluids

1.1 Glossary

Let us first introduce important symbols and terms. We will assume familiarity with these concepts, at the level of a first reading of [1]:

Bijection	A function that is both one-to-one and onto
Diffeomorphism	A differentiable bijection with differentiable inverse
$\mathrm{Diff}(M)$	Space of diffeomorphisms on M
$\Omega^k(M)$	Space of k-forms on M
$\mathrm{Den}(M)$	Space of volume forms on M
$\mathfrak{X}(M)$	Space of vector fields on M
$C^\infty(M)$	Space of real-valued, infinitely differentiable functions on M
$\mathbf{L}(E,F)$	Space of linear maps between Banach spaces E and F
$\mathbf{GL}(E,F)$	Space of invertible linear maps between Banach spaces E and F
$\varphi^*(\cdot)$	Pullback by φ
$\varphi_*(\cdot)$	Push-forward by φ
$\mathrm{div}(u)$	Divergence of vector field u
$\pounds_X(\cdot)$	Lie derivative with respect to the vector field X
$\mathbf{i}_X(\cdot)$	Interior product with the vector field X
$\mathbf{d}$	Exterior derivative
$v^\flat$	One-form associated to the vector field v
$\omega^\sharp$	Vector field associated to the one-form ω
Tf	Tangent mapping of f
$\mathbf{D}f$	Fréchet derivative of f
$\otimes$	Tensor product

1.2 The Setting

We begin with the Euler equations for a compressible fluid, written in the traditional form using Euclidean coordinates on a Euclidean space:

$$\frac{\partial \rho}{\partial t} + \operatorname{div}(\rho u) = 0 \tag{1.1a}$$

$$\frac{\partial u}{\partial t} + \operatorname{div}\left(\frac{\rho u \otimes \rho u}{\rho}\right) = -\nabla p \tag{1.1b}$$

Here ρ is the density, ρu is the momentum, and p is the pressure. As can be gleaned from (1.1), the density and pressure are scalar fields, while the momentum (and velocity u) are vector fields. Shortly, we will be much more specific about these objects. We assume that the pressure and density are related via the barotropic law

$$p = \kappa \rho^{\gamma}, \tag{1.2}$$

for constants $\kappa, \gamma > 0$. We are typically interested in (1.1) in a few different settings:

P1. The fluid lives in a container (manifold) M with smooth boundary ∂M. In this case, we assume that M is compact and that, restricted to the boundary, u is parallel to the boundary ∂M.

P2. The fluid lives in a compact manifold M without boundary, e.g., a periodic box. We may think of this as a special case of P1 in which $\partial M = \emptyset$.

P3. The fluid lives in a non-compact manifold M without boundary, e.g., all of $\mathbb{R}^3$. In this case, as the analytical details of the problem become clear, we must specify decay conditions on the density and velocity as $\|\mathbf{x}\| \to \pm\infty$.

We refer to Chapter 9, §2 of [1] and [15] for derivations of (1.1). Here we shall focus on answering the question of how to represent (1.1) as an infinite-dimensional Lagrangian dynamical system. To avoid analytical details, we shall confine ourselves to the case P2 in which the fluid lives in a compact boundaryless manifold. This will enable us to spend more time thinking about the geometric structures underlying the compressible

Euler equations. Let us remark, however, that if one is interested in P1 or P3, it is possible to modify our derivations accordingly. Along the way, we shall mention where the appropriate modifications must be made.

The variational structures we describe are an outgrowth of *topological hydrodynamics*, a discipline set into motion by the seminal papers of Arnold [3] and Ebin and Marsden [22]. Arnold's paper was the first to view the motion of an incompressible fluid as geodesic motion on an appropriate group of diffeomorphisms. The Lagrangian for the incompressible Euler equation is a quadratic, positive definite function and can therefore be used to define a metric. The ensuing Euler-Lagrange equation is then precisely the geodesic equation for this metric. The paper of Ebin and Marsden made this point of view rigorous, and demonstrated its utility by proving beautiful analytical results on the short-time well-posedness of the incompressible Euler and Navier-Stokes equations. For a recent and highly readable introduction to topological hydrodynamics, we recommend [49], while for an encyclopedic treatment of topics in this field, one must consult [4]. Finally, an overview of the field from a PDE point of view may be found in the book review [78].

Not nearly as much work has been done on the compressible side of the story. Here the dominant concern in analyzing (1.1) is the formation of shock waves, or discontinuities in the field variables. It is now well-known (see [11]) that for generic initial data $\rho(x,0) = \rho_0(x)$ and $u(x,0) = u_0(x)$, the resulting solutions $\rho(x,t)$ and $u(x,t)$ are smooth only on short time intervals $t \in [0,T]$. We refer to the maximum such T as the *break time* of the solutions. Beyond the break time, the solutions $\rho(x,t)$ and $u(x,t)$ are discontinuous, meaning that we have to work with the weak form of (1.1). This process occurs without regard for the initial smoothness of the fields, i.e., the functions ρ_0 and u_0.

Returning to geometric considerations, one finds that the compressible equations are *not* the equations for geodesic flow on $\mathrm{Diff}(M)$, the group of diffeomorphisms of M. The reason is simple: the Lagrangian is no longer positive definite, and does not define a metric. One can apply a Kaluza-Klein construction and enlarge the configuration space of the problem; then the compressible equations will describe geodesic motion on the enlarged space. However, it is unclear whether this procedure will yield any useful

information, and we will not pursue this angle here.

Nonetheless, one can still consider compressible fluid motion as either a Hamiltonian or a variational problem on an appropriate group of diffeomorphisms. For example, [21] used the variational principle for compressible flow to show that in the incompressible limit, i.e. Mach number $\to \infty$, solutions of (1.1) converge to solutions of the incompressible Euler equation. As with other papers in the early history of topological hydrodynamics, the mathematical developments take place on the group of diffeomorphisms, i.e. the *material* or Lagrangian picture. This is in contrast to the *spatial* or Eulerian picture we have assumed when writing (1.1). Let us now take a moment to review what these terms mean.

1.3 Material and Spatial Pictures

Historically, the variational structure of the equations for compressible fluid motion was first developed in the material picture, rather than the spatial picture we have assumed in (1.1). Let us examine what each picture looks like at a specific time t. For the time being, to ease the exposition, we will suppose that functions and fields are much smoother than they need to be for the mathematics to work.

In the spatial picture, at each time t, we have a vector field $u^t \in \mathfrak{X}(M)$ and a scalar function $\rho^t \in C^\infty(M)$. At each point $x \in M$, the vector $u^t(x)$ gives the velocity of the fluid at x. Similarly, $\rho^t(x)$ is the density of the fluid at x. Note that while the fluid is sloshing around in M, the points $x \in M$ themselves do not move—these points are, philosophically, just like the tick marks on a number line.

In the material picture, at each time t, we have a diffeomorphism $\eta^t \in \mathrm{Diff}(M)$. For each $X \in M$, $\eta^t(X)$ gives the deformation of X. Initially, at $t = 0$, there is no deformation, so $\eta^0(X) \equiv X$. This means we may use X as a particle label, and think of $\eta^t(X)$ as the position of particle X at time t.

Generalizing both pictures to a time interval $[0, T]$, we see that both u and ρ are

curves on the appropriate spaces. That is, the velocity field is a curve

$$u : [0, T] \to \mathfrak{X}(M)$$

that assigns, to each $t \in [0, T]$, the velocity field $u^t \in \mathfrak{X}(M)$. Similarly, the density is a curve

$$\rho : [0, T] \to C^\infty(M),$$

and the material motion η is a curve

$$\eta : [0, T] \to \text{Diff}(M).$$

From our descriptions of both pictures, we deduce the relationship between the two:

$$\frac{\partial}{\partial t}\eta(X, t) = u(\eta(X, t), t). \tag{1.3}$$

The left-hand side of (1.3) is the velocity of the particle labeled X at time t. At this time t, this particle has been deformed from its initial location at $X \in M$ to its new location at $\eta(X, t) \in M$. The right-hand side of (1.3) is the velocity at the point in space given by $\eta(X, t)$. It is clear that these two expressions refer to precisely the same measurement.

Remarks.

1. In continuum mechanics, the material picture is often called "Lagrangian" while the spatial picture is called "Eulerian." To avoid confusion with the Lagrangian functions and Euler-Lagrange equations of variational calculus, we will stick with the terms material and spatial.

2. Our description might lead one to believe that the spatial description of the fluid requires two fields, u and ρ, while the material description requires only one: η. Of course, the fluid has density in the material picture as well! The idea here is that once we specify an initial density field $\rho(X, 0)$, our full knowledge η of the subsequent deformation of the fluid should be enough to determine the density

field $\rho(X,t)$ at all later times. Intuitively, "we know where each particle is going." Indeed, suppose that the domain M is an open subset of $\mathbb{R}^n$. Then write $\mu = \rho\,\mathrm{d}^n x$, where $\mathrm{d}^n x$ is the canonical volume form on $\mathbb{R}^n$. Transforming the continuity equation (1.1a) to the material picture, we find

$$\mu(X,t) = \left(\eta^t\right)_* \mu(X,0), \tag{1.4}$$

confirming our intuition. We will often write $\mu_0(X) = \mu(X,0)$.

This brief summary of the material and spatial pictures will be enough for our purposes. For an extensive discussion of the geometry and kinematics behind continuum mechanics, we refer the reader to [55].

Passing from material to spatial coordinates is an example of symmetry reduction. In simple terms, for a barotropic compressible fluid in the material picture, we require η, $\dot{\eta}$, and μ_0 to determine the current configuration of the fluid. In the spatial point of view, we require only u and μ. To get a hint as to why one fewer variable is required in the spatial picture, let us note that the relationship between the two pictures (1.3) can be rewritten as $u(x,t) = \partial_t \eta(\eta^{-1}(x,t),t)$ or for short,

$$u = \dot{\eta} \circ \eta^{-1}. \tag{1.5}$$

In other words, the diffeomorphism η enters the material version of (1.1) only in a particular combination with $\dot{\eta}$. Suppose we now relabel each particle X by $\varphi(X)$ where φ is some arbitrary diffeomorphism. This amounts to replacing η by $\eta \circ \varphi$ in (1.5). Making the substitution, we will find that

$$\begin{aligned} \partial_t \left(\eta \circ \varphi\right) \circ \left(\eta \circ \varphi\right)^{-1} &= \dot{\eta} \circ \varphi \circ \varphi^{-1} \circ \eta^{-1} \\ &= \dot{\eta} \circ \eta^{-1} \\ &= u. \end{aligned}$$

This is the *particle relabeling symmetry* of fluid mechanics. Later, we will show that the

kinetic and potential energies of the fluid are invariant under particle relabeling as well. Before we prove this, let us provide some historical references. Our goal is not to offer an exhaustive history of the subject, but instead to point out a few sources for the main ideas that we present in this expository chapter.

The role of reduction in the geometric structures underlying compressible flow was not explained until 15-20 years after Arnold's paper [3]. Historically, the Hamiltonian point of view came first, in [57, 58, 38]. These papers lay an expansive foundation for theories of semidirect product type, of which compressible fluids is but one example. Semidirect products arise in continuum mechanics when, in addition to the primary field variables, there are other quantities, transported by the main fields, that play a dynamical role in the problem.

In the case of compressible fluids, in addition to u, we have ρ, which from (1.1a) we see is somehow being transported by the flow. Note that in the material picture, i.e., prior to reduction, the density μ_0 is simply a parameter. Once we know η and $\dot{\eta}$, we determine μ using (1.4)—we do not solve any differential equation for μ. Somehow, in the spatial picture, i.e. after symmetry reduction, the density becomes a dynamic variable which couples with the other field variables as in (1.1a). Semidirect product theory is a general toolbox for working with mechanical systems of this type.

Among their many results, the papers [57, 58, 38] explain how to derive the Poisson bracket for compressible flows in spatial variables by starting from the material picture and carrying out symmetry reduction. The resulting bracket is called the semidirect product Lie-Poisson bracket; using it, one can realize (1.1) as Hamilton's equation on the appropriate space.

The basic idea on the Lagrangian side is as follows. In the material picture, the variational principle underlying the equations of fluid mechanics is the classical stationary action principle. There are no constraints on variations $\delta\eta$, and the equation of motion corresponding to stationary points of the action is precisely the Euler-Lagrange equation. When one moves to the spatial picture, the situation changes. Now one has to guarantee

that variations δu give rise to varied vector fields u^ϵ that are indeed expressible as

$$u^\epsilon = \dot{\eta}^\epsilon \circ (\eta^\epsilon)^{-1}$$

for some diffeomorphism η^ϵ. In short, the variations δu must be constrained, meaning that the variational principle is different from the classical one. The resulting equation of motion is therefore different from the usual Euler-Lagrange equation, and it goes by the name of the Euler-Poincaré (EP) equation. Again, in the case of a compressible fluid, we must deal not only with variations δu but also variations $\delta\rho$. This gives rise to the semidirect product EP equation. All of this is explained in detail in [40], which also contextualizes the EP equation in the history of mechanics. Note that an alternative exposition of the fluid mechanical part of the story can be found in [66].

Now we shall go ahead with presenting some of the details of the theory discussed in [40]. Our aim is to be pedagogical, and make explicit many of the calculations required to view the barotropic compressible Euler equations (1.1) as a semidirect product EP system. This framework will be used in subsequent chapters to produce new equations of motion that automatically respect various symmetries.

1.4 Material Lagrangian

Our goal is to present a variational principle, in the material picture, for the Euler equations for barotropic compressible flow. Indeed, we will show that (1.1b) is precisely the Euler-Lagrange equation associated with the following Lagrangian. Let $G = \mathrm{Diff}(M)$ and $V^* = \mathrm{Den}(M)$. Then we define the Lagrangian $L : TG \times V^* \to \mathbb{R}$ by

$$L(\dot{\eta}, \mu_0) = \int_M \left[\frac{1}{2}\|\dot{\eta}\|^2 - W\Big([\eta_*\mu_0] \circ \eta\Big)\right] \mu_0, \tag{1.6}$$

where $\dot{\eta} \in T_\eta G$. The Lagrangian has the usual form of kinetic energy minus potential energy. We assume the fluid is barotropic, meaning that the pressure is a function only of the fluid's density. This forces the potential energy to also be a function only of the fluid's density. The density μ_0 is the density as a function of material points in the fluid,

at time $t = 0$.

Note that (1.6) is not positive definite; hence (1.6) does not define a metric, and we cannot interpret the Euler-Lagrange equation associated with (1.6) as a geodesic equation associated with some metric on $G \times V$.

Details. Let us explain what we mean by $\|\cdot\|$ in (1.6). We assume that the manifold M has a Riemannian metric $\langle\!\langle\, , \rangle\!\rangle_m$, and that this induces a volume form ν. In words, the metric is a smooth map $m \mapsto \langle\!\langle\, , \rangle\!\rangle_m$ that assigns to each point $m \in M$ an inner product on the tangent space $T_m M$. We use this to define a weak Riemannian structure on G. Namely, on the tangent space $T_\eta G$, we have, for $\alpha, \beta \in T_\eta G$, the inner product

$$\langle\!\langle \alpha, \beta \rangle\!\rangle = \int_M \langle\!\langle \alpha(m), \beta(m) \rangle\!\rangle_{\eta(m)}\, \nu(m). \tag{1.7}$$

If we are interested in problem P3, where the manifold M is non-compact and boundaryless, we must specify *decay conditions*. That is, unless the integrand vanishes rapidly outside a compact set K, the overall integral (1.7) may not be finite. For problems P1 and P2, where M is compact, there is no problem, since

$$\left| \int_M f(m)\nu(m) \right| \leq \|f\|_{L^\infty} \int_M \nu(m) = \|f\|_{L^\infty} \mathrm{vol}(M) < \infty.$$

We present without proof the fact that for a given diffeomorphism $\eta \in G$,

$$T_\eta G = \{ \text{ smooth maps } X : M \to TM \text{ such that } \pi_M \circ X = \eta \}.$$

With this, it is clear that for each m, the vectors $\alpha(m)$ and $\beta(m)$ belong to the tangent space $T_{\eta(m)} M$, so we can at least see that formula (1.7) makes sense. (Sometimes the elements of $T_\eta G$ are called vector fields that *cover* the diffeomorphism η. See [55].) Then we have, for all $\dot\eta \in T_\eta G$,

$$\|\dot\eta\|^2 = \langle\!\langle \dot\eta, \dot\eta \rangle\!\rangle.$$

Of course, to actually show that (1.7) is a weak Riemannian structure, we must do more work. See §9 in [22].

1.4.1 Equations of motion

We obtain equations of motion by applying Hamilton's principle directly to the Lagrangian (1.6). That is, for the action

$$S = \int_0^T L \, dt,$$

we will compute δS, assuming that the variations η^ϵ are fixed at $t = 0$ and $t = T$. By

$$\delta\eta = \left.\frac{d}{d\epsilon}\right|_{\epsilon=0} \eta^\epsilon,$$

this will give $\delta\eta = 0$ at the endpoints $t = 0$ and $t = T$.

Before computing δS, it is worth making the following point: in this thesis, we will never be able to *derive* mass conservation equations from variational principles, in either the spatial (1.1a) or material (1.4) forms. These equations will not pop out of $\delta S = 0$ the way that the momentum equation (1.1b) will. Instead, one shows (1.1a) or (1.4) from first principles; here we will take these equations as given.

Let us continue with the derivation of the Euler-Lagrange equation associated with the Lagrangian L. As a warning, the computation that follows is *not* general—in order to proceed in a general, global fashion we would have to introduce a smooth affine connection on G and compute using the covariant derivative. This calculation is carried out in detail in both [21] and §6.1 of [71]. We will not have use for such geometric machinery beyond this calculation, so we limit ourselves to the case when $M = \mathbb{R}^n$, in which case it is fine to treat ∇ as the ordinary gradient. Once we derive the Euler-Lagrange equation for L, our goal will be to change from material to spatial coordinates and recover the momentum equation (1.1b) in ordinary Euclidean space/coordinates.

Write $\mu_0 = \rho_0 \, \mathrm{d}^n x$ where $\rho_0 \in C^\infty(M)$. Then

$$\begin{aligned}
\delta S &= \int_{t_1}^{t_2} \int_M \Big[\dot\eta^i \delta\dot\eta^i \\
&\quad - W'\left((\det \nabla\eta)^{-1} \rho_0(m)\right) \left(-(\det\nabla\eta)^{-1} (\nabla\eta)^{-T} \cdot \nabla \delta\eta \rho_0\right) \Big] \rho_0 \, \mathrm{d}^n x \, dt \\
&= \int_{t_1}^{t_2} \int_M \Big[\dot\eta^i \delta\dot\eta^i \\
&\quad + W'\left((\det \nabla\eta)^{-1} \rho_0(m)\right) \left((\det\nabla\eta)^{-1} \frac{\partial (\eta^{-1})^j}{\partial x^i} \frac{\partial (\delta\eta)^i}{\partial X^j} \rho_0 \right) \Big] \rho_0 \, \mathrm{d}^n x \, dt.
\end{aligned}$$

Because the variations $\delta\eta$ vanish at the boundary, an integration by parts gives

$$\begin{aligned}
\delta S &= \int_M \Big\{ -\ddot\eta^i \delta\eta^i \rho_0 \\
&\quad - \frac{\partial}{\partial X^j} \left[W'\left((\det\nabla\eta)^{-1}\rho_0(m)\right) \left((\det\nabla\eta)^{-1} \frac{\partial (\eta^{-1})^j}{\partial x^i} \right) \rho_0^2 \right] \delta\eta^i \Big\} \, \mathrm{d}^n x.
\end{aligned}$$

If δS is to be zero for all $\delta\eta$, then η must satisfy

$$\rho_0 \ddot\eta^i = -\frac{\partial}{\partial X^j} \left[W'\left((\det\nabla\eta)^{-1}\rho_0(m)\right) \left((\det\nabla\eta)^{-1} \frac{\partial (\eta^{-1})^j}{\partial x^i} \right) \cdot \rho_0^2 \right]. \tag{1.8}$$

In the above derivation, we have assumed $\partial M = \emptyset$. If we were interested in problem P1, when M has a smooth non-empty boundary ∂M, we would have to specify boundary conditions on η. Furthermore, whenever we integrate by parts on spatial variables, we would have to treat the boundary term that arises: if we want this term to disappear, we have to choose the behavior of $\delta\eta$ on ∂M accordingly.

1.4.2 Substitution of variables

We now express (1.8) in the spatial variables u and ρ, in order to compare it with the Euler equation in spatial variables (1.1). We will need the following fact: the spatial density $\rho(x,t)$ is related to the material density $\rho_0(X)$ by

$$\rho(\eta(X,t),t) = \frac{\rho_0(X)}{\det \nabla\eta(X,t)}. \tag{1.9}$$

This follows by writing (1.4) as

$$\left(\eta^t\right)^* \mu = \mu_0$$

and carrying out the pull-back operation in the usual way. For short, we will write (1.9) as

$$\rho \circ \eta = \frac{\rho_0}{\det \nabla \eta}. \tag{1.10}$$

Differentiating the relation $u \circ \eta = \dot{\eta}$ gives

$$\ddot{\eta} = \frac{\partial u}{\partial t} \circ \eta + (\nabla u \circ \eta) \cdot u \circ \eta. \tag{1.11}$$

Substituting (1.10-1.11) into (1.8), we have

$$[(\det \nabla \eta)\rho \circ \eta] \left(\frac{\partial u}{\partial t} \circ \eta + (\nabla u \circ \eta) \cdot u \circ \eta \right)^i = -\frac{\partial}{\partial X^j} \left[W'(\rho \circ \eta) \left(\rho \circ \eta \cdot \frac{\partial (\eta^{-1})^j}{\partial x^i} (\det \nabla \eta) \rho \circ \eta \right) \right].$$

Composing on the right with η^{-1} everywhere, we obtain

$$(\det \nabla \eta \circ \eta^{-1})\rho \left(\frac{\partial u}{\partial t} + \nabla u \cdot u \right)^i = -\frac{\partial \eta^k}{\partial X^j} \frac{\partial}{\partial x^k} \left[W'(\rho) \rho^2 \frac{\partial (\eta^{-1})^j}{\partial x^i} \det \nabla \eta \circ \eta^{-1} \right]. \tag{1.12}$$

Here we have introduced the "spatial coordinates" $x^k = \eta^k(X, t)$ and used the local coordinate calculation

$$\frac{\partial}{\partial X^j} = \frac{\partial x^k}{\partial X^j} \frac{\partial}{\partial x^k} = \frac{\partial \eta^k}{\partial X^j} \frac{\partial}{\partial x^k}.$$

We expand the right-hand side of (1.12) into three terms:

$$\begin{aligned} &- \frac{\partial \eta^k}{\partial X^j} \left\{ \left(\frac{\partial}{\partial x^k} \left[W'(\rho)\rho \right] \frac{\partial (\eta^{-1})^j}{\partial x^i} \rho \det \nabla \eta \circ \eta^{-1} \right) \right. \\ &+ [W'(\rho)\rho] \left(\frac{\partial \rho}{\partial x^k} \cdot \frac{\partial (\eta^{-1})^j}{\partial x^i} \det \nabla \eta \circ \eta^{-1} \right. \\ &+ \left. \left. \rho \frac{\partial}{\partial x^k} \left(\frac{\partial (\eta^{-1})^j}{\partial x^i} \det \nabla \eta \circ \eta^{-1} \right) \right) \right\}. \end{aligned}$$

The last term is actually zero:

$$\begin{aligned}
&\frac{\partial \eta^k}{\partial X^j}\frac{\partial}{\partial x^k}\left(\frac{\partial (\eta^{-1})^j}{\partial x^i}(x)\cdot\frac{1}{\det(\nabla\eta^{-1})}\right)\\
&=\left(\frac{\partial \eta^k}{\partial X^j}\frac{\partial^2 (\eta^{-1})^j}{\partial x^k\partial x^i}\frac{1}{\det(\nabla\eta^{-1})}\right)-\frac{1}{(\det(\nabla\eta^{-1}))^2}\frac{\partial}{\partial x^i}(\det\nabla\eta^{-1})\\
&=\left(\frac{\partial \eta^k}{\partial X^j}\frac{\partial^2 (\eta^{-1})^j}{\partial x^k\partial x^i}\frac{1}{\det(\nabla\eta^{-1})}\right)-\frac{1}{(\det(\nabla\eta^{-1}))^2}\frac{\partial \eta^n}{\partial X^m}\det\nabla\eta^{-1}\frac{\partial^2 (\eta^{-1})^m}{\partial x^i\partial x^n}\\
&=0.
\end{aligned}$$

From $\eta(\eta^{-1}(x,t),t)=x$ we deduce

$$\frac{\partial \eta^i}{\partial X^j}\frac{\partial (\eta^{-1})^j}{\partial x^k}=\delta^i_k.$$

Putting everything together, we can simplify (1.12) all the way to:

$$\left(\frac{\partial u}{\partial t}+\nabla u\cdot u\right)^i=-\frac{\partial}{\partial x^i}\left[W(\rho)+\rho W'(\rho)\right]. \tag{1.13}$$

Now assuming

$$p(\rho)=\rho^2 W'(\rho),$$

we see that (1.13) is equivalent to

$$\rho\left(\frac{\partial u}{\partial t}+\nabla u\cdot u\right)^i=-\frac{\partial p}{\partial x^i}. \tag{1.14}$$

Multiplying the continuity equation (1.1a) by u and adding the result to (1.14) produces the desired momentum equation (1.1b). Note that when $p=\kappa\rho^\gamma$, the above assumption reduces to taking the potential energy

$$W=\kappa\int_{\rho_0}^{\rho}\sigma^{\gamma-2}d\sigma=\frac{\kappa}{\gamma-1}\left(\rho^{\gamma-1}-\rho_0^{\gamma-1}\right).$$

This is precisely what one finds in derivations from first principles in the barotropic regime (see, e.g., [15] or [85]).

1.4.3 Discussion

Let us summarize what we have done so far. We started with a Lagrangian in the material picture (1.6), and by computing the variation of the associated action, we derived an Euler-Lagrange equation (1.8). Next we showed that by changing variables from the material to spatial pictures, we could derive the standard form of the momentum equation (1.1b) from (1.8).

This begs the question: what happens if we make a change of variables, from material to spatial, at the level of the Lagrangian (1.6) ? Can we then directly derive (1.1b) from this new Lagrangian? It turns out that this idea is in the right direction. We will see that with respect to a certain *reduced Lagrangian* $l(v, \mu)$, the momentum equation (1.1b) will be a special case of the semidirect product Euler-Poincaré equation

$$\frac{\partial}{\partial t}\frac{\delta l}{\delta v} = -\mathrm{ad}_v^* \frac{\delta l}{\delta v} + \frac{\delta l}{\delta \mu} \diamond \mu. \tag{1.15}$$

In order to understand what all the parts of this equation mean, we now review some basic properties of the infinite-dimensional Lie groups and Lie algebras that underly the compressible Euler equations. Our aim is to present a primer for the reader interested in the variational side of semidirect product theory. We will carry out in detail a number of calculations specific to the compressible fluid context; note that all of these calculations can be generalized and adapted to a variety of continuum mechanical settings.

1.5 Spaces and Actions

We start with the group $G = \mathrm{Diff}(M)$. Before continuing, let us mention that, though G is a smooth manifold, G is not a Lie group in the ordinary sense. Multiplication in the group G is given by composition of diffeomorphisms, i.e. given $\gamma_1, \gamma_2 \in G$,

$$\gamma_1 \cdot \gamma_2 = \gamma_1 \circ \gamma_2.$$

Then, for any $\phi \in G$, the group multiplication determines

- the right translation map, $R_\phi : G \to G$, with $R_\phi(\gamma) = \gamma \circ \phi$, and

- the left translation map, $L_\phi : G \to G$, with $L_\phi(\gamma) = \phi \circ \gamma$.

Differentiating R_ϕ and L_ϕ with respect to γ, we obtain for $\xi \in T_\gamma G$,

$$TR_\phi(\gamma) \cdot \xi = \xi \circ \phi$$
$$TL_\phi(\gamma) \cdot \xi = T\phi \circ \xi.$$

For G to be a classical Lie group, both R_ϕ and L_ϕ must be C^∞ maps. We see that TR_ϕ does *not* depend on derivatives of ϕ; repeating the differentiation, it is clear that $R_\phi \in C^\infty(G, G)$. Meanwhile, TL_ϕ involves $T\phi$, which is merely continuous. Hence $L_\phi \in C^1(G, G)$, and G fails to be a Lie group in the classical sense.

However, it is well-known that G can be regarded as a certain kind of generalized Lie group—see [69] for a detailed exposition of such issues. We remark only that all results and calculations that follow do not depend on properties special to classical Lie groups. Finally, the Lie algebra of $\mathrm{Diff}(M)$ is a Lie algebra in the classical sense.

Returning to the general problem, we introduce the vector space $V = C^\infty(M)$. For our purposes, we will write the dual of V as $V^* = \mathrm{Den}(M)$, the space of all C^∞ volume forms on M. If the manifold M has dimension n, then

$$\mathrm{Den}(M) = \{\mu \in \Omega^n(M) \mid \mu(m) \neq 0 \text{ for all } m \in M\}.$$

We pair an element $f \in V$ with $\mu \in V^*$ by

$$\langle \mu, f \rangle = \int_M f\,\mu.$$

1.5.1 G action on V

Let us define an action of G on V by

$$\Phi^V_\gamma(f) = \gamma^*(f).$$

What we have here is a map which associates a linear isomorphism Φ_γ^V to each element $\gamma \in G$. We will sometimes write this action more compactly using concatenation:

$$\gamma f = \Phi_\gamma^V(f) = \gamma^*(f).$$

The action is clearly a right action:

$$\begin{aligned}\gamma_1\gamma_2 f &= (\gamma_1 \circ \gamma_2)^*(f) \\ &= {\gamma_2}^*({\gamma_1}^*(f)) \\ &= \gamma_2\gamma_1 f\end{aligned}$$

1.5.2 Induced G Action on V^*

The G action on V induces a right action on V^* via the inverse of the dual of Φ_γ^V. We now calculate the dual isomorphism $\left(\Phi_\gamma^V\right)^*$. Given $f \in V$ and $\omega \in V^*$,

$$\begin{aligned}\left\langle \left(\Phi_\gamma^V\right)^*(\omega), f \right\rangle &= \left\langle \omega, \Phi_\gamma(f) \right\rangle \\ &= \langle \omega, \gamma^*(f) \rangle \\ &= \int_M \gamma^*(f)\,\omega \\ &= \int_M \gamma^*(f\,\gamma_*(\omega)) \\ &= \int_M f\,\gamma_*(\omega) \\ &= \left\langle \gamma_*(\omega), f \right\rangle.\end{aligned}$$

This shows that $\left(\Phi_\gamma^V\right)^*(\omega) = \gamma_*(\omega)$, and with this, we can define the action of G on V^*:

$$\begin{aligned}\Phi_\gamma^{V^*}(\omega) &= \left[\left(\Phi_\gamma^V\right)^*\right]^{-1}(\omega) \\ &= (\gamma_*)^{-1}(\omega) \\ &= \gamma^*(\omega)\end{aligned}$$

The proof that this is a right action of G on V^* is identical to the case of G acting on V.

1.5.3 Semidirect Product Group and Algebra

We will be concerned with the semidirect product space $G \circledS V^*$, which is a group consisting of the set $G \times V^*$ equipped with the product

$$(\gamma_1, \omega_1) \cdot (\gamma_2, \omega_2) = (\gamma_1 \cdot \gamma_2, \omega_2 + \gamma_2 \omega_1).$$

The Lie algebra $\mathfrak{g}$ of G consists of the space T_eG equipped with the appropriate bracket. As we mentioned earlier, for a given diffeomorphism $\gamma \in G$,

$$T_\gamma G = \{ \text{ smooth maps } X : M \to TM \text{ such that } \pi_M \circ X = \gamma \}.$$

Since the identity element $e \in G$ is the identity map id_M on M

$$\begin{aligned} T_eG &= \{ \text{ smooth maps } X : M \to TM \text{ such that } \pi_M \circ X = \mathrm{id}_M \} \\ &= \mathfrak{X}(M). \end{aligned}$$

Therefore, as a vector space, $\mathfrak{g} = \mathfrak{X}(M)$. Now let us introduce two spaces

$$\begin{aligned} \mathfrak{X}^R &= \text{Lie algebra of all right-invariant vector fields on Diff}(M), \\ \mathfrak{X}^L &= \text{Lie algebra of all left-invariant vector fields on Diff}(M), \end{aligned}$$

as well as the notation

$$\begin{aligned} X_\xi^L &= \text{Left-invariant vector field associated with } \xi, \\ X_\xi^R &= \text{Right-invariant vector field associated with } \xi. \end{aligned}$$

There are several ways to define X_ξ^L and X_ξ^R—there is no unique, canonical choice. For the sake of concreteness, the reader may wish to think of

$$\begin{aligned} X_\xi^L(\phi) &= T\phi \circ \xi, \\ X_\xi^R(\phi) &= \xi \circ \phi. \end{aligned}$$

Now we may state and prove an important fact: the Lie algebra bracket on $\mathfrak{g}$ is *minus* the Jacobi-Lie bracket on M.

Proposition 1. *Suppose we define the bracket on* $\mathfrak{g}$ *via*

$$[\xi, \eta]_{\mathfrak{g}} := [X^L_\xi, X^L_\eta](e). \tag{1.16}$$

Then

$$[\xi, \eta]_{\mathfrak{g}} = -[\xi, \eta]_M,$$

where the bracket $[\xi, \eta]_M$ *is the standard Jacobi-Lie bracket of vector fields on* M*.*

Proof. Let ϕ denote the diffeomorphism on G defined by $\phi(g) = g^{-1}$. One can show that $f : \mathfrak{X}^L \to \mathfrak{X}^R$ defined by $f(X) = \phi_*(X)$ is a Lie algebra isomorphism. Furthermore one can easily compute

$$T_e\phi \cdot \xi = -\xi$$

for all $\xi \in \mathfrak{g}$. Now we have, for $\eta \in \mathfrak{g}$,

$$\begin{aligned} f(X^L_\eta)(e) &= (T\phi \circ X^L_\eta \circ \phi^{-1})(e) \\ &= T_e\phi \circ X^L_\eta(e) \\ &= T_e\phi \circ (T(\mathrm{id}_G) \circ \eta) \\ &= T_e\phi \cdot \eta \\ &= -\eta. \end{aligned}$$

Then

$$\begin{aligned}
[\xi,\eta]_M &= [f(X^L_{-\xi}), f(X^L_{-\eta})]_M \\
&= [\phi_*(X^L_{-\xi}), \phi_*(X^L_{-\eta})]_M \\
&= \phi_*[X^L_{-\xi}, X^L_{-\eta}](e) \\
&= f(X^L_{[-\xi,-\eta]_\mathfrak{g}})(e) \\
&= -[-\xi,-\eta]_\mathfrak{g} \\
&= -[\xi,\eta]_\mathfrak{g}.
\end{aligned}$$

□

The flip side of this proposition is that if we instead define the Lie algebra bracket on $\mathfrak{g}$ using *right*-invariant vector fields, the minus sign disappears. We include this result and its proof for pedagogical purposes—in practice, we *always* define the Lie algebra bracket on $\mathfrak{g}$ using *left*-invariant vector fields.

Proposition 2.

$$[X^R_\xi, X^R_\eta](e) = [\xi,\eta]_M.$$

Proof. Let us verify that X^R_ζ as defined above is right-invariant. Set $R_\gamma(\beta) = \beta\circ\gamma$, right-translation on $\mathrm{Diff}(M)$. Then, for all smooth maps $\zeta_\alpha : M \to TM$ such that $\pi_M \circ \zeta = \alpha$, it is clear that $T_\alpha R_\gamma \cdot \zeta_\alpha = \zeta \circ \gamma$. This then implies

$$\begin{aligned}
T_eR_\gamma \cdot X^R_\zeta(\eta) &= (\zeta\circ\eta)\circ\gamma \\
&= \zeta\circ\eta\circ\gamma = \zeta\circ R_\gamma(\eta) \\
&= X^R_\zeta \circ R_\gamma(\eta),
\end{aligned}$$

as required. With this in mind, we use the local formula for the bracket (again, the Jacobi-Lie bracket of vector fields on $\mathrm{Diff}(M)$):

$$\begin{aligned}
[X^R_\xi, X^R_\eta](e) &= DX^R_\xi(e)\cdot X^R_\eta(e) - DX^R_\eta(e)\cdot X^R_\xi(e) \\
&= D(\xi)\cdot\eta - D(\eta)\cdot\xi
\end{aligned}$$

and this is the local formula for the Jacobi-Lie bracket of vector fields on M. $\square$

Now that we have established the basic Lie algebra structure of $\mathfrak{g}$, we can answer a number of questions regarding induced algebras and induced actions. That is, starting with

- $\gamma \in G = \mathrm{Diff}(M)$,
- $f \in V = C^\infty(M)$,
- $\mu \in V^* = \mathrm{Den}(M)$,
- $\xi, \eta \in \mathfrak{g} = \mathfrak{X}(M)$, and
- $\theta \otimes \nu \in \mathfrak{g}^* = \Omega^1(M) \otimes \mathrm{Den}(M)$,

we will arrive at the following results:

$$
\begin{aligned}
G \text{ action on } V &: \gamma f = \gamma^*(f) \\
G \text{ action on } V^* &: \gamma \mu = \gamma^*(\mu) \\
\text{Lie algebra bracket on } \mathfrak{g} &: [\xi, \eta]_\mathfrak{g} = -[\xi, \eta]_M \\
\mathfrak{g} \text{ action on } V &: \xi f = \pounds_\xi f \\
\mathfrak{g} \text{ action on } V^* &: \xi \mu = \pounds_\xi \mu \\
\mathrm{Ad} \text{ action of } G \text{ on } \mathfrak{g} &: \mathrm{Ad}_\gamma \xi = \gamma_* \xi \\
\mathrm{Ad}^* \text{ action of } G \text{ on } \mathfrak{g}^* &: \mathrm{Ad}^*_\gamma(\theta \otimes \nu) = \gamma^*(\theta \otimes \nu) \\
\mathrm{ad} \text{ action of } \mathfrak{g} \text{ on } \mathfrak{g} &: \mathrm{ad}_\xi \eta = [\xi, \eta]_\mathfrak{g} = -[\xi, \eta]_M \\
\mathrm{ad}^* \text{ action of } \mathfrak{g} \text{ on } \mathfrak{g}^* &: \mathrm{ad}^*_\xi(\theta \otimes \nu) = (\pounds_\xi \theta + \theta \, \mathrm{div}_\nu \xi) \otimes \nu \\
\diamond : V \times V^* \to \mathfrak{g}^* &: f \diamond \mu = \mathbf{d}f \otimes \mu
\end{aligned}
$$

Now we go ahead with systematically filling in the details in this table.

1.5.4 The dual algebra $\mathfrak{g}^*$

We will take $\mathfrak{g}^*$ to be the space of one-form densities:

$$\mathfrak{g}^* = \{\omega \otimes \mu : \omega \in \Omega^1(M), \mu \in \mathrm{Den}(M)\}$$

Then for $\xi \in \mathfrak{g}$, $\theta \otimes \mu \in \mathfrak{g}^*$ the pairing between the two elements is given by

$$\langle \theta \otimes \mu, \xi \rangle = \int_M \theta(\xi)\, \mu.$$

1.5.5 Induced $\mathfrak{g}$ action on V

The action of G on V induces an action of $\mathfrak{g}$ on V. We use $\Phi : G \to \mathbf{GL}(V, V)$ to denote the original action: $\Phi(\gamma) = \gamma^*(\,\cdot\,)$. Then we can compute $T_e\Phi : \mathfrak{g} \to \mathbf{L}(V, V)$ as follows. Fix $\xi \in \mathfrak{g}$ and write ξ as a tangent vector to a curve γ^ϵ on $G = \mathrm{Diff}(M)$:

$$\xi = \left.\frac{d}{d\epsilon}\right|_{\epsilon=0} \gamma^\epsilon \text{ where } \gamma^0 = \mathrm{id}_G.$$

Then for $f \in V$,

$$(T_e\Phi \cdot \xi)(f) = \left.\frac{d}{d\epsilon}\right|_{\epsilon=0} (\gamma^\epsilon)^*(f) = \pounds_\xi f,$$

by the dynamic definition of the Lie derivative. From now on we will use concatenation to denote the action of $\mathfrak{g}$ on V, so that for $\xi \in \mathfrak{g}$ and $f \in V$,

$$\xi f = \pounds_\xi f.$$

1.5.6 Induced $\mathfrak{g}$ action on V^*

The $\mathfrak{g}$-action on V^* is defined as *minus* the dual map (i.e. linear algebraic adjoint) of the $\mathfrak{g}$-action on V. We can also start with the action of G on V^*; this action induces a $\mathfrak{g}$-action on V^*. Both $\mathfrak{g}$-actions on V^* are in fact the same, as we will now show. Take

$\mu \in V^*$, $f \in V$, and $\xi \in \mathfrak{g}$. Then

$$\begin{aligned}\langle \mu, \pounds_\xi f \rangle &= \int_M (\pounds_\xi f)\, \mu \\ &= \int_M (\pounds_\xi (f\mu) - f \pounds_\xi \mu) \\ &= \int_M (\mathbf{d}\mathbf{i}_\xi (f\mu) + \mathbf{i}_\xi \mathbf{d}(f\mu)) - \int_M f \pounds_\xi \mu\end{aligned}$$

Since $f\mu$ is an n-form, $d(f\mu) = 0$. Also, since $\partial M = \emptyset$, Stokes' theorem implies that

$$\int_M \mathbf{d}\mathbf{i}_\xi (f\mu) = 0.$$

Therefore, we have

$$\langle \mu, \pounds_\xi f \rangle = \langle -\pounds_\xi \mu, f \rangle,$$

so the $\mathfrak{g}$-action on V^* is

$$\xi\mu = \pounds_\xi \mu.$$

Verifying the calculation using the alternate approach, we start with the G-action on V^* denoted by $\Psi : G \to GL(V^*, V^*)$, where for each $\gamma \in G$, $\Psi(\gamma) = \gamma^*(\,\cdot\,)$. Then we compute $T_e\Psi : \mathfrak{g} \to L(V^*, V^*)$. As before, we fix $\xi \in \mathfrak{g}$ and write

$$\xi = \left.\frac{d}{d\epsilon}\right|_{\epsilon=0} \gamma^\epsilon \text{ where } \gamma^0 = \mathrm{id}_G.$$

Then

$$(T_e\Psi \cdot \xi)(\mu) = \left.\frac{d}{d\epsilon}\right|_{\epsilon=0} (\gamma^\epsilon)^*(\mu) = \pounds_\xi \mu$$

by the dynamic definition of the Lie derivative.

1.5.7 Adjoint action (group)

We start with the conjugation map

$$I_\gamma(\eta) = \gamma \cdot \eta \cdot \gamma^{-1}.$$

We can then write the adjoint action as

$$\begin{aligned}\mathrm{Ad}_\gamma \xi &= T_e I_\gamma \cdot \xi \\ &= T\gamma \circ \xi \circ \gamma^{-1} \\ &= \gamma_* \xi\end{aligned}$$

1.5.8 Dual of adjoint action (group)

Next on our list is the Ad^* action. We take $\omega \otimes \mu \in \mathfrak{g}^*$, $\xi \in \mathfrak{g}$, and $\gamma \in G$. Then the calculation runs as follows:

$$\begin{aligned}\left\langle \mathrm{Ad}^*_\gamma \omega \otimes \mu, \xi \right\rangle &= \left\langle \omega \otimes \mu, \mathrm{Ad}_\gamma \xi \right\rangle \\ &= \left\langle \omega \otimes \mu, \gamma_* \xi \right\rangle \\ &= \int_M \mathbf{i}_{\gamma_* \xi} \omega \, \mu \\ &= \int_M \gamma_* (\mathbf{i}_\xi \gamma^* \omega) \, \mu \\ &= \int_M \gamma_* (\mathbf{i}_\xi \gamma^* \omega \, \gamma^* \mu) \\ &= \int_M \mathbf{i}_\xi \gamma^* \omega \, \gamma^* \mu \\ &= \left\langle \gamma^* \omega \otimes \gamma^* \mu, \xi \right\rangle \\ &= \left\langle \gamma^* (\omega \otimes \mu), \xi \right\rangle,\end{aligned}$$

so

$$\mathrm{Ad}^*_\gamma (\omega \otimes \mu) = \gamma^* (\omega \otimes \mu).$$

1.5.9 Adjoint action (algebra)

Moving to the Lie algebra, we have another pair of adjoint actions. The adjoint action on the Lie algebra is simply the bracket

$$\mathrm{ad}_\xi : \mathfrak{g} \to \mathfrak{g}$$

$$\mathrm{ad}_\xi \eta = [\xi, \eta]_{\mathfrak{g}} = -[\xi, \eta]_M.$$

1.5.10 Dual of adjoint action (algebra)

Fix $\omega \otimes \mu \in \mathfrak{g}^*$ and $\eta \in \mathfrak{g}$. Then the calculation runs as follows:

$$\begin{aligned}\left\langle \mathrm{ad}^*_\xi(\omega \otimes \mu), \eta \right\rangle &= \left\langle \omega \otimes \mu, \mathrm{ad}_\xi \eta \right\rangle \\ &= -\int_M \omega \cdot [\xi, \eta]_M \, \mu.\end{aligned}$$

Now we apply the following identity:

$$\mathbf{i}_{[\xi,\eta]}\omega = \pounds_\xi \mathbf{i}_\eta \omega - \mathbf{i}_\eta \pounds_\xi \omega.$$

We apply Cartan's magic formula to the first Lie derivative:

$$\pounds_\xi \mathbf{i}_\eta \omega = \mathbf{d}\mathbf{i}_\xi \mathbf{i}_\eta \omega + \mathbf{i}_\xi \mathbf{d}\mathbf{i}_\eta \omega$$

Note that $\mathbf{i}_\eta \omega$ is a function, so $\mathbf{i}_\xi \mathbf{i}_\eta \omega = 0$. Then the divergence theorem yields

$$\int_M \mathbf{i}_\xi \mathbf{d}\mathbf{i}_\eta \omega \, \mu = -\int_M \mathbf{i}_\eta \omega \, \mathrm{div}_\mu \xi \, \mu.$$

Hence

$$-\int_M \omega \cdot [\xi, \eta]_M \, \mu = \int_M \mathbf{i}_\eta (\pounds_\xi \omega + \omega \mathrm{div}_\mu \xi) \, \mu,$$

which means that

$$\mathrm{ad}^*_\xi(\omega \otimes \mu) = (\pounds_\xi \omega + \omega \mathrm{div}_\mu \xi) \otimes \mu.$$

1.5.11 Diamond map

So far we have discussed *actions* of a group on a vector space, or to be precise, maps which give a correspondence between group elements and linear maps of the vector space to itself. Of course, there are other maps of interest. For each $f \in V$, we define $\rho_f : \mathfrak{g} \to V$ by

$$\rho_f(\xi) := \xi f = \pounds_\xi f.$$

Algebraically, what we have is a map $\Psi : V \to L(\mathfrak{g}, V)$ that assigns the linear transformation ρ_f to each element $f \in V$. Then, using the same approach as before, we may compute the dual map $\rho_f^* : V^* \to \mathfrak{g}^*$ as follows: for $a \in V^*$,

$$\begin{aligned}
\left\langle \rho_f^*(a), \xi \right\rangle &= \left\langle a, \rho_f(\xi) \right\rangle \\
&= \left\langle a, \pounds_\xi f \right\rangle \\
&= \int_M \pounds_\xi f \, a \\
&= \int_M \mathbf{d}f \cdot \xi \, a \\
&= \left\langle \mathbf{d}f \otimes a, \xi \right\rangle,
\end{aligned}$$

proving that $\rho_f^*(a) = \mathbf{d}f \otimes a$. Hence we have the associated map $\Psi' : V \to L(V^*, \mathfrak{g}^*)$ that assigns to each $f \in V$ the linear transformation ρ_f^*. In the notation of [40], we have in fact computed the diamond map

$$\begin{aligned}
\diamond &: V \times V^* \to \mathfrak{g}^* \\
f \diamond a &:= \Psi'(f)(a) = \mathbf{d}f \otimes a
\end{aligned}$$

1.6 Symmetries of the Lagrangian

Now that we have determined all the various spaces and actions that will be of relevance in the semidirect product theory, let us now revisit the Lagrangian L defined in (1.6). Recall that we showed that transforming the Euler-Lagrange equation for L from material to spatial variables gives precisely the momentum equation (1.1b). This hints at the fact that we may be able to transform the Lagrangian L itself from the material to the spatial picture. In order to do this, we will have to exploit a certain symmetry of the Lagrangian.

1.6.1 Induced action of G on TG

Before we show that L is invariant under the action of the group $G = \mathrm{Diff}(M)$, we revisit the action of G on TG. For each $\gamma \in G$, let $\Phi_g : G \to G$ denote the action of γ on G.

Specifically,

$$\Phi_\gamma(\eta) = \eta \circ \gamma.$$

This action induces an action on each tangent space $T_\eta G$. Specifically, we have the map $T\Phi_\gamma : TG \to TG$ which acts as follows. Fix $\dot{\eta} \in T_\eta G$. We can write this element of $T_\eta G$ as $\dot{\eta} = \frac{d}{d\epsilon}\big|_{\epsilon=0} \eta^\epsilon$ where $\eta^0 = \eta$. Then

$$\begin{aligned} T_\eta \Phi_\gamma \cdot \dot{\eta} &= \left.\frac{d}{d\epsilon}\right|_{\epsilon=0} \Phi_\gamma(\eta^\epsilon) \\ &= \left.\frac{d}{d\epsilon}\right|_{\epsilon=0} \eta^\epsilon \circ \gamma \\ &= \dot{\eta} \circ \gamma. \end{aligned}$$

Hence we obtain the action of γ on TG:

$$\gamma \cdot (\eta, \dot{\eta}) := T\Phi_\gamma \cdot (\eta, \dot{\eta}) = (\eta \circ \gamma, \dot{\eta} \circ \gamma).$$

Or since $\dot{\eta} \circ \gamma \in T_\eta \mathrm{Diff}(M)$, we can ignore the effect of the action on the base point of the tangent vector and write

$$\gamma \cdot \dot{\eta} = \dot{\eta} \circ \gamma.$$

The group G acts on $\mathrm{Den}(M)$ via pullback, i.e. $\gamma\mu = \gamma^*\mu$. (Though some actions considered are actually *right* actions, we will always write the group variable on the left.) Using this G-action, we may prove

Proposition 3. *The Lagrangian L defined in (1.6) by*

$$L(\dot{\eta}, \mu_0) = \int_M \left[\frac{1}{2}\|\dot{\eta}(m)\|^2 - W\Big([\eta_*\mu_0](\eta(m))\Big)\right] \mu_0.$$

is invariant under the action of $G = \mathit{Diff}(M)$.

Proof. We compute

$$\begin{aligned} L(\gamma \cdot (\dot{\eta}, \mu)) &= L(\dot{\eta} \circ \gamma, \gamma^* \mu) \\ &= \int_M \left[\frac{1}{2} \|\dot{\eta} \circ \gamma\|^2 - W\Big([(\eta \circ \gamma)_* \gamma^* \mu] \, (\eta \circ \gamma) \Big) \right] (\gamma^* \mu). \end{aligned}$$

Recall that

$$(\eta \circ \gamma)_* (\gamma^* \mu) = \eta_* \gamma_* \gamma^* \mu = \eta_* \mu,$$

so that

$$L(\gamma \cdot (\dot{\eta}, \mu)) = \int_M \left[\frac{1}{2} \|\dot{\eta} \circ \gamma(m)\|^2 - W\Big([\eta_* \mu](\eta \circ \gamma)(m) \Big) \right] (\gamma^* \mu).$$

Noting that $\gamma^*(f) = f \circ \gamma$, we see that

$$L(\gamma \cdot (\dot{\eta}, \mu)) = \int_M \gamma^* \left[\frac{1}{2} \|\dot{\eta}(m)\|^2 - W\Big([\eta_* \mu](\eta(m)) \Big) \right] (\gamma^* \mu).$$

Now applying the change of variables theorem, we have

$$L(\gamma \cdot (\dot{\eta}, \mu)) = \int_M \left[\frac{1}{2} \|\dot{\eta}(m)\|^2 - W\Big([\eta_* \mu](\eta(m)) \Big) \right] \mu = L(\dot{\eta}, \mu),$$

proving G-invariance. □

1.6.2 Reduction

Now we can write down the reduced Lagrangian. For $u \in \mathfrak{X}(M)$, and $\mu \in \mathrm{Den}(M)$, set

$$l(u, \mu) = L(u \circ \eta, \eta^* \mu). \tag{1.17}$$

This implicitly defines the spatial velocity as $u = \dot{\eta} \circ \eta^{-1}$. It is clear that u is a tangent vector based at the identity element of $\mathrm{Diff}(M)$. Also note that $\mu = \eta_* \mu_0$ or $\mu_0 = \eta^* \mu$.

$$L(u \circ \eta, \eta^* \mu) = \int_M \left[\frac{1}{2} \|u \circ \eta\|^2 - W\big(\eta^* \mu\big) \right] \eta^* \mu.$$

Using the properties of pullback, we write

$$L(u \circ \eta, \eta^* \mu) = \int_M \eta^* \left(\left[\frac{1}{2} \|u\|^2 - W(\mu) \right] \mu \right).$$

Now by the change of variables theorem,

$$l(u, \mu) = \int_M \left[\frac{1}{2} \|u\|^2 - W(\mu) \right] \mu, \tag{1.18}$$

which is our final expression for the reduced Lagrangian $l : \mathfrak{g} \times V^* \to \mathbb{R}$ where, of course, $\mathfrak{g} \times V^* = \mathfrak{X}(M) \times \mathrm{Den}(M)$. In what follows, we will use the following notation: $\mu = \rho \, \mathrm{d}^n x$, where $\mathrm{d}^n x$ is the canonical n-form on M and $\rho \in C^\infty(M)$.

1.7 Variational Principles in the Spatial Picture

The reduced Lagrangian $l(u, \mu)$ defined in (1.18) is clearly intended to be the Lagrangian for the momentum equation (1.1b) in spatial variables. To realize this intention, we must go about the business of deriving the momentum equation from the reduced Lagrangian, which we demonstrate in two ways.

1.7.1 Equations of Motion I: Variational Principle

Our first path from the reduced Lagrangian l to the equations of motion will involve the reduced variational principle. That is, given the action

$$s(u, \mu) = \int_0^T l(u, \mu) \, dt, \tag{1.19}$$

we wish to solve the variational equation

$$\delta s(u, \mu) \cdot (\delta u, \delta \mu) = 0 \tag{1.20}$$

for critical points (u, μ). Previously, in the material picture, we took *free* variations $\delta \eta$. That is, as long as the variations η^ϵ were fixed on the boundary ∂M, we were happy to

choose any tangent vector

$$\delta\eta = \left.\frac{d}{d\epsilon}\right|_{\epsilon=0} \eta^\epsilon$$

as a variation. Now that we have changed our point of view and reduced the Lagrangian, we must keep in mind the relationships between the spatial and material pictures:

$$u = \dot{\eta} \circ \eta^{-1}$$
$$\mu = \eta_* \mu_0.$$

A natural question arises: even if $\delta\eta$ is allowed to vary freely, are there restrictions on the possible variations δu and $\delta\mu$? The answer is that variations in these variables will be constrained to be of the specific form

$$\delta u = \frac{\partial w}{\partial t} + [u, w],$$
$$\delta\mu = -\pounds_w \mu,$$

where $w = \delta\eta \circ \eta^{-1}$ vanishes at the endpoints $t = 0, T$.

1.7.2 Derivation of Constraints

Before proceeding, we confirm that the above constraints are natural. They arise precisely from the relationship between the unreduced (η, ρ_0) and reduced (v, ρ) variables. Free variations of the former correspond to constrained variations of the latter.

To see this, we start with the relationship between the reduced and unreduced velocity variables, $u = \dot{\eta} \circ \eta^{-1}$, and calculate variations. First we write

$$u^\epsilon = \dot{\eta}^\epsilon \circ (\eta^\epsilon)^{-1},$$

which implies that

$$\delta u := \left.\frac{d}{d\epsilon}\right|_{\epsilon=0} u^\epsilon = \delta\dot{\eta} \circ \eta^{-1} + \nabla\dot{\eta} \cdot \left.\frac{d}{d\epsilon}\right|_{\epsilon=0} (\eta^\epsilon)^{-1}$$

Now because $(\eta^\epsilon)^{-1} \circ (\eta^\epsilon) = \mathrm{id}$, we can differentiate both sides with respect to ϵ at $\epsilon = 0$. This yields the following formula:

$$\left.\frac{d}{d\epsilon}\right|_{\epsilon=0} (\eta^\epsilon)^{-1} = -\nabla\eta^{-1} \cdot \delta\eta \circ \eta^{-1}.$$

Applying this to our earlier expression for δu, we have

$$\delta u = \delta\dot{\eta} \circ \eta^{-1} - \nabla\dot{\eta} \cdot \nabla\eta^{-1} \cdot \delta\eta \circ \eta^{-1}.$$

Defining $w = \delta\eta \circ \eta^{-1}$, we find that

$$\frac{\partial w}{\partial t} = \delta\dot{\eta} \circ \eta^{-1} + \nabla\delta\eta \cdot \frac{\partial}{\partial t}(\eta^{-1}).$$

We can calculate the very last term in this formula by differentiating $\eta^{-1} \circ \eta = \mathrm{id}$ with respect to t, resulting in

$$\frac{\partial}{\partial t}\eta^{-1} = -\nabla\eta^{-1} \cdot \dot{\eta} \circ \eta^{-1}.$$

Substituting into our earlier expression for $\partial w/\partial t$, we have

$$\begin{aligned}\frac{\partial w}{\partial t} &= \delta\dot{\eta} \circ \eta^{-1} - \nabla\delta\eta \cdot \nabla\eta^{-1} \cdot \dot{\eta} \circ \eta^{-1} \\ &= \delta\dot{\eta} \circ \eta^{-1} - \nabla w \cdot u.\end{aligned}$$

Using what we have so far, along with the fact that

$$\nabla u = \nabla\dot{\eta} \cdot \nabla\eta^{-1},$$

we have

$$\delta u = \frac{\partial w}{\partial t} + \nabla w \cdot u - \nabla u \cdot w = \frac{\partial w}{\partial t} + [u, w]. \tag{1.21}$$

Next, using basic facts about the pullback of volume forms, we write

$$\rho \circ \eta = \frac{1}{\det \nabla\eta}\rho_0.$$

Taking variations,

$$\delta\rho \circ \eta + \nabla\rho \cdot \delta\eta = -\frac{\rho_0}{\det \nabla\eta}(\nabla\eta)^{-T} \cdot \nabla\delta\eta,$$

where the final $\cdot$ denotes pairing of linear transformations. That is, $A \cdot B = \mathrm{tr}(A^T B)$. Now we solve for $\delta\rho$, obtaining

$$\delta\rho = -\nabla\rho \cdot w - \rho\,\mathrm{tr}(\nabla\eta^{-1} \cdot \nabla\delta\eta).$$

In coordinates,

$$\begin{aligned}\mathrm{tr}(\nabla\eta^{-1} \cdot \nabla\delta\eta) &= (\nabla\eta^{-1})^i_j (\nabla\delta\eta)^i_j \\ &= \frac{\partial(\eta^{-1})^i}{\partial x^j} \cdot \frac{\partial(\delta\eta)^j}{\partial X^i} \\ &= \frac{\partial(\delta\eta \circ \eta^{-1})^j}{\partial x^j} = \mathrm{div}(w).\end{aligned}$$

Therefore,

$$\delta\rho = -\nabla\rho \cdot w - \rho\mathrm{div}(w) = -\mathrm{div}(\rho w). \tag{1.22}$$

Note from (1.21) and (1.22) that when w vanishes, both δu and $\delta\rho$ vanish. Then it is sufficient, when we vary the action (1.19), to consider variations δu and $\delta\rho$ such that w vanishes at the endpoints $t = 0$ and $t = T$.

1.7.3 Critical Points of l

Having verified the constraints, let us directly compute the variational equation (1.20) with the Lagrangian and action defined in (1.18-1.19):

$$\delta s \cdot (\delta u, \delta\rho) = \int_0^T \int_M \left\{ \left[u^i \delta u^i - W'(\rho)\delta\rho \right] \rho + \left[\frac{1}{2}\|u\|^2 - W \right] \delta\rho \right\} dx\, dt$$

Using the constraints (1.21) and (1.22), we have

$$\begin{aligned}\delta s \cdot (\delta u, \delta\rho) = \int_0^T \int_M \Bigg\{ & u^i \left[\frac{\partial w^i}{\partial t} + \frac{\partial w^i}{\partial x^j} u^j - \frac{\partial u^i}{\partial x^j} w^j \right] + W'(\rho) \frac{\partial}{\partial x^i}(\rho w^i) \Bigg\} \rho \\ & - \left(\frac{1}{2}\|u\|^2 - W \right) \frac{\partial}{\partial x^i}(\rho w^i) \Bigg\}\, dx\, dt.\end{aligned}$$

Integrating by parts and assuming that w vanishes at the endpoints $t = 0$ and $t = T$, we have

$$\delta s \cdot (\delta u, \delta\rho) = \int_0^T \int_M \left\{ -\frac{\partial}{\partial t}(u^i\rho)w^i - \frac{\partial}{\partial x^j}(u^i u^j \rho)w^i - u^i \frac{\partial u^i}{\partial x^j}\rho w^j \right.$$
$$\left. -\frac{\partial}{\partial x^i}[W'(\rho)\rho]\rho w^i + \frac{\partial}{\partial x^i}\left[\frac{1}{2}\|u\|^2 - W(\rho)\right]\rho w^i \right\} dx\, dt.$$

Since $\delta s \cdot (\delta u, \delta\rho) = 0$ for all w, we must have

$$-\frac{\partial u^i}{\partial t}\rho - u^i\frac{\partial \rho}{\partial t} - u^i_{,j}u^j\rho - u^i u^j_{,j}\rho - u^i u^j \rho_{,j}$$
$$- u^j u^j_{,i}\rho - [W'(\rho)\rho]_{,i}\rho + [u^j u^j_{,i} - W'(\rho)\rho_{,i}]\rho = 0,$$

where we have used subscripts to denote spatial derivatives. Noting cancellations and regrouping the remaining terms, we have

$$\frac{\partial}{\partial t}\left(\rho u^i\right) + \frac{\partial}{\partial x^j}\left(\rho u^i u^j\right) = -\rho\frac{\partial}{\partial x^i}[W(\rho) + \rho W'(\rho)].$$

By again taking $p(\rho) = \rho^2 W'(\rho)$, we will find that this equation is equivalent to

$$\frac{\partial}{\partial t}\left(\rho u^i\right) + \frac{\partial}{\partial x^j}\left(\rho u^i u^j\right) = -\frac{\partial p}{\partial x^i}. \tag{1.23}$$

This is precisely the momentum equation (1.1b).

1.7.4 Equations of Motion II: Euler-Poincaré Equation

We can also derive the equations of motion from the reduced Lagrangian by making use of the Euler-Poincaré equation from semidirect product reduction theory:

$$\frac{\partial}{\partial t}\frac{\delta l}{\delta u} = -\mathrm{ad}^*_u\frac{\delta l}{\delta u} + \frac{\delta l}{\delta \mu}\diamond \mu.$$

1.7.5 Derivatives of l

In order to apply the Euler-Poincaré equations of motion, we must first calculate the variational derivatives of $l(u, \mu)$. First

$$\begin{aligned}
\left\langle \frac{\delta l}{\delta u}, \xi \right\rangle &= \left.\frac{d}{d\epsilon}\right|_{\epsilon=0} l(u_\epsilon, \mu) \text{ where } \left.\frac{d}{d\epsilon}\right|_{\epsilon=0} u_\epsilon = \xi \\
&= \frac{1}{2} \left.\frac{d}{d\epsilon}\right|_{\epsilon=0} \int_M \|u_\epsilon\|^2 \ \mu \\
&= \frac{1}{2} \int_M \left.\frac{d}{d\epsilon}\right|_{\epsilon=0} \|u_\epsilon\|^2 \ \mu \\
&= \int_M \langle u, \xi \rangle_{\|\cdot\|} \, \mu \\
&= \langle u^\flat \otimes \mu, \xi \rangle
\end{aligned}$$

where the $\flat$ operation is pointwise with respect to $\langle \cdot, \cdot \rangle_{\|\cdot\|}$, i.e.

$$u^\flat(\xi) = \langle u, \xi \rangle_{\|\cdot\|}.$$

Hence

$$\frac{\delta l}{\delta u} = u^\flat \otimes \mu. \tag{1.24}$$

Note that W only depends on the ρ "part" of μ, not on the canonical volume form $\mathrm{d}^n x$ part. To formalize this notion, we write

$$W(\mu) = W(\rho \, \mathrm{d}^n x) = \widehat{W}(\rho).$$

Here $W : \mathrm{Den}(M) \to C^\infty(M)$. Using $\widehat{W}$, we can write an alternate expression for $\mathbf{D}W : \mathrm{Den}(M) \to \mathbf{L}(\mathrm{Den}(M), C^\infty(M))$ as follows. Let $\mu = \rho \, \mathrm{d}^n x$ and $\mu' = \rho' \, \mathrm{d}^n x$. Then

$$\begin{aligned}
\mathbf{D}W(\mu) \cdot \mu' &= \left.\frac{d}{d\epsilon}\right|_{\epsilon=0} W(\mu + \epsilon \mu') \\
&= \left.\frac{d}{d\epsilon}\right|_{\epsilon=0} \widehat{W}(\rho + \epsilon \rho') \\
&= \frac{\partial \widehat{W}}{\partial \rho}(\rho) \cdot \rho' = \widehat{W}'(\rho) \rho'.
\end{aligned}$$

Furthermore, for densities μ, μ_1, and μ_2, we have

$$\begin{aligned}
\langle \mathbf{D}W(\mu)\cdot\mu_1,\mu_2\rangle &= \left\langle \widehat{W}'(\rho)\rho_1,\mu_2\right\rangle \\
&= \int_M \widehat{W}'(\rho)\rho_1\mu_2 \\
&= \int_M \widehat{W}'(\rho)\rho_1\rho_2\,\mathrm{d}^n x \\
&= \int_M \widehat{W}'(\rho)\rho_2\mu_1 \\
&= \left\langle \widehat{W}'(\rho)\rho_2,\mu_1\right\rangle \\
&= \langle \mathbf{D}W(\mu)\cdot\mu_2,\mu_1\rangle
\end{aligned}$$

Armed with these facts, we proceed:

$$\begin{aligned}
\left\langle \frac{\delta l}{\delta\mu},\alpha\right\rangle &= \left.\frac{d}{d\epsilon}\right|_{\epsilon=0} l(u,\mu^\epsilon) \text{ where } \left.\frac{d}{d\epsilon}\right|_{\epsilon=0}\mu^\epsilon=\alpha \\
&= \int_M \left.\frac{d}{d\epsilon}\right|_{\epsilon=0}\left[\frac{1}{2}\|u\|^2 - W(\mu^\epsilon)\right]\mu^\epsilon \\
&= \int_M \left[\frac{1}{2}\|u\|^2 - W(\mu)\right]\alpha + \mu\left[-\mathbf{D}W(\mu)\cdot\alpha\right] \\
&= \left\langle \frac{1}{2}\|u\|^2 - W(\mu),\alpha\right\rangle - \langle \mathbf{D}W(\mu)\cdot\alpha,\mu\rangle \\
&= \left\langle \frac{1}{2}\|u\|^2 - W(\mu) - \mathbf{D}W(\mu)\cdot\mu,\alpha\right\rangle,
\end{aligned}$$

so we must have

$$\frac{\delta l}{\delta\mu} = \frac{1}{2}\|u\|^2 - \widehat{W}(\rho) - \widehat{W}'(\rho)\cdot\rho. \tag{1.25}$$

1.7.6 Plugging into Euler-Poincaré

In an earlier calculation (see §4.10), we established that for $\xi\in\mathfrak{g}$ and $\theta\otimes\nu\in\mathfrak{g}^*$,

$$\mathrm{ad}^*_\xi(\theta\otimes\nu) = (\pounds_\xi\theta + \theta\mathrm{div}_\nu\xi)\otimes\nu.$$

Using this and the variational derivative (1.24) in the semidirect product EP equation (1.15) gives the evolution equation

$$\frac{\partial}{\partial t}(u^\flat \otimes \mu) = -(\pounds_u u^\flat + u^\flat \mathrm{div}_\mu u) \otimes \mu + \frac{\delta l}{\delta \mu} \diamond \mu. \tag{1.26}$$

Because $\mathrm{div}_\mu u$ is a scalar function, the two objects $-u^\flat \mathrm{div}_\mu u \otimes \mu$ and $-u^\flat \otimes \mathrm{div}_\mu u \mu$ are equivalent. Thus when we subtract from (1.26) the mass conservation equality

$$u^\flat \otimes \frac{\partial \mu}{\partial t} = -u^\flat \otimes (\mathrm{div}_\mu v)\mu,$$

we find that (1.26) simplifies to

$$\frac{\partial u^\flat}{\partial t} \otimes \mu = -\pounds_u u^\flat \otimes \mu + \mathbf{d}\left(\frac{\delta l}{\delta \mu}\right) \otimes \mu. \tag{1.27}$$

Note that we have used an earlier computation (see §4.11) to evaluate the diamond map. Now we apply Cartan's magic formula and the variational derivative (1.25) to obtain

$$\frac{\partial u^\flat}{\partial t} + \mathbf{d}\mathbf{i}_u u^\flat + \mathbf{i}_u \mathbf{d} u^\flat = \frac{1}{2}\mathbf{d}\mathbf{i}_u u^\flat - \mathbf{d}W(\mu) - \mathbf{d}(\mathbf{D}W(\mu) \cdot \mu). \tag{1.28}$$

A coordinate computation verifies that

$$\frac{1}{2}\mathbf{d}\mathbf{i}_u u^\flat + \mathbf{i}_u \mathbf{d} u^\flat = [\nabla u \cdot u]^\flat.$$

We use this in (1.28) and take sharps of both sides, resulting in

$$\frac{\partial u}{\partial t} + \nabla u \cdot u = -\left[\mathbf{d}W(\mu) + \mathbf{d}(\mathbf{D}W(\mu) \cdot \mu)\right]^\sharp. \tag{1.29}$$

As before, if we take $p(\rho) = \rho^2 W'(\rho)$ we find that the equation reduces to

$$\frac{\partial u}{\partial t} + \nabla u \cdot u = -\frac{\nabla p}{\rho}. \tag{1.30}$$

This is identical to (1.14), and as we explained earlier, taken together with the continuity equation (1.1a), it is also equivalent to the momentum equation (1.1b).

1.8 Discussion

We have shown three different ways of deriving (1.1b) from a variational principle:

A1. Start with the material Lagrangian (1.6), apply Hamilton's principle, and transform the resulting equation (1.8) to the spatial picture.

A2. Start with the spatial Lagrangian (1.18) and apply Hamilton's principle with constraints (1.21) and (1.22).

A3. Start with the spatial Lagrangian (1.18), compute variational derivatives (1.24) and (1.25), and then compute the semidirect product EP equation (1.15).

Note that this constitutes a "verification by hand" of most of the Euler-Poincaré Theorem for Continua (see Theorem 6.1 in [40]), when this theorem is applied specifically to the compressible fluid system. We remark that the general theorem is quite powerful. The reader who has followed the Lie-algebraic calculations in the previous sections is well-equipped to apply the theorem to any problem of interest in continuum mechanics.

Of the three approaches, A1 is the least preferable. The coordinate computations were very complicated, and indeed they were *not* valid except in flat Euclidean space. We could have introduced tools from Riemannian geometry to effect a global computation, and for the barotropic compressible Euler equation, this would have been reasonable. However, we can easily envision a case where the Lagrangian L may be more complicated and contain higher-derivative terms. This would be the case if say, we carried out a series expansion inside the Lagrangian (1.6). Then either the local or global approaches to writing the Euler-Lagrange equation for L would be intractable to implement.

Approaches A2 and A3 require roughly the same amount of work. Method A2 can be applied very naïvely, while A3 requires knowledge of the computations performed in Section 1.5. In the next chapters, we shall use A2 especially if we wish to change one or more of the constraints on the variations. In other cases, we shall use A3.

Both approaches A2 and A3 have many attractive features: we do not have to worry about the calculation being valid only locally. Because the semidirect product EP equation arises via symmetry reduction, the equation automatically makes sense in both the

material and spatial pictures. Finally, because we are working in the spatial picture, the variables u and ρ belong to vector spaces and we can add and multiply with impunity. Note that we *cannot* add two diffeomorphisms ϕ and ψ and hope to get a diffeomorphism back, *even* in Euclidean space[1]. As we will see, carrying out a certain kind of (additive) Taylor expansion inside the Lagrangian is one way of deriving new model equations for compressible fluid motion.

[1]Consider $\phi(x) = x$ and $\psi(x) = -x$. Both are diffeomorphisms of $\mathbb{R}$, but their sum $(\phi + \psi)(x) = 0$ is not even invertible.

Chapter 2

Lagrangian Averaging for Compressible Fluids

2.1 Introduction

Historical Remarks. The incompressible case will be discussed first. The Lagrangian averaged Euler (LAE-α) equations for average incompressible ideal fluid motion first appeared in the context of averaged fluid models in [40, 39]. Dissipation was added later to produce the Lagrangian averaged Navier–Stokes (LANS-α) equations, also known as the Navier–Stokes-α equations.[1]

Remarkably, the LAE-α equations are mathematically identical to the inviscid second grade fluid equations introduced in [72], except for the fact that the parameter α is interpreted differently in the two theories. In the case of LAE-α and LANS-α, the parameter α is a spatial scale below which rapid fluctuations are smoothed by linear and nonlinear dispersion.

As in, for example, the work of [86] on nonlinear waves, the distinctive feature of the Lagrangian averaging approach is that averaging is carried out at the level of the variational principle and not at the level of the Euler or Navier–Stokes equations, which is the traditional averaging or filtering approach used for both the Reynolds averaged Navier–Stokes (RANS) and the large eddy simulation (LES) models. As such, the variational

[1] Sometimes the term "viscous Camassa–Holm (VCH) equations" ([12]) has been used, but this terminology is a little unfortunate since the n-dimensional version of the CH equations, also known as the EPDiff equations, arise via Euler–Poincaré reduction of H^1 geodesics on the group of all diffeomorphisms and *not* the volume-preserving ones (see [37]).

procedure does not add any *artificial* viscosity, a physical reason to consider the LAE-α or LANS-α equations as good models for incompressible turbulent flow. Moreover, it has been proven that the α models are computationally very attractive (see [13, 65]).

Although sharing the same general technique (use of averaging and asymptotic methods in the variational formulation), several alternative derivations of incompressible LAE-α equations exist in the literature. One of these derivations (see [34]) uses the generalized Lagrangian mean (GLM) theory developed in [2].

An alternative derivation of the incompressible LAE-α and LANS-α equations was given in [60] by using an ensemble average over the set of solutions of the Euler equations with initial data in a phase-space ball of radius α while treating the dissipative term via stochastic variations. The derivation also uses a turbulence closure that is based on the Lagrangian fluctuations, namely a generalization of the frozen turbulence hypothesis of Taylor (see [83]).

Rigorous analysis aimed at proving global well-posedness and regularity of the three-dimensional isotropic and anisotropic LANS-α equations can be found in, for example, [27, 59, 60]. However, global existence for the inviscid three-dimensional LAE-α remains an open problem.

From a computational viewpoint, numerical simulations of the α models (see [13, 65]) show that the LANS-α equations give computational savings comparable to LES models for forced and decaying turbulent flows in periodic domains. For wall-bounded flows, it is expected that either the anisotropic model or a model with varying α needs to be used; the computational efficacy of these methods on such flows remains to be demonstrated.

As far as the compressible case is concerned, the only papers we know of are those of D. D. Holm [34, 35]. Later, we shall discuss the relation between Holm's work and the present work.

We refer the interested reader to [59, 60] for a more detailed history of the PDE analysis for LAE-α and LANS-α equations and to [65] for a survey and further references about the numerical aspects of these models.

Motivation. In compressible flows there are two major problems at higher wave numbers, or small scales, that require special attention. These are (a) turbulence for high Reynolds number flows (common with incompressible flows) and (b) strong shocks. In both cases the challenge lies in the appropriate representation of small-scale effects. For turbulence, the energy cascade to smaller scales can be balanced by viscous dissipation, resulting in the viscous regularization of the Euler equations.

Historically, viscous dissipation has been used to regularize shock discontinuities. This includes adding to the Euler equation *nonphysical and artificial viscous terms* and Fourier's law for heat transfer in the shock region (see, e.g., [53, 75]). This way, the steepening effect of the nonlinear convective term is balanced by dissipation. We believe that Lagrangian averaging is a reasonable alternative way to regularize shock waves. The net effect of Lagrangian averaging is to add dispersion instead of dissipation to the Euler equations; that is, one adds terms that redistribute energy in a nonlinear fashion. In other, rather different situations, the technique of balancing a nonlinear convective term by dispersive mechanisms was used by [52] for the KdV equation and by [47, 45] for plasma flows.

The competition between nonlinearity and dispersion has, of course, resulted in remarkable discoveries, the most famous being solitons, localized waves that collide elastically, suffering only a shift in phase. The robustness of solitons in overcoming strong perturbations is largely due to a balance between nonlinearity and linear dispersion. Note that in Lagrangian averaging the energy redistribution mechanism that is introduced is nonlinear and might yield other interesting features that warrant further investigation.

Another feature of the compressible Lagrangian averaged Euler-α equations is that in turbulent flows with shocks the effect of shocks and turbulence are simultaneously modeled by the same technique, namely the Lagrangian averaging method.

Issues Addressed in This Chapter. In this chapter we apply the averaged Lagrangian methodology to derive the isotropic and anisotropic averaged models for compressible Euler equations.

One goal of this chapter is to present a clear derivation of the averaged equations. We

are particularly interested in separating the two issues of averaging and modeling. In the derivation, a new ensemble averaging technique is proposed and investigated. Instead of taking clouds of initial conditions, as in [60], we average over a tube of trajectories η^ϵ centered around a given Lagrangian flow η. The tube is constructed by specifying the Lagrangian fluctuations $\xi^\epsilon = \eta^\epsilon \circ \eta^{-1}$ at $t = 0$ and providing a *flow rule* which evolves them to all later times. The choice of flow rule is a precise modeling assumption which brings about closure of the system.

For the incompressible case we assume that fluctuations are Lie advected by the mean flow (or frozen into the mean flow as divergence-free vector fields), and we obtain both the isotropic and the anisotropic versions of the LAE-α equations. The advection hypothesis is the natural extension to vector fields of the classical frozen turbulence hypothesis of Taylor (see [83]) stated for scalar fluctuations.

The second goal of this work is to extend the derivation to barotropic compressible flows. This problem has already been considered by Holm (see [35]) in the context of GLM motion. In this work, an alpha model appears as a GLM fluid theory with an appropriate Taylor hypothesis closure. However, even though [35] enumerates several frozen-in closure hypotheses, the averaged equations are derived only for the case when the fluctuations are parallel transported by the mean flow. In our work we will consider a more general advection hypothesis to study the compressible anisotropic case. In addition, a physically based new flow rule is introduced to deal with the isotropic case.

The averaging technique consists of expanding the original Lagrangian with respect to a perturbation parameter ϵ, truncating the expansion to $O(\epsilon^2)$ terms, and then taking the average. It turns out that the averaged compressible Lagrangian depends on the Lagrangian fluctuations ξ' only through three tensor quantities which are quadratic in ξ'. In the terminology of [35] these tensors represent the second-order statistics of the Lagrangian fluctuations. Evolution equations for these tensors are derived from a core modeling assumption: a prescribed *flow rule* for the time evolution of the fluctuations ξ'. The flow rule gives us closure, allowing us to apply Hamilton's principle to the averaged Lagrangian and thereby derive an equation for the mean velocity u.

The organization of the rest of the chapter is as follows. In section 2.2 we describe a

general procedure for Lagrangian ensemble averaging. This procedure is then applied to the action for incompressible fluids in section 2.3 to demonstrate our derivation technique. The general procedure is applied again in section 2.4, this time to the more complex case of barotropic compressible fluids. Section 2.5 is devoted to modeling issues; here the strategy of modeling the evolution of Lagrangian fluctuations ξ' using flow rules is discussed in detail. In section 2.6 we derive the averaged equations for incompressible and compressible models in both isotropic and anisotropic versions. The appendix provides technical details about the fluctuation calculus used throughout the chapter.

Main Results. The main result of this chapter is the derivation of compressible Lagrangian averaged Euler equations with

- anisotropic modeling of fluid fluctuations—see equations (2.49);
- isotropic modeling of fluid fluctuations—see equations (2.51).

In addition, we provide an improved derivation of the *incompressible* isotropic and anisotropic LAE-α equations.

2.2 General Lagrangian Averaging

A mathematical setting for a certain class of compressible fluid flow problems will be given first. After describing the general procedure for Lagrangian averaging, the specific case of the Euler action for fluids will be considered.

Let M be an open subset of $\mathbb{R}^N$ representing the containing space of a fluid. Suppose we are given a Lagrangian for a compressible fluid, $L(\psi, \dot{\psi}, \mu_0)$, where $\psi \in \mathrm{Diff}(M)$, the space of diffeomorphisms of M, $(\psi, \dot{\psi}) \in T\mathrm{Diff}(M)$, and $\mu_0 \in \Lambda^N(M)$, the space of N-forms on M. Fix a time interval $[0, T]$, and let $\mathcal{C}(\mathrm{Diff}(M))$ be the path space of smooth maps from $[0, T]$ into $\mathrm{Diff}(M)$. Then the action $S : \mathcal{C}(\mathrm{Diff}(M)) \times \Lambda^N(M) \to \mathbb{R}$ is

$$S(\eta, \mu_0) = \int_0^T L(\eta(t), \dot{\eta}(t), \mu_0)\, dt.$$

We seek an averaged action $S^\alpha(\eta, \mu_0)$, where α is a length scale characterizing the coarseness of the average. Taking η and μ_0 as given, we shall describe how to compute $S^\alpha(\eta, \mu_0)$.

Remark. It is important to emphasize that for both S and S^α, η is merely a test curve. It is *not* an extremal of the action S. We are trying to average the action S itself, not any fluid dynamical PDE or the solutions of such a PDE. Our final product S^α should not depend at all on an initial choice of the test curve η.

Tube Initialization. The first step is to take $\xi^\epsilon(x,t)$ to be a family of diffeomorphisms about the identity. That is,

$$\begin{aligned} \text{for each } \epsilon \geq 0, \quad & \xi^\epsilon(\cdot, t) \in \mathrm{Diff}(M) \text{ for all } t, \text{ and} \\ \text{at } \epsilon = 0, \quad & \xi^\epsilon(x,t) = x \text{ for all } x, t. \end{aligned}$$

Define the vector fields ξ' and ξ'' via

$$\xi' = \frac{\partial}{\partial \epsilon}\bigg|_{\epsilon=0} \xi^\epsilon \quad \text{and} \quad \xi'' = \frac{\partial^2}{\partial \epsilon^2}\bigg|_{\epsilon=0} \xi^\epsilon.$$

Use ξ^ϵ to construct a tube of material deformation maps that are close to η by letting $\eta^\epsilon(X,t) = \xi^\epsilon(\eta(X,t),t)$, or, written more compactly,

$$\eta^\epsilon = \xi^\epsilon \circ \eta. \tag{2.1}$$

Here, X is a material point in the reference configuration. Define the spatial velocity by $u^\epsilon(x,t) = \dot{\eta}^\epsilon((\eta^\epsilon)^{-1}(x,t),t)$, where η^ϵ is a given material deformation map. Compactly written, this reads as

$$u^\epsilon = \dot{\eta}^\epsilon \circ (\eta^\epsilon)^{-1}. \tag{2.2}$$

The map u^ϵ is a time-dependent vector field on M; i.e., for each $\epsilon \geq 0$, and for all t, $u^\epsilon(\cdot, t) \in \mathfrak{X}(M)$.

Averaging. The existence of an averaging operation $\langle \cdot \rangle$ will now be postulated. The properties this operation is required to satisfy and an example of such an operation will be given shortly.

Relationship Between u^ϵ and u. It is desirable to have the fluctuations ξ^ϵ centered, on average, about the identity: $\langle \xi^\epsilon(x,t) \rangle = x$ for all positions x at all times t. What is actually needed is that for $n \geq 1$,

$$\left\langle \frac{\partial^n \xi^\epsilon}{\partial \epsilon^n}\bigg|_{\epsilon=0} \right\rangle = 0. \tag{2.3}$$

In other words, the nth-order fluid fluctuation vector fields should all have mean zero. Restricting the map to be centered about the identity means simply that the average will not be skewed in an arbitrary direction. ¿From (2.2) and (2.3) one can derive

$$\langle u^\epsilon \circ \xi^\epsilon(x,t) \rangle = u(x,t). \tag{2.4}$$

Equation (2.4) shows in which sense the average of u^ϵ is u in a Lagrangian-mean theory defined by $\langle \eta^\epsilon(\cdot,t) \rangle = \eta(\cdot,t)$. This equation is closely connected with the generalized Lagrangian-mean description of [2], where the Lagrangian-mean velocity $\bar{u}^L$ and the fluctuating Eulerian velocity u^ξ are related in a similar way.

Density. For the nonaveraged Lagrangian L, μ_0 is a parameter in the sense of Lagrangian semidirect product theory; see Chapter 1 and [58, 39]. The physical interpretation of μ_0 is as follows. Since μ_0 is an N-form on M, it can be written as

$$\mu_0 = \rho_0 \, dx^1 \wedge \cdots \wedge dx^N,$$

where ρ_0 is a smooth function on M. Now $\rho_0(X)$ is the density of the fluid at the material point X in the reference configuration. This is in contrast to the spatial density $\rho^\epsilon(x,t)$, which gives us the density of the fluid at the spatial point x at time t. Defining

$$\mu^\epsilon = \rho^\epsilon \, dx^1 \wedge \cdots \wedge dx^N, \tag{2.5}$$

one has the relationship

$$(\eta^\epsilon)_*\mu_0 = \mu^\epsilon. \tag{2.6}$$

Fluctuation Calculus. Because u^ϵ and ρ^ϵ will be expanded, the ϵ-derivatives of u^ϵ and ρ^ϵ need to be calculated. First, define

$$u' = \frac{\partial}{\partial\epsilon}\bigg|_{\epsilon=0} u^\epsilon \qquad \text{and} \qquad u'' = \frac{\partial^2}{\partial\epsilon^2}\bigg|_{\epsilon=0} u^\epsilon. \tag{2.7}$$

By differentiating (2.2), one finds expressions for u' and u'' in terms of u, ξ', and ξ''. The calculations can be performed intrinsically using Lie derivative formulae—the results, as found in [60], are

$$u' = \partial_t\xi' + [u, \xi'], \tag{2.8a}$$

$$u'' = \partial_t\xi'' + [u, \xi''] - 2\nabla u' \cdot \xi' - \nabla\nabla u(\xi', \xi'). \tag{2.8b}$$

In these formulas, the bracket $[x, y] = \pounds_x y$ is the standard Jacobi–Lie bracket of vector fields on M (see, for example, [1]). Next, define

$$\rho' = \frac{\partial}{\partial\epsilon}\bigg|_{\epsilon=0} \rho^\epsilon \qquad \text{and} \qquad \rho'' = \frac{\partial^2}{\partial\epsilon^2}\bigg|_{\epsilon=0} \rho^\epsilon. \tag{2.9}$$

One obtains expressions for ρ' and ρ'' in terms of ρ, ξ', and ξ'' by differentiating (2.6) (see the appendix for the detailed calculations). The results are

$$\rho' = -\operatorname{div}(\rho\xi'), \tag{2.10a}$$

$$\rho'' = \operatorname{div}(\operatorname{div}(\rho\xi' \otimes \xi')) - \operatorname{div}(\rho\xi''). \tag{2.10b}$$

Averaging Operation. In the above development, an averaging operation has been implicitly used. The properties it is required to satisfy will now be spelled out. Let $\mathcal{F}(Y)$ mean the space of smooth, real-valued functions on a manifold Y. If Y is infinite-dimensional, then smoothness is understood in the sense of infinite-dimensional calculus with respect to, for example, suitable Sobolev topologies. These infinite-dimensional

technicalities will not be required in any detail in this chapter, and so may be treated formally.

As before, the set M is the containing space of the fluid, and α is a small positive number. Let $\mathfrak{X}$ be an appropriately chosen space of fields, designed to model "fluid fluctuations," on M, and consider the space $Y = [0, \alpha] \times \mathfrak{X}$. Assume that there is an averaging operation

$$\langle \cdot \rangle : \mathcal{F}(Y) \to \mathcal{F}(M)$$

satisfying the following properties for $f, g \in \mathcal{F}(Y)$, $a, b \in \mathbb{R}$, $\psi \in \mathcal{F}([0, \alpha])$, and $h \in \mathcal{F}(\mathfrak{X})$:

$$\text{Linearity: } \langle af + bg \rangle = a\langle f \rangle + b\langle g \rangle, \tag{2.11}$$

$$\text{Independence: } \langle \psi h \rangle = \frac{1}{\alpha} \left(\int_0^\alpha \psi(\epsilon)\, d\epsilon \right) \langle h \rangle, \tag{2.12}$$

$$\text{Commutativity: } \left\langle \int f\, dx \right\rangle = \int \langle f \rangle\, dx, \tag{2.13}$$

$$\langle \partial f \rangle = \partial \langle f \rangle, \text{ where } \partial = \partial_t \text{ or } \partial = \partial_{x^i}. \tag{2.14}$$

Here, $\psi h \in \mathcal{F}(Y)$ is defined as the pointwise product. Note that if ψ is a constant, then the first and second requirements are compatible.

For compressible flow, the space of fluid fluctuations is $\mathfrak{X} = \mathfrak{X}(M)$. For incompressible flow, the space of divergence-free vector fields is used instead, i.e., $\mathfrak{X} = \mathfrak{X}_{\text{div}}(M)$. In general, $\mathfrak{X} = T_{\text{Identity}} X$, where X is the space to which the tube maps ξ^ϵ belong.

Example. Let μ be a probability measure on the unit sphere S in $\mathfrak{X}(M)$, and define the *average* of a (vector-valued) function $f(\epsilon, w)$ on $[0, \alpha] \times S$ by

$$\langle f \rangle := \frac{1}{\alpha} \int_0^\alpha \int_S f(\epsilon, w)\, d\mu(w)\, d\epsilon.$$

One checks formally that this is an example of an averaging operation that satisfies the desired properties.

2.3 Incompressible Flow Revisited

Before applying the averaging technique to the case of compressible flow, we shall first derive averaged equations for incompressible flow, equations which have already been derived in the literature. The presentation given here has the advantage of being easily generalized to compressible flows. This advantage stems from the careful use and interpretation of modeling assumptions on the fluctuations ξ'—only intuitive assumptions are required regarding the mean behavior of the fluctuations, as well as a first-order Taylor hypothesis. Furthermore, great care has been taken to separate the algebraic issues involved with the averaging procedure from the modeling issues.

In the incompressible case, fluid fluctuations are modeled using the *volume-preserving* diffeomorphism group on M, which is denoted by $\mathrm{Diff}_{\mathrm{vol}}(M)$. Therefore, the tube construction from the previous section now reads as follows: let $\xi^\epsilon(x,t)$ be a family of volume-preserving diffeomorphisms about the identity. That is,

$$\begin{aligned} \text{for each } \epsilon \geq 0, \quad & \xi^\epsilon(\cdot,t) \in \mathrm{Diff}_{\mathrm{vol}}(M) \text{ for all } t, \text{ and} \\ \text{at } \epsilon = 0, \quad & \xi^\epsilon(x,t) = x \text{ for all } x,t. \end{aligned}$$

This forces $\xi'(\cdot,t)$ to be a divergence-free vector field for all t.

Averaged Lagrangian for Incompressible Fluids. Let us start with the standard Lagrangian

$$l(u^\epsilon) = \int_M \frac{1}{2}\|u^\epsilon\|^2 \, dx \tag{2.15}$$

and expand u^ϵ in a Taylor series about u:

$$u^\epsilon = u + \epsilon u' + \frac{1}{2}\epsilon^2 u'' + \mathcal{O}(\epsilon^3). \tag{2.16}$$

Substituting this expansion into (2.15) gives

$$l(u^\epsilon) = \int_M \frac{1}{2}\|u^2\| + \epsilon u \cdot u' + \frac{\epsilon^2}{2}\left(\|u'\|^2 + u'' \cdot u\right) + \mathcal{O}(\epsilon^3) \, dx. \tag{2.17}$$

Let $\hat{l}(u^\epsilon)$ be the truncation of l to terms of order less than ϵ^3. Using formulas (2.8), u' and u'' can be rewritten in terms of u, ξ', and ξ''. We do this in order to write $\hat{l}$ as a function only of u, ξ', and ξ''. Making the substitutions and rewriting in coordinates,

$$\begin{aligned}\hat{l}(u^\epsilon) = \int_M \frac{1}{2}u^i u^i + \epsilon\Big(u^i(\partial_t \xi'^i) + u^i u^j \xi'^i_{,j} - u^i \xi'^j u^i_{,j}\Big) \\ + \frac{\epsilon^2}{2}\Big((\partial_t\xi'^i)(\partial_t\xi'^i) + 2(\partial_t\xi'^i)\xi'^i_{,k}u^k - 2(\partial_t\xi'^i)u^i_{,k}\xi'^k + \xi'^i_{,j}u^j\xi'^i_{,k}u^k \\ - \xi'^i_{,j}u^j u^i_{,k}\xi'^k - u^i_{,j}\xi'^j\xi'^i_{,k}u^k + u^i_{,j}\xi'^j u^i_{,k}\xi'^k - 2(\partial_t\xi'^i_{,j})\xi'^j u^i - 2\xi'^i_{,jk}u^k\xi'^j u^i \\ - 2\xi'^i_{,k}u^k_{,j}\xi'^j u^i + 2u^i_{,kj}\xi'^k\xi'^j u^i + 2u^i_{,k}\xi'^k_{,j}\xi'^j u^i - u^i_{,jk}\xi'^j\xi'^k u^i\Big) \\ + \frac{\epsilon^2}{2}\Big((\partial_t\xi''^i)u^i + u^j\xi''^i_{,j}u^i - \xi''^j u^i_{,j}u^i\Big)\, dx, \end{aligned} \tag{2.18}$$

where the notation $u^i_{,j}$ means $\partial u^i/\partial x^j$. Throughout this chapter, there is an implied sum over repeated indices. The averaged Lagrangian for incompressible flow is now simply $l^\alpha_{\text{in}} = \langle\hat{l}\rangle$.

Zero-Mean Fluctuations. Before undertaking this computation, recall from section 2.2 that the fluctuation diffeomorphism maps ξ^ϵ are required to have as their average the identity map. This statistical assumption regarding the behavior of the fluctuations is the first modeling assumption:

$$\langle\xi'\rangle = 0 \quad \text{and} \quad \langle\xi''\rangle = 0. \tag{2.19}$$

This point would not be worth belaboring, except that, when combined with the properties of our averaging operation (2.11)–(2.14), assumption (2.19) forces *all* linear functions of ξ', ξ'', and their derivatives to also have zero mean. Applying this fact to (2.18) causes the entire $\mathcal{O}(\epsilon)$ group and the second $\mathcal{O}(\epsilon^2)$ group (i.e., the last line of (2.18)) to vanish inside the average.

We continue analyzing (2.18): the only remaining terms are $(1/2)u^iu^i$ and the first $\mathcal{O}(\epsilon^2)$ group. Within this $\mathcal{O}(\epsilon^2)$ group, we integrate certain terms by parts and notice

that all terms involving time derivatives of ξ' group together:

$$(\partial_t \xi'^i)(\partial_t \xi'^i) + 2(\partial_t \xi'^i)\xi'^i_{,k} u^k + \xi'^i_{,j} u^j \xi'^i_{,k} u^k \\ = \left((\partial_t \xi'^i) + \xi'^i_{,j} u^j\right)\left((\partial_t \xi'^i) + \xi'^i_{,k} u^k\right) = \left\| \frac{D\xi'}{Dt} \right\|^2, \quad (2.20)$$

where D/Dt is the material derivative:

$$\frac{D}{Dt} = (\partial_t + u \cdot \nabla). \quad (2.21)$$

We then simplify the remaining non-time-derivative terms from (2.18), integrating by parts to remove second-order spatial derivatives. The final expression for the averaged incompressible Lagrangian is

$$l^\alpha_{\text{in}}(u) = \int_M \left\{ \frac{1}{2}\|u\|^2 + \frac{\alpha^2}{2}\left[\left\langle \left\| \frac{D\xi'}{Dt} \right\|^2 \right\rangle - \frac{1}{2}\left\langle \text{tr}(\nabla \xi' \cdot \nabla \xi') \right\rangle \|u\|^2 \right] \right\} dx. \quad (2.22)$$

Modeling of $\boldsymbol{\xi'}$. Immediate application of Hamilton's principle to (2.22) does not yield a closed system of equations. Namely, we have initial ($t = 0$) data for ξ' but no way to compute this vector field for $t > 0$. Our approach in what follows will be to write down, based on physical considerations, an evolution law, or *flow rule*, for ξ'.

A flow rule consists of a prescribed choice of ϕ in the following evolution equation for ξ':

$$\frac{D\xi'}{Dt} = \phi(u, \rho, \xi'). \quad (2.23)$$

Given a choice of ξ' at $t = 0$, this equation will uniquely determine ξ' for $t > 0$. Let us assume we have a *linear flow rule*,

$$\frac{D\xi'^i}{Dt} = \Omega^{ij}\xi'^j, \quad (2.24)$$

where Ω^{ij} is allowed to depend on u and ρ but not on ξ^ϵ or its derivatives. The caveat here is that our choice of Ω must be compatible with incompressibility; in particular, $\text{div}\, \xi' = 0$ at $t = 0$, and Ω must be chosen such that ξ' remains divergence-free as it

evolves. At this stage, one might raise the issue of the tube ξ^ϵ and request a concrete description of the whole object. Such a description is unnecessary; in order to close the system of evolution equations resulting from (2.22), we need only describe the evolution of the first-order fluctuation field ξ'. Now defining the *Lagrangian covariance tensor*

$$F = \langle \xi' \otimes \xi' \rangle \tag{2.25}$$

and using the linear flow rule (2.24), the Lagrangian (2.22) can be rewritten as

$$l^\alpha_{\text{in}}(u) = \int_M \left\{ \frac{1}{2} u^i u^i + \frac{\alpha^2}{2} \left[\Omega^{ij} \Omega^{ik} F^{jk} - \frac{1}{2} F^{ij}_{,ij} u^k u^k \right] \right\} dx. \tag{2.26}$$

Here we have used the fact that ξ' must be divergence-free.

Advection Flow Rule. The first flow rule we shall consider results from setting $\Omega^{ij} = u^i_{,j}$:

$$\frac{D\xi'^i}{Dt} = u^i_{,j} \xi'^j. \tag{2.27}$$

Using the definition of the material derivative, it is trivial to see that this flow rule is equivalent to Lie advection of ξ': $\partial_t \xi' = -\pounds_u \xi'$. This advection hypothesis is the vector field analogue of the classical frozen turbulence hypothesis of Taylor introduced in [83]. This hypothesis is widely used in the turbulence community (see [16] for instance, for usage of this hypothesis even in the sense of Lie advection of vector fields). More recently, this generalized version of the Taylor hypothesis has been used to achieve turbulence closure in the derivation of incompressible LAE-α equations (see [59, 60]) or in the work of Holm (see [35]) on averaged compressible models using the GLM theory.

The advection flow rule (2.27) is perhaps the most obvious choice for Ω that is compatible with incompressibility. Note that if $\operatorname{div} \xi' = 0$ at $t = 0$, then differentiating (2.27) with respect to x^i yields

$$\partial_t \left(\operatorname{div} \xi' \right) = u^i_{,j} \xi'^j_{,i} - \xi'^i_{,j} u^j_{,i} = 0.$$

Therefore, $\operatorname{div} \xi' = 0$ for all $t > 0$. Using this flow rule, both anisotropic and isotropic

models shall be developed. For incompressible flow, no other flow rules will be considered.

Incompressible, Anisotropic, Inhomogeneous Flow. In this case, the flow rule is used to derive an evolution equation for the covariance tensor F. Time-differentiating $F^{ij} = \langle \xi'^i \xi'^j \rangle$ and using (2.27) yields the Lie advection equation $\partial_t F = -\pounds_u F$. Equipped with an evolution equation for F, we can apply Hamilton's principle to (2.26) and derive a closed system with unknowns u, the average velocity, and F, the covariance tensor.

Carrying this out, one finds that the anisotropic LAE-α equations are given by the following coupled system of equations for u and F:

$$\partial_t(1-\alpha^2 C)u + (u\cdot\nabla)(1-\alpha^2 C)u = -\operatorname{grad} p, \tag{2.28a}$$

$$\operatorname{div} u = 0, \tag{2.28b}$$

$$\partial_t F + \nabla F \cdot u - F\cdot\nabla u - \nabla u^T \cdot F = 0, \tag{2.28c}$$

where p is the fluid pressure, and the operator C is defined by

$$Cu = \operatorname{div}[\nabla u \cdot F]. \tag{2.29}$$

When $\alpha = 0$, the system (2.28a)–(2.28b) reduces to the incompressible Euler equation.

Note. Start with the generic incompressible averaged Lagrangian (2.26) and substitute the advection flow rule (2.27). Now integrate the last term by parts and use $\operatorname{div}\xi' = 0$. The result is

$$l^\alpha_{\text{in}}(u) = \int_M \left\{ \frac{1}{2}\|u\|^2 - \frac{\alpha^2}{2} u\cdot[\nabla\nabla u : F] \right\} dx, \tag{2.30}$$

which is exactly the Lagrangian used in [60] to derive the anisotropic LAE-α equations. However, in [60] the second-order Taylor hypothesis

$$\frac{D}{Dt}\langle \xi'' \rangle \perp u,$$

where the orthogonality is taken in L^2, is necessary to achieve closure. Our choice of modeling assumptions rendered unnecessary any such hypothesis on the second-order

fluctuations ξ''. Second-order Taylor hypotheses, unlike the first-order hypothesis retained from [60], do not have much precedent in the turbulence literature, as discussed above.

Incompressible, Isotropic, Homogeneous Fluids. To model the motion of an approximately isotropic fluid, we take the covariance tensor F to be the identity matrix, i.e.,

$$F^{ij} = \left\langle \xi'^i \xi'^j \right\rangle = \delta^{ij}. \tag{2.31}$$

The choice of $F^{ij} = \delta^{ij}$ is a modeling assumption, and thus will be valid only for flows which almost preserve this property. Note that (2.31) is strictly inconsistent with the advection flow rule, and thus can be regarded only as an approximation.

For the case of incompressible isotropic mean flow, we assume that (2.31) holds; then differentiating this equation with respect to x^k and x^j and using the fact that ξ' is divergence-free, we have

$$\left\langle \xi'^i_{,j} \xi'^j_{,k} \right\rangle = - \left\langle \xi'^i_{,jk} \xi'^j \right\rangle.$$

Hence

$$\left\langle \mathrm{tr}(\nabla \xi' \cdot \nabla \xi') \right\rangle = \left\langle \xi'^i_{,j} \xi'^j_{,i} \right\rangle = - \left\langle \xi'^i_{,ji} \xi'^j \right\rangle = 0,$$

and the Lagrangian (2.22) simplifies to

$$l^\alpha_{\mathrm{in,iso}}(u) = \int_M \left\{ \frac{1}{2} \|u\|^2 + \frac{\alpha^2}{2} \left\langle \left\| \frac{D\xi'}{Dt} \right\|^2 \right\rangle \right\} dx. \tag{2.32}$$

We emphasize that this is only an approximation, so that

$$l^\alpha_{\mathrm{in,iso}}(u) \approx l^\alpha_{\mathrm{in}}(u)$$

along fluid trajectories $u(t)$ for which the covariance tensor is approximately the identity. Now using the flow rule given by (2.27), the averaged Lagrangian l^α_{in} from (2.32) becomes

$$\left\langle \left\| \frac{D\xi'}{Dt} \right\|^2 \right\rangle = u^i_{,j} u^i_{,k} \left\langle \xi'^j \xi'^k \right\rangle = u^i_{,j} u^i_{,j}, \tag{2.33}$$

where we have used the isotropy assumption (2.31). Hence, (2.32) becomes

$$l^{\alpha}_{\text{in}}(u) = \int_M \left\{ \frac{1}{2}\|u\|^2 + \frac{\alpha^2}{2}\,\|\nabla u\|^2 \right\} dx. \tag{2.34}$$

This expression for the averaged Lagrangian in the isotropic case is identical to the one derived in [59]. Now applying either Hamilton's principle or Euler–Poincaré theory, we obtain the standard isotropic LAE-α equations

$$\partial_t(1-\alpha^2\Delta)u + (u\cdot\nabla)(1-\alpha^2\Delta)u - \alpha^2(\nabla u)^T\cdot\Delta u = -\operatorname{grad} p, \tag{2.35a}$$

$$\operatorname{div} u = 0, \tag{2.35b}$$

where p is the usual fluid pressure.

2.4 Averaged Lagrangian for Compressible Flow

Having understood the incompressible case, we now turn to the compressible case. The procedure is identical in all aspects, except we must now keep track of density fluctuations. Start with the reduced Lagrangian for compressible flow:

$$l(u^\epsilon, \rho^\epsilon) = \int_M \left(\frac{1}{2}\|u^\epsilon\|^2 - W(\rho^\epsilon) \right) \rho^\epsilon \; dx. \tag{2.36}$$

The fluid is assumed to be barotropic, meaning that W, the potential energy, is a function only of ρ, the fluid density. Now expand the velocity and density in Taylor series

$$\begin{aligned} u^\epsilon &= u + \epsilon u' + \frac{1}{2}\epsilon^2 u'' + \mathcal{O}(\epsilon^3), \\ \rho^\epsilon &= \rho + \epsilon\rho' + \frac{1}{2}\epsilon^2\rho'' + \mathcal{O}(\epsilon^3) \end{aligned} \tag{2.37}$$

and also expand the potential energy W:

$$W(\rho^\epsilon) = W(\rho) + \epsilon W'(\rho)\rho' + \frac{1}{2}\epsilon^2(W''(\rho)\rho'^2 + W'(\rho)\rho'') + \mathcal{O}(\epsilon^3).$$

Substituting these expansions into the reduced Lagrangian gives

$$\begin{aligned} l(u^\epsilon, \rho^\epsilon) = \int_M & \left(\frac{1}{2}\|u\|^2 - W(\rho)\right)\rho \\ & + \epsilon\left[\left(u\cdot u' - W'(\rho)\rho'\right)\rho + \left(\frac{1}{2}\|u\|^2 - W(\rho)\right)\rho'\right] \\ & + \epsilon^2\left[\frac{1}{2}\left((\|u'\|^2 + u''\cdot u) - (W''(\rho)\rho'^2 + W'(\rho)\rho'')\right)\rho\right. \\ & \left. + (u\cdot u' - W'(\rho)\rho')\rho' + \frac{1}{2}\left(\frac{1}{2}\|u\|^2 - W(\rho)\right)\rho''\right] + \mathcal{O}(\epsilon^3)\, dx. \end{aligned} \tag{2.38}$$

This expansion is now truncated, leaving out all terms of order ϵ^3 and higher. Denote the truncated Lagrangian by $\hat{l}(u^\epsilon, \rho^\epsilon)$, and define the averaged Lagrangian l^α by

$$l^\alpha(u, \rho) = \langle \hat{l}(u^\epsilon, \rho^\epsilon)\rangle. \tag{2.39}$$

We now outline the procedure by which we arrive at a final written expression for the averaged Lagrangian l^α. The algebra is straightforward but tedious, so details will be omitted.

1. Use (2.8) and (2.10) to rewrite (2.38) in terms of only u, ρ, and the fluctuations ξ', ξ''.

2. Remove two kinds of terms that vanish inside the average:

   (a) linear functions of ξ' or ξ'',

   (b) linear functions of derivatives (either spatial or temporal) of ξ' or ξ''.

   Note: see "zero-mean fluctuations" in section 2.3 for justification.

3. Carry out the averaging operation. As in the incompressible case, the only quantities left inside the average should be nonlinear functions of ξ'.

The end result for the averaged Lagrangian for compressible flow is

$$l^{\alpha}_{\text{comp}}(u,\rho) = \int_M \left\{ \frac{1}{2}\rho\|u\|^2 - \rho W(\rho) + \alpha^2 \left[\frac{1}{2}\rho \left\langle \left\| \frac{D\xi'}{Dt} \right\|^2 \right\rangle \right.\right. \\ \left.\left. - \frac{1}{2}w'(\rho)\left\langle \operatorname{div}(\rho\xi')^2 \right\rangle - \frac{1}{2}w(\rho)\left\langle \operatorname{div}\operatorname{div}(\rho\xi'\otimes\xi') \right\rangle \right] \right\} dx. \tag{2.40}$$

We have introduced w, the enthalpy,[2] defined by

$$w(\rho) = W(\rho) + \rho W'(\rho). \tag{2.41}$$

2.5 Flow Rule Modeling

In deriving the expressions (2.40) and (2.22) for the averaged Lagrangians, no assumptions were made regarding how the Lagrangian fluctuations ξ' evolve. In this section we describe one possible strategy for modeling ξ'. Note that such a strategy is necessary to achieve closure for the evolution equations associated with the Lagrangians (2.40) or (2.22).

Preliminary Observation. Assuming ξ' evolves via a linear flow rule, as in (2.24), the vector field ξ' appears in the averaged Lagrangian (2.40) only as part of the following three expressions:[3]

$$F^{ij} = \left\langle \xi'^i \xi'^j \right\rangle, \tag{2.42a}$$

$$G^i = \left\langle \xi'^i \xi'^j_{,j} \right\rangle, \tag{2.42b}$$

$$H = \left\langle \xi'^i_{,i} \xi'^j_{,j} \right\rangle. \tag{2.42c}$$

Note that F is the same *Lagrangian covariance tensor* from the incompressible derivation. In terms of these quantities, the averaged compressible Lagrangian is given in coordinates

[2] Any function w satisfying $\nabla w = (\nabla p)/\rho$, where p is pressure, is called enthalpy. Our definition of w implies $w_{,i} = 2W'(\rho)\rho_{,i} + \rho W''(\rho)\rho_{,i} = (\rho^2 W'(\rho))_{,i}/\rho = p_{,i}/\rho$, as required.

[3] Similar tensors appear in [34]; they are referred to as second-order statistics of the Lagrangian fluctuations.

by

$$l^{\alpha}_{\text{comp}}(u,\rho) = \int_M \left\{ \frac{1}{2}\rho u^i u^i - \rho W(\rho) + \alpha^2 \left[\frac{1}{2}\rho \Omega^{ij}\Omega^{ik}F^{jk} \right.\right. \\ \left.\left. - \frac{1}{2}w'(\rho)\left(\rho_{,i}\rho_{,j}F^{ij} + 2\rho\rho_{,j}G^j + \rho^2 H\right) - \frac{1}{2}w(\rho)\left(\rho F^{ij}\right)_{,ij} \right] \right\} dx. \tag{2.43}$$

Time-differentiating (2.42a)–(2.42c) and using the linear flow rule (2.24) results in evolution equations for F, G, and H:

$$\partial_t F^{ij} = \Omega^{ik}F^{kj} + \Omega^{jk}F^{ki} - u^k F^{ij}_{,k}, \tag{2.44a}$$

$$\partial_t G^i = \Omega^{ik}G^k - u^k G^i_{,k} + F^{ij}\Omega^{kj}_{,k} + \left\langle \xi'^i \xi'^j_{,k} \right\rangle (\Omega^{kj} - u^k_{,j}), \tag{2.44b}$$

$$\partial_t H = 2\Omega^{ik}_{,i}G^k - u^k H_{,k} + 2\left\langle \xi'^j_{,k}\xi'^i_{,i} \right\rangle (\Omega^{kj} - u^k_{,j}). \tag{2.44c}$$

Flow Rules. For compressible flows, two flow rules will be considered. We define them first and then go on to consider their relative merits and demerits:

(I) Advection: $\Omega^{ij} = u^i_{,j}$.

(II) Rotation: $\Omega^{ij} = \frac{1}{2}(u^i_{,j} - u^j_{,i})$.

Advection. For our anisotropic model, we shall advect ξ' and treat the quantities F, G, and H as parameters in the final system, each of which will have its own evolution equation. Substituting $\Omega^{ij} = u^i_{,j}$ into the system (2.44) gives

$$\partial_t F = -\pounds_u F, \tag{2.45a}$$

$$\partial_t G = -\pounds_u G + F \cdot \operatorname{grad}(\operatorname{div} u), \tag{2.45b}$$

$$\partial_t H = 2\operatorname{grad}(\operatorname{div} u) \cdot G - u \cdot \operatorname{grad} H. \tag{2.45c}$$

One advantage of the advection flow rule is that it automatically closes the system (2.44). For a general choice of Ω, the system involves $\langle \xi'^i \xi'^j_{,k} \rangle$ and $\langle \xi'^j_{,k}\xi'^i_{,i} \rangle$, which cannot be expressed solely in terms of F, G, and H.

Rotation. For our isotropic model, we want to know whether the evolution equation (2.44a) for F preserves the isotropy relationship $F = \text{Identity}$. Suppose $F^{ij} = \delta^{ij}$ at $t = 0$. Then substituting into (2.44a) reveals that

$$\partial_t|_{t=0} F^{ij} = \Omega^{ij} + \Omega^{ji}. \tag{2.46}$$

If Ω is antisymmetric, we have $\partial_t|_{t=0} F = 0$, and $F(x,t) = \text{Identity}$ solves (2.44a) for all t. We wish to know whether this solution is unique. This is guaranteed by a straightforward generalization of the results concerning linear hyperbolic systems of first-order equations from [24], assuming sufficient smoothness of u.

We conclude that antisymmetry of Ω is sufficient to guarantee that the initial data $F = \text{Identity}$ is, in fact, preserved for all t. Then an immediate choice of a tensor Ω that is antisymmetric is given by the rotation flow rule (II). This form has a very attractive physical interpretation. Putting the linear flow rule equation (2.24) together with (II) gives us

$$\frac{D\xi'}{Dt} = \omega \times \xi', \tag{2.47}$$

where $\omega = \operatorname{curl} u$ is the vorticity vector. The last equation can be interpreted in the sense that fluctuations are rigidly transported by the mean flow, with a local angular velocity given by the vorticity vector.

Finally, the rotation flow rule (II) does not by itself close the system (2.44). When using this flow rule, we shall assume that $G = 0$ and $H = \beta^2$.

2.6 Equations for Averaged Dynamics

Here we shall write down two systems of coupled PDEs which describe the evolution of the average velocity and density in a compressible flow. Each PDE is derived from an associated averaged Lagrangian.

Compressible, fully Anisotropic, Inhomogeneous Fluids. By substituting (I) into the Lagrangian (2.40), we obtain closure: the Lagrangian no longer depends ex-

plicitly on ξ', instead depending on the tensors F, G, and H, for which a self-contained system of evolution equations (2.45) has already been derived—see section 2.5 for details. Applying Hamilton's principle directly to (2.43) yields an evolution equation for u, the average fluid velocity. We write this equation using the operator $\mathcal{A}$, which is defined as

$$(\mathcal{A}v)^i = \frac{1}{\rho}\left(\rho v^i_{,j} F^{jk}\right)_{,k}. \tag{2.48}$$

We also write $\tilde{w} = \rho w'(\rho)$, where $'$ means $d/d\rho$ as usual. The anisotropic compressible LAE-α equations are

$$\begin{aligned}\left(\partial_t u^n + u^n_{,i} u^i\right) = (1-\alpha^2\mathcal{A})^{-1}\frac{1}{\rho}\Bigg\{ &-\rho w_{,n} - \frac{\alpha^2}{2}\Bigg[\rho\left(F^{ij}u^k_{,i}u^k_{,j}\right)_{,n} + F^{ij}_{,ij}\rho\tilde{w}_{,n} \\ &+ F^{ij}_{,n}\rho_{,i}\tilde{w}_{,j} + \left(F^{ij}_{,n}\rho\right)_{,ij}\tilde{w} + 2G^i_{,n}\rho\tilde{w}_{,i} + 2G^i\left(\rho\tilde{w}_{,n}\right)_{,i} + \left(H\rho^2\tilde{w}'\right)_{,n}\Bigg]\Bigg\},\end{aligned} \tag{2.49a}$$

$$\partial_t \rho = -\operatorname{div}(\rho u), \tag{2.49b}$$

$$\partial_t F = -\nabla F \cdot u + F\cdot\nabla u + \nabla u^T \cdot F, \tag{2.49c}$$

$$\partial_t G = -u\cdot\nabla G + G\cdot\nabla u + F\cdot\operatorname{grad}(\operatorname{div} u), \tag{2.49d}$$

$$\partial_t H = 2\operatorname{grad}(\operatorname{div} u)\cdot G - u\cdot\operatorname{grad} H. \tag{2.49e}$$

Well-posedness. We now sketch a rough well-posedness argument for the system (2.49). Assume that the tensor F is positive definite. By this it is meant, since F is a $(2,0)$ tensor, that for any one-form θ, the contraction $F:(\theta\otimes\theta)$ is positive everywhere. Given the ρ-weighted inner product $\langle f, g\rangle = \int f\, g\, \rho$, we have $\langle f, -\mathcal{A}f\rangle = -\int f\, \frac{1}{\rho}\left(\rho f_{,j}F^{jk}\right)_{,k}\rho = \int f_{,j}F^{jk}f_{,k}\,\rho > 0$. Since $-\mathcal{A}$ is a positive definite linear operator, the kernel of $(1-\alpha^2\mathcal{A})$ is trivial, and we expect that (2.49) is well-posed.

It would be of analytical interest to see to what extent the "geodesic part" of these equations defines a smooth spray in the sense of [22], and which holds for the EPDiff equations (that is, the n-dimensional CH equations), as explained in [37].

Compressible, isotropic, inhomogeneous. For this case we use flow rule (II), which can be written in vector notation as

$$\Omega = \frac{1}{2}\left(\nabla u - \nabla u^T\right).$$

Recall that this flow rule is compatible with an isotropic choice of the covariance tensor, i.e., $F^{ij} = \delta^{ij}$. We further assume that $G = 0$ and $H = \beta^2$ for some constant β. Using flow rule (II) along with these extra assumptions in the general Lagrangian expression (2.43) gives us a Lagrangian in only two variables:

$$l(u,\rho) = \int_M \left(\frac{1}{2}\rho\|u\|^2 - \rho W(\rho) + \alpha^2\left[\frac{1}{4}\rho\big(\|\nabla u\|^2 - \operatorname{tr}(\nabla u\cdot\nabla u)\big) - \frac{1}{2}w'(\rho)\left(\|\nabla\rho\|^2 + \rho^2\beta^2\right) - \frac{1}{2}w(\rho)\Delta\rho\right]\right) d^N x, \quad (2.50)$$

where $w(\rho) = W(\rho) + \rho W'(\rho)$ is the enthalpy introduced in (2.41). Regarding this as a Lagrangian in u and $\mu = \rho\, d^N x$, one uses the semidirect product Euler–Poincaré equations (1.15) to derive the system

$$\partial_t(\rho v) + (u\cdot\nabla)(\rho v) + \alpha^2 \operatorname{div}(\rho\Omega\cdot\nabla u) + \rho v \operatorname{div} u = -\nabla\tilde{p}, \quad (2.51a)$$

$$\partial_t\rho + \operatorname{div}(\rho u) = 0, \quad (2.51b)$$

with the modified momentum ρv and modified pressure $\tilde{p}$ given by

$$\rho v = \rho u + \alpha^2 \operatorname{div}(\rho\Omega), \quad (2.52)$$

$$\nabla\tilde{p} = \nabla p + \alpha^2\beta^2\rho\,\nabla\left(\rho w' + \frac{1}{2}\rho^2 w''\right). \quad (2.53)$$

Here are explicit coordinate expressions for two slightly complicated objects:

$$\rho v^i = \rho u^i + \frac{1}{2}\alpha^2\left(\rho\Big(u^j_{,i} - u^i_{,j}\Big)\right)_{,j},$$

$$\operatorname{div}(\rho\Omega\cdot\nabla u) = \left(\rho\Omega^{ki}u^i_{,j}\right)_{,k}.$$

The following convention for divergences of tensors has been used: given a 2-tensor A^{ij}, we set

$$(\operatorname{div} A)^j = A^{ij}_{,i}.$$

That is, the contraction implicit in the divergence operation always takes place on the *first* index.

Observations.

- In the case of homogeneous incompressible flow, where ρ is constant and $\operatorname{div} u = 0$, the definition of ρv in (2.52) reduces to

$$v = \left(1 - \frac{1}{2}\alpha^2\Delta\right) u,$$

  which after rescaling α to get rid of the factor of $1/2$ is precisely the v one finds in treatments of the incompressible LAE-α and LANS-α equations.

- The above does not work in one spatial dimension. The problem is that here Ω reduces to $(u_x - u_x)/2 = 0$, which clearly does not describe transport at all. For a one-dimensional isotropic model one may very well want to forget about antisymmetry of Ω and instead use something such as the advection flow rule. One may, quite reasonably, conclude that the only meaning of isotropy in one dimension should be reflection symmetry.

2.7 Future Directions

The Initialization Problem. Perhaps the largest unsolved problem for the Lagrangian averaged equations is the initialization problem. A concise statement of the problem reads as follows:

> Given initial data $u_0(x)$ for the Euler equation, how does one obtain initial data $U_0(x)$ for the LAE-α equation?

Let us look at this problem in slightly more detail. Let u denote the solution of the incompressible Euler equations for initial data u_0, i.e., $u(x,0) = u_0(x)$. Similarly, let U denote the solution of the incompressible, isotropic LAE-α equations (2.35) for initial data U_0.

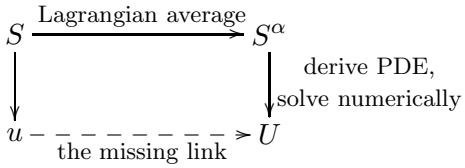
Now U should be, in some sense, the mean flow of the fluid. This means that U_0 should be the mean flow of the fluid at time $t = 0$, implying that U_0 should be, in some sense, an "averaged" or "filtered" version of u_0. The question is, How does one derive U_0 from u_0? Another way of phrasing this question is, How do we describe (approximately) the initial state of the fluid (given exactly, for our purposes, by the field u_0) using only the mean flow variable U_0?

Numerous methods have been used to initialize the LAE-α equations for use in numerical simulations, but none of these methods has any theoretical foundation. There is also no theory regarding how one should filter a full Euler flow u, or even a family of flows u^ϵ, in order to obtain a mean flow that could be compared with the full LAE-α trajectory U. In this respect, (2.4), which states that

$$\langle u^\epsilon \circ \xi^\epsilon(x,t) \rangle = u(x,t),$$

is not helpful: we have no way to compute the fluctuation diffeomorphism group ξ^ϵ. Therefore, we have no way to compute the left-hand side $\langle u^\epsilon \circ \xi^\epsilon \rangle$.

The difficulty can be summarized in the following commutative diagram. Here S is the standard Euler action and S^α is the Lagrangian averaged action.



Solid arrows represent steps that we know how to carry out. The dashed arrow represents the one step that we do not know how to carry out. Our strategy for this problem will be to develop methods by which we can test different filters for obtaining U_0 from u_0 in practice.

Treatment of Densities. Another area for further investigation involves our treatment of the density tube μ^ϵ. There are two questions to ground us:

1. We have tacitly assumed that at $t = 0$, for all ϵ and all x,

$$\mu(x,t) = \mu^\epsilon(x,t).$$

   An argument similar to the one made above in our discussion of the initialization problem can be made here. Namely, $\mu(x,0)$ represents the mean density at time $t = 0$. Meanwhile, $\mu^\epsilon(x,0)$ represents the true density of the fluid at time $t = 0$. These two quantities need not be equal. This prompts the following question: how would we carry out the procedure from sections 2.2 and 2.4 with tubes in which each trajectory does *not* have the same initial density $\mu(x,0)$?

2. As our derivation of the averaged compressible equations stand, we have derived the fact that the "mean" density μ was advected by the mean flow U: $\partial_t \mu = -\pounds_U \mu$. Substituting $\mu = \rho\, d^N x$ and using the definition of divergence yields the standard continuity equation

$$\partial_t \rho + \operatorname{div}(\rho U) = 0.$$

   In both RANS and LES treatments of averaged/filtered flow, the mean flow U satisfies a *modified* continuity equation rather than the standard one. Therefore, why does the Lagrangian averaged mean density μ satisfy the usual continuity equation?

The two questions regarding densities are, in fact, related. To see this, let us suppose that, given the initial density μ_0 associated with the center line of our tube η, we have a method for constructing a family of initial densities μ_0^ϵ for each of the other curves in the tube η^ϵ. Now defining[4]

$$\mu^\epsilon(t) = (\eta_t^\epsilon)_* \mu_0^\epsilon \quad \text{and} \quad \bar{\mu}(t) = \langle \mu^\epsilon(t) \rangle,$$

[4]Note that $\langle \mu^\epsilon(x,t) \rangle \neq \mu(x,t)$.

we will find that $\bar{\mu}(t) = \bar{\rho}(t)\, d^N x$ satisfies a modified continuity equation

$$\partial_t \bar{\rho}(t) + \operatorname{div}(\bar{\rho} u) + \operatorname{div} \left\langle \rho^\epsilon \left(\epsilon u' + \frac{1}{2}\epsilon^2 u'' \right) \right\rangle = 0.$$

To close this equation, we either must carry out the average directly or must expand ρ^ϵ about a suitable trajectory and make modeling assumptions.

Filtered Lagrangians. We have seen that the current averaging procedure leads to complicated averaged equations. Furthermore, there is no clear way to evaluate numerically the flow rules we have proposed on physical grounds. One of our immediate goals is to investigate a filtering approach, still at the level of the Lagrangian, which will lead to simpler averaged models that can be tested numerically. The filtering approach we have in mind begins with a decomposition of the velocity field

$$u = \bar{u} + u' \quad \text{and} \quad \rho = \bar{\rho} + \rho' \tag{2.54}$$

into mean and fluctuating components. This would replace the Taylor expansion (2.37) of u^ϵ and ρ^ϵ that we carried out in the present work and would therefore lead to Lagrangians and equations with much less algebraic complexity. As opposed to the axiomatic averaging operation $\langle \cdot \rangle$, the filter shall be specified concretely. We expect this to help greatly with the initialization and density problems discussed above; furthermore, the filtering approach leads naturally to questions about the relationship between LES and LAE-α models.

Simpler Models. As we previously noted, the flow rule approach developed in this chapter does not yield a one-dimensional compressible averaged model. By employing the filtering approach, one may derive the grangian

$$l(\rho, u) = \int \left(\frac{1}{2} u v - W(\rho) \right) \rho \, dx, \tag{2.55}$$

where $v = \left(1 - \alpha^2 \partial_x^2\right) u$. To derive this Lagrangian, we filter only the velocity, leaving density and potential energy alone. This is the compressible analogue of the filtered

Lagrangian used in deriving the CH equation [9]. The analysis and numerical simulation of the new equations presented in section 2.6 of this work will be difficult. The analysis of the PDE associated with (2.55) is much easier. In particular, we expect that numerical studies of this one-dimensional model will yield insight into the dynamics of the higher-dimensional equations.

Entropy. In the derivation of our compressible averaged models, we have made the barotropic assumption $W = W(\rho)$. We expect the resulting barotropic model to be useful in computing mean flow quantities in regimes where we are not concerned with strong physical shocks, for example in climate models. The next major step forward will be to remove the barotropic assumption and derive a model that is valid in regimes where we are concerned with shocks.

To this end, we have derived an averaged model for the general case, where the potential energy has the form $W(\rho, S)$, where S is the entropy. This model, which consists of a system of equations for ρ, u, and S, also involves the pressure p. Therefore, in order to close the system, we require an equation of state relating p to ρ and S. The open question now is as follows: given an equation of state for the compressible Euler system, what is the equation of state relating the averaged variables to one another? In other words, how does Lagrangian averaging interact with the thermodynamics of the system? We hope that analyzing a finite-dimensional case of this interaction will shed light on this issue.

Connections with Kevrekidis' Coarse/Fine Methods. Given a description of any mechanical system, not necessarily involving fluids, in the form of a Lagrangian ℓ, we can carry out the procedure described in section 2.2 to find an averaged Lagrangian $\langle \ell \rangle$. From this we can derive equations of motion for the average dynamics of the original system. Changing our language slightly, we say that we have a general method for extracting the "coarse" dynamics of a mechanical system whose full description involves motions on both fine and coarse scales.

Another method for computing the coarse-scale dynamics of a mechanical system

has been put forth in [48]. Kevrekidis's method does not involve trying to write down equations of motion which govern the coarse dynamics. Instead, he offers an algorithmic approach, the crux of which is as follows. The coarse dynamics of a system are found by *lifting* the initial ($t = t_0$) state to an ensemble of initial states, *integrating* each using the full equations until some small final time $t = \epsilon$ has been reached, and *projecting* the resulting $t = \epsilon$ states onto a single state. This $t = \epsilon$ state is then *extrapolated* to a state at some desired $t = t_f > 0$. By iterating this process and tuning the lifting, projection, and extrapolation operations, this method can be used to recover the coarse dynamics of the system.

Now the question that begs to be asked is as follows: for the case of fluid dynamics, how different are the coarse dynamics provided by the LANS-α equation from the coarse dynamics one would obtain by following Kevrekidis? The difficulty in answering this question lies in implementing a full fine-scale integrator for fluids that one could successfully embed inside Kevrekidis's coarse-scale algorithm. We look forward to tackling this task soon.

2.8 Appendix: Fluctuation Calculus Details

Before proceeding with any derivations, we state the Lie derivative theorem for time-dependent vector fields: if the vector field X_λ has flow F_λ, then

$$\frac{d}{d\lambda} F_\lambda^* Y_\lambda = F_\lambda^* \left(\frac{\partial Y_\lambda}{\partial \lambda} + \pounds_{X_\lambda} Y_\lambda \right). \tag{2.56}$$

Our task now is to derive equations (2.10). Starting with (2.6), let us move η^ϵ to the right-hand side of the equation:

$$\mu_0 = (\eta^\epsilon)^* \, \mu^\epsilon. \tag{2.57}$$

The strategy is to differentiate with respect to ϵ and use the Lie derivative theorem (2.56). The intrinsic definition of divergence

$$\pounds_\zeta(\nu) = (\mathrm{div}_\nu \, \zeta) \, \nu \tag{2.58}$$

and the canonical volume form $d^N x = dx^1 \wedge \cdots \wedge dx^N$ will both be used in what follows. Note that $\operatorname{div} \zeta$ with no subscript on the div means $\pounds_\zeta(d^N x)$. Before applying the Lie derivative theorem, note that the vector field

$$W^\epsilon = \frac{\partial}{\partial \epsilon} \eta^\epsilon \circ (\eta^\epsilon)^{-1} \tag{2.59}$$

has flow η^ϵ. A simple computation yields

$$\left.\frac{\partial}{\partial \epsilon}\right|_{\epsilon=0} W^\epsilon = \xi'' - \nabla \xi' \cdot \xi'. \tag{2.60}$$

Then we start with ρ':

$$\begin{aligned}
\frac{\partial}{\partial \epsilon}\mu_0 = 0 &= \frac{\partial}{\partial \epsilon} \left(\eta^\epsilon\right)^* \mu^\epsilon && \text{by differentiating (2.57)} \\
&= (\eta^\epsilon)^* \left(\frac{\partial \mu^\epsilon}{\partial \epsilon} + \pounds_{W^\epsilon} \mu^\epsilon \right) && \text{by (2.56)} \\
&= \eta^* \left(\mu' + \pounds_{\xi'} \mu \right) && \text{at } \epsilon = 0 \\
\implies \mu' &= -\pounds_{\xi'} \mu, \\
\rho'\, d^N x &= -\left(\pounds_{\xi'} \rho\right) d^N x - \rho \left(\pounds_{\xi'} d^N x\right) && \text{by (2.5)}, \\
\rho'\, d^N x &= -\left(\nabla \rho \cdot \xi' + \rho \operatorname{div} \xi'\right) d^N x && \text{by (2.58)} \\
\implies \rho' &= -\operatorname{div}\left(\rho \xi'\right).
\end{aligned}$$

Next we compute ρ'':

$$\begin{aligned}
\frac{\partial^2}{\partial\epsilon^2}\mu_0 = 0 &= \frac{\partial^2}{\partial\epsilon^2}\left(\eta^\epsilon\right)^*\mu^\epsilon \\
&= \left(\eta^\epsilon\right)^*\left(\frac{\partial^2}{\partial\epsilon^2}\mu^\epsilon + \pounds_{W^\epsilon}\frac{\partial\mu^\epsilon}{\partial\epsilon} + \frac{\partial}{\partial\epsilon}\left(\pounds_{W^\epsilon}\mu^\epsilon\right) + \pounds_{W^\epsilon}\pounds_{W^\epsilon}\mu^\epsilon\right) \\
\implies 0 &= \eta^*\left(\mu'' + 2\pounds_{\xi'}\mu' + \pounds_{\xi''-\nabla\xi'.\xi'}\mu + \pounds_{\xi'}\pounds_{\xi'}\mu\right) \\
\implies \mu'' &= -\pounds_{\xi''}\mu + 2\pounds_{\xi'}\pounds_{\xi'}\mu - \pounds_{\xi'}\pounds_{\xi'}\mu + \pounds_{\nabla\xi'.\xi'}\mu, \\
\rho''\,d^Nx &= -\left(\operatorname{div}(\rho\xi'')\right)d^Nx + \operatorname{div}(\operatorname{div}(\rho\xi')\,\xi')\,d^Nx + \operatorname{div}(\rho\nabla\xi'\cdot\xi')\,d^Nx \\
\implies \rho'' &= -\left(\operatorname{div}(\rho\xi'')\right) + \left(\left(\rho\xi'^i\right)_{,i}\xi'^j\right)_{,j} + \left(\rho\xi'^j_{,i}\xi'^i\right)_{,j} \\
&= -\left(\operatorname{div}(\rho\xi'')\right) + \left(\rho\xi'^i\xi'^j\right)_{,ij} \\
&= -\left(\operatorname{div}(\rho\xi'')\right) + \operatorname{div}\operatorname{div}\left(\rho\xi'\otimes\xi'\right).
\end{aligned}$$

Chapter 3

A Dispersive Regularization of 1D Gas Dynamics

3.1 Towards Simpler Models

The model equations introduced in the previous chapter were designed to cope with both turbulence and shock formation in compressible fluids. However, these lengthy equations involved high-order derivatives and inverse elliptic operators in many dimensions. For example, the anisotropic model (2.49) is a system of five equations, but the unknown tensor F defined in (2.42a) may contain up to nine scalar entries. For this reason, either numerical or analytical investigations of the model equations is difficult.

More importantly, an analysis of the full models we wrote down is not necessary at this stage. Before analyzing the full equations of motion, we should look into basic features of Lagrangian-averaged compressible models that may be unique to the compressible regime. Therefore, in this chapter we consider one-dimensional models. By doing so, we extract the dynamics of wave propagation from the Lagrangian-averaged compressible models, and minimize the influence of turbulent dynamics. Note that two- and three-dimensional Lagrangian averaged compressible models will feature all the complexity of incompressible turbulence *together* with shock formation. A one-dimensional (1D) model is something unique to compressible flow: if we try to enforce incompressibility, i.e.

$$\operatorname{div} u = 0,$$

in a 1D model, we end up with $u_x = 0$ and a constant velocity field. Hence 1D com-

pressible fluid models do not have any incompressible counterpart. Examining one-dimensional (1D) models will enable us to answer the following questions:

1. What types of wave motions are supported by Lagrangian-averaged models for compressible flow?

2. Can Lagrangian-averaged models for compressible flow be used to approximate shock wave solutions of the compressible Euler equations?

Let us first present the simplest possible 1D Lagrangian averaged model for barotropic compressible flow. Here ρ is density, u is velocity, and p is pressure. We assume that the $p(\rho)$ is given by the state equation $p = \kappa\rho^\gamma$. Also, we use the definition of v given by

$$v := u - \alpha^2 u_{xx}. \tag{3.1}$$

The system, which we shall call System I, is derived from the Lagrangian

$$l = \int \left[\frac{1}{2}\left(\rho u^2 + \alpha^2 \rho u_x^2\right) - \rho W(\rho)\right] dx. \tag{3.2}$$

The semidirect product Euler-Poincaré equations for this Lagrangian are

$$\rho_t + (\rho u)_x = 0 \tag{3.3a}$$

$$w_t + (uw)_x - \frac{1}{2}\left(u^2 + \alpha^2 u_x^2\right)_x = -\frac{p_x}{\rho} \tag{3.3b}$$

where we define w via

$$\rho w := \rho v - \alpha^2 \rho_x u_x. \tag{3.4}$$

Using (3.4), we can show that (3.3b) is equivalent to the following equation for u only:

$$\rho u_t + \rho u u_x - \alpha^2\left(\rho_x u_{xt} + \rho_x u u_{xx} + \rho u_x u_{xx} + \rho u_{xxt} + \rho u u_{xxx}\right) = -p_x. \tag{3.5}$$

Remark. If we set $\alpha = 0$, the system (3.3) reduces to the compressible barotropic Euler equations. However, the behavior of solutions of (3.3) as $\alpha \to 0$ may be very subtle, as we will explain later.

Outline of this chapter. In §2, we will motivate System I in a few different ways, showing how it can be derived from the full-blown models of the previous chapter or from a simple filtering argument. We also compare System I with a similar model that has appeared in the literature. Next, in §3, we describe various wave propagation properties of System I, including a large class of traveling wave solutions. In §4, we numerically study the initial-value problem for System I, with an eye towards checking its shock-approximation qualities. Finally, in §5, we compare System I with other models for wave propagation in compressible flows. We conclude with §6, which explains future projects involving System I.

Summary of results. Regarding the two questions posed above, we will show that:

1. System I supports intriguing wave phenomena and pattern formation, raising many questions of mathematical interest. Also, System I is, in a certain sense, a "compressible" generalization of a well-known dispersive wave equation.

2. System I does not approximate shock solutions of the compressible Euler equations. However, the wave solutions of System I indicate where/how we should look for a Lagrangian-averaged model for compressible flow that is capable of capturing shocks.

3.2 Motivation and Derivation

Let us explain how we arrived at the kinetic energy terms in the Lagrangian (3.2). It should be understood that the kinetic energy in (3.2) is a replacement for the usual compressible Euler kinetic energy, which is

$$\int \rho u^2 \, dx. \tag{3.6}$$

The potential energy term that we use,

$$\int \rho W(\rho) \, dx, \tag{3.7}$$

is precisely the potential energy for compressible Euler.

Favre Filtering. The kinetic energy term for System I is

$$\frac{1}{2}\int \rho\left(u^2+\alpha^2 u_x^2\right)\,dx. \tag{3.8}$$

We arrive at this term by filtering (3.6) via a simple procedure that we shall now describe. Suppose there is a filter $f \mapsto \overline{f}$ which satisfies four basic properties:

P1. $\overline{af+bg} = a\overline{f} + b\overline{g}$ for scalar constants a and b

P2. $\overline{\overline{f}} = \overline{f}$.

P3. $\overline{\overline{f}g} = \overline{f}\,\overline{g}$.

P4. $(\overline{f})_x = \overline{(f_x)}$ and $(\overline{u})_t = \overline{(u_t)}$

Following [25] or [80], we define the Favré, or density-weighted, filter as follows:

$$f \mapsto \tilde{f} = \frac{\overline{\rho f}}{\overline{\rho}}. \tag{3.9}$$

The reader can easily verify that properties P1-P3 of the $f \mapsto \overline{f}$ filter are true for the Favré filter as well. Note that if we decompose $f = \overline{f} + f'$, we must have $\overline{f} = \overline{\overline{f}} + \overline{f'}$. Then property P2 forces $\overline{f'} = 0$. The same holds for the Favré filter.

Mass transport. One might ask why we use two different filters. With our setup, we filter the continuity equation and find

$$0 = \overline{\rho_t + (\rho u)_x} = \overline{\rho}_t + (\overline{\rho u})_x = \overline{\rho}_t + (\overline{\rho}\tilde{u})_x. \tag{3.10}$$

Therefore, the filtered variables $\overline{\rho}$ and $\tilde{u}$ satisfy the ordinary continuity equation. Note that if we define the volume form $\overline{\mu} = \overline{\rho}\,dx$, then we find that (3.10) implies

$$\frac{\partial \overline{\mu}}{\partial t} = -\pounds_{\tilde{u}}\overline{\mu}.$$

This means that if we define the filtered material motion $\tilde{\eta}$ by $\dot{\tilde{\eta}} = \tilde{u} \circ \tilde{\eta}$, the filtered material density will satisfy

$$\overline{\nu} = \tilde{\eta}_* \overline{\nu}_0,$$

where $\overline{\nu}(X,t) = \overline{\mu}(\tilde{\eta}(X,t),t)$ and $\overline{\nu}(X,0) = \overline{\nu}_0$. We conclude that Favré filtering respects standard mass transport in both the spatial and material frames. This is a desirable feature of the approach.

Filtered kinetic energy. Introducing the decomposition $u = \tilde{u} + u'$, we filter the kinetic energy (3.6):

$$\int \overline{\rho u^2}\, dx = \int \overline{\rho}\widetilde{u^2}\, dx = \int \overline{\rho}\left(\widetilde{\tilde{u}^2} + 2\widetilde{\tilde{u}u'} + \widetilde{u'^2}\right) dx. \tag{3.11}$$

Note that the cross-term $\widetilde{\tilde{u}u'}$ vanishes, by application of P3 and P2:

$$\widetilde{\tilde{u}u'} = \tilde{u}\widetilde{u'} = 0.$$

By applying P3 and P2 in the same way, we derive $\widetilde{\tilde{u}^2} = \tilde{u}^2$. Now we make one assumption:

$$\int \overline{\rho}\widetilde{u'^2}\, dx \approx \int \alpha^2 \overline{\rho}\tilde{u}_x^2\, dx. \tag{3.12}$$

Here α is a small parameter with units of length. Justification of (3.12) stems from two arguments:

1. The left-hand side of (3.12) is a weighted L^2 norm of the fluctuations u'. In order to produce a closed model, we must estimate this norm in terms of $\tilde{u}$ only. A reasonable candidate for the estimate is a measure of how wiggly $\tilde{u}$ is—and this is precisely the right-hand side of (3.12). We introduce α, a small parameter with units of length, as the effective filter width. That is, we want our estimate to respect that fact that as the filter becomes finer, i.e., as $\alpha \to 0$, less mass from u goes into u'. That is,

$$\int \overline{\rho}\widetilde{u'^2}\, dx \to 0 \quad \text{as} \quad \alpha \to 0.$$

2. If we carry out the same filtering procedure in the incompressible regime, we will find that the modeling assumption

$$\int \|u'\|^2 \, \mathrm{d}^n x \approx \int \alpha^2 \|\nabla u\|^2 \, \mathrm{d}^n x.$$

yields the H^1 kinetic energy found in LAE-α models.

Putting all of this together, we see that (3.11) becomes

$$\int \overline{\rho u^2} \, dx \approx \int \overline{\rho}\tilde{u}^2 + \alpha^2 \overline{\rho}\tilde{u}_x^2 \, dx. \tag{3.13}$$

Remarks.

1. In the previous chapter, we derived a rather general form of an averaged Lagrangian for compressible flow. The averaging in that derivation took place over small-scale material-frame fluctuations ξ^ε. These fluctuations in the material picture induce fluctuations u', u'' and ρ', ρ'' in the spatial picture. If we ignore the density fluctuations, then the final averaged Lagrangian reduces to

$$l^\alpha_{\text{comp}}(u,\rho) = \int_M \left\{ \frac{1}{2}\rho\|u\|^2 - \rho W(\rho) + \frac{\alpha^2}{2}\rho \left\langle \left\| \frac{D\xi'}{Dt} \right\|^2 \right\rangle \right\} dx. \tag{3.14}$$

If we now assume that the fluctuations ξ' are Lie-advected by the mean flow u and that the Lagrangian covariance tensor is the identity, i.e.,

$$\begin{aligned} \partial_t \xi' &= -\pounds_u \xi' \\ \left\langle \xi'^j \xi'^k \right\rangle &= \delta^{jk}, \end{aligned}$$

we will find that

$$\left\langle \left\| \frac{D\xi'}{Dt} \right\|^2 \right\rangle = u^i_{,j} u^i_{,k} \left\langle \xi'^j \xi'^k \right\rangle = \|\nabla u\|^2 ,$$

and the kinetic energy from (3.14) becomes

$$\frac{1}{2}\int_M \rho \, \|u\|^2 + \alpha^2 \rho \, \|\nabla u\|^2 \, dx.$$

Restricting to one spatial dimension yields the kinetic energy for System I.

2. Also in the previous chapter[1], we mentioned that we would consider the 1D model with Lagrangian

$$l = \int \frac{1}{2}\rho uv - \rho W(\rho)\, dx,$$

where $v = u - \alpha^2 u_{xx}$. Unfortunately, the "kinetic energy" ρuv suffers a fatal defect since it fails to be positive definite. Note that when ρ is not present, we have, by integration by parts,

$$\int uv\, dx = \int u^2 + \alpha^2 u_x^2\, dx.$$

However, when ρ is present, one obtains instead

$$\int \rho uv\, dx = \int \left[\rho - \frac{1}{2}\alpha^2 \rho_{xx}\right] u^2 + \alpha^2 \rho u_x^2\, dx. \tag{3.15}$$

The most we can guarantee about ρ is that $\rho(x,t) > 0$ always. This, however, is *not* sufficient to enforce positivity of the quantity in square brackets in (3.15). Indeed, for any $M < 0$ it is possible to find $\rho > 0$ such that $\rho(x) - (1/2)\alpha^2 \rho_{xx}(x) < M$ for all x in a set of positive measure. Hence $\int \rho uv\, dx$ is not bounded from below and cannot be used as a kinetic energy.

3. Using different formalism and an averaging procedure that differs from both that of Chapter 2 and the Favré-filtering approach, D. D. Holm has derived the "Eulerian mean polytropic gas equations" (see [34]). Holm's potential energy is precisely the same as ours, and his kinetic energy is

$$\frac{1}{2}\int \rho \left(u^2 + w u_x^2\right)\, dx,$$

where $w = \langle \xi' \xi' \rangle$ is an "evolving width," i.e., mean covariance of material frame fluctuations. Our kinetic energy corresponds to taking this mean covariance to be constant and positive: $w = \alpha^2$. Holm derives the following evolution equation for

[1] See "filtered Lagrangians," "simpler models," and equation (2.55) in Chap. 2, §7.

w:

$$w_t + uw_x = 0.$$

Hence $w(x,t) = \alpha^2$ is the global unique solution corresponding to the initial data $w(x,0) = \alpha^2$. So $w = \alpha^2$ is in fact an invariant manifold of initial conditions for the general Eulerian mean model. Holm points this out and refers to our model as the "polytropic gas alpha model." No analysis of solutions of the model is carried out.

3.2.1 Equations of Motion

We regard the Lagrangian (3.2) as a functions of the vector field u and the volume form $\mu = \rho\, dx$. The equations of motions corresponding to these Lagrangians are the semidirect product Euler-Poincaré equations:

$$\frac{\partial}{\partial t}\frac{\delta l}{\delta u} = -\operatorname{ad}^*_u \frac{\delta l}{\delta u} + \frac{\delta l}{\delta \mu} \diamond \mu. \tag{3.16}$$

For a detailed derivation of these equations of motion and general definitions of all the terms involved, we refer to [40]. In the case we consider, where all functions, fields, and forms are defined over a one-dimensional manifold, we only need the forms of the ad^* and $\diamond$ operators as specified in Chapter 1. For a $1-$form θ and a vector field u, we have

$$\begin{aligned} \operatorname{ad}^*_u(\theta \otimes dx) &= (\pounds_u \theta + \theta \operatorname{div} u) \otimes dx \\ &= (\mathbf{d}\mathbf{i}_u\theta + \mathbf{i}_u\mathbf{d}\theta + u_x\theta) \otimes dx \\ &= (\mathbf{d}\mathbf{i}_u\theta + u_x\theta) \otimes dx. \end{aligned} \tag{3.17}$$

Also, for a scalar function f and a $1-$form μ, we have

$$f \diamond \mu = \mathbf{d}f \otimes \mu. \tag{3.18}$$

We compute the variational derivatives $\delta l/\delta u$ and $\delta l/\delta \mu$ for the Lagrangian (3.2), and then use these expressions in (3.16).

3.2.2 Derivation of System I

We begin computing the Euler-Poincaré equations for the Lagrangian (3.2). The computations lead to

$$\frac{\delta l}{\delta u} = \left(\rho u - \alpha^2 \rho u_{xx} - \alpha^2 \rho_x u_x\right) dx \otimes dx. \tag{3.19}$$

The quantity inside parentheses is actually the momentum for System I, which we denote

$$\rho w = \rho v - \alpha^2 \rho_x u_x. \tag{3.20}$$

Using this definition, the formula (3.19) becomes

$$\frac{\delta l}{\delta u} = \rho w \, dx \otimes dx. \tag{3.21}$$

By applying formula (3.17) to (3.21), we get:

$$\mathrm{ad}^*_u \frac{\delta l}{\delta u} = \left[(\rho u w)_x + \rho u_x w\right] dx \otimes dx. \tag{3.22}$$

The computation of $\delta l/\delta \mu$ gives

$$\frac{\delta l}{\delta \mu} = \frac{1}{2}\left(u^2 + \alpha^2 u_x^2\right) - s(\rho), \tag{3.23}$$

where $s(\rho)$ is the enthalpy, defined by

$$s(\rho) = W(\rho) + \rho W'(\rho). \tag{3.24}$$

We introduce the pressure function p using

$$s_x = \frac{p_x}{\rho}. \tag{3.25}$$

By using (3.18) and (3.25) we can write

$$\frac{\delta l}{\delta \mu} \diamond \mu = \left[\frac{1}{2}\rho\left(u^2 + \alpha^2 u_x^2\right)_x - p_x\right] dx \otimes dx. \tag{3.26}$$

Plugging (3.21), (3.22) and (3.26) into (3.16) we obtain the momentum equation for System I:

$$\frac{\partial}{\partial t}(\rho w) + (\rho u w)_x + \rho u_x w - \frac{1}{2}\rho(u^2 + \alpha^2 u_x^2)_x = -p_x, \tag{3.27}$$

where ρw is given by (3.20). We also recall that the semidirect product theory treats the volume form $\rho\, dx$ as an advected quantity. This assumption leads to the continuity equation:

$$\frac{\partial \rho}{\partial t} + (\rho u)_x = 0. \tag{3.28}$$

By combining (3.27) and (3.28) we can derive:

$$w_t + (uw)_x - \frac{1}{2}\left(u^2 + \alpha^2 u_x^2\right)_x = -\frac{p_x}{\rho}. \tag{3.29}$$

In terms of u, the PDE (3.29) reads

$$\rho u_t + \rho u u_x - \alpha^2\left(\rho_x u_{xt} + \rho_x u u_{xx} + \rho u_x u_{xx} + \rho u_{xxt} + \rho u u_{xxx}\right) = -p_x. \tag{3.30}$$

3.2.3 Conserved Energies

The Lagrangian (3.2) is of the form $K - V$, where K and V stand for kinetic and potential energy, respectively. As one might expect, the system conserves an energy of the form $K + V$. In fact, more is true. We state without proof that starting from the Hamiltonian

$$h = \int \left[\frac{1}{2}\left(\rho u^2 + \alpha^2 \rho u_x^2\right) + \rho W(\rho)\right] dx, \tag{3.31}$$

one may apply the semidirect product Lie-Poisson equations (see [58]) to derive System I. Since h contains no explicit time-dependence, we have by Noether's theorem the easy consequence that h is conserved in time along solutions of System I. Of course, because we are dealing with an infinite-dimensional Hamiltonian system, this energy conservation is, at this stage, purely formal. To rigorously verify this, we would have to prove that the initial-value problem for System I is well-posed in the appropriate function spaces.

As an example, we might seek a proof that given $\rho(x,0) \in W^{1,1}(\mathbb{R}) \cap W^{1,\infty}(\mathbb{R})$ and $u(x,0) \in H^2(\mathbb{R})$, the solution $\rho(\cdot,t) \in W^{1,1} \cap W^{1,\infty}$ and $u(\cdot,t) \in H^2$ for all $t > 0$. If

we had this proof then, trivially, $dh/dt = 0$ along the solution. We will not prove any results on the initial-value problem for System I here. We merely note that when we solve System I numerically, we expect that wherever the solution is smooth, the energy h will be conserved.

3.3 Traveling Waves

Typically, one searches for traveling wave solutions by way of the ansatz

$$u = u\left(\frac{x-ct}{\alpha}\right) \quad \text{and} \quad \rho = \rho\left(\frac{x-ct}{\alpha}\right). \tag{3.32}$$

However, System I is Galilean invariant. By this we mean that the equations (3.3a) and (3.5) are unchanged by the transformation

$$\tilde{x} \mapsto x + u_0 t$$
$$\tilde{u} \mapsto u + u_0,$$

for any $u_0 \in \mathbb{R}$. This implies that if the system has traveling wave solutions $u(x-ct)$ for some fixed $c = c_0$, then it has solutions for all c. Hence for System I it is sufficient to consider $c = 0$, so we take the ansatz

$$u = u(x/\alpha) \quad \text{and} \quad \rho = \rho(x/\alpha). \tag{3.33}$$

This traveling wave form includes a factor of α^{-1} in the argument, which we shall see eliminates α from the resulting traveling wave ode. In what follows, the primes represent differentiation with respect to the variable $z = x/\alpha$.

Continuity equation. We start with (3.3a) and introduce the ansatz (3.33), resulting in

$$u\rho' + \rho u' = 0,$$

or

$$\rho'/\rho = -u'/u. \tag{3.34}$$

This can be trivially integrated, and the answer is

$$\rho = B/u. \tag{3.35}$$

for an arbitrary constant B.

Momentum equation. Next we consider (3.3b), with pressure term given by $p = \kappa\rho^\gamma$. Using the ansatz (3.33) and equation (3.34) in (3.4), we find that w is related to u by

$$w = u - u'' + \frac{u'^2}{u}. \tag{3.36}$$

The barotropic state equation becomes

$$\frac{p'}{\rho} = \kappa\gamma\rho^{\gamma-2}\rho'. \tag{3.37}$$

Using (3.34-3.37) in equation (3.3b), we can derive

$$u' - u''' = \kappa\gamma B^{\gamma-1}u^{-\gamma-1}u'. \tag{3.38}$$

Trivially integrating both sides, we obtain

$$u'' = u + \kappa B^{\gamma-1}u^{-\gamma} - C_1, \tag{3.39}$$

where $C_1 \in \mathbb{R}$ is an arbitrary constant. As this is a single second-order equation, we may chose values for the various constants and plot the phase portrait numerically—see Figure 3.1 for the specific case $\kappa = 0.4$, $\gamma = 1.4$, $B = 1$, $C_1 = 2$. The phase portrait features two fixed points marked in red. The left fixed points is clearly a nonlinear center surrounded by periodic orbits. The right fixed point is a saddle with a homoclinic orbit plotted in green. Note that the only trajectories that remain bounded as $z \to \pm\infty$ are the homoclinic orbit and the periodic orbits contained inside.

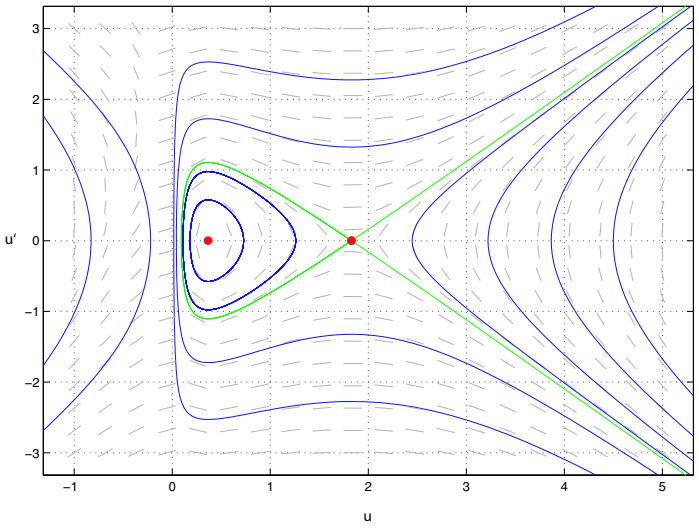


Figure 3.1: Sample phase portrait u vs. u' for ODE (3.39), with $\kappa = 0.4$, $\gamma = 1.4$, $B = 1$, $C_1 = 2$.

There is a line of essential singularities at $u = 0$. Existence and uniqueness of solutions hold for initial conditions in the set $\{(u, u') \in \mathbb{R}^2 \text{ s.t. } u \neq 0\}$, including points arbitrarily close to $u = 0$. However, it is possible for two or more different trajectories to have α- or ω-limit sets[2] consisting of the same point on the line $u = 0$.

Now let us attempt to show that the general phase portrait of (3.39) resembles what we saw numerically in Figure 3.1.

It is obvious that (3.39) is Hamiltonian. Let $K(u') = u'^2/2$ denote the kinetic energy; then, with the potential energy

$$V(u) = -\frac{1}{2}u^2 + C_1 u + \frac{\kappa B^{\gamma-1}}{\gamma - 1}u^{1-\gamma}, \tag{3.40}$$

we may write the Hamiltonian for (3.39) as

$$H(u, u') = K(u') + V(u). \tag{3.41}$$

[2]Here we mean α- and ω-limit sets in the dynamical systems sense (see [1]). There is no connection between this and the α that we use elsewhere.

Note: for $\gamma = 1$, the potential $V(u)$ must be modified, as must the following discussion. We tacitly assume $\gamma \neq 1$ in what follows. Continuing, the fixed-energy level set $H(u, u') = C_2$ is given by the locus of points (u, u') that satisfy

$$\frac{1}{2}u'^2 - \frac{1}{2}u^2 + C_1 u + \frac{\kappa B^{\gamma-1}}{\gamma - 1}u^{1-\gamma} = C_2. \tag{3.42}$$

The equilibria of (3.39) are the extrema of V. Hence we search for zeros of

$$V'(u) = -u + C_1 - \kappa B^{\gamma-1}u^{-\gamma}. \tag{3.43}$$

Let us keep in mind that C_1, C_2, and B are constants that appear in the ODE only—they do not appear anywhere in System I, as is obvious from (3.3). By adjusting the values of these parameters, one can obtain an infinite number of traveling wave solutions for System I. To show this, we first prove

Lemma 1. *Given any pair of distinct reals u_1, u_2 with the same sign, there exist B and C_1 such that $V'(u_1) = V'(u_2) = 0$. This is true for all $\kappa, \gamma > 0$.*

Proof. Regard u_1, u_2, κ, and γ as given and solve the system

$$-u_1 + C_1 - \kappa B^{\gamma-1}u_1^{-\gamma} = 0$$
$$-u_2 + C_1 - \kappa B^{\gamma-1}u_2^{-\gamma} = 0$$

for B and C_1. One finds that

$$B = \left[\frac{1}{\kappa}\left(\frac{u_2 - u_1}{u_2^\gamma - u_1^\gamma}\right)(u_1 u_2)^\gamma\right]^{1/(\gamma-1)}$$
$$C_1 = \frac{u_1^{1+\gamma} - u_2^{1+\gamma}}{u_1^\gamma - u_2^\gamma}.$$

C_1 is well-defined since $u_1 \neq u_2$. Because u_1 and u_2 have the same sign, for all $\kappa, \gamma > 0$, the argument inside square brackets in the expression for B is always positive. Hence we may always raise this argument to the $1/(\gamma - 1)$ power, so B is well-defined. □

Now that we have established that the equilibria of (3.39) may be arbitrarily chosen,

we should examine their stability.

Lemma 2. *Suppose u_1, u_2 satisfy the hypotheses of Lemma 1, and suppose $|u_1| < |u_2|$. Then u_1 is stable and u_2 is unstable for all $\kappa, \gamma > 0$.*

Proof. First suppose $0 < u_1 < u_2$. Differentiating (3.43), we have

$$V''(u) = -1 + \kappa\gamma B^{\gamma-1} u^{-\gamma-1}. \tag{3.44}$$

We evaluate this at $u = u_1$ and $u = u_2$ and use the formula for B given by Lemma 1 to obtain

$$V''(u_1) = -1 + \gamma \frac{u_2}{u_1} \frac{1 - (u_1/u_2)}{1 - (u_1/u_2)^\gamma} \tag{3.45a}$$

$$V''(u_2) = -1 + \gamma \frac{u_1}{u_2} \frac{1 - (u_2/u_1)}{1 - (u_2/u_1)^\gamma}. \tag{3.45b}$$

The signs of (3.45a-3.45b) are determined by the behavior of the function

$$f(x) = -1 + \frac{\gamma}{x} \frac{1 - x}{1 - x^\gamma}. \tag{3.46}$$

Let us list some facts about f that can be shown using elementary calculus:

$$\lim_{x \to 0} f(x) = +\infty$$

$$\lim_{x \to 1} f(x) = 0$$

$$\lim_{x \to +\infty} f(x) = -1$$

$$f'(x_0) = 0 \iff 1 + \gamma(1 - x_0) = x_0^{-\gamma}$$

$$f'(x_0) = 0 \Longrightarrow f''(x_0) < 0.$$

It is impossible for a function to have *only* local maxima while decreasing from $+\infty$ to zero, and then from zero to -1. Hence f does not have any critical points on $(0, \infty)$, implying $f'(x) < 0$ for all $x > 0$. Now we return to (3.45a-3.45b). Because u_1 and u_2

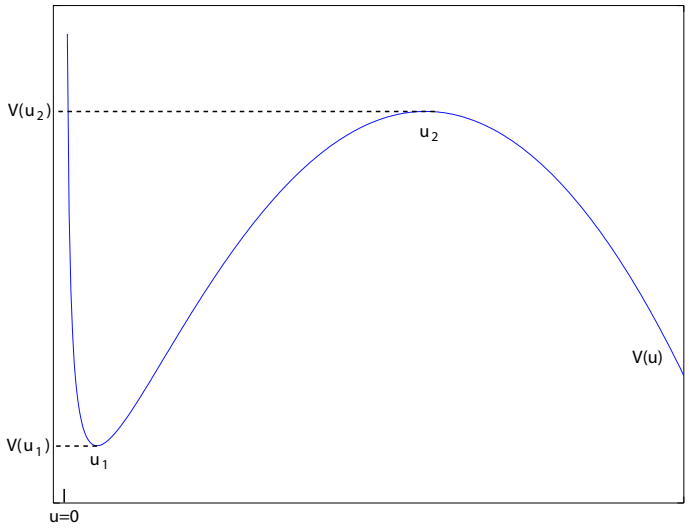


Figure 3.2: Potential energy for the Hamiltonian ODE (3.39).

have the same sign,

$$\frac{|u_1|}{|u_2|} = \left|\frac{u_1}{u_2}\right| = \frac{u_1}{u_2}.$$

As $|u_1|/|u_2| < 1$ and $f(x) > 0$ for $x < 1$, we have $V''(u_1) > 0$, so u_1 is stable. Similarly, since $f(x) < 0$ for $x > 1$ and $|u_2|/|u_1| > 1$, we have $V''(u_2) < 0$, so u_2 is unstable. □

Using Lemmas 1 and 2, for any $\kappa, \gamma > 0$, and for any u_1, u_2 satisfying $0 < u_1 < u_2$, we may fix various constants so that the potential energy V is as shown in Figure 3.2. To answer the question of what happens when we choose $u_2 < u_1 < 0$, an inspection of the potential (3.40) reveals that we must specify γ, at least in the exponent of u. The constant γ in the barotropic state law is the ratio of specific heats for the compressible fluid. For an ideal gas, we may derive (see [85]) the relationship $\gamma = 1 + 2/n$, where n is the total number of degrees of freedom for each gas molecule. With this in mind, it is clear that $1 - \gamma = -2/n$ and hence the potential V is invariant under the reflection $u \mapsto -u$, $C_1 \mapsto -C_1$. Then V satisfies

$$\lim_{u \to 0} V(u) = +\infty$$

$$\lim_{u \to \pm\infty} V(u) = -\infty,$$

and the behavior of V for $u < 0$ is given by the reflection of Figure 3.2 across the $u = 0$ line.

For the purposes of solving the System I PDE, we are interested in solutions $u(z)$ of the ODE (3.39) that stay bounded as $z \to \pm\infty$. From the above description of the potential together with the energy integral (3.42), it is clear that such solutions exist when

$$V(u_1) \leq C_2 \leq V(u_2).$$

Indeed, if we set $C_2 = V(u_2)$, we find that the solution of

$$\frac{1}{2}u'^2 = V(u_2) - V(u)$$

is the homoclinic orbit for the fixed point $(u, u') = (u_2, 0)$. Taking C_2 such that $V(u_1) < C_2 < V(u_2)$, we find that (3.42) describes periodic trajectories of the system. Let us quickly summarize this result:

Lemma 3. *Let $\kappa > 0$ be arbitrary and let $\gamma = 1 + 2/n$ for some positive integer n. Given any u_1, u_2 of the same sign, we may choose B and C_1 such that the Hamiltonian ODE (3.39) has a homoclinic orbit $u(z)$ that connects the u_2 equilibrium with itself. The solution $u(z)$ is a real-analytic function of z.*

Proof. Without loss of generality, suppose $|u_1| < |u_2|$. Examine Figure 3.2 and denote by u^* the first point of intersection of the upper dashed line with the blue $V(u)$ curve. That is, define u^* such that $|u^*| < |u_1|$ and

$$V(u^*) = V(u_2).$$

Imagine a fictitious particle situated on the graph, that starts from rest at the point $(u^*, V(u_2))$. This particle will reach $(u_2, V(u_2))$ in an infinite amount of "time." That

is, the trajectory $(u(z), u'(z))$ that emanates from $(u(0), u'(0)) = (u^*, 0)$ satisfies

$$\lim_{z\to\pm\infty} u(z) = u_2$$
$$\lim_{z\to\pm\infty} u'(z) = 0.$$

This is the homoclinic trajectory in question, connecting the fixed point $(u_2, 0)$ to itself. Note that we may force $|u^*|$ to be as small as we want by choosing $|u_1|$ sufficiently small. That the solution $u(z)$ is real-analytic follows[3] from the fact that the right-hand side of (3.39) is a real-analytic function of u away from the singularity line $u = 0$. □

Solving the PDE. The technique of finding traveling wave solutions of a PDE by searching for homoclinic orbits of an associated ODE can be found in Chapter 3 of [56]. In our case, the analogy is trivial, because the homoclinic solutions $u(z)$ of the Hamiltonian ODE (3.39) are related to solutions of the System I PDE by the ansatz (3.32).

Because of Galilean invariance, we investigated only the $c = 0$ case of this ansatz. For arbitrary c, it is a simple matter to see that (3.39) would now possess a line of singularities at $u = c$ instead of at $u = 0$. Adjusting the results accordingly, we arrive at the following:

Theorem 1. *Let $\alpha > 0$ and $c \in \mathbb{R}$ be arbitrary. Let $\kappa > 0$ be arbitrary and let $\gamma = 1+2/n$ for some positive integer n. Given any u_1, u_2 such that $u_1 - c$ and $u_2 - c$ have the same sign, the System I PDE (3.3) considered on the domain $(x, t) \in \mathbb{R} \times \mathbb{R}$ has either of two possible traveling wave solutions.*

1. *In case $u_2 > u_1 > c$, we have the* depression *wave solution $u(x, t), \rho(x, t)$ such that*

$$\lim_{x\to\pm\infty} u(x, t) = u_2$$
$$u(ct, t) = u^*$$
$$\rho(x, t) = \frac{B}{u(x, t) - c},$$

[3] For a proof that a real-analytic vector field possesses real-analytic integral curves, see [6, Chap. 6.10].

where $c < u^* < u_1$.

2. *In case* $u_2 < u_1 < c$, *we have the* elevation *wave solution* $u(x,t)$ *such that*

$$\lim_{x\to\pm\infty} u(x,t) = u_2$$
$$u(ct,t) = u^*$$
$$\rho(x,t) = \frac{-B}{u(x,t)-c},$$

where $u_1 < u^* < c$.

Proof. In what follows, we label the points such that $c < |u_1| < |u_2|$. As all the constants in the theorem satisfy the conditions of Lemmas 1, 2, and 3, we may solve the Hamiltonian ODE (3.39) for a homoclinic orbit $u^0(z)$ that connects u_2 to itself. We take[4]

$$\rho^0(z) = \frac{\pm B}{u^0(z)-c},$$

and choose either the plus sign if $u(z) > c$ or the minus sign if $u^0(z) < c$. The solution then satisfies $\rho^0 > 0$ everywhere. Now it is clear from the ansatz (3.33) that the functions

$$u(x,t) = u^0\left((x-ct)/\alpha\right)$$
$$\rho(x,t) = \rho^0\left((x-ct)/\alpha\right)$$

satisfy System I, i.e., the PDE system 3.3, for all x and all t. Then we may compute the movement of the trough/peak of the wave:

$$u(ct,t) = u^0(0) = u^*$$

as in Lemma 3. □

Physical solutions. In many physical differential equations, one seeks solutions with a finite amount of energy, at least in some appropriate norm. Not all the solutions we just

[4]Note that the constant B entered the ODE analysis only through its $\gamma - 1$ power. Since $\gamma - 1 = 2/n$, none of the ODE results change if we replace B with $-B$.

described satisfy this criterion. Indeed, if one fixes $c > 0$ and considers the depression wave solution from Theorem 1, one finds that as $x \to \pm\infty$, $u(x,t) \to u_2 > 0$. The L^p norm of this solution is unbounded for any $p < \infty$.

Let us pick out the *solitary wave* solutions of System I. These are traveling wave solutions $u(x,t)$ that decay to zero as $x \to \pm\infty$. For $c > 0$, we take $u_2 = 0$ and u_1 such that $0 < u_1 < c$. Then the elevation wave solution is a solitary wave. For $c < 0$, we again take $u_2 = 0$ and u_1 such that $c < u_1 < 0$. Then the depression wave solution is a solitary wave.

To summarize, for System I, we can have localized waves of elevation ($u > 0$) that travel to the right at speed c, and localized waves of depression ($u < 0$) that travel to the left at speed c. The amplitudes of both types of waves are bounded above by c, the absolute value of the wave speed.

Because these solitary wave solutions decay to zero as $x \to \pm\infty$, we expect that they are bounded in certain L^p norms. We will not make this statement precise here. It would be interesting to check this, and also check whether the solitary waves are local minima of the Hamiltonian functional (3.31) associated with System I. This would be a first step towards proving stability of the solitary waves, an issue we will not address in this thesis.

Periodic boundary conditions. Let us reconsider System I on the finite domain $x \in [0,T]$, and impose periodic boundary conditions:

$$u(x,0) = u(x,T) \tag{3.47a}$$

$$\rho(x,0) = \rho(x,T). \tag{3.47b}$$

Now when we examine the traveling wave ODE (3.39), we must search for periodic solutions. However, it is already clear from Figure 3.2 that there are an infinite number of periodic solutions. Just imagine a fictitious particle starting from rest at any point on the $V(u)$ curve strictly between the dashed lines $V(u_1)$ and $V(u_2)$. The particle will oscillate (forever) inside the potential well with some period $T < \infty$. As we did before,

we can use this periodic solution of the ODE and the ansatz (3.33) to generate a traveling wave solution of the System I PDE, now posed with periodic boundary conditions (3.47).

Note that on the periodic interval, all traveling wave solutions are bounded in any L^p norm. This is a trivial consequence of the fact that on a compact interval, all smooth functions are integrable. We conjecture that the periodic version of System I may support a wealth of phenomena that may not occur for System I posed on the real line.

Dispersion relation. For any wave equation, knowledge of the dispersion relation $\omega(k)$ is quite important. If $\omega(k)$ is linear, then the phase speed ω/k is constant with respect to k, implying that waves of different wave number all propagate at the same speed. In such a case, we refer to the dynamics as *hyperbolic*. If $\omega(k)$ is nonlinear, then the phase speed ω/k is not constant with respect to k, implying that waves of different wave number propagate at different speeds. In such a case, if we start with a packet consisting of several wavenumbers and let is propagate under the dynamics of the wave equation, we will find that the packet will disperse, or spread out, in space. In this case, we refer to the dynamics as *dispersive*.

With this terminology, now widely used but originally due to Whitham (see [86]), the 1D barotropic compressible Euler equation is a nonlinear hyperbolic equation. Typical behavior in such a system is steepening of wave profiles and the formation of sharp discontinuities, or shock waves. As we have seen, System I has a large family of solitary wave solutions, which is typical of nonlinear dispersive equations such as the Korteweg-de Vries (KdV), Nonlinear Schrödinger (NLS), and sine-Gordon equations. As we shall now see, System I is indeed a dispersive system, but this is not the full story.

In order to compute the dispersion relation for System I, we shall have to deal with both the continuity equation (3.3a) and the velocity equation (3.5). The first equation is unchanged from the compressible Euler context, and therefore retains its hyperbolic

character. Indeed, linearizing (3.3a) about a constant solution[5] via

$$\rho(x,t) = \rho_0 + \varepsilon \exp i(kx - \omega t) \tag{3.48}$$

$$u(x,t) = u_0 + \varepsilon \exp i(kx - \omega t), \tag{3.49}$$

we will obtain, at first-order in ε, the linear dispersion relation

$$\omega(k) = (u_0 + \rho_0)k.$$

This is the standard approach for obtaining dispersion relations, but it simply does not work for System I. The reason is that System I consists of a coupled hyperbolic-dispersive system. If we use the naïve approach of (3.48) and (3.49) in both the hyperbolic part (3.3a) and the dispersive part (3.5), we will not produce anything meaningful.

Therefore, we generalize the usual procedure. We retain the usual expression (3.49) for u consisting of a small sinusoidal perturbation from a constant state. Note that this is a traveling wave solution, with phase velocity $c = \omega/k$, so we expect from our earlier analysis that $\rho(x,t)$ will take the form

$$\rho(x,t) = \frac{B}{u(x,t) - \omega/k}, \tag{3.50}$$

for some constant B to be specified later. For now, we assume the existence of δ such that

$$u_0 - \frac{\omega}{k} = \delta = O(\varepsilon^0) \gg \varepsilon. \tag{3.51}$$

Later we shall determine a closed-form expression for δ. Now we must check that (3.49-

[5]The reader will verify by inspection that System I possesses the constant solution $\rho(x,t) = \rho_0$ and $u(x,t) = u_0$.

3.50) solves the continuity equation (3.3a). We compute

$$\rho_x = -\frac{B\varepsilon ik \exp i(kx - \omega t)}{(u - \omega/k)^2} \tag{3.52}$$

$$\rho_t = \frac{B\varepsilon i\omega \exp i(kx - \omega t)}{(u - \omega/k)^2} \tag{3.53}$$

$$u_x = \varepsilon ik \exp i(kx - \omega t), \tag{3.54}$$

and the reader may verify that using this together with (3.49-3.50), we have

$$\rho_t + \rho_x u + \rho u_x = 0.$$

Hence we move on to the velocity equation (3.5), which we write as follows:

$$u_t + uu_x - \alpha^2 \left(\frac{\rho_x}{\rho} (u_{xt} + uu_{xx}) + u_x u_{xx} + u_{xxt} + uu_{xxx} \right) = -\frac{p_x}{\rho}. \tag{3.55}$$

We are interested in the $O(\varepsilon)$ terms from (3.55). First let us put (3.49) together with (3.51) to get

$$u - \frac{\omega}{k} = \delta + O(\varepsilon) = O(\varepsilon^0). \tag{3.56}$$

At this point, it is clear that $u_x, u_{xx} = O(\varepsilon)$ and that

$$\frac{\rho_x}{\rho} = -\frac{u_x}{u - \omega/k} = \frac{O(\varepsilon)}{O(\varepsilon^0)} = O(\varepsilon).$$

Already we may eliminate three $O(\varepsilon^2)$ terms from (3.55) and rewrite it as

$$u_t + uu_x - \alpha^2 (u_{xxt} + uu_{xxx}) = -\frac{p_x}{\rho}. \tag{3.57}$$

Now we choose the scalings $B = O(\varepsilon)$ and $\kappa = O(\varepsilon^{1-\gamma})$ in order to achieve the following

balance:

$$\rho = \frac{B}{u - \omega/k} = O(\varepsilon),$$
$$p = \kappa\rho^{\gamma} = O(\varepsilon),$$
$$\frac{p_x}{\rho} = \kappa\gamma\rho^{\gamma-1}\frac{\rho_x}{\rho} = O(\varepsilon).$$

Hence all terms from (3.57) are $O(\varepsilon)$, so we may proceed to compute the dispersion relation. Note that from (3.56) we have

$$\frac{1}{u - \omega/k} = \frac{1}{\delta + O(\varepsilon)} = \frac{1}{\delta} + O(\varepsilon),$$

and raising both sides to the γ power, we obtain

$$(u - \omega/k)^{-\gamma} = \delta^{-\gamma}$$

where we are ignoring a term of order $O(\varepsilon^{\gamma})$. In light of this, we may compute p_x/ρ:

$$\frac{p_x}{\rho} = -\kappa\gamma B^{\gamma-1}\delta^{-\gamma}\varepsilon ik \exp i(kx - \omega t).$$

Remembering that the right-hand side of (3.57) has a minus sign, we obtain, after substitution of our p_x/ρ result and (3.49),

$$-\omega(1 + \alpha^2k^2) + u_0k(1 + \alpha^2k^2) = \kappa\gamma B^{\gamma-1}\delta^{-\gamma}k.$$

This leads to

$$\omega(k) = u_0k - \kappa\gamma B^{\gamma-1}\delta^{-\gamma}\frac{k}{1 + \alpha^2k^2}.$$

Now we may attempt to "back out" an expression for δ. Let us compute

$$\delta = u_0 - \frac{\omega}{k} = \kappa\gamma B^{\gamma-1}\delta^{-\gamma}(1 + \alpha^2k^2)^{-1}.$$

Solving for δ and raising the answer to the γ power, we obtain

$$\delta^\gamma = \left(\frac{\kappa\gamma B^{\gamma-1}}{1+\alpha^2 k^2} \right)^{\frac{\gamma}{\gamma+1}}.$$

Now the dispersion relation becomes

$$\frac{\omega(k)}{k} = u_0 - \left(\frac{\kappa\gamma B^{\gamma-1}}{1+\alpha^2 k^2} \right)^{1/(\gamma+1)}. \tag{3.58}$$

Camassa-Holm. Interestingly, the dispersion relation for System I bears a resemblance to that of the Camassa-Holm equation:

$$v_t + uv_x + 2vu_x = 0 \tag{3.59a}$$

$$u - \alpha^2 u_{xx} = v. \tag{3.59b}$$

This equation was originally derived in [9] by vertically averaging the Hamiltonian for shallow water waves. Equation (3.59) is a completely integrable equation with a bi-Hamiltonian structure. Its geometric and analytical properties have been extensively studied. To compute its dispersion relation, we follow the usual procedure of taking sinusoidal perturbations about a constant solution as in (3.49). We obtain

$$\frac{\omega(k)}{k} = u_0 + \frac{2u_0}{1+\alpha^2 k^2}. \tag{3.60}$$

Also note that the Camassa-Holm equation (3.59) is the Euler-Poincaré equation for the Lagrangian

$$l(u) = \int u^2 + \alpha^2 u_x^2 \, dx.$$

Comparing this with the System I Lagrangian (3.2), we see that both in terms of geometry and wave dynamics, System I can be considered a "compressible" version of Camassa-Holm.

Zero-α limit of solutions. It is clear from the dispersion relation (3.58) that α is not only a length scale that arises in the filtering/averaging derivation, but also a measure

of dispersion in the model. For $\alpha > 0$, the dispersion relation $\omega(k)$ for System I depends *nonlinearly* on k. This means that both the phase velocity ω/k and the group velocity $d\omega/dk$ depend on the wavenumber k. System I, and more specifically the momentum equation (3.5), is a dispersive regularization of the 1D barotropic compressible Euler equation.

Suppose we solve System I, on either the real line or a periodic interval, with initial data $\rho(x,0) = \rho_0$ and $u(x,0) = u_0$. Let us explicitly label the α-dependence of the resulting solution by writing it as $u^\alpha(x,t)$, $\rho^\alpha(x,t)$. The next question to ask is: what happens to these solutions as $\alpha \to 0$, or in other words, how do solutions of System behave in the zero-dispersion limit?

Let us take a slight digression and discuss the zero-dispersion limit of the Korteweg-de Vries (KdV) equation:

$$\begin{aligned} u_t + uu_x + \epsilon u_{xxx} &= 0 \\ u(x,0) &= u_0(x) \end{aligned} \tag{3.61}$$

The zero-ϵ limits of solutions of this equation is a zero-dispersion limit, and it has been pursued quite vigorously in the literature (see [52]). It is trivial to take $\epsilon \to 0$ in the KdV equation itself, and one obtains the (inviscid) Burgers equation

$$\begin{aligned} u_t + uu_x &= 0 \\ u(x,0) &= u_0(x) \end{aligned} \tag{3.62}$$

Taking $\epsilon \to 0$ in the *solutions* of the KdV equation is a completely different matter. The basic idea is that globally in time, as $\epsilon \to 0$, the solutions $u^\epsilon(x,t)$ of (3.61) equations do *not* converge to solutions of the limiting equation, which is (3.62). There is, of course, much more to the story than that:

- Assuming that $u_0(x)$ contains at least one point x_0 at which $u_0'(x) < 0$, the resulting solution $u(x,t)$ of (3.62) will develop a discontinuity in finite time. This happens regardless of how smooth $u_0(x)$ is. Let T denote the earliest time at which the

solution $u(x,t)$ breaks—we refer to T as the *break time*.

- Now, for $t < T$, solutions $u^\epsilon(x,t)$ of the KdV equation (3.61) converge strongly to solutions $u(x,t)$ of the Burgers equation (3.62) with the same initial data.

- However, for $t > T$, solutions $u^\epsilon(x,t)$ of the KdV equation do not converge in *any* sense to weak solutions $u(x,t)$ of the Burgers equation. The solution $u^\epsilon(x,t)$ becomes highly oscillatory as ϵ vanishes—these oscillations are bounded in L^∞ but their frequency increases. In fact, it can be shown that $u^\epsilon(x,t)$ converges weakly to a solution of the Whitham modulation equation for KdV, which is quite different from the Burgers equation.

The above results summarize the content of the series of papers by Lax and Levermore ([52]) on the zero-dispersion limit of the KdV equation. One should also mention the contributions of Venakides, who established many results on the fine structure of the oscillations that develop as $\epsilon \to 0$, as well as Deift and Zhou, who have framed the zero-dispersion limit as a Riemann-Hilbert problem ([18]). The Riemann-Hilbert approach has also been used (see [46] and [84]) to analyze the semiclassical, or zero-dispersion, limit of the Nonlinear Schrödinger equation:

$$i\epsilon u_t + \frac{1}{2}\epsilon^2 u_{xx} + |u|^2 u = 0 \tag{3.63}$$
$$u(x,0) = u_0(x).$$

Qualitatively identical phenomena, including high-frequency oscillations and weak limits, have been discovered in this case.

It is reasonable to conclude that analyzing the zero-α limit of System I is most likely a subtle and mathematically challenging problem. Furthermore, it seems unlikely that, as $\alpha \to 0$, we can extract information about gas dynamics from the solution $u^\alpha(x,t)$, $\rho^\alpha(x,t)$ of System I. That is, even if u^α and ρ^α converge to some functions as $\alpha \to 0$, this convergence is likely to be weak, and the limit functions are not likely to be solutions of the 1D barotropic compressible Euler equations.

3.4 Initial-value Problem

We wish to study not only the traveling wave solutions of System I, but also the general initial-value problem for unknown functions $\rho : U \times [0, \infty) \to \mathbb{R}$ and $u : U \times [0, \infty) \to \mathbb{R}$:

$$\rho_t + (\rho u)_x = 0 \tag{3.64a}$$

$$\rho u_t + \rho u u_x - \alpha^2 \left(\rho_x u_{xt} + \rho_x u u_{xx} + \rho u_x u_{xx} + \rho u_{xxt} + \rho u u_{xxx}\right) = -p_x \tag{3.64b}$$

$$\rho(x, 0) = \rho_0(x) \tag{3.64c}$$

$$u(x, 0) = u_0(x). \tag{3.64d}$$

As usual, we take $p = \kappa\rho^\gamma$. The analytical problem consists of finding an appropriate spaces of functions such that if ρ_0 and u_0 are chosen from those spaces, then the solutions $\rho(x, t)$, $u(x, t)$ of (3.64) remain in those spaces for all $t > 0$. We will not attempt a theoretical investigation of this problem here, but instead pursue a numerical treatment. In order to sidestep issues regarding the decay of solutions as $x \to \pm\infty$, we will take the domain U to be a compact interval equipped with the standard periodic boundary conditions (3.47).

We wish to test whether solutions of (3.64) approximate the shock wave solutions of the 1D barotropic compressible Euler equations. Of course, we already have evidence that this approximation is not going to be terribly good. From our results on traveling wave solutions, we know that there exists an initial condition u_0 (and an associated initial condition ρ_0) consisting of an upward-pointing pulse. This initial condition u_0 is rigidly transported to the right at a speed c by System I. No steepening, and indeed no change whatsoever in the shape of the wave occurs as it propagates.

This is already in sharp contrast with the solution of the 1D compressible Euler equation with the same initial data. In this case, the velocity field u would steepen and eventually form a shock wave at the point of inflection of u_0. At the instant at which this shock forms, the L^2 energy of the solution would drop. By comparison, the L^2 energy of the traveling wave solution stays constant for all time.

Description of the numerical method. We describe how to numerically solve System I (3.3a, 3.5) on the interval $[0,1]$ with periodic boundary conditions. We use the barotropic law $p = \kappa\rho^\gamma$, with $\kappa = 0.4$, $\gamma = 1.4$. Pseudospectral techniques, as described in [28], play a large role in the method. First, we write the systems in semidiscrete form

$$\boldsymbol{\rho}_t = \mathbf{F}(\boldsymbol{\rho}, \mathbf{u}) \tag{3.65a}$$

$$\mathbf{u}_t = \mathbf{G}(\boldsymbol{\rho}, \mathbf{u}), \tag{3.65b}$$

where $\boldsymbol{\rho} = (\rho_1, \ldots, \rho_N)$ and $\mathbf{u} = (u_1, \ldots, u_N)$. Here, $\rho_i(t) \approx \rho(x_i, t)$ and $u_i(t) \approx u(x_i, t)$, where x_i, $i = 1, \ldots, N$, are grid points in the interval $[0,1]$. We use the equispaced grid given by $x_i = (i-1)\Delta x$ with $\Delta x = 1/N$.

The method carries the quantities $\boldsymbol{\rho}$ and $\mathbf{u}$ in physical space. We pass to Fourier space using the FFT only when we take derivatives. Before applying the inverse FFT to any quantity in Fourier space, we always use a two-thirds dealiasing rule. By this rule, the highest one-third wavenumber components of the spectrum are set to zero.

We see from (3.3a) that the ρ dynamics for System I is given by the standard continuity equation. Then, our function $\mathbf{F}(\boldsymbol{\rho}, \mathbf{u})$ is merely a pseudospectral approximation of the derivative $-(\rho u)_x$. In words, we first compute the product $(\boldsymbol{\rho}\mathbf{u})_i = \rho_i u_i$ and take the FFT of the result. After multiplying by $2\pi i\mathbf{k}$, where $\mathbf{k} = (-N/2+1, \ldots, N/2)$, and dealiasing, we take the inverse FFT, and multiply by -1.

Now we describe the construction of $\mathbf{G}$ for System I. Let us group all time-derivative terms from equation (3.5) in the following way:

$$\mathcal{A}u_t = -uu_x + \alpha^2\left(\frac{\rho_x}{\rho}uu_{xx} + u_x u_{xx} + uu_{xxx}\right) - \frac{p_x}{\rho}, \tag{3.66}$$

where $\mathcal{A} = \left[\mathrm{Id} - \alpha^2(\rho_x/\rho)\partial_x - \alpha^2\partial_{xx}\right]$. Again we use standard pseudospectral techniques to discretize and compute the right-hand side of (3.66); let us write the result of this as $\mathbf{b}$. To discretize the operator $\mathcal{A}$, we use standard high-order finite-difference approximations

for ∂_x and ∂_{xx}. These approximations are, respectively (see [43, Chap. 7.1]):

$$\mathcal{D}^1 = \frac{1}{\Delta x}\Gamma_0 \left[\mathrm{Id} - \frac{1}{6}\left(\Delta_0^2\right) + \frac{1}{30}\left(\Delta_0^2\right)^2\right] + \mathcal{O}(\Delta x^5) \tag{3.67}$$

$$\mathcal{D}^2 = \frac{1}{\Delta x^2}\left[\left(\Delta_0^2\right) - \frac{1}{12}\left(\Delta_0^2\right)^2 + \frac{1}{90}\left(\Delta_0^2\right)^3\right] + \mathcal{O}(\Delta x^6), \tag{3.68}$$

where Γ_0 and Δ_0^2 are operators on the space of vectors $\mathbf{z} = (z_1, \ldots, z_N)$, defined by

$$\left(\Delta_0^2 \mathbf{z}\right)_k = z_{k+1} - 2z_k + z_{k-1} \tag{3.69}$$

$$\left(\Gamma_0 \mathbf{z}\right)_k = \frac{1}{2}\left(z_{k+1} - z_{k-1}\right). \tag{3.70}$$

Because we impose periodic boundary conditions on the system, we use the convention that $z_k = z_{k+N}$ for all k. With this convention, the operators $\mathcal{D}^1$ and $\mathcal{D}^2$ can be written as matrices that include the periodic boundary conditions. We denote these matrices by D^1 and D^2, respectively.

Let $\mathbf{r}$ denote the result of computing ρ_x/ρ pseudospectrally. Now we are ready to write the discretization A of the operator $\mathcal{A}$:

$$A = \mathrm{Id} - \alpha^2 \,\mathrm{diag}(\mathbf{r}) D^1 - \alpha^2 D^2. \tag{3.71}$$

Here $\mathrm{diag}(\mathbf{r})$ is the diagonal matrix with $r_1, \ldots, r_N$ on the diagonal. We remark that the component-wise product of the vectors $\mathbf{r}$ and $\mathbf{z}$ can be written as the matrix-vector product $\mathrm{diag}(\mathbf{r})\mathbf{z}$.

Then the discretization of (3.66) becomes

$$A\mathbf{u}_t = \mathbf{b}, \tag{3.72}$$

which implies $\mathbf{G}(\boldsymbol{\rho}, \mathbf{u}) = A^{-1}\mathbf{b}$.

Having now described the construction of $\mathbf{F}$ and $\mathbf{G}$, we solve the ODE system (3.65) with the fourth-order Runge-Kutta method. We remark that in practice, for System I, the presence of a third-derivative term on the right-hand side of (3.66) necessitates a CFL condition of $\Delta t \propto (\Delta x)^2$.

Initial data. Using the method just described, we performed numerical experiments for System I with the following initial conditions: the `gauss` data

$$\rho(x,0) = \frac{1}{100}\sin(2\pi x) + \frac{1}{2},$$
$$u(x,0) = \frac{1}{2}\exp(-25(x-1/2)^2),$$

and the `sine` data

$$\rho(x,0) = \frac{1}{100}\sin(2\pi x) + \frac{1}{2},$$
$$u(x,0) = \frac{1}{2}\sin 2\pi x + \frac{3}{2}.$$

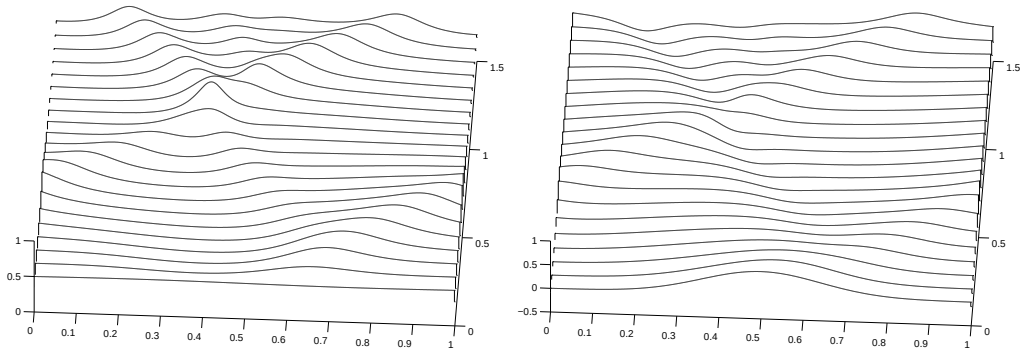
We take $\alpha = 0.03$. We solve until $t = 5$ at resolution $N = 1024$. As the solutions evolved in time, we kept track of the energy (3.31) of the solution. All solutions presented in this work preserved their initial energy to within 0.1%. See Figure 3.3 for the solutions to System I for `gauss` and `sine` initial data. Movies of these solutions are available at

`http://www.cds.caltech.edu/~bhat/pub/.pde1/`

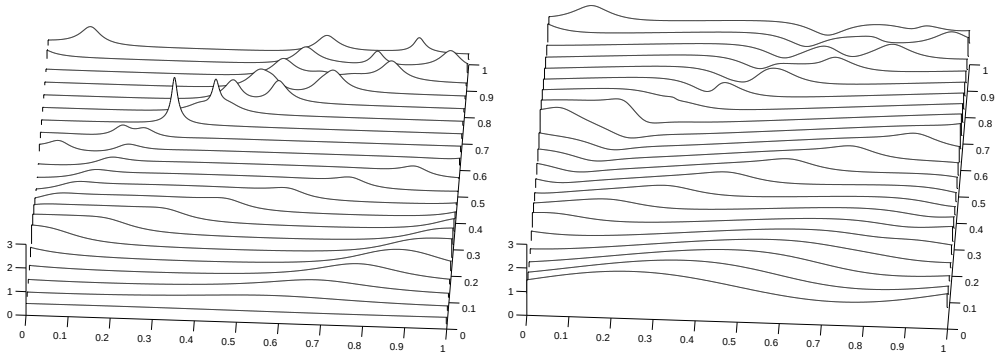
Unfortunately, our studies of the initial-value problem for System I indicate that it is not well-suited for the approximation of shock solutions of the compressible Euler equations. If one used either the `gauss` or `sine` initial conditions in the compressible Euler equations, the resulting solutions would contain shock waves at points of inflection of u. In all of our numerical experiments on System I, we have not seen any wave profiles that resemble smoothed or approximate shock waves.

For both ρ and u, the numerical solution retains its smoothness and stays bounded. Various wave-like structures of similar width appear and propagate to the right. Based on this and the fact that the energy stays very nearly constant, we conjecture that the systems are well-posed for sufficiently smooth initial data. However, analytic results on well-posedness need to be further investigated.

We note that especially in the evolution of $\rho(x,t)$ given the `sine` initial data, we see the emergence of numerous symmetric patterns and oscillations that hint at deep, fundamental structures for System I. This includes geometric structures such as higher-



(a) ρ - `gauss` (b) u - `gauss`



(c) ρ - `sine` (d) u - `sine`

Figure 3.3: The solution of System I with `gauss` and `sine` initial data, respectively. The figure shows the solution until $t = 1.5$ and $t = 1$, respectively. We performed runs until $t = 5$, and the solution manifests the same wave-like behavior, while preserving the energy h to within 0.1% of its initial value.

order symmetries for System I that are waiting to be discovered, as well as algebraic structures such as Lax pairs and other objects related to complete integrability. System I already has one Hamiltonian structure, the semidirect product Euler-Poincaré structure mentioned above (see (3.31)). If System I possessed another compatible Hamilonian structure, then it would be completely integrable in the sense of Lax.

3.5 Comparisons with Other Models

Solitary waves in compressible fluids. There have been just a handful of published studies on solitary waves in compressible fluids, featuring two main ideas. The first main idea is to use asymptotic expansions of the compressible Euler equations. At first order, such an expansion yields the linear wave equation of acoustics. The next step is to carry the expansion to higher order, and then apply the usual assumptions of weakly nonlinear theories to derive wave amplitude equations of Korteweg-de Vries (KdV) type. This procedure is carried out in [76], for a compressible fluid with free boundary, in [73], for a vertically unbounded compressible fluid, and in [77], for a two-layer compressible fluid bounded above and below by rigid plates. It is interesting to note that the resulting models are not precisely the classical KdV equation but instead feature time-dependent coefficients:

$$A_t + f(t)AA_x + g(t)A_{xxx} = 0. \tag{3.75}$$

We may infer that these models are different from System I in one important sense: the effective dispersion relation for (3.75) is time-dependent. This implies that it would be highly unlikely for (3.75) to have *any* traveling wave solutions. Furthermore, (3.75) is unlikely to have a Lagrangian structure or a conserved energy, as its derivation proceeded along the lines of classical asymptotics at the level of the equation of motion.

Of course, models such as (3.75) may be used in the regime where f and g are slowly varying function of time. In this case, an adiabatic approximation may be used, with some care, to show that the solutions of (3.75) consist of the usual sech^2 pulse solutions of KdV with a trailing shelf that degenerates into a train of dispersive waves as $t \to \infty$. This procedure is carried out in [32], and the resulted presented there do not match the

behaviors we see in numerical simulations of the initial-value problem for System I.

The second main idea is to generalize, to the case of compressible flow, a solitary wave model that has already been derived for incompressible flows. This is the approach of [62, 63] and [31], who seek "compressible" version of either the KdV or Benjamin-Davis-Ono (BDO) equation. The BDO equation models interval waves in a stratified incompressible fluid of infinite depth. These derivations are different from those described above in that new model equations are derived from first principles rather than by carrying out asymptotic expansions on old models. As such, the resulting equations may possess conserved quantities or even Lagrangian/Hamiltonian structures. These papers do not analyze the dynamics of the new model equations in any detail, and we are not aware of studies of these equations anywhere else in the literature. A hint as to why this must be the case is considered next.

Acoustic solitary waves, morning glories and undular bores. We are aware of two physical scenarios in which solitary wave formation occurs in a gas.

The first is the recently discovered acoustic solitary wave, which was predicted theoretically in 1992 and verified experimentally in 2004 (see [81] and references therein). The situation described is quite different from open air: Sugimoto considers a tube of air, with a periodic array of resonators attached to the tube. Each resonator is a cavity of a certain size that sits transverse to the tube. Using a piston, the air is forced at the left end of the tube, generating a solitary wave that propagates to the right. If the resonators were not present, the initial compression wave would break and the resulting shock wave would propagate down the tube. This suggests the application for Sugimoto's discovery: the suppression of shocks in tunnels used by high-speed trains.

As interesting as this is, the only connection to System I is that the mathematical model of the acoustic solitary waves is a coupled system of non-local, nonlinear dispersive wave equations. By non-local it is meant that writing the equations in the abstract form

$$u_t = N(u, u_x, u_{xx}, \dots), \tag{3.76}$$

we find that N depends on the values of its arguments at only one instant of time t,

but at *more* than one point x. Note that System I is also non-local: as we saw in (3.66), writing the u-equation in the form (3.76) requires inverting the elliptic operator $\mathcal{A}$. Inverting this operator requires knowledge of the right-hand side of (3.66) at all points in the domain. These sorts of nonlinear, non-local, dispersive wave equations are exceedingly rare in the literature; hence we point out this similarity between the model equations of acoustic solitary waves and System I.

The second situation in which solitary wave formation occurs in a gas is the occurrence of so-called "morning glory" waves in the atmosphere, especially above the Gulf of Carpenteria in Australia. As described in [29], these are long, nonlinear internal waves that occur close to the ground. The internal atmospheric wave can be seen when sufficient moisture exists so that the wave formation occurs together with the formation of a roll cloud. This roll cloud can be more than 100 km long, though it is typically 1-2 km wide and 1 km deep. In [61], field measurements of morning glories are used to infer that the internal waves are undular bores.

As described in the recent study [23], an undular bore is a type of "dissipationless shock wave." Mathematically, this kind of solution arises in the following way. Given certain initial conditions for the shallow water equations, the solution breaks; one or more of the spatial derivatives of the solution blows up in finite time, producing a sort of shock wave. However, beyond the break time and behind the point of blow-up, high-order dispersion effects remain significant. The discontinuity propagates with a train of dispersive oscillations. This sort of wave solution consisting of a discontinuity plus dispersive oscillations is what we would expect from the zero-α limit of System I and other conservative, dispersive α models derived using the Euler-Poincaré machinery.

These physical situations are far removed from the concerns that motivated the derivation and analysis of System I in the first place. Therefore, we do not expect System I to magically turn out to be an accurate model for any of these phenomena. However, we do expect that because other model equations are close to System I in a strictly mathematical sense, the techniques used to analyze System I will be applicable elsewhere.

3.6 Future Projects

First we briefly list projects of mathematical interest involving System I:

- One project is to study the Casimirs of System I and use them to determine the stability of the traveling wave solutions.

- The Lagrangian (3.2) consists of a kinetic energy minus a potential energy. One can apply a Kaluza-Klein construction and turn this Lagrangian into a positive-definite metric on a certain enlarged space. It would be of analytical interest to study the geodesic spray associated with this metric. For one, it may yield a new method of proving existence/uniqueness for regularized gas dynamical equations.

- Another project is to apply a Painlevé test to System I to determine whether it is completely integrable in the sense of Lax pairs. We may also study the relationship between System I and a completely integrable hierarchy of equations derived from the barotropic Euler equations by [7].

Leaving behind the mathematical issues, we next list projects of physical interest that seek to use the knowledge gained from the System I analysis:

- We found that the phase portrait for the Hamiltonian ODE (3.39) contains one saddle and one center. The homoclinic trajectory associated with the saddle corresponds to the traveling wave solution of System I. Now, what if instead of one saddle and one center, we had two saddles? Then we might expect a heteroclinic connection between the two saddles. This would correspond to a different kind of traveling wave solution: a traveling front, or smoothed shock profile.

- As we saw, System I is a dispersive wave equation. Generally speaking, the zero-dispersion limits of such equations feature high-frequency oscillations and weak convergence. What if, after our averaging procedure and subsequent application of the variational principle, we had derived a non-dispersive regularized equation, i.e., one with the linear dispersion relation

$$\omega(k) = ck,$$

for some constant c? In this case, we might legitimately expect that in the zero-α limit, solutions of the regularized equation converge strongly to solutions of the compressible Euler equations.

- Finally, instead of working with the $1 + 1$ system of density and momentum or density and velocity, we might profit from taking a step back and working with the inviscid Burgers equation (3.62). This is the simplest mathematical model for shock formation. If there is no way to regularize (3.62) using some sort of Hamiltonian/Lagrangian equation, then that indicates there may be no way to regularize gas dynamics either.

These tasks are of short-term interest, and we consider one possible answer to these questions in Chapter 4.

As a longer-term project of physical interest, we must extend our range of models to the full compressible system of equations, i.e., not just the barotropic case. In the full Euler system, one has conservation laws for mass, momentum, and energy. The energy equation is coupled together with a thermodynamic equation of state. The full Euler system is the one whose shock waves matter for most physical applications. Assuming we have answered, in a conclusive and satisfactory way, the questions of short-term interest, we may consider the full Euler system.

Chapter 4

A Hamiltonian Regularization of the Inviscid Burgers Equation

4.1 Introduction

In this chapter, we consider the following quasilinear evolution equation:

$$u_t + uu_x - \alpha^2 u_{txx} - \alpha^2 uu_{xxx} = 0, \tag{4.1}$$

with $\alpha > 0$. By introducing the Helmholtz operator,

$$\mathcal{H} = \mathrm{Id} - \alpha^2 \partial_x^2, \tag{4.2}$$

we may rewrite (4.1) as

$$v_t + uv_x = 0, \tag{4.3}$$

where

$$v = \mathcal{H}u. \tag{4.4}$$

The main goal of this chapter is to show that (4.1) represents a valid regularization of the Burgers equation. That is, the solutions $u^\alpha(x,t)$ of (4.1) with initial data

$$u^\alpha(x,0) = u_0(x),$$

converge strongly, as $\alpha \to 0$, to the unique, weak entropy solution of the Cauchy problem for the inviscid Burgers equation

$$u_t + uu_x = 0, \tag{4.5a}$$

$$u(x,0) = u_0(x). \tag{4.5b}$$

For the sake of exposition, let us mention without proof a few well-known facts. We are primarily concerned with the case when $u_0'(x) < 0$ for at least one point $x \in \mathbb{R}$. In this case, regardless of how smooth $u_0(x)$ is, the classical solution $u(x,t)$ of the Burgers equation exists only until a *break time* T (see Section 4.3 and [44]). Therefore, we work with the weak form of (4.5). When u_0' has mixed sign, there exists a global weak solution $u(x,t)$ of (4.5). The solution is not unique unless we impose an additional constraint, which is called an entropy inequality by analogy with gas dynamics. Here we will work with the Oleinik inequality given by

$$\frac{u(x+a,t) - u(x,t)}{a} < \frac{C}{t}, \tag{4.6}$$

for every $a > 0$, $t > 0$ and $x \in \mathbb{R}$, where C is a constant that depends only on u_0. Taken together, the system (4.5-4.6) has a unique weak solution $u(x,t)$ globally in time, for a large class of initial data u_0.

Viscous regularizations. It is well known that if one desires to capture the physically relevant solutions of (4.5), one can solve the viscous Burgers equation

$$u_t + uu_x = \nu u_{xx}, \tag{4.7a}$$

$$u(x,0) = u_0(x). \tag{4.7b}$$

Let $u^\nu(x,t)$ denote the solution of the Cauchy problem (4.7). We emphasize that even for rough initial data $u_0 \in L^\infty$, the solution $u^\nu(x,t)$ exists uniquely and globally in time, in the *classical* sense, for each $\nu > 0$. Now suppose we fix u_0 and repeatedly solve (4.7a) with a sequence of values for ν, i.e. $\{\nu_k\}$ such that $\lim_{k\to\infty} \nu_k = 0$. Then we will find

that the resulting solutions u^{ν_k} converge strongly to a function u that is the unique weak entropy solution of the Cauchy problem (4.5).

Suppose we ask whether there are useful viscous mechanisms besides u_{xx}. One candidate is filtered viscosity, e.g. $\mathcal{H}^{-1}u_{xx}$, which has been analyzed in [74, 54, 51]. It is shown that this sort of filtered viscosity leads to solutions u^ν that enjoy uniform L^∞ bounds and uniform L^1 contraction properties. Strong convergence to the weak entropy solution in the $\nu \to 0$ limit is proved via the Kruzhkov BV theory [50].

More recently, the use of hyper-viscosity (e.g. $(-1)^{n+1}\partial_x^{2n}u$) has been analyzed in [82]. Here it is shown that under the assumption that the solution u^ν stays bounded in L^∞, there is strong convergence as $\nu \to 0$ to the weak entropy solution.

Compared with regularizations that have appeared in the literature, the regularization (4.1) does not involve standard viscous mechanisms. Using the definition (4.2) of $\mathcal{H}$, we may verify that

$$u_t + uu_x = -\frac{3}{2}\alpha^2\mathcal{H}^{-1}\left(u_x^2\right)_x. \tag{4.8}$$

is formally equivalent to (4.1). The right-hand side of (4.8) represents a *nonlinear* smoothing term that differs from standard viscosities, which are all linear in the unknown u. Currently, we are unaware of previous works which have used the right-hand side term in (4.8) as a regularizing term for shock-forming hyperbolic equations.

Previous results on equation (4.1). Equation (4.1) has appeared previously in the literature, as the $b = 0$ member of the b-family proposed in [17]:

$$v_t + uv_x + bu_xv = 0. \tag{4.9}$$

Various results regarding the complete integrability (for $b = 2$ and $b = 3$) and traveling wave solutions of (4.9) may be found in [17, 42, 36, 41, 19, 20, 64, 10]. In what follows, we will discuss the results from this collection that specifically deal with the $b = 0$ case of (4.9).

Physical motivation for the b-family is provided in [19, 20], which show that (4.9) is an asymptotically equivalent approximation of the shallow water equations. That is,

suppose we write the classical equations for surface waves in shallow water, and then apply a weakly nonlinear expansion in the regime where (I) the waves' amplitude is much smaller than their length and (II) the waves' amplitude is much smaller than the mean depth of the water. Restricting the model to unidirectional waves and truncating at quadratic order in the perturbation parameters, we obtain a family of pde's that can be transformed into the b-family via certain Kodama transformations. In [86] we find this same technique applied at linear order in the perturbation parameters to recover the Korteweg-de Vries equation.

In [17], the b-family is realized as the Euler-Lagrange equation corresponding to a certain Lagrangian density. As the authors point out, this Lagrangian structure breaks down when $b = 0$. The authors do propose a Hamiltonian structure that appears well-defined for the $b = 0$ case, though they do not prove here that the proposed structure in fact satisfies the requirements for Hamiltonian operators as described in, e.g., [68].

The Hamiltonian structure of the b-family (4.9) is given by (see [42]):

$$v_t = -b^2 \mathcal{B} \frac{\delta H}{\delta v}, \qquad H = \frac{1}{b-1} \int v \, dx. \tag{4.10a}$$

$$\mathcal{B} = v^{1-1/b} \partial_x v^{1/b} (\partial_x - \alpha^2 \partial_x^3)^{-1} v^{1/b} \partial_x v^{1-1/b}. \tag{4.10b}$$

As the authors of [42] state, "when $b = 1$ the Hamiltonian must be modified; for $b = 0$ the operator $\mathcal{B}$ can be redefined." In [42], they prove that except in these special cases, the functional/operator pair given in (4.10) satisfies the Jacobi identity and is a valid Hamiltonian structure for the b-family (4.9). In this chapter, we will show that the proper redefinition of (4.10) in the $b = 0$ case is also a valid Hamiltonian structure. Hence (4.1) is Hamiltonian in a certain sense.

Classical traveling wave solutions for the $b = 0$ equation are briefly discussed in [41]. It is shown that the $b = 0$ equation has peakon solutions of the form

$$u(x,t) = \pm c \exp(-|x - ct|/\alpha),$$

but numerical experiments reveal that these solutions are unstable. There is another

class of traveling wave solutions, the so-called "ramps" and "cliffs," which look like viscous shock profiles. The authors provide numerical evidence that these traveling wave solutions are stable.

Formally similar non-viscous regularizations. In various other pde's that have appeared in the literature, we may find the terms $\alpha^2 u_{txx}$ and $\alpha^2 u u_{xxx}$ from (4.1) either separately or in combination with other terms. Let us discuss a few such cases, with an eye on the extent to which these pde's constitute a regularization of the Burgers equation.

Goodman and Lax considered in [30] a dispersive finite-difference scheme whose solutions behave, as the meshsize goes to 0, like solutions of the zero dispersion limit of the KdV equation [52]. By ignoring higher order terms in Δ (the meshsize), the difference scheme investigated in [30] approximates the solution of

$$u_t + u u_x + \frac{1}{6}\Delta^2 u u_{xxx} = 0. \tag{4.11}$$

In [30], it is shown that, as long as the solution u of the Cauchy problem (4.5) is smooth, the solution of the difference scheme with the same initial data converges strongly to u as $\Delta \to 0$. However, beyond the time at which the classical solution of (4.5a) breaks, the solution of the finite difference scheme develops high-frequency oscillations and ceases to converge strongly in the $\Delta \to 0$ limit. The solution of the difference scheme does converge weakly, but its weak limit—in the sense of distributions—is not a weak solution of the Burgers equation.

Compared to (4.1), the equation (4.11) does not contain the mixed derivative term u_{txx}. Also, in (4.11), the nonlinear term $u u_{xxx}$ enters with a plus sign. A pde that contains the mixed derivative term u_{txx}, but not $u u_{xxx}$ is the RLW (regularized long wave) or BBM (Benjamin-Bona-Mahoney) equation,

$$u_t + u_x + u u_x - \alpha^2 u_{xxt} = 0, \tag{4.12}$$

introduced in [70, 5] as a model for the unidirectional propagation of long waves in certain

nonlinear dispersive systems. By dropping the term u_x in (4.12), one can consider the following regularization of the Burgers equation:

$$u_t + uu_x - \alpha^2 u_{xxt} = 0. \tag{4.13}$$

The solutions of (4.13) and (4.12) behave similarly. It is well-known that, generically, initial data for (4.13) breaks up into a sequence of solitary waves, followed by a dispersive tail. As $\alpha \to 0$, numerical experiments indicate that the solutions of (4.13) do not converge (in any sense) to a weak solution of the Burgers equation.

Whitham's family. We use the Green's function of the Helmholtz operator (4.2) given by

$$G(x) = \frac{1}{2\alpha} \exp\left(-\frac{|x|}{\alpha}\right), \tag{4.14}$$

and write (4.8) as

$$u_t + uu_x = -\frac{3}{2}\alpha^2 \int_{\mathbb{R}} G(x-y) \left(u_y^2\right)_y (y,t) dy. \tag{4.15}$$

Now let us turn our attention to the general equation proposed by Whitham (see [86, Section 13.14]), as the simplest equation that combines dispersive and nonlinear effects:

$$u_t + uu_x + \int_{\mathbb{R}} K(x-y) u_y(y,t) dy = 0, \tag{4.16}$$

where $K(x)$ is a convolution kernel. When we compare Whitham's equation[1] with our equation we notice that the right-hand side of (4.16) is linear in $u_y(y,t)$, while the right-hand-side of (4.15) is quadratic in $u_y(y,t)$.

In writing (4.16), one of Whitham's objectives was to model the breaking and peaking of shallow water waves. The specific form of (4.16) is motivated by the fact that the

[1] Many beautiful results on the initial-value problem and long-time asymptotics of (4.16) may be found in the monograph [67].

dispersion relation for this equation is given by

$$\omega(k) = k\widehat{K}(k), \tag{4.17}$$

where $\widehat{K}$ is the Fourier transform of K. Of course, one can invert this, and design an equation of the form (4.16) with any prescribed dispersion relation. The idea now is to find K such that: (1) the right-hand-side of (4.17) closely approximates the true dispersion relation for shallow water waves, at least in the regime of interest, and (2) the resulting equation (4.16) is easier to analyze than the full shallow water equation.

In our case, by considering sinusoidal perturbations about a constant solution $u = u_0$, we find the dispersion relation for (4.1):

$$\omega(k) = u_0 k. \tag{4.18}$$

Hence (4.1) is not a dispersive wave equation. Returning for a moment to the b-family (4.9), we see that the dispersion relation for the whole family is

$$\omega(k) = u_0 k + u_0 \frac{bk}{1 + \alpha^2 k^2}. \tag{4.19}$$

Hence the b-family is dispersive for $b > 0$. We expect that small-dispersion oscillations typical of dispersive equations such as KdV or BBM will play a role in the zero-α limits of the b-family for all $b > 0$. Again, the $b = 0$ equation under consideration in this chapter does not produce oscillations in the zero-α limit.

The equation that is obtained from (4.16) with the Helmholtz kernel $K = G$ is called the Burgers-Poisson equation and is formally equivalent to the system

$$v_t + vv_x = u_x,$$
$$u_{xx} = u + v.$$

As shown in [26], the Burgers-Poisson equation features wave breaking in finite time. Furthermore, [26] shows that the $\alpha \to 0$ limit of the Burgers-Poisson equation is similar

to the zero dispersion limit of the KdV equation (see [52]). That is, the limits of solutions of the Burgers-Poisson equation are weak limits; these weak limits satisfy the limiting equation (Burgers) only in the regime where Burgers has smooth, classical solutions. In other words, one cannot recover shock solutions of Burgers by considering zero-dispersion limits, in the strong or weak sense, of solutions of Burgers-Poisson.

Summary. Numerical simulations reveal that though (4.1) is Hamiltonian, solutions of (4.1) dissipate energy in all L^p norms for $p > 1$. In comparison, a smooth solution of the Burgers equation (4.5) preserves the L^2 energy of u until the break time T. For the Burgers equation (4.5), shock formation, and subsequent evolution of discontinuous weak solutions, occurs with a corresponding decay in the L^2 energy of u.

Roughly speaking, equation (4.1) is dissipative enough to approximate shock solutions of the Burgers equation, but conservative enough to retain a certain non-canonical Hamiltonian structure. Though (4.1) is formally similar to certain shallow water model equations, it is neither dispersive nor completely integrable. This hints at why, when α goes to zero, we see the strong convergence typical of zero-viscosity limits instead of the oscillations and weak convergence typical of zero-dispersion limits.

Outline of this Paper. In Section 4.2, we prove the existence of a large class of traveling front solutions for (4.1). We study the initial-value problem of (4.1) in Section 4.3 and prove well-posedness of solutions assuming that $v(x, 0) > 0$ for all x. Next, we consider the $\alpha \to 0$ limit. In Section 4.4 we show that solutions of (4.1) converge strongly in the zero-α limit to weak solutions of the inviscid Burgers equation. Next, in Section 4.5, we provide numerical evidence that the weak solution that is selected in the $\alpha \to 0$ limit is indeed the entropic, or physically relevant, solution. We discuss the Hamiltonian structure of (4.1) in Section 4.6. Finally, in Section 4.7, we discuss future directions in which we plan to take this line of research.

4.2 Traveling Front Waves

We now demonstrate that (4.1) has traveling wave solutions of the form

$$u(x,t) = u\left(\frac{x-ct}{\alpha}\right).$$

These solutions consists of fronts, or shock-like profiles, connecting two different states:

$$\lim_{x\to-\infty} u(x,t) = u_L, \qquad \lim_{x\to+\infty} u(x,t) = u_R. \tag{4.20}$$

Remark. Equation (4.1) is Galilean invariant, i.e. invariant under the mapping

$$x \mapsto x + u_0 t,$$
$$u \mapsto u + u_0.$$

In precise terms, if $u(x,t)$ solves (4.1), then so does $\overline{u}(\overline{x},t) = u(\overline{x} - u_0 t, t) + u_0$, for any $u_0 \in \mathbb{R}$. Hence we could eliminate the wave speed c from the proceedings and search for stationary solutions $u(x,t) = u(x/\alpha)$ only. We choose not to do this in order to show that the Rankine-Hugoniot relationship,

$$c = \frac{1}{2}\left(u_L + u_R\right), \tag{4.21}$$

holds for all traveling wave solutions.

Traveling waves. Let us search for solutions of the form

$$u(x,t) = u\left(\frac{x-ct}{\alpha}\right).$$

This choice of u forces $v(z) = u(z) - u''(z)$, with $z = (x-ct)/\alpha$. Substituting this into (4.3), we obtain the equation

$$-cv' + uv' = 0. \tag{4.22}$$

This equation is third-order in u'. However, we can find an integral of motion, reducing the equation to second-order in u'. Noting that $u'v$ is an exact differential, i.e.

$$2u'v = \left(u^2 - u'^2\right)',$$

we add $u'v$ to both sides of (4.22) and integrate:

$$-cv + uv = -\frac{1}{2}u'^2 + \frac{1}{2}u^2 + C_1, \tag{4.23}$$

for arbitrary $C_1 \in \mathbb{R}$. Substituting for v, we obtain a second-order equation in u, which can be rewritten as the first-order system

$$\begin{cases} u' = w \\ w' = \dfrac{u^2 - 2cu - 2C_1 + w^2}{2(u-c)}. \end{cases} \tag{4.24}$$

At first glance, it appears that system (4.24) blows up on the line $u = c$. As we show below, there do exist trajectories which cross this line in phase space. However, the line of apparent singularities does have implications for the uniqueness of such trajectories. We set aside such issues for the moment and move ahead to the phase portrait of this ode system.

To guide our study, we numerically plot a sample phase portrait in Figure 4.1 for the specific values $c = 1$, $C_1 = 0$. Note the two fixed points connected by the "patched" heteroclinic orbit, marked by arrows. Each half of the heteroclinic orbit is a straight line segment. Note also the failure of uniqueness at the two points $(u, w) = (1, \pm 1)$. We shall now explain these phenomena analytically.

Fixed Points. System (4.24) is invariant under the reflections $w \mapsto -w$ and $u \mapsto (2c - u)$. Hence the phase portrait is symmetric across the $w = 0$ and $u = c$ lines.

The fixed points of (4.24) are located at $(u_-, 0)$ and $(u_+, 0)$ where

$$u_- = c - \sqrt{c^2 + 2C_1}, \qquad u_+ = c + \sqrt{c^2 + 2C_1}. \tag{4.25}$$

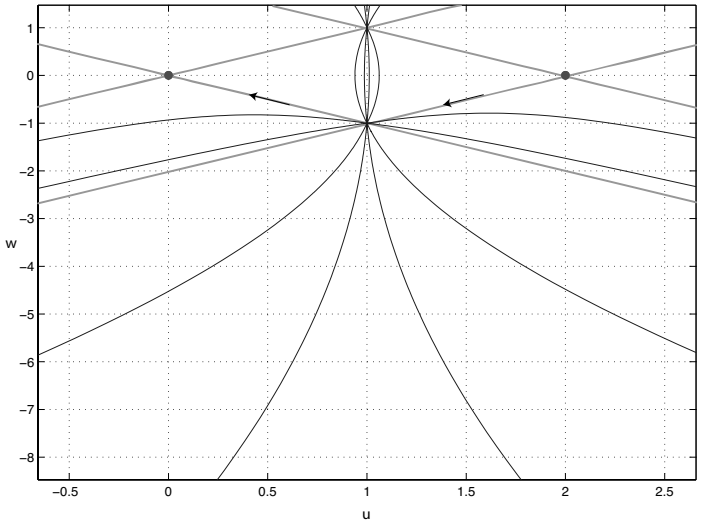


Figure 4.1: Phase portrait of (4.24) with $c = 1$, $C_1 = 0$.

Take $C_1 > -(1/2)c^2$ in order for the two fixed points to exist. The fixed point are always saddles, since the linearization of (4.24) about either fixed point is

$$\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}. \tag{4.26}$$

Second integral of motion. The only missing pieces in the phase portrait are the stable and unstable manifolds of the two saddle points. In particular, we would like to prove the existence of a heteroclinic orbit that connects $(u_+, 0)$ to $(u_-, 0)$. In order to do this, we further reduce (4.23) from a second-order equation to a first-order equation, by finding another integral of motion.

Assume $u'(z) = f(u(z))$ and differentiate both sides with respect to z:

$$u'' = f'(u)u' = f'(u)f(u) = \frac{1}{2}\frac{d}{du}f(u)^2.$$

We know from (4.23) that

$$u'' = \frac{u^2 - 2cu - 2C_1 + u'^2}{2(u - c)}.$$

Equating both expressions for u'' and setting $g = f^2$ yields the linear ode

$$\frac{dg}{du} = \frac{u^2 - 2cu - 2C_1 + g}{u - c}.$$

Searching for solutions of the form $g(u) = (u - c)h(u)$, we find that $h(u)$ has to satisfy

$$h'(u) = \frac{u^2 - 2cu - 2C_1}{(u - c)^2}.$$

Integrating, we find $h(u)$. Then, by $g(u) = (u - c)h(u)$, we arrive at

$$g(u) = u^2 + (C_2 - c)u + c^2 + 2C_1 - cC_2,$$

for an arbitrary constant of integration $C_2 \in \mathbb{R}$. Recalling $g = f^2 = (u')^2$, we have

$$u' = \pm\sqrt{u^2 + (C_2 - c)u + c^2 + 2C_1 - cC_2}. \tag{4.27}$$

As explained above, the constants c and C_1 determine the location of the fixed points in the phase plane. Suppose the fixed points have been determined, and that we wish to plot a trajectory, say $u' = \gamma(u)$, passing through a given point (u_0, u_0'). We plug this point into (4.27) and solve for C_2. Using this value of C_2, equation (4.27) now gives us γ, and hence the trajectory we wished to plot, wherever it is defined. In this way, we may use (4.27) to plot the phase portrait of (4.24) without resorting to numerical integration.

Line of singularities. The only place where this procedure might break down is along the line $u = c$. Specifically, inserting $u = c$ in (4.27), we find that C_2 cancels out of the expression, leaving only

$$u' = \pm\sqrt{c^2 + 2C_1}.$$

Therefore, all trajectories in the $u > c$ half-plane that cross into the $u < c$ half-plane (or vice versa) must do so at one of these two points:

$$(c, s_\pm) = (c, \pm\sqrt{c^2 + 2C_1}). \tag{4.28}$$

Hence uniqueness of solutions for (4.24) fails at the line $u = c$.

Examining (4.24), it is then clear that for $\delta > 0, h > 0$, all trajectories starting from $(u, u') = (c - \delta, h)$ will move to the right towards the $u = c$ line, intersecting at the point $(c, \sqrt{c^2 + 2C_1})$. By symmetry, this picture may be appropriately extended to the cases where $\delta < 0$ and/or $h < 0$.

Stable/unstable manifolds. We return to the question of a heteroclinic orbit connecting the two fixed points $(u_\pm, 0)$ (see (4.25)). First, we substitute $(u, u') = (u_+, 0)$ in (4.27) to obtain

$$C_2 = -c - 2\sqrt{c^2 + 2C_1}.$$

Then, for general (u, u'), we use this value of C_2 in (4.27) and obtain

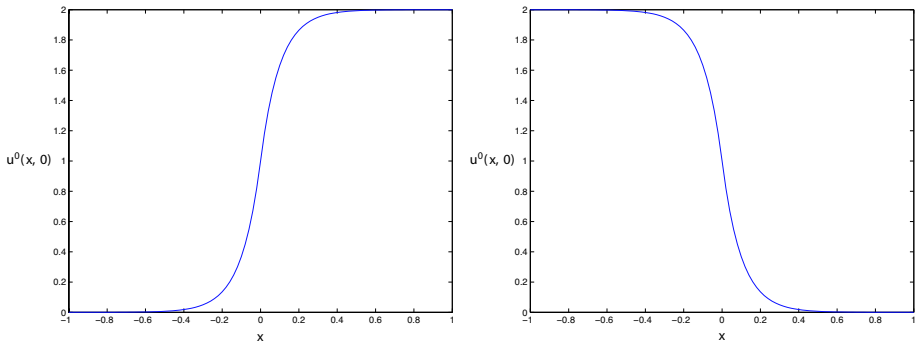
$$u' = \pm(u - (c + \sqrt{c^2 + 2C_1})) = \pm(u - u_+).$$

This gives the stable/unstable manifolds for $(u_+, 0)$. By symmetry, we obtain for $(u_-, 0)$ the analogous expression

$$u' = \pm(u - u_-).$$

Surprisingly, the stable/unstable manifolds for the nonlinear system (4.24) consist of straight lines connecting the fixed points $(u_\pm, 0)$ with the singular points $(c, s_\pm)$, given by (4.28). Because these manifolds are straight lines, we may use the eigenvectors and eigenvalues of the linearization (4.26) to determine, for each fixed point, which line is stable and which is unstable.

PDE solutions. We may stitch together the unstable manifold of $(u_+, 0)$ and the stable manifold of $(u_-, 0)$, producing the piecewise differentiable curve $z \mapsto (u(z), u'(z))$, which we call a patched heteroclinic orbit. By symmetry, the same construction works in the $u' \geq 0$ half-plane. This patched heteroclinic orbits correspond to a solution $u(x, t) : \mathbb{R} \times \mathbb{R} \to \mathbb{R}$ of the pde (4.1) with boundary conditions given by (4.20). Integrating along the stable/unstable manifolds, we find the exact formulas for the traveling front



(a) $d > 0$ (b) $d < 0$

Figure 4.2: Traveling wave solutions of (4.1) at fixed t, with $\alpha = 0.1$, $d = \pm 1$.

solutions:

$$u^0(x,t) = \begin{cases} u_R - d\exp(-(x-ct)/\alpha) & x > ct \\ u_L + d\exp((x-ct)/\alpha) & x < ct, \end{cases}$$

where $u_R = c + d$ and $u_L = c - d$. As can be verified directly, $u^0(\cdot, t) \in H^2(\mathbb{R})$. However, the second derivative u^0_{xx} is discontinuous and therefore u^0_{xxx} exists only in the sense of a Dirac δ distribution. Nevertheless, we may verify directly that u^0 solves either (4.8) or the weak form of (4.1). Note also that $u^0(ct, t) = c = (u_R + u_L)/2$, and that we have the correspondence

- $d > 0$ means $u_R > u_L$ and $u^0_x \geq 0$ everywhere.
- $d < 0$ means $u_R < u_L$ and $u^0_x \leq 0$ everywhere.

See Figure 4.2 for plots of both the $d > 0$ and $d < 0$ exact solution, plotted at a fixed instant in time. The width of these front solutions is governed only by α. Taking $\alpha \to 0$ in any of the front solutions $u(x,t) = u((x-ct)/\alpha)$, we arrive at the function

$$u(x,t) = \begin{cases} u_R & x > ct \\ u_L & x < ct. \end{cases}$$

Interestingly enough, this function is a global (weak) solution of the inviscid Burgers equation.

4.3 Initial-value Problem

In this section we will study the regularity of solutions of (4.1) (or equivalently, (4.3)) for certain classes of initial data.

The material picture. Let us recall that the Burgers equation as it is usually written,

$$u_t + uu_x = 0,$$

is, from the fluid dynamical point of view, an equation in spatial coordinates. That is, the points $x \in \mathbb{R}$ are fixed measuring points—at each time t, the velocity of fluid moving past x is recorded as $u(x,t)$.

Another standard way of viewing continuum mechanical problems is the material picture. Here we track particle positions: let $\eta(X,t)$ denote the position at time t of the particle which was initially at $\eta(X,0) = X$. The relationship between the material and spatial pictures is given by

$$\partial_t \eta(X,t) = u(\eta(X,t),t). \tag{4.29}$$

Differentiating both sides of this expression in time, we obtain

$$\partial_t^2 \eta(X,t) = \Big(u(x,t)u_x(x,t) + u_t(x,t)\Big)_{x=\eta(X,t)}.$$

We see that for any classical solution u of the Burgers equation, the material map η, defined as the solution of (4.29), must satisfy

$$\partial_t^2 \eta(X,t) = 0. \tag{4.30}$$

Let us examine what happens when we try to go the other way, namely when we start with the initial conditions $\eta(X,0) = X$, $\partial_t \eta(X,0) = V(X)$, and find the resulting unique

solution of (4.30):

$$\eta(X,t) = X + tV(X).$$

So far, there is no problem: we have a smooth, global solution $\eta(X,t)$. Problems arise when we try to define u by formally inverting (4.29):

$$u(x,t) = \partial_t \eta \left(\eta^{-1}(x,t), t\right). \tag{4.31}$$

If $V'(X) \geq 0$ for all X, then $\partial_X \eta(X,t) > 0$ for all X and all $t > 0$. Then it is clear that for each fixed t, the map $X \mapsto \eta(X,t)$ is a diffeomorphism of $\mathbb{R}$. Indeed we can show (using the Inverse Function Theorem) that η^{-1} will be as smooth as η. Then u defined by (4.31) must be a global smooth solution of the Burgers equation.

However, if there exists X_0 such that $V'(X_0) < 0$, then at the break time $t_0 = -1/V'(X_0)$, we have $\partial_X \eta(X_0, t_0) = 0$. This means that two fluid particles have collided; in this case, for $t > t_0$, we cannot solve for η^{-1}.

We have shown that solving the material problem (4.30) eventually produces a solution of the spatial problem, as long as we can guarantee that

$$\partial_X \eta(X,t) \neq 0,$$

for all X and all t. With this in mind, we shift our attention from the Burgers equation to the initial-value problem

$$v_t + uv_x = 0, \tag{4.32a}$$

$$u - u_{xx} = v, \tag{4.32b}$$

$$v(x,0) = v_0(x). \tag{4.32c}$$

Note that we have taken $\alpha = 1$. Given a solution $u(x,t)$ of (4.32a)-(4.32b), we may construct

$$\tilde{u}(x,t) = u\left(\frac{x}{\alpha}, \frac{t}{\alpha}\right). \tag{4.33}$$

Then it is easy to check that $\tilde{u}$ solves (4.1).

Remark. The curves $\eta(X,t)$ are commonly called "characteristics" and the condition

$$\partial_X \eta(X,t) \neq 0$$

is simply the statement that characteristics do not cross. We offer the above explanation to point out that employing the "method of characteristics" for the Burgers equation is nothing more than shifting one's view from the spatial to the material picture.

Material version of the regularized equation. Let us suppose that (4.32) holds for a smooth function $v(x,t)$. Then we may solve (4.32b) for u and define the associated material map η as the solution of

$$\partial_t \eta(X,t) = u(\eta(X,t),t), \tag{4.34}$$

subject to $\eta(X,0) = X$. Then the usual calculation shows

$$\frac{d}{dt}\left[v(\eta(X,t),t)\right] = 0,$$

which implies

$$v(\eta(X,t),t) = v(\eta(X,0),0) = v_0(X). \tag{4.35}$$

We will now use this fact to show that η is determined completely as the solution of a certain ordinary differential equation that does not involve $v(\cdot,t)$, except at its initial value v_0. We begin by using the Green's function of $\mathcal{H} = \mathrm{Id} - \partial_x^2$ to invert (4.32b):

$$u(x,t) = \frac{1}{2}\int_{\mathbb{R}} \exp(-|x-y|)v(y,t)\,dy. \tag{4.36}$$

Let us exploit (4.35) by changing variables with $y = \eta(Y,t)$ and $x = \eta(X,t)$, resulting in

$$u(\eta(X,t),t) = \frac{1}{2}\int_{\mathbb{R}} \exp(-|\eta(X,t)-\eta(Y,t)|)v_0(Y)\partial_Y \eta(Y,t)\,dY. \tag{4.37}$$

We define the map ψ via

$$\psi(X,t) = v_0(X)\partial_X \eta(X,t). \tag{4.38}$$

Using (4.34), we then have

$$\partial_t \eta(X,t) = \frac{1}{2}\int_{\mathbb{R}} \exp(-|\eta(X,t)-\eta(Y,t)|)\psi(Y,t)\,dY. \tag{4.39}$$

We may derive an equation for ψ simply by differentiating (4.39) with respect to X:

$$\partial_t \psi(X,t) = -\frac{1}{2}\psi(X,t)\int_{\mathbb{R}} \operatorname{sgn}(\eta(X,t)-\eta(Y,t))\exp(-|\eta(X,t)-\eta(Y,t)|)\psi(Y,t)\,dY. \tag{4.40}$$

The pair (4.39-4.40) is an infinite-dimensional dynamical system for η and ψ. We concern ourselves with the initial-value problem where $\eta(X,0) = X$ and, by (4.38), $\psi(X,0) = v_0(X)$.

***A priori* results.** Wherever η and ψ are defined, they must satisfy two properties that we present below. These results are adapted from the work of R. Camassa in [8].

Lemma 4. *Suppose that $v_0 \in L^1(\mathbb{R})$ and suppose $\eta(X,0) = X$. Then for all $t \geq 0$ such that (4.39-4.40) hold,*

$$\int_{\mathbb{R}} \psi(X,t)\,dX = \int_{\mathbb{R}} v_0(X)\,dX.$$

Proof. Define

$$\phi(t) = \int_{\mathbb{R}} \psi(X,t)\,dX. \tag{4.41}$$

By differentiating both sides of (4.41) in time and using (4.40), we obtain

$$-2\dot{\phi}(t) = \int_{\mathbb{R}}\int_{\mathbb{R}} \psi(X,t)\psi(Y,t)\operatorname{sgn}(\eta(X,t)-\eta(Y,t))\exp(-|\eta(X,t)-\eta(Y,t)|)\,dY\,dX.$$

Antisymmetry of the integrand forces $\dot{\phi}(t) = 0$. Then, using definition (4.38) and $\partial_X \eta(X,0) = 1$, we may conclude that for all t:

$$\phi(t) = \int_{\mathbb{R}} \psi(X,0)\,dX = \int_{\mathbb{R}} v_0\,dx. \qquad \square$$

Proposition 4. *Suppose that $v_0 \in L^1(\mathbb{R})$ is everywhere positive, and suppose that*

$\eta(X,0) = X$. *Then for all* X *and all* $t \geq 0$ *such that (4.39-4.40) hold,*

$$\partial_X \eta(X,t) \neq 0.$$

Proof. Suppose there exists (X_0, t_0) such that $\partial_X \eta(X_0, t_0) = 0$. Without loss of generality, assume that $\partial_X \eta(X,t) > 0$ for all X and all $t \in [0, t_0)$. By definition (4.38), we have $\psi(X,t) \geq 0$ for $t \in [0, t_0]$. Because ψ and v_0 are nonnegative, applying Hölder's inequality to (4.39) gives

$$\partial_t \eta(X,t) \leq \frac{1}{2}\|\psi(\cdot,t)\|_{L^1} = \frac{1}{2}\|v_0\|_{L^1},$$

by Lemma 4. It follows that

$$\eta(X,t) \leq \frac{t}{2}\|v_0\|_{L^1} + X, \tag{4.42}$$

for all X and all $t \in [0, t_0]$. Now let $\zeta(X,t) = 1/\partial_X \eta(X,t)$. Then straightforward computations and the sign-definiteness of ψ give, for all $t \in [0, t_0]$,

$$\begin{aligned}\frac{\partial_t \zeta(X,t)}{\zeta(X,t)} &= \frac{1}{2}\int_{\mathbb{R}} \operatorname{sgn}(\eta(X,t) - \eta(Y,t)) \exp(-|\eta(X,t) - \eta(Y,t)|)\psi(Y,t)\, dY \\ &\leq \frac{1}{2}\int_{\mathbb{R}} \exp(-|\eta(X,t) - \eta(Y,t)|)\psi(Y,t)\, dY = \partial_t \eta(X,t).\end{aligned}$$

Now Gronwall's inequality, $\zeta(X,0) = 1$, and (4.42) give

$$\zeta(X,t) \leq \exp\left(\eta(X,t) - X\right) \leq \exp\left(\frac{t}{2}\|v_0\|_{L^1}\right),$$

for $t \in [0, t_0]$. Clearly $\zeta(X_0, t_0)$ is finite, contradicting $\partial_X \eta(X_0, t_0) = 0$.

Note that we have divided by $\zeta(X,t)$ and thus tacitly assumed $\zeta(X,t) \neq 0$ for $t \in [0, t_0]$. Suppose instead that $\zeta(X_1, t_1) = 0$ for some X_1 and some $t_1 \leq t_0$. Then $\zeta(X,t) = 1/\partial_X \eta(X,t) > 0$ for all X and all $t \in [0, t_1)$. Again, straightforward computations and

the sign-definiteness of ψ give, for all $t \in [0, t_1]$,

$$\frac{\partial_t \zeta(X,t)}{\zeta(X,t)} \geq -\frac{1}{2}\int_{\mathbb{R}} \exp(-|\eta(X,t)-\eta(Y,t)|)\psi(Y,t)\,dY = -\partial_t \eta(X,t).$$

Again using Gronwall's inequality, $\zeta(X,0)=1$, and (4.42), we have

$$\zeta(X,t) \geq \exp(-\eta(X,t)+X) \geq \exp\left(-\frac{t}{2}\|v_0\|_{L^1}\right) > 0,$$

for $t \in [0,t_1]$, contradicting $\zeta(X_1,t_1)=0$. □

Global well-posedness. Using these *a priori* results, we will now show that for certain initial data, we have existence and uniqueness of η and ψ globally in time. We will use standard theory for ordinary differential equations on Banach spaces (see [1, Chap.4]). Here we sidestep the issue of the space of optimal well-posedness and consider the Banach space $C_b(\mathbb{R})$ of continuous and bounded functions on $\mathbb{R}$, endowed with the supremum norm.

Theorem 2. *Define the vector field F by*

$$F\begin{bmatrix}\eta\\ \psi\end{bmatrix}(X) = \begin{bmatrix} \frac{1}{2}\int_{\mathbb{R}} \exp(-|\eta(X,t)-\eta(Y,t)|)\psi(Y,t)\,dY \\ -\frac{1}{2}\psi(X,t)\int_{\mathbb{R}} \operatorname{sgn}(\eta(X,t)-\eta(Y,t))\exp(-|\eta(X,t)-\eta(Y,t)|)\psi(Y,t)\,dY \end{bmatrix}. \tag{4.43}$$

Then, for any $v_0 \in C_b(\mathbb{R}) \cap L^1(\mathbb{R})$ such that $v_0 > 0$, there exists a unique solution pair $\eta : [0,\infty) \to C_b(\mathbb{R})$, $\psi : [0,\infty) \to C_b(\mathbb{R})$, that solves the initial-value problem

$$\begin{aligned} \frac{d}{dt}\begin{pmatrix}\eta\\ \psi\end{pmatrix} &= F\begin{pmatrix}\eta\\ \psi\end{pmatrix}, \\ \eta(0) &= \mathrm{Id}, \\ \psi(0) &= v_0. \end{aligned} \tag{4.44}$$

Proof. The vector field F is clearly a Lipschitz continuous map from $C_b(\mathbb{R}) \times C_b(\mathbb{R})$ to itself. It is clear that $\eta(0) = \mathrm{Id} \in C_b(\mathbb{R})$. Choose $\psi(0) = v_0 \in C_b(\mathbb{R})$. Then, since $C_b(\mathbb{R})$ is a Banach space, the standard existence/uniqueness theorem for ordinary differential

equations implies the existence of $T > 0$ and a unique pair

$$\eta \in C^1([0,T], C_b(\mathbb{R})),$$
$$\psi \in C^1([0,T], C_b(\mathbb{R})),$$

such that η and ψ solve (4.44). We shall now extend this to the time interval $[0,\infty)$. We label the components of F by

$$F = \begin{bmatrix} F_1 \\ F_2 \end{bmatrix},$$

and first estimate

$$\begin{aligned}
\sup_X \Big| F_1[\eta,\psi](X) \Big| &\leq \frac{1}{2} \sup_X \left| \int_{\mathbb{R}} \exp(-|\eta(X) - \eta(Y)|)\psi(Y)\, dY \right| \\
&\leq \frac{1}{2} \sup_X \|\exp(-|\eta(X) - \eta|)\|_{L^\infty} \|\psi\|_{L^1} \\
&\leq \frac{1}{2} \|v_0\|_{L^1}.
\end{aligned}$$

Here, the last inequality is obtained using Lemma 4 and the fact that $\psi > 0$ (see equation (4.38) and Proposition 4).

Estimating in a similar fashion, we have

$$\begin{aligned}
\sup_X \Big| F_2[\eta,\psi](X) \Big| &= \frac{1}{2} \sup_X \left(\psi(X) \left| \int_{\mathbb{R}} \operatorname{sgn}(\eta(X) - \eta(Y)) \exp(-|\eta(X) - \eta(Y)|)\psi(Y)\, dY \right| \right) \\
&\leq \frac{1}{2} \sup_X \left(\psi(X) \|\operatorname{sgn}(\eta(X) - \eta) \exp(-|\eta(X) - \eta|)\|_{L^\infty} \|\psi\|_{L^1} \right) \\
&\leq \frac{1}{2} \|v_0\|_{L^1} \sup_X \psi(X) \\
&\leq \frac{1}{2} \|v_0\|_{L^1} \|v_0\|_{L^\infty} \sup_X \partial_X \eta(X,t)
\end{aligned}$$

In the proof of Proposition 4, we showed that it is impossible for $\eta(X,t_0) = \infty$ for any finite time $t_0 \in [0,\infty)$. Hence for any $T > 0$, we have

$$\sup_{t\in[0,T]} \sup_X \Big| F_2[\eta,\psi](X) \Big| \leq \frac{1}{2} \|v_0\|_{L^1} \|v_0\|_{L^\infty} \sup_{t\in[0,T]} \sup_X \partial_X \eta(X,t) < \infty.$$

Our estimates of F_1 and F_2 together show that for any $T > 0$,

$$\sup_{t\in[0,T]} \|F[\eta,\psi]\|_{C_b\times C_b} < \infty.$$

Then by a standard theorem in ODE theory (see [1, Proposition 4.1.22]), we may extend the solutions η and ψ for all time. □

The solution η of the material form of the equations can now be used to construct solutions v and u of the spatial form of the equations. We first note that under the hypotheses of Theorem 2, we have enough information to show that $\partial_X\eta(\cdot,t) \in C_b(\mathbb{R})$ for each fixed $t \geq 0$. To see that $\partial_X\eta(\cdot,t)$ is continuous, note that $\partial_X\eta(X,t) = \psi(X,t)/v_0(X)$ and $v_0 > 0$. Furthermore, the proof of Proposition 4 gives us the bound

$$0 < \partial_X\eta(X,t) \leq \exp\left(\frac{t}{2}\|v_0\|_{L^1}\right).$$

So it is clear that for each fixed $t \geq 0$, $\sup_X |\partial_X\eta(X,t)| < \infty$.

With this enhanced regularity of η, we can now prove global well-posedness of the spatial problem.

Theorem 3. *Given v_0 bounded and positive such that $v_0 \in C^1(\mathbb{R}) \cap L^1(\mathbb{R})$, there exists a unique global solution $v(x,t)$ of (4.32) such that*

$$v \in C^1([0,\infty), C^1(\mathbb{R}) \cap L^1(\mathbb{R})).$$

Proof. We have taken v_0 satisfying the hypotheses of Theorem 2. Because $\eta(\cdot,t) \in C^1(\mathbb{R})$ and $\partial_X\eta(X,t) > 0$ for all $t \geq 0$, the Inverse Function Theorem guarantees that $\eta(\cdot,t)$ is a diffeomorphism, i.e., there exists $\eta^{-1}(x,t)$ such that

$$\eta(\eta^{-1}(x,t),t) = x,$$

for all $x \in \mathbb{R}$, and all $t \geq 0$. Furthermore, $\eta^{-1}(\cdot,t) \in C^1(\mathbb{R})$. So we may invert the relationship

$$v(\eta(X,t),t) = v_0(X),$$

and obtain, globally in time,

$$v(x,t) = v_0(\eta^{-1}(x,t)).$$

Since $v_0 \in C^1(\mathbb{R})$, we see that $v(\cdot,t) \in C^1(\mathbb{R})$ for all $t \geq 0$. It is also clear that $v_0 > 0$ forces $v(x,t) > 0$ for all x and all $t \geq 0$. Note also that because η is a diffeomorphism, we may use it as a change of variables as follows:

$$\begin{aligned}\int_{\mathbb{R}} v(x,t)\,dx &= \int_{\mathbb{R}} v(\eta(X,t),t)\partial_X \eta(X,t)\,dX \\ &= \int_{\mathbb{R}} \psi(X,t)\,dX \\ &= \int_{\mathbb{R}} v_0(X)\,dX,\end{aligned}$$

by Lemma 4. Hence, $v > 0$, $v_0 > 0$ and $v_0 \in L^1(\mathbb{R})$ imply $v(\cdot,t) \in L^1(\mathbb{R})$ for all $t \geq 0$. Then we may define u globally in time by

$$u(x,t) = \frac{1}{2}\int_{\mathbb{R}} e^{-|x-y|} v(y,t)\,dy.$$

Young's inequality then gives $u(\cdot,t) \in L^1(\mathbb{R})$ for all $t \geq 0$, and indeed we have

$$u \in C^1([0,\infty), C^3(\mathbb{R}) \cap L^1(\mathbb{R})).$$

The reader may verify, using the fact that η and ψ solve the material form of the equations, that v and u solve the spatial equation $v_t(x,t) + u(x,t)v_x(x,t) = 0$ for all x and all $t \geq 0$. Finally, note that $\eta^{-1}(x,0) = x$ so $v(x,0) = v_0(x)$. □

4.4 The $\alpha \to 0$ Limit.

Let us now examine in a different context the Cauchy problem

$$v_t + uv_x = 0, \tag{4.45a}$$

$$u - \alpha^2 u_{xx} = v, \tag{4.45b}$$

$$v(x,0) = v_0(x). \tag{4.45c}$$

Suppose we fix initial data v_0 bounded and positive such that

$$v_0 \in C^1(\mathbb{R}) \cap L^1(\mathbb{R}), \tag{4.46a}$$

$$v_0' \in L^1(\mathbb{R}). \tag{4.46b}$$

We will refer to this setup as the *standard data* for the Cauchy problem (4.45). Then let $v^\alpha(x,t)$ denote the unique solution to the Cauchy problem, which exists based on the above conditions and Theorem 3 from Section 4.3. Now we can formulate the question: what happens to $u^\alpha(x,t) = \mathcal{H}^{-1}v^\alpha(x,t)$ in the limit as $\alpha \to 0$? Again, we may think of this limiting process as repeatedly solving the Cauchy problem with fixed initial data v_0 while taking values of α from a sequence $\{\alpha_n\}$, where $\alpha_n \downarrow 0$ as $n \to \infty$.

Initial data. It is important to remember that as we repeatedly solve (4.45) with decreasing values of α, the initial data v_0 stays fixed. How does this affect $u^\alpha(x,0)$? To answer this, we introduce the Fourier transform

$$\widehat{\psi}(k) = \int_{\mathbb{R}} e^{-2\pi i k x}\psi(x)\,dx.$$

Since $v_0 \in L^1(\mathbb{R})$, we may compute the Fourier transform $\widehat{v}_0$. Using this and (4.45b), we have

$$\widehat{u}_0^\alpha(k) = \frac{\widehat{v}_0(k)}{1 + 4\pi^2\alpha^2 k^2},$$

where $u_0^\alpha(x) = u^\alpha(x,0)$. It is clear that as $\alpha \to 0$, we have $\widehat{u}_0^\alpha(k) \to \widehat{v}_0(k)$ for each k. Moreover, for all $\alpha \geq 0$, we have $\widehat{u}_0^\alpha(k) \leq |\widehat{v}_0(k)|$ for all k. Then, using the dominated

convergence theorem and the Fourier inversion formula, we deduce pointwise convergence: for each x,

$$u_0^\alpha(x) \to v_0(x) \text{ as } \alpha \to 0.$$

Conservation properties/estimates. Recall from the previous section that η gives particle trajectories corresponding to the velocity field u. We established in (4.35) that v is constant along η. Three simple consequences of this fact will be very useful in analyzing the zero-α limit.

Proposition 5. *Given standard data for the Cauchy problem (4.45), the resulting solution $v^\alpha(x,t)$ satisfies*

$$\|v^\alpha(\cdot,\cdot)\|_{L^\infty} = \|v_0(\cdot)\|_{L^\infty}, \tag{P1}$$

$$\|v_x^\alpha(\cdot,t)\|_{L^1} = \|v_0'(\cdot)\|_{L^1} \tag{P2}$$

$$\text{T. V. } v^\alpha(\cdot,t) = \text{T. V. } v_0 \tag{P3}$$

Proof. Given standard data, the Cauchy problem (4.45) has a unique smooth solution $v^\alpha(x,t)$ defined for all $t \geq 0$. In this proof, we will omit the superscript α. From the previous section, we know that if η describes integral curves of u, then we must have

$$v(\eta(X,t),t) = v_0(X) \tag{4.47}$$

for all X and all $t \geq 0$. Note also that for each t, the map $X \mapsto \eta(X,t)$ is a diffeomorphism of $\mathbb{R}$. Hence, if we define $x = \eta(X,t)$, we see that

$$\operatorname*{ess\,sup}_x |v(x,t)| = \operatorname*{ess\,sup}_X |v(\eta(X,t),t)| = \operatorname*{ess\,sup}_X |v_0(X)|,$$

which then implies (P1). Now let us differentiate (4.47) with respect to X:

$$\partial_x v(\eta(X,t),t)\partial_X \eta(X,t) = v_0'(X).$$

Because $\eta(X,0) = X$, we know $\partial_X \eta(X,0) = 1$. Proposition 4 then gives us $\partial_X \eta(X,t) > 0$

for all X and all $t > 0$. So we may write

$$|\partial_x v(\eta(X,t),t)|\,\partial_X \eta(X,t) = \left|v_0'(X)\right|.$$

Integrating both sides, we obtain

$$\int_{\mathbb{R}} |\partial_x v(\eta(X,t),t)|\,\partial_X \eta(X,t)\,dX = \int_{\mathbb{R}} \left|v_0'(X)\right|\,dX.$$

Now we change variables via $x = \eta(X,t)$ and obtain

$$\int_{\mathbb{R}} |\partial_x v(x,t)|\,dx = \int_{\mathbb{R}} \left|v_0'(X)\right|\,dX,$$

which is precisely (P2). Now (P3) follows immediately from (P2) together with the fact that for a smooth function f,

$$\mathrm{T.V.}\, f = \int_{\mathbb{R}} \left|f'(x)\right|\,dx. \qquad \square$$

Already we can conclude based on Helley's theorem (see Corollary A.7 in [33]) that there exists a subsequence α_j with $\lim_{j\to\infty} \alpha_j = 0$ such that $v^{\alpha_j}(x,t)$ converges almost everywhere to a function $v(x,t)$. To prove convergence in L^1, we must do more work.

Proposition 6. *Given standard data for the Cauchy problem (4.45), the solution $v^\alpha(x,t)$ may be used to define the function $u^\alpha(x,t) = \mathcal{H}^{-1}v^\alpha(x,t)$. Then u satisfies*

$$\|u^\alpha(\cdot,\cdot)\|_{L_\infty} \le M_1, \tag{H1}$$

$$\|u^\alpha(x+h,t) - u^\alpha(x,t)\|_{L^1} \le \omega(|h|) \quad \text{for any } h \in \mathbb{R}, \tag{H2}$$

$$\|u^\alpha(\cdot,t+k) - u^\alpha(\cdot,t)\|_{L_1} \le M_3 k, \quad \text{for any } k > 0, \tag{H3}$$

for $t \in [0,T]$. Here, M_1 is independent of α, M_3 is independent of t, k and α and ω is a nonnegative continuous function on $[0,\infty)$ with $\omega(r) \downarrow 0$ as $r \downarrow 0$. Such a function ω is called a modulus of continuity.

Proof. Starting with

$$u^\alpha(x,t) = \frac{1}{2\alpha}\int_{\mathbb{R}} e^{-\frac{|x-y|}{\alpha}} v^\alpha(y,t)\,dy, \tag{4.48}$$

we use (P1) to estimate

$$\begin{aligned}|u^\alpha(x,t)| &\le \frac{1}{2\alpha}\int_{\mathbb{R}} e^{-\frac{|x-y|}{\alpha}}|v^\alpha(y,t)|\,dy \\ &\le ||v_0(\cdot)||_{L_\infty}\frac{1}{2\alpha}\int_{\mathbb{R}} e^{-\frac{|x-y|}{\alpha}}\,dy = ||v_0(\cdot)||_{L_\infty},\end{aligned}$$

proving (H1). Here, we also used

$$\frac{1}{2\alpha}\int_{\mathbb{R}} e^{-\frac{|x-y|}{\alpha}}\,dx = 1. \tag{4.49}$$

To prove (H2), we estimate

$$\begin{aligned}\int_{\mathbb{R}} |u^\alpha(x+h,t) - u^\alpha(x,t)|\,dx &\le \frac{1}{2\alpha}\int_{\mathbb{R}}\int_{\mathbb{R}} e^{-\frac{|x-y|}{\alpha}}|v^\alpha(y+h,t) - v^\alpha(y,t)|\,dy\,dx \\ &= \int_{\mathbb{R}} |v^\alpha(y+h,t) - v^\alpha(y,t)|\,dy\frac{1}{2\alpha}\int_{\mathbb{R}} e^{-\frac{|x-y|}{\alpha}}\,dx.\end{aligned}$$

Hence,

$$\int_{\mathbb{R}} |u^\alpha(x+h,t) - u^\alpha(x,t)|\,dx \le \int_{\mathbb{R}} |v^\alpha(x+h,t) - v^\alpha(x,t)|\,dx. \tag{4.50}$$

Then we use (P3) to conclude

$$\int_{\mathbb{R}} |v^\alpha(x+h,t) - v^\alpha(x,t)|\,dx \le |h|\,\mathrm{T.V.}\,v^\alpha(\cdot,t) = \mathrm{T.V.}\,v_0. \tag{4.51}$$

Finally, to prove (H3) we start from the following estimate, derived in the same way as (4.50):

$$\int_{\mathbb{R}} |u^\alpha(x,t+k) - u^\alpha(x,t)|\,dx \le \int_{-\infty}^{\infty} |v^\alpha(x,t+k) - v^\alpha(x,t)|\,dx. \tag{4.52}$$

By integrating (4.3) from t to $t+k$ ($k>0$), we have

$$\int_{\mathbb{R}} |v^\alpha(x,t+k)-v^\alpha(x,t)|\,dx \le \int_{\mathbb{R}}\int_t^{t+k} |u^\alpha(x,s)v_x^\alpha(x,s)|\,ds\,dx$$
$$\le ||u^\alpha||_{L_\infty}\int_t^{t+k} \|v_x^\alpha(\cdot,s)\|_{L^1}\,ds. \tag{4.53}$$

Using (P2) and (H1), we may estimate the right-hand side of (4.53), resulting in

$$\int_{\mathbb{R}} |v^\alpha(x,t+k)-v^\alpha(x,t)|\,dx \le M_1\|v_0'\|_{L^1}k.$$

Now combining this with (4.52), we have the desired result. □

Strong convergence to a weak solution of Burgers. Using the estimates given above, we may prove the following

Theorem 4. *Suppose we solve the Cauchy problem (4.45) with standard data. Using the solution v^α, let us define $u^\alpha = \mathcal{H}^{-1}v^\alpha$ in the usual way. Then, as $\alpha\to 0$, passing if necessary to a subsequence, there exists a function $u(x,t)$ such that*

$$u^\alpha \to u \text{ in } C([0,\infty);L^1_{loc}(\mathbb{R})).$$

The function u is a global weak solution of the initial-value problem (4.5) for the inviscid Burgers equation.

Proof. The first part of the theorem concerns compactness, i.e. strong convergence of u^α in the zero-α limit. The three uniform estimates proved in Proposition 6 are precisely the conditions of the L^1 compactness theory for conservation laws. (See Theorem A.8 in [33] or Theorem 19.9 in [79] for modern accounts of this.) The specific result is that there exists a subsequence $\alpha_j\to 0$ such that $\{u^{\alpha_j}(t)\}$ converges strongly to a function $u(x,t)$, where $u(\cdot,t)\in L^1_{\rm loc}(\mathbb{R})$ for each $t\ge 0$. The convergence is in $C([0,\infty);L^1_{\rm loc}(\mathbb{R}))$.

For the second half of the theorem, we go back to equation (4.1), which we repeat here:

$$u_t^\alpha + u^\alpha u_x^\alpha - \alpha^2 u_{txx}^\alpha - \alpha^2 u^\alpha u_{xxx}^\alpha = 0. \tag{4.54}$$

We wish to prove that the α^2 terms

$$\alpha^2 u^\alpha_{txx} + \alpha^2 u^\alpha u^\alpha_{xxx},$$

converge weakly to 0 as $\alpha \to 0$. Suppose we have shown this; then, we may multiply (4.54) by a test function φ that is compactly supported in $\mathbb{R} \times [0, \infty)$ and integrate in space and time. Now taking $\alpha \to 0$, we will find that the order α^2 terms vanish, and we are left with a function u that satisfies

$$\int_0^\infty \int_{\mathbb{R}} u\varphi_t + \frac{1}{2}u^2\varphi_x \, dx \, dt = 0,$$

for all compactly supported φ. This is precisely the statement that u is a global weak solution of the inviscid Burgers equation, and would prove the theorem.

For the first α^2 term from (4.54), we have, for any compactly supported φ,

$$\alpha^2 \int_0^T \int_{\mathbb{R}} u^\alpha_{txx}\varphi \, dx \, dt = -\alpha^2 \int_0^T \int_{\mathbb{R}} u^\alpha \varphi_{txx} \, dx \, dt.$$

Using the convergence of the sequence u^α, it is clear that this term converges to 0 as $\alpha \to 0$. For the second α^2 term from (4.54), we may derive using integration by parts

$$\alpha^2 \int_0^T \int_{\mathbb{R}} u^\alpha u^\alpha_{xxx}\varphi \, dx \, dt = \frac{1}{4}\alpha^2 \int_0^T \int_{\mathbb{R}} (u^\alpha)^2 \varphi_{xxx} \, dx \, dt - \frac{3}{2}\alpha^2 \int_0^T \int_{\mathbb{R}} u^\alpha u^\alpha_{xx}\varphi_x \, dx \, dt. \tag{4.55}$$

By using the boundedness and the convergence of u^α, we conclude that the first term on the right-hand side of (4.55) vanishes in the $\alpha \to 0$ limit. Regarding the second term, by considering the boundedness of u^α, it is enough to show that

$$\alpha^2 \int_0^T \int_K |u^\alpha_{xx}| dx \to 0,$$

for any compact K. We have

$$\begin{aligned}\int_0^T \int_K |\alpha^2 u^\alpha_{xx}|\,dx &= \int_0^T \int_K |u^\alpha - v^\alpha|\,dx \\ &= \int_0^T \int_K \left| \frac{1}{2\alpha} \int_{\mathbb{R}} e^{-\frac{|x-y|}{\alpha}} v^\alpha(y,t)\,dy - v^\alpha(x,t) \right| dx\,dt.\end{aligned}$$

Here, we used (4.48) to obtain the second equality. Integrating by parts, we get

$$\frac{1}{2\alpha} \int_{\mathbb{R}} e^{-\frac{|x-y|}{\alpha}} v^\alpha(y,t)\,dy = v^\alpha(x,t) + \frac{1}{2} \int_{\mathbb{R}} \operatorname{sgn}(y-x) e^{-\frac{|y-x|}{\alpha}} v^\alpha_y(y,t)\,dy.$$

Continuing, we find

$$\begin{aligned}\int_0^T \int_K |\alpha^2 u^\alpha_{xx}|\,dx\,dt &\le \frac{1}{2} \int_0^T \int_K \int_{\mathbb{R}} e^{-\frac{|y-x|}{\alpha}} |v^\alpha_y(y,t)|\,dy\,dx\,dt \\ &= \frac{1}{2} \int_0^T \int_K |v^\alpha_y(y,t)|\,dy\,dt \int_{\mathbb{R}} e^{-\frac{|y-x|}{\alpha}}\,dx \\ &= \alpha \int_0^T \int_K |v^\alpha_y(y,t)|\,dy\,dt,\end{aligned}$$

where we used (4.49) for the last equality. From (P2), we conclude that the term

$$\alpha \int_0^T \int_K |v^\alpha_y(y,t)|\,dy\,dt$$

is of order $O(\alpha)$ and hence, goes to 0 as $\alpha \to 0$. This completes the argument. □

4.5 Entropy/Numerics

Using a standard finite-difference scheme, we solve the initial-value problem (4.45) numerically with an eye towards checking the Oleinik entropy inequality (4.6). Here we simply describe the numerical scheme, deferring discussion of its convergence properties to future work. Then we discuss various numerical results for both short- and long-time simulations.

Numerical scheme. Beginning with system (4.45), we truncate the spatial domain to $[-a, a]$. Because the domain is now finite, we must impose artificial boundary conditions;

we impose the condition that v vanishes for $|x| > a$. We discretize the domain $[-a, a]$ using an equispaced grid with N grid points. Let us denote this grid by $x_i = -a + (i - 1)\Delta x$, where $i = 1, \ldots, N$, and the grid spacing is given by $\Delta x = 2a/(N - 1)$.

On this discrete domain, we consider the evolution in time of the vector $\mathbf{v}(t) = (v_1(t), \ldots, v_N(t))$. We will suppose that $v_i(t) \approx v(x_i, t)$.

Following [43], we define two basic operators on $\mathbb{R}^N$:

$$\mathbf{z} \mapsto \Delta_0^2 \mathbf{z}, \quad \left(\Delta_0^2 \mathbf{z}\right)_k = z_{k+1} - 2z_k + z_{k-1}, \tag{4.56}$$

$$\mathbf{z} \mapsto \Gamma_0 \mathbf{z}, \quad (\Gamma_0 \mathbf{z})_k = \frac{1}{2}\left(z_{k+1} - z_{k-1}\right). \tag{4.57}$$

Here we use the convention that $z_k = 0$ for $k < 1$ and for $k > N$. This corresponds to the artificial boundary conditions discussed above. Note that the operators (4.56-4.57) are in fact linear transformations of $\mathbb{R}^N$ and may be written in matrix form. Now in terms of the operators (4.56-4.57), we may write the standard finite-difference approximations to the first- and second-derivative operators ∂_x and ∂_x^2:

$$D^1 = \frac{1}{\Delta x}\Gamma_0 \left[\mathrm{Id} - \frac{1}{6}\left(\Delta_0^2\right) + \frac{1}{30}\left(\Delta_0^2\right)^2\right] + \mathcal{O}(\Delta x^5), \tag{4.58}$$

$$D^2 = \frac{1}{\Delta x^2}\left[\left(\Delta_0^2\right) - \frac{1}{12}\left(\Delta_0^2\right)^2 + \frac{1}{90}\left(\Delta_0^2\right)^3\right] + \mathcal{O}(\Delta x^6). \tag{4.59}$$

With this notation, it is clear that the semidiscrete form of (4.45) is

$$\mathbf{v}_t = -\left[\left(\mathrm{Id} - \alpha^2 D^2\right)^{-1} \mathbf{v}\right] D^1 \mathbf{v}, \tag{4.60a}$$

$$v_j(0) = v(x_j, 0), \tag{4.60b}$$

where $v(x, 0)$ is the initial data for the continuum problem, and where concatenation of vectors means component-wise multiplication, i.e., $(\mathbf{ab})_k = a_k b_k$. The first-order ODE (4.60) can now be solved numerically using the time-stepping algorithm of one's choice—we used a high-order explicit Runge-Kutta method.

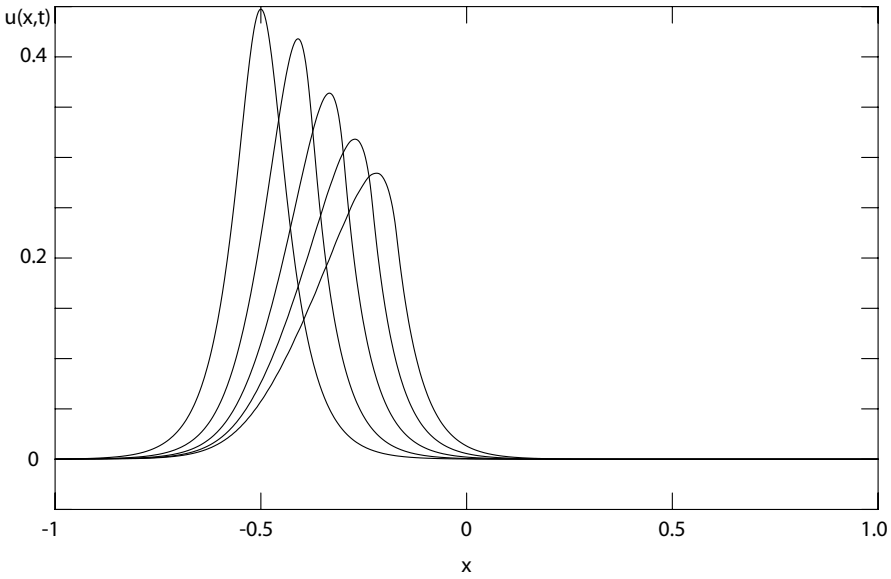


Figure 4.3: The numerical solution of (4.45) for $\alpha = 0.3$ with initial data (4.61). The tallest curve, with a peak at $x = -0.5$, is the solution at $t = 0$. From left to right, we then have the solutions at $t = 1.25$, $t = 2.5$, $t = 3.75$, and $t = 5$. As time passes, the height of the pulse decays while its width increases.

Norm decay of solutions. First we present results for the following choice of initial data

$$v(x,0) = \operatorname{sech}^2\left(\frac{x+1/2}{1/5}\right), \tag{4.61}$$

for $\alpha = 0.3$. With this choice of initial data, $v(x,0) > 0$ so we are within the bounds of our well-posedness and convergence theory. We solve the problem using $N = 1024$ grid points. See Figure 4.3 for snapshots of the solution $u(x,t)$ at $t = 0$, $t = 1.25$, $t = 2.5$, $t = 3.75$, and $t = 5$. The initial profile does not shock or develop any singularities. Instead, it decays steadily in a rather similar fashion as the solution of the viscous Burgers equation (4.7a). To see this decay in three norms, we use the numerically

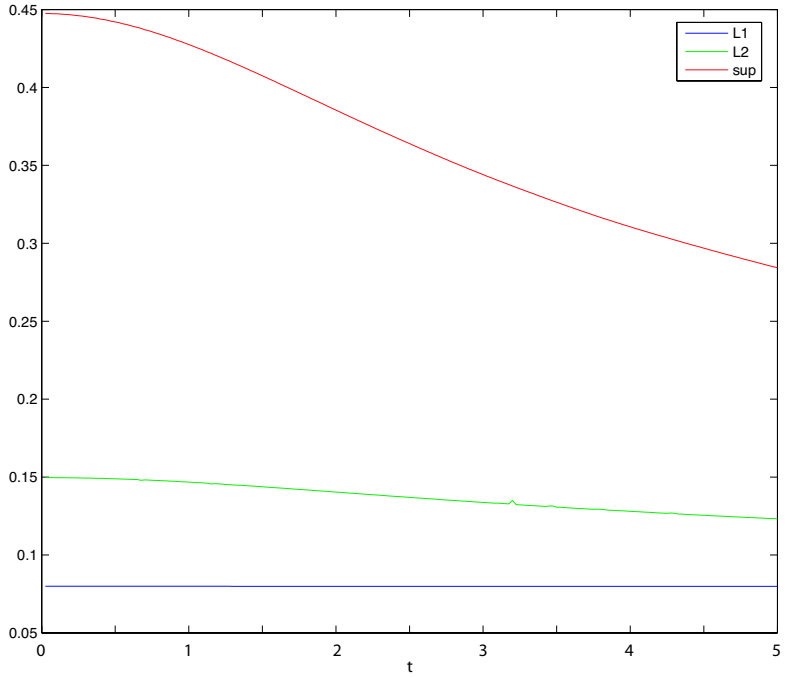


Figure 4.4: $\|u(\cdot,t)\|_{L^1}$, $\|u(\cdot,t)\|_{L^2}$, and $\|u(\cdot,t)\|_{L^\infty}$ as functions of t for the solution to (4.45) with initial data (4.61) and $\alpha = 0.3$.

computed solution to compute

$$\|u(\cdot,t)\|_{L^1},$$
$$\|u(\cdot,t)\|_{L^2},$$
$$\|u(\cdot,t)\|_{L^\infty},$$

as functions of time t. The results are plotted in Figure 4.4, clearly showing the decay. Here we see that the L^1 norm of u stays constant in time, i.e.,

$$\|u(\cdot,t)\|_{L^1} = \|u_0\|_{L^1}.$$

This is a simple consequence of the fact that we chose $v_0 > 0$. For when $v_0 > 0$, we know that $v(x,t) > 0$ for all $x \in \mathbb{R}$, $t > 0$. Then, using the Green's function of Helmholtz operator as in (4.48), we may deduce that $u(x,t) > 0$ for all $x \in \mathbb{R}$, $t > 0$ as well. Then

we have

$$\int_{\mathbb{R}} |u(x,t)|\,dx = \int_{\mathbb{R}} u(x,t)\,dx = \int_{\mathbb{R}} u_0(x)\,dx = \int_{\mathbb{R}} |u_0(x)|\,dx,$$

showing that $\|u\|_{L^1}$ must stay constant in time. Note also from Figure 4.4 that both the L^2 and L^∞ norms of u are strictly decreasing in time. We showed the L^∞ decay property in the previous section—see Proposition 6. However, at this time, we have no analytical method for deriving a uniform L^2 decay law such as what is seen in Figure 4.4.

Entropy inequality: numerical evidence. In the previous section, we established a basic convergence theory for (4.1). That is, we choose initial data $v_0 > 0$, solve (4.1), and label the solution as u^α. Then we know that a subsequence of u^α converges, in the zero-α limit, to a function u. We know one more thing: this function u is a weak solution of the inviscid Burgers equation (4.5) with initial data v_0.

At the time of writing, this is where rigorous analysis ends. This is unfortunate, in light of the fact that there are many weak solutions of the inviscid Burgers equation (4.5) with initial data v_0—the unique, physically relevant solution, is the one that satisfies the Oleinik inequality (4.6).

We will investigate numerically the validity of

$$\sup_x u_x^\alpha(x,t) < \frac{C}{t}, \tag{4.62}$$

where C does not depend on α. Fortunately, there is plenty of numerical evidence that (4.62) holds uniformly in α. By implication, this is evidence that the strong limit u of solutions of (4.1) does in fact satisfy the Oleinik inequality (4.6). Let us discuss some of this numerical evidence.

We repeatedly solve (4.45) with the initial data

$$v_0(x) = \operatorname{sech}^2\left(\frac{x+1/2}{1/5}\right),$$

using successively smaller α values: $\alpha = 0.4$, $\alpha = 0.3$, and $\alpha = 0.2$. We then plot the

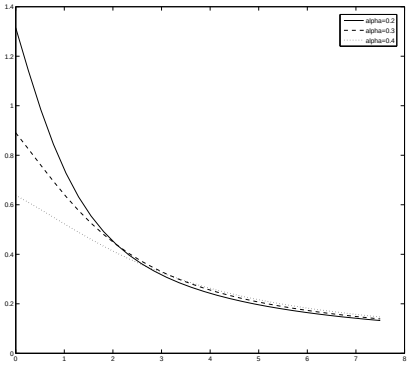


Figure 4.5: Plot of $\sup_x u_x^\alpha(x,t)$ as a function of t for three decreasing values of α, from $t = 0$ to $t = 7$.

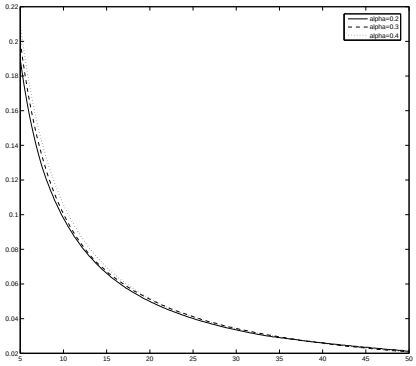


Figure 4.6: Plot of $\sup_x u_x^\alpha(x,t)$ as a function of t for three decreasing values of α, from $t = 5$ until $t = 50$.

quantity

$$m^\alpha(t) := \sup_x u_x^\alpha(x,t) \tag{4.63}$$

as a function of t for each of the three values of α. First we present Figure 4.5 which shows (4.63) from $t = 0$ until $t = 7$. The same quantity (4.63) from $t = 5$ until $t = 50$ is plotted in Figure 4.6. Both plots lend one to believe that as $\alpha \to 0$, the curves $m^\alpha(t)$ are uniformly bounded by a curve of the form C/t. The evidence becomes clearer when we consider the same data on logarithmic axes. Taking the logarithm of both sides of (4.62), we obtain for $t > 1$,

$$\frac{\log\left(\sup_x u_x^\alpha(x,t)\right)}{\log t} < \frac{\log C}{\log t} - 1, \tag{4.64}$$

where C must not depend on α. With this in mind, we examine Figure 4.7, which shows the same data as Figure 4.6 now plotted on a log-log scale. The numerically computed slope of the linear part of this plot is less than -1.25, meaning that the numerically computed solutions $u^\alpha(x,t)$ all satisfy

$$\frac{\log\left(\sup_x u_x^\alpha(x,t)\right)}{\log t} < -1 < -1 + \frac{\log C}{\log t}, \tag{4.65}$$

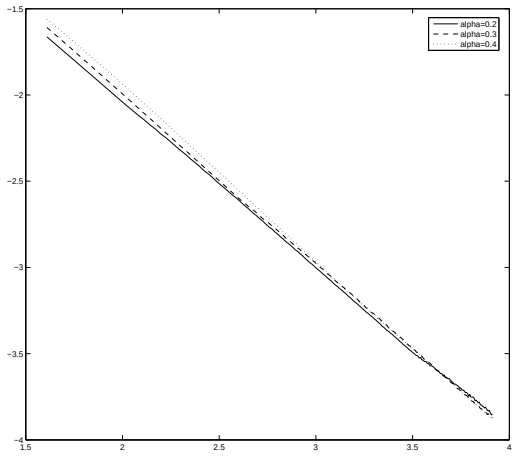


Figure 4.7: Plot of $\log\left(\sup_x u_x^\alpha(x,t)\right)$ as a function of $\log t$ for three decreasing values of α, from $t = 5$ until $t = 50$.

for any $C > 1$. Let us remark that we have run the same numerical test with different choices of initial data, resolution, and values of α. In all cases, we find that the numerically computed solutions satisfy (4.65).

There is solid numerical evidence that the solutions $u^\alpha(x,t)$ satisfy (4.62). Because we have not found any evidence that falsifies this claim, we theorize that the limit function $u(x,t)$ is indeed a weak entropy solution of the inviscid Burgers equation.

4.6 Geometric Structure

Consider the functional $H : L^1(\mathbb{R}) \to \mathbb{R}$ defined by

$$H(v) = \int_{\mathbb{R}} v\,dx, \tag{4.66}$$

and the operator

$$\mathfrak{D} = -v_x\left(\partial_x - \alpha^2\partial_x^3\right)^{-1} v_x. \tag{4.67}$$

This functional/operator pair is the $b = 0$ case of the Hamiltonian structure proposed in [17] for the b-family (see (4.9)). Using these two objects, we write the infinite-dimensional generalization of Hamilton's equation:

$$v_t = \mathfrak{D}\frac{\delta H}{\delta v}. \tag{4.68}$$

Here $\delta H/\delta v$ is the *functional derivative*, defined by

$$\left\langle \frac{\delta H}{\delta v}, \delta v \right\rangle (v) = \left. \frac{d}{d\epsilon} \right|_{\epsilon=0} H\left(v + \epsilon \delta v\right),$$

where $\langle \cdot\,, \cdot \rangle$ is the natural pairing.

Let us now show that (4.68) is precisely (4.1). Define

$$u(x,t) = \frac{1}{2\alpha} \int_{\mathbb{R}} e^{-|x-y|/\alpha} v(y,t)\, dy.$$

so that $\mathcal{H}u = v$ where $\mathcal{H} = \mathrm{Id} - \alpha^2 \partial_x^2$ as in (4.2). It is clear from (4.66) that $\delta H/\delta v = 1$. Using this in (4.68) yields

$$v_t = -v_x \left(\partial_x - \alpha^2 \partial_x^3\right)^{-1} v_x = -v_x u,$$

which was what was desired. This calculation shows that the regularized equation (4.1) is Hamiltonian, assuming of course that $\mathfrak{D}$ is a valid Hamiltonian operator. The operator $\mathfrak{D}$ is Hamiltonian if the induced bracket $\{\cdot, \cdot\}$, defined by

$$\{F, G\} = \int_{\mathbb{R}} \frac{\delta F}{\delta v} \mathfrak{D} \frac{\delta G}{\delta v}\, dx, \tag{4.69}$$

is a Poisson bracket.

Definition 1. *A **Poisson bracket** on a manifold M is a skew-symmetric, bilinear operation $\{\,,\}$ on $C^\infty(M)$ satisfying both*

1. *the Jacobi identity $\{F, \{G, H\}\} + \{H, \{F, G\}\} + \{G, \{H, F\}\} = 0$; and*
2. *the Leibniz identity $\{FG, H\} = \{F, H\}G + F\{G, H\}$,*

for all F, G, and $H \in C^\infty(M)$.

Lemma 5. *The bracket $\{\,,\}$ induced by $\mathfrak{D}$ is skew-symmetric.*

Proof. Because the operator $\mathfrak{L} = \partial_x - \alpha^2\partial_x^3$ has only odd-ordered derivatives, integration by parts gives

$$\langle f, \mathfrak{L}g\rangle = -\langle \mathfrak{L}f, g\rangle,$$

where $\langle\,,\rangle$ is the natural pairing

$$\langle f, g\rangle = \int_{\mathbb{R}} fg\,dx.$$

Hence we write $\mathfrak{L}^* = -\mathfrak{L}$, which implies $\left(\mathfrak{L}^{-1}\right)^* = -\mathfrak{L}^{-1}$. We use this and the definition of $\mathfrak{D}$ to obtain

$$\begin{aligned}
\{F, G\} &= -\int \frac{\delta F}{\delta v} v_x \mathfrak{L}^{-1}\left(v_x \frac{\delta G}{\delta v}\right) dx \\
&= -\int \left(\mathfrak{L}^{-1}\right)^*\left(\frac{\delta F}{\delta v} v_x\right) v_x \frac{\delta G}{\delta v}\, dx \\
&= \int \mathfrak{L}^{-1}\left(v_x \frac{\delta F}{\delta v}\right) v_x \frac{\delta G}{\delta v}\, dx \\
&= -\{G, F\}.
\end{aligned}$$

□

Lemma 6. *The bracket $\{\,,\}$ induced by $\mathfrak{D}$ satisfies the Jacobi identity.*

Proof. Directly verifying the Jacobi identity for (4.69) requires copious amounts of algebra, so we use the multi-vector formalism and Schouten bracket described in [56, Chap. 10]. Let us give a sketch of the proof first: using the bi-vector B defined by

$$B = \frac{1}{2}\partial_x \wedge \mathfrak{D}\partial_x, \tag{4.70}$$

we realize the Poisson bracket as

$$\{F, G\} = \mathbf{i}_B\left(\mathbf{d}F \wedge \mathbf{d}G\right). \tag{4.71}$$

We will prove that the Schouten bracket of B with itself is zero. Then, by the Jacobi-

Schouten identity (see [56, Thm. 10.6.2]), we know that the Jacobi identity holds for the bracket $\{\,,\}$. First let us verify (4.71) by direct computation:

$$\begin{aligned}\mathbf{i}_B\,(\mathbf{d}F \wedge \mathbf{d}G) &= \frac{1}{2}\int \frac{\delta F}{\delta v}\mathfrak{D}\frac{\delta G}{\delta v}\,dx - \frac{1}{2}\int \frac{\delta G}{\delta v}\mathfrak{D}\frac{\delta F}{\delta v}\,dx \\ &= \int \frac{\delta F}{\delta v}\mathfrak{D}\frac{\delta G}{\delta v} = \{F, G\},\end{aligned}$$

where we have used skew-symmetry (Lemma 5). Next we use the definition of $\mathfrak{D}$ and $\mathfrak{L} = \partial_x - \alpha^2\partial_x^3$ to write

$$\mathfrak{D}\partial_x = -v_x\mathfrak{L}^{-1}(v_x\partial_x),$$

which implies

$$-v_x^{-1}\mathfrak{L}v_x^{-1}\mathfrak{D}\partial_x = \partial_x.$$

Here we mean simply $v_x^{-1} = 1/v_x$. Using this, we compute the Schouten bracket:

$$\begin{aligned}&[\partial_x \wedge \mathfrak{D}\partial_x, \partial_x \wedge \mathfrak{D}\partial_x] = [\partial_x \wedge \mathfrak{D}\partial_x, -v_x^{-1}\mathfrak{L}v_x^{-1}\mathfrak{D}\partial_x \wedge \mathfrak{D}\partial_x] \\ &\quad = -\int_{\mathbb{R}} -v_x^{-2}\partial_x(\mathfrak{D}\partial_x) \wedge \mathfrak{L}v_x^{-1}\mathfrak{D}\partial_x \wedge \mathfrak{D}\partial_x + v_x^{-1}\mathfrak{L}(-v_x^{-2})\partial_x(\mathfrak{D}\partial_x) \wedge \mathfrak{D}\partial_x \wedge \mathfrak{D}\partial_x\,dx.\end{aligned}$$

The second term vanishes because $\mathfrak{D}\partial_x \wedge \mathfrak{D}\partial_x = 0$. For the first term, we evaluate

$$\begin{aligned}\mathfrak{L}v_x^{-1}\mathfrak{D}\partial_x &= \partial_x(v_x^{-1}\mathfrak{D}\partial_x) - \alpha^2\partial_x^3(v_x^{-1}\mathfrak{D}\partial_x) \\ &= -v_x^{-2}v_{xx}\mathfrak{D}\partial_x + v_x^{-1}\partial_x(\mathfrak{D}\partial_x) - \alpha^2\partial_x^3(v_x^{-1}\mathfrak{D}\partial_x).\end{aligned}$$

Since $\mathfrak{D}\partial_x \wedge \mathfrak{D}\partial_x = 0$ and $\partial_x(\mathfrak{D}\partial_x) \wedge \partial_x(\mathfrak{D}\partial_x) = 0$, we are left with

$$[\partial_x \wedge \mathfrak{D}\partial_x, \partial_x \wedge \mathfrak{D}\partial_x] = -\alpha^2\int_{\mathbb{R}} v_x^{-2}\partial_x(\mathfrak{D}\partial_x) \wedge \partial_x^3(v_x^{-1}\mathfrak{D}\partial_x) \wedge \mathfrak{D}\partial_x\,dx.$$

The only contributions from the ∂_x^3 term that will matter are those that involve either

$\partial_x^2(\mathfrak{D}\partial_x)$ or $\partial_x^3(\mathfrak{D}\partial_x)$. With this in mind, we continue the computation:

$$\begin{aligned}
&= -\alpha^2 \int_{\mathbb{R}} v_x^{-2}\partial_x(\mathfrak{D}\partial_x) \wedge (-3v_x^{-2}v_{xx}\partial_x^2(\mathfrak{D}\partial_x) + v_x^{-1}\partial_x^3(\mathfrak{D}\partial_x)) \wedge \mathfrak{D}\partial_x \, dx \\
&= -\alpha^2 \int_{\mathbb{R}} -3v_x^{-4}v_{xx}\partial_x(\mathfrak{D}\partial_x) \wedge \partial_x^2(\mathfrak{D}\partial_x) \wedge (\mathfrak{D}\partial_x) \\
&\quad + v_x^{-3}\partial_x(\mathfrak{D}\partial_x) \wedge \partial_x^3(\mathfrak{D}\partial_x) \wedge (\mathfrak{D}\partial_x)\, dx.
\end{aligned}$$

Integrating the second term by parts to move one derivative *off* the $\partial_x^3(\mathfrak{D}\partial_x)$ term, we find that the entire expression cancels, proving that $[B, B] = 0$ as required. □

Lemma 7. *The bracket $\{\,,\}$ induced by $\mathfrak{D}$ satisfies the Leibniz identity.*

Proof. We will use the fact that the Leibniz rule holds for functional derivatives:

$$\frac{\delta(FG)}{\delta v} = \frac{\delta F}{\delta v}G + F\frac{\delta G}{\delta v}.$$

The proof of this consists of a simple calculation combined with the observation that $F(v)$ and $G(v)$ do not depend[2] explicitly on x. Using this, we evaluate

$$\begin{aligned}
\{FG, H\}(v) &= \int_{\mathbb{R}} \frac{\delta(FG)}{\delta v}(v)\mathfrak{D}\frac{\delta H}{\delta v}(v)\, dx \\
&= \int_{\mathbb{R}} \frac{\delta F}{\delta v}(v)G(v)\mathfrak{D}\frac{\delta H}{\delta v}(v) + F(v)\frac{\delta G}{\delta v}(v)\mathfrak{D}\frac{\delta H}{\delta v}(v)\, dx \\
&= \{F, H\}(v)G(v) + F(v)\{G, H\}(v),
\end{aligned}$$

where again we have used the fact that $F(v)$ and $G(v)$ are x-independent real numbers. As v was arbitrary, we have shown that the Leibniz identity holds. □

Theorem 5. *The bracket $\{\,,\}$ induced by $\mathfrak{D}$ is a Poisson bracket.*

Proof. By linearity of functional derivatives, it is clear that (4.69) is bilinear. Then the preceding lemmas have established that the bracket is skew-symmetric and satisfies the Jacobi and Leibniz identities. □

[2]Note that this statement does not hold for the functional derivatives. For example, take $F(v) = \int v^2\, dx$. Then clearly $\partial_x[F(v)] = 0$. But $(\delta F/\delta v)(v) = 2v$ and so in general $\partial_x[(\delta F/\delta v)(v)] \neq 0$.

Remark. The functional plus operator style of describing infinite-dimensional Hamiltonian systems generalizes certain objects that arise naturally for Hamiltonian systems with finite degrees of freedom. Let us briefly review Hamilton's equations for a classical particle in $\mathbb{R}^n$. Let $\mathbf{q} = (q^1, \ldots, q^n)$ denote the particle's position, and let $\mathbf{p} = (p_1, \ldots, p_n)$ denote its momentum. Then, given the function $H : \mathbb{R}^{2n} \to \mathbb{R}$, Hamilton's equations are

$$\dot{q}^i = \frac{\partial H}{\partial p_i}, \tag{4.72a}$$

$$\dot{p}_i = -\frac{\partial H}{\partial q^i}. \tag{4.72b}$$

Let us write this more compactly: we introduce the the skew-symmetric matrix

$$\mathbf{J} = \begin{bmatrix} \mathbf{0} & \mathbf{1} \\ -\mathbf{1} & \mathbf{0} \end{bmatrix},$$

and the variable $\mathbf{z} = \big(q^1, \ldots, q^n, p_1, \ldots, p_n\big)$. Then we may write (4.72) as

$$\dot{z} = \mathbf{J}\nabla H(z).$$

Comparing with (4.68), we see two similarities: $\delta H/\delta v$ is the infinite-dimensional version of ∇H, while $\mathfrak{D}$ can be thought of as a generalization of $\mathbf{J}$. Moreover, we see that the bracket (4.69) is inspired by the classical Poisson bracket in finite dimensions:

$$\{F, G\} = (\nabla F)^T \mathbf{J} \nabla G.$$

See [56, Chap. 3], [68, Chap. 7], and [14, 55] for further information on infinite-dimensional Hamiltonian mechanics.

Casimirs. It happens to be the case that the bracket (4.69) defined using the operator (4.67) has no non-trivial Casimirs. Let us quickly verify this. Suppose there exists a

function G such that for all F, we have

$$\{F, G\} = 0.$$

By definition (4.69) of the bracket, this would imply that

$$\int_{\mathbb{R}} \frac{\delta F}{\delta v} \mathfrak{D} \frac{\delta G}{\delta v}\, dx = 0,$$

for all F. The only way this can happen is if in fact

$$\mathfrak{D} \frac{\delta G}{\delta v} = 0.$$

This reads

$$\left(\partial_x - \alpha^2 \partial_x^3\right)^{-1} \left(v_x \frac{\delta G}{\delta v} \right) = 0.$$

Now applying $\left(\partial_x - \alpha^2 \partial_x^3\right)$ to both sides, we obtain

$$\frac{\delta G}{\delta v} = 0,$$

so the only Casimirs are trivial. This is unfortunate—if we had even one non-trivial Casimir, we could use it to prove stability of the traveling wave solutions via the energy-Casimir method. As things stand, deciding the stability of the traveling wave solutions of (4.1) is likely to be very challenging.

4.7 Future Directions

At this point, it should be clear that there are at least three problems of immediate interest:

1. We should prove either that the limit $u = \lim_{\alpha \to 0} u^\alpha$ is an entropy solution of the inviscid Burgers equation, or we should demonstrate why the entropy condition fails.

2. We should explore well-posedness for general initial data in a particular function

space, extending the $v_0 > 0$ result that we have already.

3. We should determine the stability of the traveling wave front solutions. In particular, it would be interesting if the upwardly-sloped ($d > 0$) traveling wave is unstable while the downwardly-sloped ($d < 0$) traveling wave is stable.

Besides these immediate issues, there are several projects of longer-term interest suggested by the current work.

Extension to higher dimensions. Consider the "vector Burgers equation:"

$$\mathbf{u}_t + (\mathbf{u} \cdot \nabla)\,\mathbf{u} = 0. \tag{4.73}$$

Is it possible that the system

$$\begin{aligned} \mathbf{v}_t + (\mathbf{u} \cdot \nabla)\,\mathbf{v} &= 0, \\ \mathbf{u} - \alpha^2 \Delta \mathbf{u} &= \mathbf{v}. \end{aligned}$$

regularizes the vector Burgers equation (4.73), just as (4.3) regularizes the scalar Burgers equation?

Extension to one-dimensional gas dynamics. Consider one-dimensional isentropic gas dynamics:

$$\rho_t + (\rho u)_x = 0, \tag{4.74a}$$

$$(\rho v)_t + \left(\rho u^2 + p\right)_x = 0, \tag{4.74b}$$

where $p = p(\rho)$. Might it be possible to regularize this system using a mechanism similar to that of (4.1) ? One candidate system that comes to mind is

$$\rho_t + (\rho u)_x = 0, \tag{4.75a}$$

$$(\rho u)_t + (\rho u v + p)_x = 0, \tag{4.75b}$$

$$u - \alpha^2 u_{xx} = v. \tag{4.75c}$$

Suppose there is zero-α convergence of (ρ^α, u^α) to weak solutions of (4.74). Then is (4.75) Hamiltonian in some sense?

Rough initial data. It is important to note that both standard and filtered viscosities select the correct entropy solution even in the case of discontinuous initial data, where $u_0 \in L^\infty$ only. This allows one to legitimately use these viscous regularizations to solve Riemann problems. It would be interesting to see whether we can solve Riemann problems directly using (4.1). That is, what happens to (4.3) for initial data $v_0 \in L^\infty$?

Other smoothing kernels. Another idea is to replace the Helmholtz operator with another operator. For example, we could attempt to regularize Burgers' equation via

$$v_t + uv_x = 0, \tag{4.76}$$

where u and v are related in any number of ways. One interesting possibility is

$$\widehat{u} = \frac{\widehat{v}}{1 + \alpha|k|}.$$

Now v is, roughly speaking, the "square root" of the Helmholtz operator $\mathcal{H}$ applied to u. Hence u is only *one* derivative smoother than v, whereas in (4.1), u is *two* derivatives smoother than v. How smooth does u have to be, relative for v, for (4.76) to genuinely regularize the Burgers equation?

Geometric structures. Finally, it would be interesting to determine where the Hamiltonian structure (4.66-4.67) comes from. The Hamiltonian functional (4.66) is linear in

the field variable and therefore does not have the meaning of a kinetic energy. Similarly, what is the meaning of the non-local operator in (4.67)? Is there a Lagrangian structure that yields (4.1)? Answering these questions will give us physical insight into why our model works the way it does.

Bibliography

[1] R. Abraham, J. E. Marsden, and T. S. Ratiu. *Manifolds, tensor analysis, and applications*, volume 75 of *Applied Mathematical Sciences*. Springer-Verlag, New York, 2004. Draft edition.

[2] D. Andrews and M. E. McIntyre. An exact theory of nonlinear waves on a Lagrangian-mean flow. *J. Fluid Mech.*, 89:609–646, 1978.

[3] V. I. Arnold. Sur la géométrie différentielle des groupes de Lie de dimension infinie et ses applications à l'hydrodynamique des fluides parfaits. *Ann. Inst. Fourier (Grenoble)*, 16:319–361, 1966.

[4] V. I. Arnold and B. A. Khesin. *Topological methods in hydrodynamics*, volume 125 of *Applied Mathematical Sciences*. Springer-Verlag, New York, 1998.

[5] T. B. Benjamin, J. L. Bona, and J. J. Mahoney. Model equations for long waves in nonlinear dispersive systems. *Phil. Trans. R. Soc. London, Ser. A*, 272:47–78, 1972.

[6] G. Birkhoff and G.-C. Rota. *Ordinary differential equations*. Blaisdell, Waltham, MA, second edition, 1969.

[7] J. C. Brunelli and A. Das. On an integrable hierarchy derived from the isentropic gas dynamics. *J. Math. Phys.*, 45:2633–2645, 2004.

[8] R. Camassa. Characteristics and the initial value problem of a completely integrable shallow water equation. *Discr. Cont. Dyn. Sys. Ser. B*, 3(1):115–139, 2003.

[9] R. Camassa and D. D. Holm. An integrable shallow water equation with peaked solitons. *Phys. Rev. Lett.*, 71:1661–1664, 1993.

[10] C. Cao, D. D. Holm, and E. S. Titi. Traveling wave solutions for a class of one-dimensional nonlinear shallow water wave models. *J. Dynam. Differential Equations*, 16:167–178, 2004.

[11] G.-Q. Chen and D. Wang. The Cauchy problem for the Euler equations for compressible fluids. In *Handbook of mathematical fluid dynamics, Vol. I*, pages 421–543. North-Holland, Amsterdam, 2002.

[12] S. Chen, C. Foias, D. D. Holm, E. Olson, E. S. Titi, and S. Wynne. Camassa-Holm equations as a closure model for turbulent channel and pipe flow. *Phys. Rev. Lett.*, 81:5338–5341, 1998.

[13] S. Y. Chen, D. D. Holm, L. G. Margolin, and R. Zhang. Direct numerical simulation of the Navier-Stokes alpha model. *Phys. D*, 133:66–83, 1999.

[14] P. R. Chernoff and J. E. Marsden. *Properties of infinite dimensional Hamiltonian systems*, volume 425 of *Lecture Notes in Mathematics*. Springer-Verlag, Berlin, 1974.

[15] A. J. Chorin and J. E. Marsden. *A mathematical introduction to fluid mechanics*, volume 4 of *Texts in Applied Mathematics*. Springer-Verlag, New York, third edition, 1993.

[16] W. J. Cocke. Turbulent hydrodynamic line stretching: Consequences of entropy. *Phys. Fluids*, 12:2488–2492, 1969.

[17] A. Degasperis, D. D. Holm, and A. N. W. Hone. Integrable and non-integrable equations with peakons. In *Nonlinear physics: theory and experiment, II (Gallipoli, 2002)*, pages 37–43. World Sci. Publishing, River Edge, NJ, 2003.

[18] P. Deift, S. Venakides, and X. Zhou. An extension of the steepest descent method for Riemann-Hilbert problems: The small dispersion limit of the Korteweg-de Vries (KdV) equation. *Proc. Nat. Acad. Sci*, 95:450–454, 1998.

[19] H. R. Dullin, G. A. Gottwald, and D. D. Holm. Camassa-Holm, Korteweg-de Vries-5 and other asymptotically equivalent equations for shallow water waves. *Fluid. Dynam. Res.*, 33:73–95, 2003.

[20] H. R. Dullin, G. A. Gottwald, and D. D. Holm. On asymptotically equivalent shallow water wave equations. *Phys. D*, 190:1–14, 2004.

[21] D. G. Ebin. The motion of slightly compressible fluids viewed as a motion with strong constraining force. *Ann. of Math.*, 105:141–200, 1977.

[22] D. G. Ebin and J. E. Marsden. Groups of diffeomorphisms and the notion of an incompressible fluid. *Ann. of Math.*, 92:102–163, 1970.

[23] G. A. El, R. H. J. Grimshaw, and A. M. Kamchatnov. Wave breaking and the generation of undular bores in an integrable shallow water system. *Stud. Appl. Math.*, 114:395–411, 2005.

[24] L. C. Evans. *Partial differential equations*, volume 19 of *Graduate Studies in Mathematics*. American Mathematical Society, Providence, RI, 1998.

[25] A. Favre. Équations des gaz turbulents compressibles I. *J. Méc.*, 4:361–390, 1965.

[26] K. Fellner and C. Schmeiser. Burgers-Poisson: A nonlinear dispersive model equation. *SIAM J. Appl. Math.*, 64:1509–1525, 2004.

[27] C. Foias, D. D. Holm, and E. S. Titi. The three dimensional viscous Camassa-Holm equations and their relation to the navier-stokes equations and turbulence theory. *J. Dynam. Differential Equations*, 14:1–36, 2002.

[28] B. Fornberg. *A practical guide to pseudospectral methods*, volume 1 of *Cambridge monographs on applied and computational mathematics*. Cambridge University Press, New York, 1996.

[29] R. A. Goler and M. J. Reeder. The generation of the morning glory. *J. Atmos. Sci.*, 61:1360–1376, 2004.

[30] J. Goodman and P. D. Lax. On dispersive difference schemes. I. *Comm. Pure Appl. Math.*, 41(5):591 – 613, 1988.

[31] R. H. J. Grimshaw. Solitary waves in a compressible fluid. *Pure Appl. Geophys.*, 119:780–797, 1981.

[32] R. H. J. Grimshaw and S. R. Pudjaprasetya. Generation of secondary solitary waves in the variable-coefficient Korteweg-de Vries equation. *Stud. Appl. Math.*, 112:271–279, 2004.

[33] H. Holden and N. H. Risebro. *Front tracking for hyperbolic conservation laws*, volume 152 of *Applied Mathematical Sciences*. Springer-Verlag, New York, 2002.

[34] D. D. Holm. Fluctuation effects on 3d Lagrangian mean and Eulerian mean fluid motion. *Physica D*, 133:215–269, 1999.

[35] D. D. Holm. Averaged Lagrangians and the mean effects of fluctuations in ideal fluid dynamics. *Phys. D*, 133:215–269, 2002.

[36] D. D. Holm and A. N. W. Hone. A class of equations with peakon and pulson solutions. *J. Nonlinear Math. Phys.*, 12(suppl. 1):380–394, 2005. With an appendix by H. W. Braden and J. G. Byatt-Smith.

[37] D. D. Holm and J. E. Marsden. Momentum maps and measure-valued solutions (peakons, filaments, and sheets) for the EPDiff equation. In *The breadth of symplectic and Poisson geometry*, volume 232 of *Progr. Math.*, pages 203–235. Birkhäuser Boston, Boston, MA, 2005.

[38] D. D. Holm, J. E. Marsden, and T. S. Ratiu. The Hamiltonian structure of continuum mechanics in material, inverse material, spatial and convective representations. In *Hamiltonian structure and Lyapunov stability for ideal continuum dynamics*, volume 100 of *Sém. Math. Sup.*, pages 11–124. Presses Univ. Montréal, Montreal, Canada, 1986.

[39] D. D. Holm, J. E. Marsden, and T. S. Ratiu. Euler–poincaré models of ideal fluids with nonlinear dispersion. *Phys. Rev. Lett.*, 349:4173–4177, 1998.

[40] D. D. Holm, J. E. Marsden, and T. S. Ratiu. The Euler-Poincaré equations and semidirect products with applications to continuum theories. *Adv. Math.*, 137:1–81, 1998.

[41] D. D. Holm and M. F. Staley. Wave structure and nonlinear balances in a family of evolutionary pdes. *SIAM J. Appl. Dyn. Sys.*, 2:323–380, 2003.

[42] A. N. W. Hone and J. P. Wang. Prolongation algebras and Hamiltonian operators for peakon equations. *Inverse Problems*, 19:129–145, 2003.

[43] A. Iserles. *A first course in the numerical analysis of differential equations*. Cambridge University Press, New York, 1996.

[44] F. John. *Nonlinear wave equations, formation of singularities*, volume 2 of *University Lecture Series*. American Mathematical Society, Providence, RI, 1990.

[45] T. Kakutani and T. Kawahara. Weak ion-acoustic shock waves. *J. Phys. Soc. Japan*, 29:1068–1073, 1970.

[46] S. Kamvissis, K. D. T.-R. McLaughlin, and P. D. Miller. *Semiclassical soliton ensembles for the focusing nonlinear Schrödinger equation*, volume 154 of *Annals of Mathematics Studies*. Princeton University Press, Princeton, NJ, 2003.

[47] T. Kawahara. Weak nonlinear magneto-acoustic waves in a cold plasma in the presence of effective electron-ion collisions. *J. Phys. Soc. Japan*, 29:1321–1329, 1970.

[48] I. G. Kevrekidis, C. W. Gear, J. M. Hyman, P. G. Kevrekidis, O. Runborg, and C. Theodoropoulos. Equation-free, coarse-grained multiscale computation: Enabling microscopic simulators to perform system-level analysis. *Commun. Math. Sci.*, 1:715–762, 2002.

[49] B. A. Khesin. Topological fluid dynamics. *Notices Amer. Math. Soc.*, 52:9–19, 2005.

[50] S. N. Kružkov. First order quasilinear equations with several independent variables. *Mat. Sb. (N.S.)*, 81 (123):228–255, 1970.

[51] C. Lattanzio and P. Marcati. Global well-posedness and relaxation limits of a model for radiating gas. *J. Diff. Eqns.*, 190:439–465, 2003.

[52] P. D. Lax and C. D. Levermore. The small dispersion limit of the Korteweg-de Vries equation. *Comm. Pure Appl. Math.*, 36:I 253–290, II 571–593, III 809–829, 1983.

[53] H. W. Liepmann and A. Roshko. *Elements of gasdynamics*. Wiley, New York, 1957.

[54] H. Liu and E. Tadmor. Critical thresholds in a convolution model for nonlinear conservation laws. *SIAM J. Math. Anal.*, 33:930–945, 2001.

[55] J. E. Marsden and T. J. R. Hughes. *Mathematical foundations of elasticity*. Dover Publications Inc., New York, 1994. Corrected reprint of the 1983 original.

[56] J. E. Marsden and T. S. Ratiu. *Introduction to mechanics and symmetry*, volume 17 of *Texts in Applied Mathematics*. Springer-Verlag, New York, second edition, 1999.

[57] J. E. Marsden, T. S. Ratiu, and A. Weinstein. Reduction and Hamiltonian structures on duals of semidirect product Lie algebras. In *Fluids and plasmas: geometry and dynamics (Boulder, Colo., 1983)*, volume 28 of *Contemp. Math.*, pages 55–100. Amer. Math. Soc., Providence, RI, 1984.

[58] J. E. Marsden, T. S. Ratiu, and A. Weinstein. Semidirect products and reduction in mechanics. *Trans. Amer. Math. Soc.*, 281:147–177, 1984.

[59] J. E. Marsden and S. Shkoller. Global well-posedness for the lans-α equations on bounded domains. *R. Soc. London Philos. Trans. Ser. A Math. Phys. Eng. Sci.*, 359:1449–1468, 2001.

[60] J. E. Marsden and S. Shkoller. The anisotropic averaged euler equations. *Arch. Ration. Mech. Anal.*, 166:27–46, 2003.

[61] A. Menhofer, R. K. Smith, M. J. Reeder, and D. R. Christie. The bore-like character of three morning glories observed during the central australian fronts experiment. *Aust. Meteor. Mag.*, 46:277–285, 1997.

[62] R. H. M. Miesen, L. P. J. Kamp, and F. W. Sluijter. Long solitary waves in compressible shallow fluids. *Phys. Fluids A*, 2:359–370, 1990.

[63] R. H. M. Miesen, L. P. J. Kamp, and F. W. Sluijter. Solitary waves in compressible deep fluids. *Phys. Fluids A*, 2:1404–1411, 1990.

[64] A. V. Mikhailov and V. S. Novikov. Perturbative symmetry approach. *J. Phys. A*, 35:4775–4790, 2002.

[65] K. Mohseni, B. Kosović, S. Shkoller, and J. E. Marsden. Numerical simulations of the Lagrangian averaged Navier-Stokes equations for homogeneous isotropic turbulence. *Phys. Fluids*, 15:524–544, 2003.

[66] P. J. Morrison. Hamiltonian description of the ideal fluid. *Rev. Modern Phys.*, 70:467–521, 1998.

[67] P. I. Naumkin and I. A. Shishmarëv. *Nonlinear nonlocal equations in the theory of waves*, volume 133 of *Translations of Mathematical Monographs*. American Mathematical Society, Providence, RI, 1994.

[68] P. J. Olver. *Applications of Lie groups to differential equations*, volume 107 of *Graduate Texts in Mathematics*. Springer-Verlag, New York, second edition, 1993.

[69] H. Omori. *Infinite-dimensional Lie groups*, volume 158 of *Translations of Mathematical Monographs*. Amer. Math. Soc., Providence, RI, 1997.

[70] D. H. Peregrine. Calculations of the development of an undular bore. *J. Fluid. Mech.*, 25:321–330, 1966.

[71] S. C. Preston. *Eulerian and Lagrangian stability of fluid motions*. Ph.D. thesis, State University of New York at Stony Brook, 2002.

[72] R. Rivlin and J. L. Erickson. Stress-deformation relations for isotropic materials. *J. Rational Mech. Anal.*, 4:323–425, 1955.

[73] P. L. Sachdev and V. S. Seshadri. Long waves in inviscid compressible atmospheres: Isothermal atmosphere. *Phys. Fluids*, 22:60–67, 1979.

[74] S. Schochet and E. Tadmor. The regularized Chapman-Enskog expansion for scalar conservation laws. *Arch. Rat. Mech. Anal.*, 119:95–107, 1992.

[75] A. H. Shapiro. *The dynamics and thermodynamics of compressible fluid flow*. Ronald Press, New York, 1953.

[76] M. C. Shen and J. B. Keller. Ray method for nonlinear wave propagation in a rotating fluid of variable depth. *Phys. Fluids*, 16:1565–1572, 1973.

[77] M. C. Shen and S. M. Sun. Asymptotic method for interfacial solitary waves in a compressible fluid. *Meth. Appl. Anal.*, 3:135–156, 1996.

[78] S. Shkoller. *Bull. Amer. Math. Soc.*, 37:175–181, 2000.

[79] J. Smoller. *Shock Waves and Reaction-Diffusion Equations*, volume 258 of *Grundlehren der Mathematischen Wissenschaften*. Springer-Verlag, New York, 1983.

[80] C. G. Speziale, G. Erlebacher, T. A. Zang, and M. Y. Hussaini. The subgrid-scale modeling of compressible turbulence. *Phys. Fluids*, 31:940–942, 1988.

[81] N. Sugimoto, M. Masuda, K. Yamashita, and H. Horimoto. Verification of acoustic solitary waves. *J. Fluid Mech.*, 504:271–299, 2004.

[82] E. Tadmor. Burgers' equation with vanishing hyper-viscosity. *Comm. Math. Sci.*, 2:317–324, 2004.

[83] G. I. Taylor. The spectrum of turbulence. *Proc. R. Soc. London Ser. A*, 164:476–490, 1938.

[84] A. Tovbis, S. Venakides, and X. Zhou. On semiclassical (zero dispersion limit) solutions of the focusing nonlinear Schrodinger equation. *Comm. Pure Appl. Math.*, 57:877–985, 2004.

[85] W. G. Vincenti and C. H. Kruger. *Introduction to physical gas dynamics*. Wiley, New York, 1965.

[86] G. B. Whitham. *Linear and Nonlinear Waves*. Wiley-Interscience, New York, 1974.