Mathematical perspectives on waves and currents

a thesis presented for the degree of

Doctor of Philosophy of Imperial College London

and the

Diploma of Imperial College

by

Oliver D. Street

Department of Mathematics

Imperial College

180 Queen's Gate, London SW7 2BZ

August 2022

I certify that this thesis, and the research to which it refers, are the product of my own work, and that any ideas or quotations from the work of other people, published or otherwise, are fully acknowledged in accordance with the standard referencing practices of the discipline.

Copyright

Thesis advisor: Professor Dan Crisan Oliver D. Street
Additional advisor(s): Professor Darryl D. Holm

Mathematical perspectives on waves and currents

Abstract

The behaviour of waves on the surface of a fluid has fascinated scientists for centuries. Attempts to describe the problem mathematically have revealed a rich geometric structure, as well as a number of celebrated equations. When considering a stochastic theory of water waves, it is therefore sensible to begin with a structure preserving methodology of introducing a stochastic noise into a fluid model. Within this thesis, the mathematical framework of *semi-martingale driven variational principles* is introduced, which reveals a new methodology of formulating problems for which we have a stochastic action integral. The inclusion of *stochastic advection by Lie transport* into the underlying fluid momentum equation will allow us to achieve a novel stochastic perturbation of water wave theory which preserves its geometric properties. A number of phenomena observable on the free surface of a fluid are challenging to describe using existing modelling approaches. In particular, the classical modelling approach requires modification to permit the introduction of thermal gradients or rotational flows. Through a new variational perspective involving the composition of two maps, the interaction between waves and thermal fronts in the upper ocean is studied. This approach involves a natural separation of waves and currents on the free surface as vertical oscillations around a horizontal two dimensional flow, and allows the consideration of wave-current interactions. Separately, currents are responsible for the advection of material suspended within the fluid. We consider the procession of an inertial object through a fluid domain, which involves fluid equations with the structure of a fractional order differential equation. It is shown that the most commonly applied such equation, the *Maxey-Riley equation*, is globally well-posed, a fact which was absent from the literature prior to this thesis.

To my family.

Acknowledgments

Having the opportunity to conduct this research over the past few years has been an enormous privilege, for which I am indebted to a number of people.

Without the careful supervision and friendship of both Professors Dan Crisan and Darryl Holm, it would not have been possible for me to undertake this journey. Dan has an astonishing eye for detail, true leadership skills, and a remarkable ability to see the path through a problem. Darryl is a rare example of someone whose passion for their work is evident to all, and whose flair for sharing the beauty of mathematics with others is something I aspire to replicate.

I would like to thank my examiners, Colin Cotter and Jacques Vanneste, as well as Bertrand Chapron and Baylor Fox-Kemper, whose expert advice I have valued greatly. I thank Tom Bendall for supervising my summer project at the Met Office, and for ensuring that it was a success in challenging circumstances. I am grateful for the guidance of Robert Bowles, for his assistance with my development during my undergraduate degree. I would also like to express my gratitude towards a number of my colleagues who I have enjoyed conversing with and learning from, including Erwin Luesink, Wei Pan, Oana Lang, So Takao, Ruiao Hu, Stuart Patching, Theo Diamantakis, James-Michael Leahy, and Aythami Bethencourt de León.

I must also acknowledge the support of everyone involved with the EPSRC

Centre for Doctoral Training in the Mathematics of Planet Earth, which is a community in which I have found a home. In particular, to my colleagues in Cohort Echo, and to Justine Jones, for teaching my valuable communication skills.

I acknowledge the unwavering support and loving companionship of my partner Maria, proofreader extraordinaire. I thank her for sharing my highs and lows, and her parents for supporting our living in London. I owe a debt of gratitude to my family, who gave me the self belief necessary to embark on this project. In particular, to my mother and brother, Sue and Josh, for their encouragement and for ensuring that home is never far away. Finally, I thank my late father, John, from whom I have learned the most.

List of Figures

Contents

1

Introduction

During the course of this project it has become apparent that research in mathematics can be divided into two broad themes. One can attempt to either solve problems, or to create problems. Whilst the answers to some questions will be answered rigorously here, the introduction and exploration of new concepts will also be a key theme. Considering this, the reader should not expect this thesis to be a closed story which leaves few questions unanswered. Instead, we will pose new problems, introduce new modelling frameworks, and provide new perspectives. In the following chapters, a number of related concepts will be explored and we will be discussing mathematics arising from a selection of physical problems in fluid mechanics. This content will be arranged into three main sections.

Variational calculus is at the heart of contemporary mathematical interpretations of classical mechanics. The principle of stationary action may be mathematically posed in an abstract setting, and has found application to a plethora of problems. By considering the action of a Lie group on the configuration manifold, it has been shown that this description contains a rich geometric structure. The features of this structure which are relevant to the work in this thesis will be summarised in Chapter 2. In particular, the application of these ideas to fluid dynamics, initially made by Arnold [4], is introduced in an abstract setting. This description of the dynamics of a fluid allows us to summarise a methodology of introducing stochastic noise into a fluid model which preserves its geometric structure, as introduced by Holm [53] and named *stochastic advection by Lie transport* (SALT). In particular, a stochastic perturbation of the transport vector field is made as a fundamental assumption of the model. As will become apparent, the development of stochastic models will be a recurring theme in this thesis and, as such, it is worth motivating their existence.

A numerical simulation of a fluid can inherit uncertainty from a number of sources. Whilst the equations themselves involve making approximations and assumptions which deviate from physical reality, operational models of planetary scale fluid motion also develop errors as a result of technological limitations. The assimilation of data introduces uncertainty as a result of the quality or sparseness of the available data, and current computational limitations are such that the grid on which the equations are discretised is larger than ideal. Both of these examples apply to oceanography and atmospheric physics, and are particularly prevalent in weather forecasting. Since uncertainty is unavoidable in such models, it is worth considering methods of uncertainty quantification. Common in weather forecasting, ensemble simulations are used to account for error in the initial conditions. A stochastic perturbation of the model equations themselves allows the calibration of the model to the data, thus allowing the consideration of model uncertainty. This can be achieved through an intelligent, data-driven, selection of the stochastic terms in the model, where SALT ensures that the resulting equations obey the fundamental geometry of fluid dynamics.

In Chapter 3, we will turn our focus to the dynamics of the surface of a fluid. Due to its immediate visibility in the natural world, the free upper surface of a body of water has long attracted attention from leading thinkers. This system has an abundance of nonlinear features and interactions between variables and represents a surprisingly challenging problem for mathematicians. As we will see explicitly, the projection of three dimensional dynamics onto its upper boundary is more involved than considering the evaluation of the variables on the free surface. In order to derive a closed system of equations, approximations are needed such as irrotational flow or vertical averaging. For the case of irrotational flow, the evaluation of the Euler momentum equation onto the free surface, together with the kinematic boundary, comprise a closed pair of boundary equations known as the *classical water wave equations*. Despite the fact that free boundary problems such as these have been extensively studied mathematically, there are still a number of physical examples of wave current interaction for which we have no adequate mathematical model. Indeed, satellite oceanography observations provide information on the interactions between physical variables on the surface of the ocean. Whilst classical theory provides a methodology to close the problem in terms of quantities on the surface only, the modification of existing methods of closure to include additional physics, such as inhomogeneous density and salinity, is a significant obstacle. The limitations of existing models are particularly stark when interpreting regions of the sea surface with notice-able thermal gradients, including those exhibiting frontogenesis. With this in mind, new approaches to free surface dynamics are necessary to overcome some of these modelling challenges.

The movement of inertial particles through a fluid domain has many applications in environmental sciences. Often, it is the debris near or on the surface of a fluid which has the most devastating impact [102]. The development of equations to understand the influence of mass on the dynamics of such objects has a long history, which is summarised in Chapter 4. The *Maxey-Riley equation* [77] has emerged as the most popular equation amongst the literature surrounding this. Despite this, its mathematical properties have received little attention. In Chapter 4, we give an in-depth analysis of the Maxey-Riley equation. Here, we expand on known results, as well as state and prove a number of previously unknown properties. Inertial particle equations with memory, such as this one, are nonlinear fractional order differential equations. Whilst this inhibits them from having a variational structure, as the other models featured in this thesis do, it is noteworthy that their analytical properties do not differ from those of an ordinary differential equation as much as one might expect.

Contribution of the thesis

Here, we will precisely describe the new results presented in this thesis, and how they relate to the above story. Many of these have appeared in recent papers together with my collaborators D. Crisan [34, 95], D. D. Holm [33], and R. Hu [57].

- In Section 2.2, we introduce a new framework [95] which clarifies the proper formulation of a stochastic action integral. This will involve the definition of a *semimartingale driven variational principle*.

- We will prove that the fundamental lemma of calculus of variations extends to the case of stochastic time integration, in Lemma 2.31. Thus, it is mathematically well defined to derive stochastic models in this way.

- In Theorem 2.35, we give the Euler-Poincaré equations for incompressible fluids with stochastic advection by Lie transport. Here, the semimartingale form of the pressure follows from its role as a Lagrange multiplier within a semimartingale driven variational principle. This extends the work of Holm [53], and allows us to infer the stochastic Euler equations as an example.

- In Section 3.3.1, we will present a new stochastic perturbation of the *classical water wave equations*, using the approach of stochastic advection by Lie transport. Here, the stochastic terms will not be bound by the same potential flow assumption as the deterministic part of the transport.

- In Section 3.3.2, we will show that our stochastic classical water wave equations have a Hamiltonian structure analogous to that found by Zakharov in 1968 [106] for the deterministic theory. This further demonstrates the fact that the SALT noise added to the Euler equations in Chapter 2 is structure preserving.

- In Proposition 3.13, we will show that, as was claimed by Zakharov [106], a variation of the free surface variable induces a variation in the velocity potential. We will show that this is in fact a consequence of the *Lie chain rule*.

- In Section 3.4, we will present a novel modelling approach to derive closed systems of equations for free surface fluid problems which encompass a wider collection of physical variables. This will be based on contributions made during a collaboration with D. D. Holm and R. Hu [57], and with D. Crisan and D. D. Holm [33]. This is founded upon a characterisation of the dynamics as a composition of two maps, which is made stochastic in Section 3.4.4.

- In Sections 3.4.2 and 3.4.3, model equations are derived which show potential to model relationships between waves and thermal gradients observable in the upper ocean.

- In Chapter 4, we analyse the Maxey-Riley equation. In particular, in Theorem 4.13, the Maxey-Riley equation is shown to have global in time weak solutions. In Theorem 4.17, we prove the conditions under which these solutions are strong solutions.

- In Section 4.8, we go on to prove some properties of the solution of the Maxey-Riley equation as a function of its initial conditions. Namely, in Proposition 4.18 we prove that the distance between a pair of trajectories is limited by their initial conditions, as well as prove a non-collision property in Proposition 4.22.

2

A variational approach to deterministic and stochastic modelling

Lagrangian mechanics formulates classical mechanics in terms of the principle of least action, associating a *configuration space* and a *Lagrangian* to a physical system. The mathematical description of the physical system corresponds to critical points of the action functional, defined as the time integral of the Lagrangian. In one of the many graceful interrelations in applied mathematics, these critical points may be identified as being equivalent to solutions of the much celebrated Euler-Lagrange equations. Rather than Newton's laws, this may be considered to be the crux of the mathematical framework of mechanics. Indeed, when considering systems with conservative forces, the Euler-Lagrange equations are analogous to Newton's second law, in that they are equivalent when the Lagrangian is taken to be the difference between kinetic and potential energies. This concept was, in the 19th century, reformulated into Hamiltonian mechanics through the Legendre transform, whereby the problem is written in terms of phase space coordinates. Here, rather than 'velocities', the motion is considered in terms of 'momenta' and corresponds to solutions of Hamilton's equations.

As these ideas have evolved with time, the natural connection between geometry and mechanics has been developed mathematically. We may view Lagrangian and Hamiltonian mechanics through the lens of differential geometry, thus describing mechanics in a generalised coordinate-free manner. The resulting framework is both powerful and beautiful, and the corresponding field of study is known as *geometric mechanics*.

In contemporary applied mathematics, stochastic equations of motion are used to represent uncertainty in the modelling process. Whilst a large number of papers published in this direction incorporate the noise into the model at the level of the equations, in an unphysical manner, geometric mechanics gives us a framework through which we can incorporate a stochastic noise in a way which preserves the mathematical structure of the model. The earliest instance of a stochastic geometric framework is attributed to Bismut in 1981 [13], which was done through the Hamiltonian formulation. Indeed, stochastic Hamiltonian mechanics has been thus far developed more thoroughly than the Lagrangian case [68]. Using a stochastic Lagrangian approach, a framework for deriving stochastic models for fluid mechanics which preserve the Kelvin-Noether circulation theorem was introduced by Holm in his seminal 2015 paper [53]. This approach, known as *stochastic advection by Lie trans-*

port, or SALT, has since been much studied and will form the backbone of the stochastic models we will consider here.

This chapter will be broken into two main sections:

- In Section 2.1, we will give a summary of some key results from geometric mechanics, which we will subsequently make use of in the following sections on wave current interactions. This summary, whilst not fully comprehensive, will include some elementary notions before giving a discussion of semidirect product Euler-Poincaré theory and fluid mechanics. An introduction to stochastic fluid mechanics will be given from a variational perspective.

- Secondly, in Section 2.2, we will report on the findings of original work conducted for this thesis [95], where a framework for formulating stochastic action principles is introduced. This framework will allow us to properly formulate stochastic Lagrange multipliers in a manner which ensures that the resulting equations make sense mathematically. In particular, this allows for the proper formulation of pressure in stochastic fluid models.

2.1 A review of Geometric Mechanics

In this section, together with Appendix A, we will summarise the main results from geometric mechanics necessary for us to later interpret fluid mechanics problems from this perspective, and we will attempt for this to be as self-contained as possible. Naturally, a complete summary of geometric mechanics, where everything is carefully defined, is neither possible nor appropriate here and it is recommended that the reader consult some of the more comprehensive texts [51, 52, 63, 73]. The concepts we will introduce represent tools with which we can interpret nature, much in the same way as paints and a canvas. Since one must paint a picture with these tools to fully appreciate their potential, we will be providing some examples along the way to illustrate the geometric mechanics framework. The examples we will be using are classical, and can be found in many papers and summaries, including [62] for example. Those examples which are relevant to the subsequent chapters will be contained here, and those which are purely illustrative are found in Appendix A. Furthermore, we will introduce a framework through which some of the ideas from geometric mechanics can be made stochastic. For elaboration on any concepts from stochastic calculus used, it is recommended that the reader consult a textbook on stochastic calculus, such as Karatzas and Shreve [66]. Whilst the full history of the development of these ideas is outside of the scope of this work, we will begin with a brief summary of some of the more important developments below. For a more thorough overview of the development of geometric mechanics, see one of the comprehensive texts mentioned above, or [72] for a summary of geometric mechanics for fluids.

The concept behind geometric mechanics dates back to a famous note by Henri Poincaré in 1901 [86], in which the foundations for the Euler-Poincaré equations were laid in just two pages. Given a mechanical system where a Lie algebra acts transitively on the configuration space, the note shows that rather than the governing equations being on the tangent bundle to the configuration space, they can be rewritten on the product of the configuration space with the Lie algebra. These equations are equivalent to the Euler-Lagrange equations, but are written in terms of different variables on a different space. This structure underpins what is today known as the Euler-

Poincaré theorem. A key feature of geometric mechanics is the identification of conserved quantities of the physical system, the identification of a conservation law corresponding to each differential symmetry of the action was made by Emmy Noether in 1918 [80].

There are many problems which have been considered from the perspective of geometric mechanics, ranging from magnetohydrodynamics to quantum chemistry, and certainly many more which have not yet been imagined. An early application, which has received much attention, was Arnold's remarkable observations on fluid mechanics in 1966 [4]. Arnold noticed a profound connection between the motion of an ideal incompressible fluid and the geometry of infinite dimensional Lie groups. Namely, the flow given by the Euler equations for an incompressible fluid is a *geodesic* on the Lie group of volume preserving diffeomorphisms, with respect to the metric given by the kinetic energy. Shortly after, Ebin and Marsden [38] carefully analysed the diffeomorphism group needed for formulations such as Arnold's, making some revealing analytical remarks about the Euler equations. Following decades of progress from a large community of authors, the Euler-Poincaré equations for continuum mechanics with advected quantities was formulated in 1998 [62]. This provided the general framework needed for the full power of the geometric perspective to be unleashed.

2.1.1 A variational formulation of classical mechanics

This section will briefly describe many of the ideas discussed thus far, and lay the foundations for us to apply these ideas to problems in fluid mechanics. Firstly, we will define the degrees of freedom necessary to describe a physical system.

Space and time. We first introduce the idea of the *space* in which the dynamics occurs. By *space*, we mean a manifold, Q, with points denoted by $q \in Q$. The manifold may, on occasion, be considered as a *Lie group*, G, in which case configurations are given by the Lie group action on a reference configuration $G \times Q_0 \mapsto Q$. This gives us a mechanism to describe the state, or *configuration*, of a physical system at a given moment in time. We must define a notion of *time* in order to describe the dynamics on this space. In the most general sense, *time* is represented by a point in a manifold, T. In most cases, and in all of the examples described in this thesis, the manifold will be taken to be $\mathbb{R}$, though it is worth noting that other manifolds are possible and the mathematical framework is valid with more general notions of time.

Motion. By a *motion*, we mean a time dependent map into the configuration manifold, i.e. $\phi_t : T \times Q \mapsto Q$. The motion is a curve, parametrised by time, given by $q_t = \phi_t q_0$. The motion is called a *flow* if it satisfies the so called *flow property*. That is, if $\phi_0 = \mathrm{Id}$ and if, for any $s, t \in \mathbb{R}$, we have $\phi_{t+s} = \phi_t \circ \phi_s$. If the configuration is identified with a Lie group, G, then the motion is a curve, $T \times G \to G$, on the Lie group manifold.

Given some curve, $q(t)$ in Q, there is a space of vectors, T_qQ, which sit atop of this curve. The velocity along this flow is given by the *tangent lift vector*, $\dot{q}(t) = v_q \in T_qQ$, which, at time t, is one particular vector from this space.

Hamilton's principle and the Euler-Lagrange equations

In order to derive governing equations for the dynamics of a system, we may apply *Hamilton's principle* to an action. We will now define the objects necessary to substantiate this statement mathematically.

A *Lagrangian* is a functional $L : TQ \times \mathbb{R} \mapsto \mathbb{R}$, where TQ denotes the tangent bundle to the configuration manifold. The Lagrangian is commonly chosen to be the kinetic energy minus the potential energy, and the choice of the Lagrangian dictates the form of the model which will result from the application of Hamilton's principle. The *action*, S, is defined by the time integral of the Lagrangian, with the addition of any desired constraints enforced by Lagrange multipliers. We will here introduce the mathematical definitions necessary to apply Hamilton's principle, $\delta S = 0$, in practice. In particular, variational derivatives of functionals are necessary.

THE MATHEMATICS OF VARIATIONAL CALCULUS. In order to define the variational derivative, we must first understand the concept of a Radon-Nikodym derivative [12].

Definition 2.1. Let (X, Σ) be a measurable space and suppose we have two σ-finite measures μ and ν. We say that ν is *absolutely continuous* with respect to μ if

$$\mu(A) = 0 \qquad \Longrightarrow \qquad \nu(A) = 0\,,$$

for all $A \in \Sigma$, in which case we write $\nu \ll \mu$.

Theorem 2.2. *Let μ and ν be two measures as defined in Definition 2.1. If ν is absolutely continuous with respect to μ, then there exists some measurable non-negative function $f : X \to [0, \infty)$ such that for any measurable set $A \subseteq X$, we have*

$$\nu(A) = \int_A f \, d\mu\,.$$

Definition 2.3. In the above theorem, the function f is called the *Radon-Nikodym derivative* of ν with respect to μ and is denoted by $\frac{d\nu}{d\mu}$.

Let X be a compact Hausdorff space and let F be a functional $F : C(X) \to \mathbb{R}$. We can now define the variational derivative.

Definition 2.4. Suppose $\rho, \phi \in C(X)$, then we define the *Gâteaux variation* of F at ρ by

$$\delta F(\rho; \phi) = \lim_{\varepsilon \to 0} \frac{F(\rho + \varepsilon\phi) - F(\rho)}{\varepsilon} = \frac{d}{d\varepsilon}\bigg|_{\varepsilon=0} F(\rho + \varepsilon\phi) \,. \tag{2.1}$$

Definition 2.5. We can think of $\delta F(\rho; \phi)$ as a continuous functional with respect to ϕ, hence by the Riesz representation theorem there exists some measure, μ, such that

$$\delta F(\rho; \phi) = \int \phi \, d\mu \,. \tag{2.2}$$

We then define the *variational derivative* of F with respect to ρ to be the Radon-Nikodym derivative of this measure with respect to the Lebesgue measure. We call the function ϕ the 'variation of ρ' and denote this by $\delta\rho$.

Remark 2.6. This definition implies the following relation,

$$\delta F(\rho) = \left\langle \frac{\delta F}{\delta \rho}, \delta\rho \right\rangle,$$

where the angled brackets denote an L_2 pairing. This can be interpreted as an operational definition of the functional derivative δ, which we will later apply in practice when considering variations of an action.

Remark 2.7. A rigorous analysis of this definition and the conditions under which it holds is presented in [47].

These compacts allow us to introduce the Euler-Lagrange equations.

Theorem 2.8. *For any differentiable $L(q, \dot{q})$, we have that Hamilton's principle implies the Euler-Lagrange equations*

$$\frac{d}{dt}\frac{\partial L}{\partial \dot{q}_j} = \frac{\partial L}{\partial q_j}\,, \quad \textit{for} \quad j = 1, \ldots, n\,, \tag{2.3}$$

where $q \in Q$ is assumed to be n dimensional.

In the proof of the above theorem, found in Appendix A, we conveniently ignore the detail of what happens to the endpoint term. An identical endpoint term will occur when we derive Hamilton's canonical equations. Noether's theorem will explain what happens in relation to this term.

Remark 2.9. Hamilton's principle means, informally, that the *physical* path that the system takes is such that an infinitesimal variations of that path does not change the action, up to first order. Suppose that the Lagrangian does not depend on some component of q, call this q_j. Since the Euler-Lagrange equations (2.3) result from Hamilton's principle, we have that

$$\frac{d}{dt}\frac{\partial L}{\partial \dot{q}_j} = \frac{\partial L}{\partial q_j} = 0\,,$$

and the momentum $p_j = \partial L/\partial \dot{q}_j$ is conserved. The symmetry of the Lagrangian in this case is the simplest kind, where it does not depend on a variable. Noether's theorem is a more abstract generalisation of this idea, where a far richer class of symmetries is permitted.

Theorem 2.10 (Noether's theorem)**.** *Suppose that the Lagrangian is invariant under a smooth infinitesimal transformation*

$$\delta q = \Phi_\xi(q) \in TQ\,, \tag{2.4}$$

of a 1-parameter Lie group acting on the manifold, Q, taken at some point $q \in Q$. This transformation depends on an element of the Lie algebra[*]*, $\xi \in \mathfrak{g}$, since it is a transformation infinitesimally close to the identity. The Noether quantity, defined by*

$$J_\xi(q,p) := \langle p, \Phi_\xi(q) \rangle_{T^*Q \times TQ} \,, \tag{2.5}$$

where the angled brackets denote a pairing on $T^*Q \times TQ$, is a constant of motion.

Remark 2.11. By associating the Lagrangian with the equations which correspond to them through Hamilton's principle, Noether's theorem says that the Noether quantity is a conserved function of the variables (p, q), the *phase space* variables, when the equations of motion hold.

Proof (of Noether's theorem). By assumption, the Lagrangian is invariant under $q \to q + \epsilon\Phi_\xi(q)$. We have that $\delta S = 0$ for δq of this form, since the Lagrangian is invariant. By Hamilton's principle, the Euler-Lagrange equations hold and thus the endpoint term $\langle p, \Phi_\xi(q) \rangle |_{t_0}^{t_1}$ is zero. $\mathscr{Q.E.D.}$

In the statement of Noether's theorem, we have used the idea of a *Lie group*. This will be expanded upon in Section 2.1.2, and the Noether quantity will be used to segue into the action of a Lie group on the configuration manifold.

The Hamiltonian formulation and Hamilton's equations

In order to transition to the Hamiltonian formulation of classical mechanics, we must introduce the *Legendre transform*. This will enable us to formu-

[*]The Lie algebra can be interpreted as the tangent space at the identity of the Lie group, as we will see later.

late the problem on *phase space*. The Lagrangian perspective enables us to interpret the dynamics as solutions of the Euler-Lagrange equations, which are equations on TQ written in terms of coordinates $(q, \dot{q})$. The Legendre transformation will take us onto the phase space, T^*Q, which is defined to be the cotangent bundle of the configuration manifold.

The Legendre transformation may be considered as a map, $(q, \dot{q}) \mapsto (q, p)$, from the tangent bundle to the cotangent bundle, where p is a momentum defined through the Lagrangian, $L(q, \dot{q})$, as follows

$$p := \frac{\partial L}{\partial \dot{q}} \,. \tag{2.6}$$

This expression is the *fibre derivative* of L, and is on T^*Q. This allows us to define the Hamiltonian as follows.

Definition 2.12 (The Hamiltonian)**.** In terms of the variables $(q, \dot{q}, p)$ defined above, the *Hamiltonian* of a physical system is a mapping, $H : T^*Q \mapsto \mathbb{R}$, which may be defined in terms of the Lagrangian as follows

$$H(q, p) := \langle p, \dot{q} \rangle - L(q, \dot{q}) \,, \tag{2.7}$$

where the angled brackets denotes a pairing which maps $T^*Q \times TQ$ to $\mathbb{R}$.

Theorem 2.13 (Hamilton's canonical equations)**.** *Applying Hamilton's principle to the action written on the phase space in terms of the Hamiltonian, as follows,*

$$0 = \delta \int_{t_0}^{t_1} \langle p, \dot{q} \rangle - H(q, p) \, dt \,, \tag{2.8}$$

implies Hamilton's canonical equations

$$\dot{q} = \frac{\partial H}{\partial p} \,, \quad \textit{and} \quad \dot{p} = -\frac{\partial H}{\partial q} \,. \tag{2.9}$$

For a proof and discussion of this, see Appendix A.

2.1.2 Lie groups and reduction by symmetry

Whilst the previous section summarised the variational approach to classical mechanics, we will now discuss the lift of the mechanics on the manifold, Q, to the mechanics on a Lie group, G, which acts on Q. This is the idea introduced by Poincaré [86] and is at the heart of geometric mechanics.

Definition 2.14. A *Lie group*, G, is a manifold which also has a group structure.

Definition 2.15. A *Lie algebra*, $\mathfrak{g}$, is a vector space together with a *Lie bracket*, $[\cdot,\cdot] : \mathfrak{g} \times \mathfrak{g} \mapsto \mathfrak{g}$, which is a bilinear, skew-symmetric operation which satisfies the Jacobi identity

$$[f,[g,h]] + [g,[h,f]] + [h,[f,g]] = 0\,, \quad \text{for all} \quad f,g,h \in \mathfrak{g}\,.$$

Remark 2.16. A Lie algebra, $\mathfrak{g}$, is (isomorphic to) the tangent space of the corresponding Lie group, G, at its identity element. For a description of this, see Lee [69] or a similar text.

For our purposes, we may think of a Lie group as a group of transformations which depend smoothly on a set of parameters

$$q(t) \to q(t,\epsilon) \in Q\,, \quad \text{where} \quad q(t,0) = q(t)\,,$$

which defines the group action of the Lie group on the configuration manifold, $G \times Q \mapsto Q$. We can denote this group action by concatenation,

$$q_{t,\epsilon} = g_\epsilon q_t\,, \quad \text{where} \quad g_\epsilon \in G\,, \text{ and } g_0 = Id\,.$$

The *infinitesimal transformation* of q under this group is the tangent at the identity

$$\delta q = \frac{d}{d\epsilon}\bigg|_{\epsilon=0} g_\epsilon q_t \,, \tag{2.10}$$

since this term multiplies the group parameter, ϵ, in the Taylor expansion of $q_{t,\epsilon}$ around $\epsilon = 0$. Notice that such a variation is taken at fixed time t. This enables us to return to the Noether quantity, and interpret it in terms of the Lie algebra corresponding to the Lie group action.

Definition 2.17 (Cotangent lift momentum map)**.** With the introduction of a pairing, $\langle \cdot, \cdot \rangle : \mathfrak{g}^* \times \mathfrak{g} \mapsto \mathbb{R}$, we may define the Noether quantity from Theorem 2.10 in terms of this pairing,

$$J_\xi = \langle p, \Phi_\xi(q) \rangle_{T^*Q \times TQ} =: \langle J(q,p), \xi \rangle_{\mathfrak{g}^* \times \mathfrak{g}} \,, \tag{2.11}$$

for any fixed $\xi \in \mathfrak{g}$. The map $J(q,p)$ is the *cotangent lift momentum map* associated with the infinitesimal transformation $\delta q = \Phi_\xi(q)$ and its cotangent lift.

Note that we will often employ the useful notation of $J(q,p) := -p \diamond q \in \mathfrak{g}^*$. When we consider partial differential equations for continuum dynamics, the Lie algebra becomes the space of vector fields on the domain and the above infinitesimal transformation will be identified as the *Lie derivative.*

For physical systems where such a Lie group action defines the motion, we may consider an element of the Lie group, g_t, as a map from some reference configuration q_0, to the current configuration q_t. The configuration of the system is therefore described by an element of the Lie group, and the dynamics can be considered to be a curve in the Lie group parametrised by time.

Adjoint actions. The *adjoint representation* of a Lie group provides a representation of the elements of the group as linear transformations of the Lie algebra corresponding to the tangent space of the Lie group at its identity element. We can consider the adjoint action of a Lie group on itself, on its Lie algebra, and the action of a Lie algebra on itself, which we will denote by AD, Ad, and ad respectively. For simplicity, we will define these actions for matrix Lie algebras which, by Ado's theorem, is equivalent to the class of all finite dimensional Lie algebras over a field of characteristic zero.

The adjoint action of a Lie group on itself, $\mathrm{AD} : G \times G \mapsto G$, is defined, for $g, h \in G$, by the *inner automorphism* associated with g

$$\mathrm{AD}_g h := g h g^{-1} \,. \tag{2.12}$$

To define $\mathrm{Ad} : G \times \mathfrak{g} \mapsto \mathfrak{g}$, we take the derivative of this with respect to h at the identity

$$\mathrm{Ad}_g \eta := g \eta g^{-1} \,, \quad \text{where} \quad \eta = h'(0) \,. \tag{2.13}$$

Similarly, we may define $\mathrm{ad} : \mathfrak{g} \times \mathfrak{g} \mapsto \mathfrak{g}$ by differentiating as follows

$$\begin{aligned} \mathrm{ad}_\xi \eta &= \frac{d}{dt}\bigg|_{t=0} \mathrm{Ad}_{g_t} \eta = \frac{d}{dt}\bigg|_{t=0} \left(g_t \eta g_t^{-1} \right) \\ &= \left(\frac{d}{dt}\bigg|_{t=0} g_t \right) \eta g_0^{-1} - g_0 \eta g_0^{-1} \left(\frac{d}{dt}\bigg|_{t=0} g_t \right) g_0^{-1} \\ &= \xi\eta - \eta\xi =: [\xi, \eta] \,, \quad \text{where} \quad g_0 = \mathrm{Id} \quad \text{and} \quad \xi = \frac{d}{dt}\bigg|_{t=0} g_t \,. \end{aligned} \tag{2.14}$$

Reduction by symmetry. We will look at reduction by symmetry on the Lagrangian side, as was done by Marsden and Scheurle [74]. We will be first discussing the concepts informally, before stating the Euler-Poincaré theorem. In this case, the dynamics on the configuration manifold, $q_t = g_t q_0$, are lifted to a curve, $g_t \in G$, with tangent $\dot{g} := v \in T_gG$. The Euler-Lagrange equations naturally hold for the Lagrangian $L(g, v) : TQ \mapsto \mathbb{R}$, as well as Hamilton's canonical equations. Reduced Euler-Poincaré dynamics occur on the dual to the Lie algebra, $\mathfrak{g}^*$, and correspond to the case when the Lagrangian is invariant under the action of G. We say that a Lagrangian, L, is *left invariant* under the action of any element $\tilde{g} \in G$ if

$$L(g, v) = L(\tilde{g}g, \tilde{g}v)\,. \tag{2.15}$$

Right invariance is also possible, and will be used in the case of fluid dynamics. Choosing $\tilde{g} = g^{-1}$, we can define the restriction of the Lagrangian to $\mathfrak{g}$ by

$$\ell(\xi) := L(e, \xi)\,, \quad \text{where} \quad \xi = g^{-1}\dot{g}\,. \tag{2.16}$$

The application of Hamilton's principle to the reduced Lagrangian implies the *Euler-Poincaré equations* for the *left* invariant case

$$\frac{d}{dt}\left(\frac{\delta \ell}{\delta \xi}\right) = \mathrm{ad}^*_\xi \frac{\delta \ell}{\delta \xi}\,. \tag{2.17}$$

For a discussion of this, and how this changes in the case of right invariance, see Appendix A, where an illustrative example may also be found.

Semidirect product reduction. We now consider the extension of the Euler-Poincaré equations to the case where the invariance of the Lagrangian L under the action of G is broken by the dependence of the Lagrangian on some parameter $a_0 \in V^*$, where V is a vector space. In this case we say that the symmetry of the Lagrangian is broken, since it is only invariant on a subset of the Lie group. As before, we have a curve $g_t \in G$ and define

$\xi = g^{-1}\dot{g}$. There exists a left[†] representation of G on V, and G acts on $TG \times V^*$ by $h(\dot{g}, a) = (h\dot{g}, ha)$. Note that, for simplicity, we have denoted the actions of groups and Lie algebras by concatenation. For example, the action of $\mathfrak{g}$ on V^*, which is minus the dual of the action of $\mathfrak{g}$ on V, is denoted by ξa for $\xi \in \mathfrak{g}$ and $a \in V^*$. This provides a definition of the diamond operator, which can be identified as the momentum map from Definition 2.17, through this notation as

$$\langle -\xi a, v \rangle = \langle v \diamond a, \xi \rangle \,, \tag{2.18}$$

for $v \in V$.

Suppose we have a Lagrangian, $L : TQ \times V^* \mapsto \mathbb{R}$, which is left invariant on G. We then define family of Lagrangians, $L_{a_0} : TG \mapsto \mathbb{R}$, parametrised by a_0, by

$$L_{a_0}(g, \dot{g}) = L(g, \dot{g}, a_0) \,. \tag{2.19}$$

The symmetry of this Lagrangian is thus reduced from the group G to the isotropy subgroup, which is defined to leave a_0 invariant under left action. By the G invariance of L, we can define a reduced Lagrangian

$$\ell(\xi, a_t) := L(e, g^{-1}\dot{g}, g^{-1}a_0) \,, \tag{2.20}$$

where a_t is the solution of $\dot{a}_t = -\xi_t a_0$, which may be written as $a_t = g_t^{-1} a_0$. This relationship for a_t is inherited from the left invariance of L. As we will see in the following theorem, by applying Hamilton's principle to the Lagrangian 2.20 we derive Euler-Poincaré equations which are *not* simply the standard Euler-Poincaré equations (2.17) on the semidirect product Lie

[†]We could similarly define this for the right invariant case.

algebra $\mathfrak{g} \ltimes V^*$. We have the following theorem.

Theorem 2.18 (Semidirect product Euler-Poincaré theorem [62])**.** *With the notation introduced above, the following are equivalent:*

a. Hamilton's principle

$$0 = \delta \int_{t_0}^{t_1} L_{a_0}(g, \dot{g})\, dt\,,$$

holds for fixed a_0 and for variations of g vanishing at the endpoints.

b. The curve g_t satisfies the Euler-Lagrange equations for L_{a_0} on G.

c. The variational principle on $\mathfrak{g} \times V^*$

$$0 = \delta \int_{t_0}^{t_1} \ell(\xi, a_t)\, dt\,,$$

holds for variations of the form

$$\delta\xi = \dot{\eta} + [\xi, \eta]\,, \quad \delta a = -\eta a\,,$$

where η vanishes at the endpoints.

d. The following Euler-Poincaré equations hold

$$\frac{d}{dt}\frac{\delta \ell}{\delta \xi} = \mathrm{ad}^*_\xi \frac{\delta \ell}{\delta \xi} + \frac{\delta \ell}{\delta a} \diamond a\,. \tag{2.21}$$

For a full proof, see Holm, Marsden and Ratiu [62]. We will here illustrate the relationship $c) \implies d)$. This follows from taking variations of the action corresponding to c) as follows

$$\begin{aligned}
\delta \int_{t_0}^{t_1} \ell(\xi, a)\, dt &= \int_{t_0}^{t_1} \left\langle \frac{\delta \ell}{\delta \xi}, \delta\xi \right\rangle + \left\langle \delta a, \frac{\delta \ell}{\delta a} \right\rangle dt \\
&= \int_{t_0}^{t_1} \left\langle \frac{\delta \ell}{\delta \xi}, \dot{\eta} + [\xi, \eta] \right\rangle - \left\langle \eta a, \frac{\delta \ell}{\delta a} \right\rangle dt \\
&= \int_{t_0}^{t_1} \left\langle -\frac{d}{dt}\frac{\delta \ell}{\delta \xi} + \mathrm{ad}^*_\xi \frac{\delta \ell}{\delta \xi} + \frac{\delta \ell}{\delta a} \diamond a, \eta \right\rangle dt\,.
\end{aligned}$$

Euler-Poincaré equations for deterministic and stochastic fluid dynamics

After the remarkable observation by Arnold in 1966 [4], described earlier in this chapter, the mathematical perspective on the dynamics of a fluid was forever changed. Following Holm, Marsden, and Ratiu [62], we will demonstrate how the machinery of geometric mechanics (in the semidirect product case) can be applied to this problem. For a more complete summary of how the above definitions are modified for the specific case of fluid dynamics, including the definitions of the adjoint actions of the Lie group and algebra, see Luesink [72].

We take a domain, $\mathcal{D}$, which is an n-dimensional compact manifold with boundary, and let $\mathrm{Diff}(\mathcal{D})$ denote the group of diffeomorphisms from $\mathcal{D}$ to itself. Whilst there are a number of technicalities concerning the form of this group [38, 72], the above framework is valid for *right* invariant Lagrangians and right representations. The motion will be considered as a curve $g_t \in \mathrm{Diff}(\mathcal{D})$, which maps an initial, or *reference*, configuration $\boldsymbol{x}_0 \in \mathcal{D}$ to $\boldsymbol{x}_t$ as

$$\boldsymbol{x}_t = g_t \boldsymbol{x}_0\,. \tag{2.22}$$

The time derivative of this allows us to define the velocity field, u, as follows

$$\dot{\boldsymbol{x}}_t = \dot{g}_t \boldsymbol{x}_0 =: u_t(g_t \boldsymbol{x}_0)\,.$$

Since this is true for any $\boldsymbol{x}_0$, the velocity field is related to the flow map by

$$u_t = \dot{g}_t g_t^{-1}\,. \tag{2.23}$$

We denote by $\mathfrak{X}(\mathcal{D})$ the space of vector fields on $\mathcal{D}$, which is the algebra corresponding to our group. The adjoint action of this Lie algebra on itself is given, for $u_1, u_2 \in \mathfrak{X}$, by $\mathrm{ad}_{u_1} u_2 = -[u_1, u_2]$, and the dual of ad is the Lie derivative, defined as follows.

Definition 2.19 (Lie derivative)**.** Suppose we have a k-form, f, on the configuration manifold, Q, and let u be a vector field with the associated smooth flow, ϕ_ϵ. Then the *Lie derivative* of f along u is given by

$$\mathcal{L}_u f = \frac{d}{d\epsilon}\bigg|_{\epsilon=0} \phi_\epsilon^* f\,, \tag{2.24}$$

where ϕ_ϵ^* denotes the pull back of f by ϕ_ϵ. This a particular case of the *Lie chain rule*, which states that

$$\phi_\epsilon^* \mathcal{L}_u f = \frac{d}{d\epsilon} \phi_\epsilon^* f\,, \tag{2.25}$$

which is equivalent to the definition of the Lie derivative when $\epsilon = 0$.

This is known as the *dynamic* definition of a Lie derivative, it may be defined instead by Cartan's formula from exterior calculus as follows

$$\mathcal{L}_u f = u \,\lrcorner\, df + d(u \,\lrcorner\, f)\,, \tag{2.26}$$

where d and $\lrcorner$ denote the exterior derivative and interior product respectively. It's important to note the Cartan's formula can be used only as a definition of the lie derivative on the space of differential forms.

Using the same notation as in Definition 2.19, note that the pullback relation, known as the *Lie chain rule* is given by

$$\frac{d}{d\epsilon}(\phi_\epsilon^* f) = \phi_\epsilon^* \mathcal{L}_u f \,.$$

Evaluating this at $\epsilon = 0$ gives the definition of a Lie derivative.

The semidirect product group for fluid dynamics with advected quantities is this group structure together with a vector space of advected quantities[‡]. As in the general case, we have a representation of G on V, which induces a representation on V^*. The advected quantities are taken to be in V^*, which is a representation space of $\mathrm{Diff}(\mathcal{D})$. The representation space of the diffeomorphism group is a subset of tensor field densities [62] and the representation is by pullback. Comparing the dynamic definition of the Lie derivative with equation (2.10), we see that the Lie derivative is the infinitesimal transformation associated with the differential of the representation of $\mathrm{Diff}(\mathcal{D})$ on tensor fields, and is the right action of the Lie algebra $\mathfrak{X}(\mathcal{D})$ on V^*. The diamond operation becomes a mapping, $\diamond : V \times V^* :\mapsto \mathfrak{X}(\mathcal{D})^*$, defined, for $v \in V$, by

$$\langle v \diamond a, u \rangle_{\mathfrak{X}^* \times \mathfrak{X}} = -\langle \mathcal{L}_u a, v \rangle_{V^* \times V} \,. \qquad (2.27)$$

The diamond operator may be interpreted as the dual of the Lie derivative when considered as a map $\mathcal{L}_{(\cdot)} a : \mathfrak{X}(\mathcal{D}) \mapsto V^*$.

We will thus have a reduced Lagrangian, $\ell : \mathfrak{X}(\mathcal{D}) \times V^* \mapsto \mathbb{R}$, and the

[‡]For a thorough overview of this group and its actions, see [72]

evolution of an advected quantity, $a \in V^*$, is given by

$$\dot{a} = -\mathcal{L}_u a \,. \tag{2.28}$$

Note that the mass density, which we will discuss later, will always be an advected variable in V^*. The Euler-Poincaré theorem does not differ radically from Theorem 2.18, so will not be repeated in full here. However it should be noted that the Euler-Poincaré equation itself becomes

$$\frac{\partial}{\partial t}\frac{\delta \ell}{\delta u} + \mathcal{L}_u \frac{\delta \ell}{\delta u} = \frac{\delta \ell}{\delta a} \diamond a \,, \tag{2.29}$$

where a is advected according to equation (2.28).

Notice that we have presented this structure in a coordinate free manner. For our purposes, we will be working in Euclidian space with the corresponding Euclidian metric. The basis for the vector space corresponding to the tangent space to the manifold at some point is given by partial derivatives with respect to the local coordinates x^i. We may therefore express a vector field $u \in \mathfrak{X}(\mathcal{D})$ as

$$u = u^1 \frac{\partial}{\partial x^1} + \cdots + u^n \frac{\partial}{\partial x^n} =: \boldsymbol{u} \cdot \nabla \,, \tag{2.30}$$

where $\nabla := (\frac{\partial}{\partial x^1}, \ldots, \frac{\partial}{\partial x^n})$ is the aforementioned local coordinate basis. Here, we note that $\boldsymbol{u}$ represents the vector field relative to the coordinate system. When taking variations, we will see that the velocity also appears as a 1-form. Indeed, we have an isomorphism, $\flat : T\mathcal{D} \mapsto T^*\mathcal{D}$ and its inverse, $\sharp : T^*\mathcal{D} \mapsto T\mathcal{D}$, known as the *musical* isomorphisms. The definition of these is, in general, given in terms of the pseudo-Riemannian metric, g, corresponding to the manifold. Here we will discuss these ideas in terms of Euclidian space, where the components of the (Euclidian) metric, g_E, are given by the Kronecker delta, $g_{Eij} = \delta_{ij}$. Note that the components of the Euclidian metric are related to the dual basis, $\boldsymbol{dx} = (dx^1, \ldots, dx^n)$, by $g_E = \delta_{ij} dx^i \otimes dx^j$. In Euclidean space we have that, for the notation as described above, a 1-form may be associated with the vector field (2.30) by

$$u^\flat := g_{Eij} u^i dx^j =: u_i dx^j \,. \tag{2.31}$$

This is known as *lowering the index*, since the index has been lowered in the final line of the above definition. For the Euclidean metric, we further note that lowering the index has no effect on the functions u^i, indeed

$$u_i dx^j = u^i dx^j \,, \tag{2.32}$$

due to equation (2.31) and the definition of the Kronecker delta. We therefore see that for the vector field $u = \boldsymbol{u} \cdot \nabla$, the corresponding 1-form is $u^\flat = \boldsymbol{u} \cdot \boldsymbol{dx}$. The inverse isomorphism, $\sharp$, is defined similarly and associates a vector field with a given 1-form.

An important concept in fluid theories is that of *mass density*. The manifold of diffeomorphisms, when the specific Sobolev regularity of the mapping is considered, inherits a Riemannian structure from the metric of the underlying manifold, $\mathcal{D}$, as shown by Ebin and Marsden [38]. The measure which corresponds to the bilinear form on the tangent space of $\mathrm{Diff}(\mathcal{D})$ is expressed in terms of the mass density in the reference configuration, D_0, which we take to be strictly positive. The Eulerian description then corresponds to the pushforward by the flow, and we have a measure $\bar{D} = D d^n x$. Here, the notation D is used for the mass density, rather than the commonly used notation ρ, since we can consider D to be the determinant of the Lagrange to Euler map. The mass density is therefore an intrinsic part of the formulation of a fluid theory and, since it is an element of V^*, the semidirect product formulation is natural for a variational continuum theory.

The definition of this measure allows us to elaborate on the form of the variation of the reduced Lagrangian with respect to the vector field, which is a 1-form density as follows

$$\frac{\delta \ell}{\delta u} = \frac{\delta \ell}{\delta \boldsymbol{u}} \cdot d\boldsymbol{x} \otimes d^n x \in \mathfrak{X}^*(\mathcal{D})\,. \tag{2.33}$$

We define by $\mathcal{C}$ the space of continuous closed loops in $\mathcal{D}$. For an arbitrary 1-form density, we may produce a 1-form by dividing it by a volume form. This division is defined as follows.

Definition 2.20. The 1-form corresponding to the division of a 1-form density, $m = \boldsymbol{m} \cdot d\boldsymbol{x} \otimes d^n x \in \mathfrak{X}^*$, by a volume form, $\bar{D} = D d^n x$, is denoted by $m/\bar{D}$ and defined as

$$m = \frac{1}{D}\boldsymbol{m} \cdot d\boldsymbol{x} \otimes D\, d^n x = \frac{m}{\bar{D}} \otimes \bar{D}\,.$$

We define the *circulation map*, $\mathcal{K} : \mathcal{C} \times V^* \mapsto \mathfrak{X}^{**}$, by the integral of the 1-form, $m/\bar{D}$, around a material loop moving with the flow as

$$\langle \mathcal{K}(c(t), a(t)), m \rangle := \oint_{c(t)} \frac{m}{\bar{D}}\,. \tag{2.34}$$

Theorem 2.21 (Abstract Kelvin-Noether circulation theorem)**.** *For $c(t) \in \mathcal{C}$, let u and a satisfy the Euler-Poincaré equation* (2.29) *and the advection equation* (2.28)*. The Kelvin-Noether circulation map, $\mathcal{I} : \mathcal{C} \times V^* \times \mathfrak{X} \mapsto \mathbb{R}$,* *is defined by*

$$\mathcal{I}(c,a,u) := \left\langle \mathcal{K}(c,a), \frac{\delta \ell}{\delta u} \right\rangle , \tag{2.35}$$

and satisfies the following equation

$$\frac{d}{dt}\mathcal{I}(t) = \oint\limits_{c(t)} \frac{1}{\bar{D}} \frac{\delta \ell}{\delta a} \diamond a \,. \tag{2.36}$$

Proof. Since the loop, $c(t)$, is moving with the flow, the time derivative of the loop integral is equal to the loop integral of the advective derivative of the integrand. We thus have

$$\frac{d}{dt}\langle \mathcal{K}(c(t),a(t)), m \rangle = \frac{d}{dt} \oint\limits_{c(t)} \frac{m}{\bar{D}} = \oint\limits_{c(t)} \left(\partial_t + \mathcal{L}_u\right) \frac{m}{\bar{D}} \,.$$

The volume form, $\bar{D}$, passes through the operator, $(\partial_t + \mathcal{L}_u)$, since it is advected. Indeed

$$\begin{aligned}
(\partial_t + \mathcal{L}_u)(\boldsymbol{m} \cdot d\boldsymbol{x} \otimes d^n x) &= (\partial_t + \mathcal{L}_u) \left(\frac{1}{D} \boldsymbol{m} \cdot d\boldsymbol{x} \otimes D \, d^n x \right) \\
&= (\partial_t + \mathcal{L}_u) \left(\frac{1}{D} \boldsymbol{m} \cdot d\boldsymbol{x} \right) \otimes D \, d^n x \,.
\end{aligned}$$

Considering Definition 2.20, the above equations imply

$$\frac{d}{dt}\langle \mathcal{K}(c(t),a(t)), m \rangle = \oint\limits_{c(t)} \frac{1}{\bar{D}} \left(\partial_t + \mathcal{L}_u\right) m \,.$$

This calculation with $m = \frac{\delta \ell}{\delta u}$ gives our result. $\mathscr{Q}.\mathscr{E}.\mathscr{D}.$

Remark 2.22. Advected quantities can take many structural forms, our examples will include advected scalar functions, b, and volume forms[§], $Dd^n x$.

[§]We have used notation corresponding to an n dimensional fluid for the volume form, It is therefore worth commenting on the form of the Lie derivative of these objects, and the definition of the 'diamond' term in the Euler-Poincaré equation (2.29). Beginning with Lie derivatives, the Lie derivative of each of these quantities with respect to u is given by

$$\mathcal{L}_u b = \boldsymbol{u} \cdot \nabla b\,,$$
$$\mathcal{L}_u (D d^n x) = \nabla \cdot (D\boldsymbol{u}) d^n x\,.$$

The diamond terms are defined by integrating by parts, for a vector field $\eta \in \mathfrak{X}$, as

$$\left\langle \frac{\delta \ell}{\delta b} \diamond b, \eta \right\rangle = -\int \left(\frac{\delta \ell}{\delta b} \nabla b \right) \cdot \boldsymbol{\eta}\, d^n x = -\left\langle \frac{\delta \ell}{\delta b} db \otimes d^n x, \eta \right\rangle\,,$$
$$\left\langle \frac{\delta \ell}{\delta D} \diamond D, \eta \right\rangle = \int D \nabla \frac{\delta \ell}{\delta D} \cdot \boldsymbol{\eta}\, d^n x = \left\langle D d \frac{\delta \ell}{\delta D} \otimes d^n x, \eta \right\rangle\,.$$

We will also take a Lie derivative of 1-forms, $\boldsymbol{A} \cdot d\boldsymbol{x}$, which has the form

$$\mathcal{L}_u (\boldsymbol{A} \cdot d\boldsymbol{x}) = \big(\boldsymbol{u} \cdot \nabla \boldsymbol{A} + A_j \nabla u^j \big) \cdot d\boldsymbol{x}\,.$$

This will be necessary since a 1-form will result from taking variational derivatives of the kinetic energy term, which in the examples to follow will be written as the spatial integral of $|\boldsymbol{u}|^2 = \boldsymbol{u} \cdot \boldsymbol{u}$. At face value, the kinetic energy in this may give the impression that the coordinate free description has been abandoned. Instead, this form of the kinetic energy results from the following relationship between the basis of the tangent space and its dual basis

$$\frac{\partial}{\partial x^i} \,\lrcorner\, dx^j = \delta_{ij}\,. \tag{2.37}$$

though it is worth noting that n is usually two or three.

The kinetic energy can be associated to an inner product through its bilinearity as

$$
\begin{aligned}
u \lrcorner u^\flat &= (\boldsymbol{u}\cdot\nabla) \lrcorner (\boldsymbol{u}\cdot d\boldsymbol{x}) \\
&= \left(u^1\frac{\partial}{\partial x^1} + \cdots + u^n\frac{\partial}{\partial x^n}\right) \lrcorner (u_1 dx^1 + \cdots u_n dx^n) \\
&= u^1u_1 + \cdots + u^n u_n = \boldsymbol{u}\cdot\boldsymbol{u}\,.
\end{aligned}
\tag{2.38}
$$

That is, the interior product between a vector field and the one form associated to another vector field is equivalent to the inner product between the two vector fields, where this inner product corresponds to the metric tensor. When taking variational derivatives of the kinetic energy with respect to the vector field u, we see that the result will, formally, be in terms of the 1-form $u^\flat$. This will not necessarily be made explicit through the notation at every step, but should be evident from the context.

Example (The Euler equations for incompressible flow [4])**.** For the flow to be incompressible, we may use a Lagrange multiplier, π, to enforce the volume element to be constant, $D = 1$. This Lagrange multiplier, which is the pressure, will itself play a central role in the dynamics since it is maintaining an algebraic relationship between the variables. The action, complete with this constraint, is

$$
\int_{t_0}^{t_1}\int_{\mathcal{D}} \frac{D}{2}|\boldsymbol{u}|^2 - \pi(D-1)\, d^n x dt\,. \tag{2.39}
$$

The variational derivatives of the Lagrangian are given by

$$
\frac{\delta \ell}{\delta u} = D u^\flat \otimes d^n x\,, \quad \frac{\delta \ell}{\delta D} = \frac{1}{2}|\boldsymbol{u}|^2 - \pi\,, \quad \frac{\delta \ell}{\delta \pi} = D - 1\,. \tag{2.40}
$$

The left and right hand sides of the Euler-Poincaré momentum equation are given by

$$\begin{aligned}\frac{\partial}{\partial t}\frac{\delta \ell}{\delta u} + \mathcal{L}_u \frac{\delta \ell}{\delta u} &= D\left(\partial_t + \mathcal{L}_u\right)(\boldsymbol{u}\cdot d\boldsymbol{x}) \otimes d^n x\\ \frac{\delta \ell}{\delta D} \diamond D &= D\nabla\left(\frac{1}{2}|\boldsymbol{u}|^2 - \pi\right)\cdot d\boldsymbol{x} \otimes d^n x\,.\end{aligned}$$

Combining these gives the Euler momentum equation for incompressible fluids

$$\partial_t \boldsymbol{u} + \boldsymbol{u}\cdot\nabla\boldsymbol{u} = -\nabla\pi\,, \tag{2.41}$$

where we have used the relationship

$$\mathcal{L}_u(\boldsymbol{u}\cdot d\boldsymbol{x}) = (\boldsymbol{u}\cdot\nabla\boldsymbol{u})\cdot d\boldsymbol{x} + (\nabla\boldsymbol{u})^T\cdot\boldsymbol{u})\cdot d\boldsymbol{x} =: \left(\boldsymbol{u}\cdot\nabla\boldsymbol{u} + \sum_{i=1}^n u_i\nabla u_i\right)\cdot d\boldsymbol{x}\,.$$

The momentum equation is closed by incompressibility, which follows from the advection of the mass density together with the constraint that it is constant, i.e.

$$\left.\begin{aligned}\partial_t D + \operatorname{div}(D\boldsymbol{u}) &= 0\\ D &= 1\end{aligned}\right\} \implies \nabla\cdot\boldsymbol{u} = 0\,.$$

Notice also that the Kelvin-Noether circulation theorem for the Euler equations is

$$\frac{d}{dt}\oint_{c(t)} \boldsymbol{u}\cdot d\boldsymbol{x} = 0\,. \tag{2.42}$$

Stochastic Advection by Lie Transport (SALT). A reason for which stochastic equations of motion are desirable is their ability to represent uncertainty that exists within a model of fluid dynamics, as well as their capacity to model the mean effect of fluctuating phenomena. Uncertainty when modelling can be inherited from inexact observations or, more relevantly to this approach, from 'subgridscale' motions, which occur on a scale smaller than the resolution of a simulation and thus cannot be modelled using a discrete deterministic system. Upon inspection, data for Lagrangian trajectories takes the form of a deterministic mean flow with fluctuations around this. Following Holm [53], we may find stochastic equations of motion for a fluid system by creating a stochastic analogue of the semidirect product Euler-Poincaré framework above. For a discussion of stochastic geometric mechanics for fluids, see Holm and Luesink [60]. This can be considered preferable to adding noise to the system at the level of the equations since it ensures that the system has a Kelvin-Noether circulation theorem, as we will see. Motivated by data, we make the system by stochastically perturbing the velocity field, u, by introducing a collection of vector field ξ_i. This will lead to equations which feature a transport type noise, and we will be assuming from the outset that fluid particles follow stochastic trajectories. We do this by assuming that $\boldsymbol{x}_t$ is a solution of

$$
\begin{aligned}
\mathrm{d}\boldsymbol{x}_t = \mathrm{d}g_t\boldsymbol{x}_0 &= (u_t \circ g_t)\boldsymbol{x}_0\,dt + \sum_{i=1}^{\infty}(\xi_i \circ g_t)\boldsymbol{x}_0 \circ dW_t^i \\
&= u_t(\boldsymbol{x}_t)\,dt + \sum_{i=1}^{\infty}\xi_i(\boldsymbol{x}_t) \circ dW_t^i\,,
\end{aligned}
$$

where $\circ$ denotes Fisk-Stratonovich integration and W_t^i are independent, identically distributed Brownian motions. We can therefore define the Lie derivatives with respect to a stochastic vector field, for which we will use the notation $\mathrm{d}x_t = u\,dt + \sum \xi_t \circ dW_t^i$.

Suppose a_t is invariant under the flow, meaning that it is an advected quantity. Then

$$
a_0(x_0) = a_t(x_t) = (a_t \circ g_t)x_0 = (g_t^* a_t)x_0,
$$

and hence, by an application of the stochastic Kunita-Itô-Wentzell formula, as in Bethencourt de Léon et al. [11],

$$
\begin{aligned}
0 &= \mathrm{d}a_0(\boldsymbol{x}_0) = \mathrm{d}(a_t \circ g_t)\boldsymbol{x}_0 \\
&= \mathrm{d}(g_t^* a_t)\boldsymbol{x}_0 \\
&= g_t^*(\mathrm{d}a_t(\boldsymbol{x}_0) + \mathcal{L}_u a_t(\boldsymbol{x}_0)\, dt + \sum_i \mathcal{L}_{\xi_i} a_t(\boldsymbol{x}_0) \circ dW_t^i) \\
\implies \quad 0 &= \mathrm{d}a_t(\boldsymbol{x}_t) + \mathcal{L}_u a_t(\boldsymbol{x}_t)\, dt + \sum_i \mathcal{L}_{\xi_i} a_t(\boldsymbol{x}_t) \circ dW_t^i \\
&:= \mathrm{d}a_t(\boldsymbol{x}_t) + \mathcal{L}_{\mathrm{d}x_t} a_t \, .
\end{aligned}
\tag{2.43}
$$

We have introduced a new notation in the final line of (2.43) for a more convenient way of writing the Lie derivative terms. Namely, we absorb the notation for temporal integration into the vector field with which we are taking a Lie derivative with respect to

$$
\mathcal{L}_{\mathrm{d}x_t} a_t := \mathcal{L}_u a_t(\boldsymbol{x}_t)\, dt + \sum_i \mathcal{L}_{\xi_i} a_t(\boldsymbol{x}_t) \circ dW_t^i \, .
$$

We will now derive the Euler-Poincaré equations in the stochastic case. We will not be concerned with the reduction by symmetry for the stochastic case, instead only considering reduced Lagrangians. For a discussion on stochastic Euler-Poincaré reductions for the Hamilton-Pontryagin principle, see Takao [96]. The action in this case will be formulated using the Clebsch approach, where the advection of the advected quantities is constrained by a Lagrange multiplier. Note that, formally, the collection of advected quantities must be sufficiently large to ensure that we have a complete Clebsch representation of the velocity field, we will return to this point following the theorem.

Theorem 2.23 (A stochastic Euler-Poincaré theorem)**.** *The following are equivalent, where there are implicit sums over the terms corresponding to* *advected quantities.*

a) The variational principle

$$0 = \delta \int_{t_0}^{t_1} \ell(u, a)\, dt + \langle \lambda, \mathrm{d}a + \mathcal{L}_{\mathrm{d}x_t} a \rangle\,,$$

holds on $\mathfrak{X} \times V^*$, where the collection of advected quantities is of dimension greater than or equal to n, the dimension of the spatial domain.

b) The stochastic Euler-Poincaré equations

$$\mathrm{d}\frac{\delta \ell}{\delta u} + \mathcal{L}_{\mathrm{d}x_t} \frac{\delta \ell}{\delta u} = \frac{\delta \ell}{\delta a} \diamond a\, dt\,, \tag{2.44}$$

hold on $\mathfrak{X} \times V^*$, together with the advection equation

$$\mathrm{d}a + \mathcal{L}_{\mathrm{d}x_t} a = 0\,.$$

Proof. Taking variations of the action integral yields

$$\begin{aligned} \delta u: &\quad \frac{\delta \ell}{\delta u} - \lambda \diamond a = 0\,, \\ \delta a: &\quad \frac{\delta \ell}{\delta a}\, dt - \mathrm{d}\lambda + \mathcal{L}^*_{\mathrm{d}x_t} \lambda = 0\,, \\ \delta \lambda: &\quad \mathrm{d}a + \mathcal{L}_{\mathrm{d}x_t} a = 0\,. \end{aligned}$$

For an arbitrary vector field, $\eta \in \mathfrak{X}$, we have

$$\left\langle \mathrm{d}\frac{\delta \ell}{\delta u} - \frac{\delta \ell}{\delta a} \diamond a\, dt, \eta \right\rangle = \left\langle \mathrm{d}\lambda \diamond a + \lambda \diamond \mathrm{d}a - \frac{\delta \ell}{\delta a} \diamond a\, dt, \eta \right\rangle\,,$$

where we have used the relationship corresponding to the variation in u.

Making use of the remaining relationships then gives

$$\left\langle \mathrm{d}\frac{\delta \ell}{\delta u} - \frac{\delta \ell}{\delta a} \diamond a\, dt, \eta \right\rangle = \left\langle \mathcal{L}^*_{\mathrm{d}x_t} \lambda \diamond a - \lambda \diamond \mathcal{L}_{\mathrm{d}x_t} a, \eta \right\rangle$$
$$= \left\langle -\mathcal{L}_{\mathrm{d}x_t}\mathcal{L}_\eta a + \mathcal{L}_\eta \mathcal{L}_{\mathrm{d}x_t} a, \lambda \right\rangle = -\left\langle \mathcal{L}_{[\mathrm{d}x_t, \eta]} a, \lambda \right\rangle$$
$$= \left\langle \lambda \diamond a, -\mathrm{ad}_{\mathrm{d}x_t} \eta \right\rangle = \left\langle -\mathrm{ad}^*_{\mathrm{d}x_t} \frac{\delta \ell}{\delta u}, \eta \right\rangle ,$$

and, since coajoint action is equivalent to the Lie derivative for vector fields acting on 1-form densities, this completes our proof. $\mathscr{Q.E.D.}$

Remark 2.24. As noted previously, the mass density is a crucial advected quantity in fluid dynamics, As we can see from the variation in the velocity field, u, in the above proof, if we include only one advected quantity, then we obtain only one term featuring the diamond operator in the variational derivative of the Lagrangian with respect to the velocity field. This would reduce the model to a potential flow, and we would have an incomplete Clebsch representation. We need at least n scalar Lagrange multipliers enforcing the Clebsch constraints for our variational principle to encapsulate all possible solutions to the resulting momentum equation. A more in depth discussion of this can be found in Cendra and Marsden [20]. It has been noted that a Clebsch variational principle can describe fluid motion fully by using Lagrange multipliers to enforce the advection of the Lagrangian labels, following the approach of Holm and Kupershmidt [59]. For a more contemporary discussion of this issue which, as in this thesis, illustrates this principle using the Euler equations, see Cotter and Holm [28].

Let it be noted that, whilst additional advected quantities are needed for the variational principle to formally represent the full collection of possible solutions to the model, the Euler-Poincaré equation itself only features advected quantities which appear in the Lagrangian, $\ell(u, a)$. Thus, when working with the Euler-Poincaré equations in practice, we need only consider the advected quantities present in the Lagrangian to write down the equation of motion.

In the deterministic case, we used an alternative approach where the form of the variations was externally determined. The resulting constraints on the form of permissible variations are known as Lin constraints. This approach can also be used in the stochastic case, as has been done in multiple other studies [11, 60, 72]. For the stochastic case, the Euler-Poincaré equations are identical to those found in Theorem 2.23, the action, $\int \ell(u, a)dt$, is analogous to the deterministic action, and the variations are constrained to be of the form

$$\begin{aligned}\delta u\, dt &= \mathrm{d}v + [\mathrm{d}x_t, v]\,,\\ \delta a &= -\mathcal{L}_v a\,,\end{aligned}$$

where $v \in \mathfrak{X}$ is arbitrary and vanishes at the endpoints. See Holm, Marsden, and Ratiu [62] for details.

In the upcoming section, we will discuss the mathematics of variational principles of this type in more detail, as well as properly formulate how such a stochastic model can be constrained to be incompressible.

2.2 Semimartingale driven variational principles

In order to derive the Euler-Poincaré equation for the case of stochastic advection by Lie transport (SALT), we considered a stochastic version of a variational principle. The addition of stochasticity into the process of applying Hamilton's principle to some action has attracted attention elsewhere in the literature also, and there are numerous ways of achieving this. For example, one may take the action to be the expected value of an integral [2] or, as in the above section, the action can be defined as a stochastic integral. This second approach has attracted interest from a number of research groups (see e.g. [15, 22, 53, 103]), and is the methodology we will be considering in this section. This typically involves action integrals involving stochastic constraints or, equivalently, one for which variations are assumed to have a particular stochastic form. Here, we will properly formulate this and prove that it is permitted to take variational derivatives of such actions in the standard way. We will be working within the framework of semimartingale driven variational principles, first introduced by Crisan and Street [95], with a particular focus on reduced variational principles with Lagrangians which take the form of a spatial integral, as found in continuum dynamics. As such, this section will closely follow the form of this paper [95] and use much of the same notation. For a discussion of the unreduced Euler-Lagrange equations corresponding a semimartingale driven variational principle, see Takao [96].

The purpose of this section is to introduce a rigorous theoretical framework for adding stochasticity into a system through Hamilton's principle, using concepts from stochastic analysis. Lagrange multipliers, which must correspond to the stochastic noise chosen, may be used to constrain a physical model to behave in a certain way as observed in the physical system of choice. The action function appearing in Hamilton's principle can be defined as an integral with respect to a given measure which in the stochastic case can be chosen to be random. The framework introduced below clarifies how to make these concepts stochastic in a consistent manner, i.e., the dynamical variables, the Lagrange multipliers, as well as the integrator measure can all be chosen to be random as long as they remain compatible with an exogenously chosen semimartingale. Subsequently, this will define what we mean when we say that the variational principle is *semimartingale driven*.

As a consequence, we will show that, in the case of stochastic fluid dynamics, pressure must be thought of as a *stochastic* Lagrange multiplier which has the role of ensuring that the volume element remains constant and thus the flow is incompressible. This will work in a manner analogous to the example of the incompressible Euler equations found in Section 2.1.2. We will demonstrate the case of the stochastic Lagrange multiplier by again considering the Euler equations for incompressible fluids, and deduce an explicit stochastic differential equation for the pressure.

The introduction of stochasticity presented here generalises the approach taken by Holm [53]. After the initial publication of this framework, the driving semimartingale was replaced by a rough path [32], leading to the introduction of a new class of rough path driven variational principles.

The form of the action. We can consider the action to be an integral of a space-time domain, which we call $\bar{\mathcal{D}}$ with elements $\bar{\boldsymbol{x}} \in \bar{\mathcal{D}}$. The action then takes the form

$$\int_{\bar{\mathcal{D}}} l(v(\bar{\boldsymbol{x}}))\, d\bar{\mu}(\bar{\boldsymbol{x}})\,, \tag{2.45}$$

where $\bar{\mu}(\bar{\boldsymbol{x}})$ is a measure on the given domain $\bar{\mathcal{D}}$ and v encompasses the physical variables as well as the Lagrange multipliers used to constrain them. Outside of the examples and particular cases, to preserve generality we will not be discussing the specific space that elements of v are in. We can unravel this action into a more recognisable form in the following way. Typically one separates $\bar{\mathcal{D}}$ into a time component and a space component and considers it to be a product space. This space is most commonly chosen to be of the form $[t_0, t_1] \times \mathcal{D}$ for $0 \leqslant t_0 < t_2 \leqslant \infty$, where $\mathcal{D}$ is the spatial manifold in which our dynamics occurs. Similarly, we unpack $\bar{\mu}$ into a product between a measure μ on $\mathcal{D}$ and the Lebesgue measure, dt, for the time variable. The action integral can then be identified with the familiar form found in the case of continuum dynamics, where the Lagrangian, ℓ, is the spatial integral of some Lagrangian 'density', l, as follows

$$\ell(v) = \int_{\mathcal{D}} l(v)\, \mu(dx)\,,$$

and the action is

$$\int_{t_0}^{t_1} \ell(v)\, dt\,. \tag{2.46}$$

Note that we have used the notation $\mu(dx)$ to denote integration in the spatial variable $\boldsymbol{x} \in \mathcal{D}$ with respect to the measure μ. The alternative notation, $d\mu(\boldsymbol{x})$, where the 'd' is positioned outside of the measure, will be reserved for the temporal integration. The reason for this will become clear, where it will be particularly necessary to distinguish between spatial and temporal integration.

The action can be made stochastic through the time integral. This has the effect of altering how the system evolves in time, changing from deterministic integration to stochastic, without changing the definition of space. Regardless of the source of stochasticity, the action function will be assumed to be *compatible* with a given semimartingale S, see Definition 2.25. The form of the action will become

$$\int_{t_0}^{t_1} \ell(v) \circ dS_t\,, \tag{2.47}$$

where both $[t_0, t_1] \ni t \mapsto v$ and $[t_0, t_1] \ni t \mapsto \ell(v)$ will be assumed to be semimartingales. This choice of the action integral (2.47) is justified for the following reasons.

- The action defined by (2.47) is a natural generalisation of (2.46) in that, by choosing the semimartingale to be given by $S_t \equiv t, \;\; t \in [t_0, t_1]$, the stochastic case reduces to the classical deterministic case.

- The stochastic calculus rules governing the Stratonovitch integral, as opposed to the Itô integral, coincide with the classical rules of deterministic calculus. We may therefore expect that the technical details introduced through the stochasticity will be natural extensions of their deterministic counterparts.

- This generalisation is a natural extension of the approach introduced by Holm [53], since the actions needed for deriving equations of this type are special cases of the action integral (2.47), where the driving semimartingale is given by $S_t = (t, W_t^1, ..., W_t^n, ...)$. In other words, S_t is an infinite dimensional stochastic process with the first component identically equal with the time variable and the rest of the components being given by independent Brownian motions. In particular, it incorporates models where the advected quantities are constrained to follow stochastically perturbed trajectories.

- By a judicious choice of the driving semimartingale, one can introduce non-independent noise increments (for example through an Ornstein-Uhlenbeck process, as in [54]) in order to incorporate effects, such as memory, into the fluid dynamic model.

- The new framework lends itself easily to extensions to manifolds, where the Itô based stochastic calculus does not have an intrinsic development, see Hsu [64].

- It is a natural precursor of a new class of rough path driven variational principles, see Crisan et al. [32].

We will now introduce the mathematical concepts necessary to understand stochastic action integrals of this form, from the perspective of the mathematics of stochastic processes.

Let $(\Omega, \mathcal{F}, P)$ be a probability space endowed with a filtration $\{\mathcal{F}_t\}_{t\geqslant 0}$ that satisfies the usual conditions. For an arbitrary Banach space, B, a $\{\mathcal{F}_t\}_{t\geqslant 0}$-adapted stochastic process, X, is said to be B-valued if it is a family of random variables $(X_t)_{t\geqslant 0}$, where $X_t : \Omega \mapsto B$, parameterised by $t \geqslant 0$ such that for all $t \geqslant 0$, X_t is $\mathcal{F}_t$ measurable. The standard notions of probability theory (integration, conditional expectation, etc.) are easily extended to B-valued random variables and processes. Similarly, all notions of stochastic calculus are extended from finite dimensional Euclidean spaces to B-valued stochastic processes. In order to generalise our action integral we will consider function valued stochastic processes, which can be thought of as a specific class of B-valued stochastic processes, and measure valued processes.

We introduce a suitably chosen, but here left arbitrary, space of functions, $\mathcal{F}(\mathcal{D})$, over our spatial domain. This space is a Banach space when equipped with a norm, $\|\cdot\|_{\mathcal{F}(\mathcal{D})}$. The space of n-dimensional versions of these functions is denoted by $\mathcal{F}(\mathcal{D})^n$ where, by this notation, we mean that the functions have n components rather than referring to the dimension of their domain. The 'driving' semimartingale, from which all others will inherit their stochasticity, will be taken to be $\mathbb{R}^{\mathbb{N}}$-valued, meaning that it has an unspecified number of real components which may be infinite, as indeed is the case with stochastic advection by Lie transport.

Definition 2.25. An $\mathcal{F}(\mathcal{D})^n$-valued semimartingale, g_t, is said to be *compat-*

ible with respect to a (continuous) $\mathbb{R}^{\mathbb{N}}$-valued semimartingale, $S_t = \{S_t^j,\ j \geqslant 1\}$, if there exists a set of $\mathcal{F}(\mathcal{D})$-valued continuous semimartingales $G_t = \{G_t^{i,j} : i = 1, \ldots, n,\ j = 1, 2, \ldots\}$ such that

$$g_t^i = g_0^i + \sum_j \int_0^t G_s^{i,j} \circ dS_s^j, \quad i = 1, \ldots, n, \tag{2.48}$$

and, in the case where the sums in (2.48) are infinite sums, the semimartingales $G^{i,j}$ are such that the sums converge.

Remark 2.26. The system of identities (2.48) is written componentwise in integral form, and can be compactly re-written in differential form as

$$\mathrm{d}g_t = G_t \circ dS_t, \tag{2.49}$$

where the above equation encompasses all the relevant summation and all components. Note that it is not necessary that the continuous semimartingales $G_t^{i,j}$ are also compatible with respect to S_t. We can think of G_t as a stochastic generalisation of the derivative of g_t, where the time evolution of the process g_t is defined by a stochastic integral with respect to S_t, rather than a Lebesgue integral.

The notation found in (2.48) and (2.49), where upper cases and lower cases are used, will consistently be used for objects which are compatible with a semimartingale S_t.

Remark 2.27. In the case where the sums in (2.49) are infinite sums, we need to impose constraints on the choice of the semimartingales $G^{i,j}$, $j = 1, 2, \ldots$ $i = 1, \ldots, n$ to ensure that the sums make sense. Let us identify the finite variation parts and the martingale parts of S^j $j = 1, 2, \ldots,$ through the

Doob-Meyer decomposition of each component, as

$$S^j = B^j + M^j \,, \tag{2.50}$$

where B^j and M^j are the finite variation and martingale parts of S^j respectively. We will assume that $G^{i,j}$ will be integrable with respect to B^j for all $j \geqslant 1$ and that

$$\mathbb{E}\left[\left(\sum_{i=1}^{\infty}\int_0^t \|G^{i,j}\|_{\mathcal{F}(\mathscr{D})}\, dV_{B^j}\right)^2\right] < \infty \,, \tag{2.51}$$

where V_{B^j} is the variation process corresponding to B^j, and $\mathbb{E}$ denotes the expectation. Separately we will assume that

$$\mathbb{E}\left[\sum_{i=1}^{\infty}\int_0^t \|G_s^{i,j}\|^2_{\mathcal{F}(\mathcal{D})}\, d[M^j]_s\right] < \infty \,, \tag{2.52}$$

where $[M^j]$ is the quadratic variation of the martingale M^j. Considering the equations (2.51) and (2.52), together with the assumption that $\|g_0^i\|_{\mathcal{F}(\mathcal{D})} < \infty$, we have that the semimartingales g_s^i are well defined and are square integrable. More precisely

$$\mathbb{E}\left[\sup_{s\in[0,t]} \|g^i\|_{\mathcal{F}(\mathscr{D})}\right] < \infty \,.$$

We will henceforth refer to conditions such as these as 'integrability constraints' on $G^{i,j}$, since we will need similar conditions on other objects.

Definition 2.28. The process S_t in Definition 2.25, which we assume to be continuous, will be called the *driving semimartingale* of the system.

In the following, when formulating the stochastic action integral, we will assume, without loss of generality, that all the finite variation terms of the semimartingale S_t are collected into one of its components, with all other components consisting of (continuous) martingales.

To relate these ideas to the stochastic action integral, we first notice that, for g_t as defined in Definition 2.25, the process

$$t \to \int_0^t g_t^j \circ dS_t^i \,,$$

is a well defined square-integrable one-dimensional semimartingale, for each i and j. Here we have assumed that g_t^j each satisfy similar integrability conditions to those in Remark 2.27. The stochastic action takes a similar form. For a driving semimartingale S_t, we assume that $v(t)$ is compatible with S_t. The Lagrangian, ℓ, is then assumed to be a $\mathbb{R}^{\mathbb{N}}$-valued semimartingale such that the process $t \mapsto \ell(v)$ is compatible with S_t, then

$$\int_{t_0}^{t_1} \ell(v) \circ dS_t, \tag{2.53}$$

is a well defined square integrable one-dimensional semimartingale, where ℓ is assumed to satisfy the relevant integrability conditions.

Definition 2.29. By a *semimartingale driven variational principle*, we mean the application of the principle of stationary action to a well defined stochastic action integral of the form (2.53) in order to derive the corresponding stochastic governing equation for the chosen physical system.

In the case of continuum dynamics, where the Lagrangian takes the form of a spatial integral, the above formulation of the stochastic integral is adequate for the case where the Lagrangian takes the form of a deterministic spatial integral and the stochasticity appears in the time integration. This case can, however, be thought of as a particular case of the more general framework in which the *measure* on the space-time domain is itself taken to be a stochastic process. We will define here the class of measures which are suitable for our requirements, for a thorough summary of the theory of random measures, see Kallenberg [65]. Using a notation analogous to the function valued semimartingales, let $\mathcal{M}(\mathcal{D})$ be the space of finite measures over $\mathcal{D}$ endowed with the total variation norm and $\mathcal{M}(\mathcal{D})^n$ denotes the space of n-dimensional versions of such measures. Note that this methodology can be extended to spaces of non-finite measures.

Definition 2.30. An $\mathcal{M}(\mathcal{D})^n$-valued semimartingale ν_t, is called compatible with respect to a given $\mathbb{R}^{\mathbb{N}}$-valued semimartingale, S_t, if $\nu_t = (\nu_t^i)_{i=1}^n$ has a representation of the form

$$\nu_t^i = \nu_0^i + \sum_j \int_0^t \mu_s^{i,j} \circ dS_s^j \,, \quad i = 1, \ldots, n, \tag{2.54}$$

where $\mu_t^{i,j}$ are continuous $\mathcal{M}(\mathcal{D})$-valued semimartingales for every $i = 1, \ldots, n$, $j = 1, 2, \ldots$. As in Definition 2.25, the system of equations (2.54) can be re-written in the following compact form:

$$\mathrm{d}\nu_t = \mu_s \circ dS_s \,. \tag{2.55}$$

Similar integrability constraints to those imposed on $G^{i,j}$ in Remark 2.27 are needed here to ensure that the ν_i are well defined.

Now suppose that g_t and ν_t are function and measure valued semimartingales respectively, which are compatible with the driving semimartingale S_t and are as in Definitions 2.25 and 2.30. Then

$$\int_{\mathcal{D}} g_t^i(x)\nu_t^i(dx) = \int_{\mathcal{D}} g_0^i(x)\nu_0^i(dx) + \sum_j \int_0^t \int_{\mathcal{D}} G_t^{i,j}\nu_t^i(dx) \circ dS_t^j$$
$$+ \sum_j \int_0^t \int_{\mathcal{D}} g_s^i \mu_s^{i,j}(dx) \circ dS_t^j \,,$$

where we have assumed that the relevant integrability conditions are satisfied. We may interpret this as a stochastic version of the product rule for differentiation $\mathrm{d}(g_t^i\nu_t^i) = (\mathrm{d}g_t^i)\nu_t^i + g_t^i\mathrm{d}(\nu_t^i)$. This implies that the processes

$$t \to \int_0^t \int_{\mathcal{D}} g_t(x)^i \circ d\nu_t^i(dx)\,, \tag{2.56}$$

and

$$t \to \int_0^t \int_{\mathcal{D}} g_t(x)^i \mu_t^{i,j}(dx) \circ dS_t^j\,, \tag{2.57}$$

are equivalent and are well defined one dimensional semimartingales. When reading (2.56), recall that spatial integration is represented by writing 'dx' as the argument of the measure ν. The integration in the time variable is achieved by integrating with respect to ν_t as a stochastic process. We thus have defined a stochastic version of the action integral,

$$\int_{t_0}^{t_1} \int_{\mathcal{D}} l(v) \circ d\nu_t(dx) = \int_{t_0}^{t_1} \int_{\mathcal{D}} l(v)\, \mu_t(dx) \circ dS_t\,. \tag{2.58}$$

Comparison between deterministic and stochastic cases. The semimartingale driven action principle defined above is a true generalisation of the classical deterministic case since, when the driving semimartingale is taken to be $S_t = t$, the system reverts back to the standard deterministic theory. Although it should be noted that the addition of stochasticity introduces subtleties and complexities into the problem which will become evident when looking at examples and Lagrange multipliers. For a differentiable function of time, $v(t)$, the object $\mathrm{d}v$ is a stochastic generalisation of the object $\dot{v}\,dt$, where $\dot{v}$ is the time derivative of v. This can be thought of as a variable together with the measure with respect to which the time integral is defined, and is an object which can be integrated. In the deterministic case, this is an uninteresting and trivial object, since the integration is standard Lebesgue integration and the equations are most commonly written in a differential form which negates the need for this notation. In the stochastic framework, the variables are not necessarily differentiable in time, thus it is essential that our equations are interpreted in integral form. The framework of semimartingale driven variational principles enables the clear understanding of how the time integration is defined, which permeates each part of the model.

In the case where the driving semimartingale is $S_t = t$, the definition of compatibility has an interesting deterministic analogue. In particular, it reduces to

$$v_t = v_0 + \int_0^t V_s\,ds\,,$$

where v is some function and V is continuous. Thus, in this case, it is apparent that we have $\partial v/\partial t = V$. The compatibility of the function valued process v with respect to t is therefore equivalent to the requirement that the function is differentiable in time and that the derivative is continuous. The compatibility of a process with the driving semimartingale can therefore be thought of as a stochastic generalisation of this concept.

THE VALIDITY OF STOCHASTIC VARIATIONAL CALCULUS. Recall that Hamilton's principle enables us to find equations from an action by taking variations and considering the relations that arise from the condition that the first variation of the action is zero. Recall that, in this section, we are considering the addition of stochasticity into this process by taking the action integral to be a stochastic integral, rather than any other methodology. We will present here a stochastic version of the so-called 'fundamental lemma of the calculus of variations', designed for stochastic variational methods corresponding to our approach. If the action is semimartingale driven and takes the same form as that in equation (2.58), then Hamilton's principle implies

$$
\begin{aligned}
0 = \delta \int_{t_0}^{t_1} \ell(v) \circ dS_t &= \delta \int_{t_0}^{t_1} \int_{\mathcal{D}} l(v)\, \mu_t(dx) \circ dS_t \\
&= \int_{t_0}^{t_1} \int_{\mathcal{D}} \frac{\delta \ell}{\delta v} \delta v\, \mu_t(dx) \circ dS_t \,.
\end{aligned}
$$

In order to derive equations from this concept, we need to prove that the fundamental lemma of the calculus of variations is valid in this framework. This lemma is the following.

Lemma 2.31 (A fundamental lemma of stochastic calculus of variations [95])**.** *Suppose $f(t,x)$ is a $\mathcal{F}(\mathcal{D})^{\mathbb{N}}$-valued semimartingale. If for any $\mathcal{F}(\mathcal{D})^{\mathbb{N}}$-valued semimartingale, $\psi(t,x)$, we have*

$$
\int_{t_0}^{t_1} \int_{\mathcal{D}} f(t,x)\psi(t,x)\, \mu(dx) \circ dS_t = 0 \,, \tag{2.59}
$$

then, for any α, β such that $t_0 \leqslant \alpha < \beta \leqslant t_1$, we have

$$
\int_{\alpha}^{\beta} f(t,x) \circ dS_t = 0 \,, \tag{2.60}
$$

μ-almost everywhere on the set $\mathcal{D}$.

Remark 2.32. When applying this in practice, the equations (2.33) which result from Hamilton's principle can be stochastic partial differential equations, or algebraic relationships between variables. This will be evident from the examples.

Proof. The proof of this is as in Crisan and Street [95]. Since (2.59) holds for any such ψ, it holds for semimartingales of the form

$$\psi(t,x) = (\psi^0(t,x), \psi^1(t,x), \dots) = (\phi(x)\varphi^0(t), \phi(x)\varphi^1(t), \dots) = \phi(x)\varphi(t)\,,$$

where $\varphi : [t_0, t_1] \to \mathbb{R}^{\mathbb{N}}$ is a smooth function such that $\varphi(t_0) = \varphi(t_1) = 0$ and $\phi : \mathcal{D} \to \mathbb{R}$ belongs to a class of functions, $\mathscr{S}$, which represents a total set. By this we mean that the class of functions, $\mathscr{S}$, is such that if, for some function $g : \mathcal{D} \to \mathbb{R}$, we have

$$\int_{\mathcal{D}} g(x)\phi(x)\,\mu(dx)\,, \quad \forall\,\phi \in \mathscr{S}\,,$$

then μ-almost surely we have that $g = 0$.

We now define $F_\phi : [t_0, t_1] \to \mathbb{R}^{\mathbb{N}}$ by

$$F_\phi(t) = \int_{\mathcal{D}} f(t,x)\phi(x)\,\mu(dx)\,.$$

Then F_ϕ is a semimartingale and, moreover, the covariation process $[F_\phi, S]_t$ is well defined. In particular,

$$[F_\phi, S]_t = \int_{\mathcal{D}} [f, S]_t(x)\phi(x)\,\mu(dx)\,.$$

Since φ is smooth, $F_\phi\varphi$ is also a semimartingale and $[F_\phi\varphi, S]_t$ is well defined with

$$[F_\phi\varphi, S]_t = \int_{t_0}^{t_1} \varphi(t)\,d[F_\phi, S]_t\,.$$

It follows that

$$\begin{aligned}0 &= \int_{t_0}^{t_1} \int_{\mathcal{D}} f(t,x)\psi(t,x)\,\mu(dx) \circ dS_t = \int_{t_0}^{t_1} F_\phi(t)\varphi(t) \circ dS_t \\ &= \int_{t_0}^{t_1} F_\phi(t)\varphi(t)\,dS_t + \frac{1}{2}\int_{t_0}^{t_1} F_\phi(t)\varphi(t)\,d[F_\phi, S]_t\,.\end{aligned}$$

For arbitrary α and β such that $t_0 \leqslant \alpha < \beta \leqslant t_1$, choose $\varphi = \mathbb{1}_{[\alpha,\beta]}$ where $\mathbb{1}$ is the indicator function. Let $(\varphi_n)_{n=1}^{\infty}$ be a uniformly bounded sequence of smooth functions such that

$$\varphi_n \to \mathbb{1}_{[\alpha,\beta]}\,,$$

where this convergence is pointwise. Then, by the Itô isometry and the bounded convergence theorem, we have

$$\begin{aligned}&\mathbb{E}\left[\left(\int_{t_0}^{t_1} F_\phi(t)(\varphi_n(t) - \mathbb{1}_{[\alpha,\beta]}(t))dS_t\right)^2\right] \\ &\qquad = \mathbb{E}\left[\int_{t_0}^{t_1} \left(F_\phi(t)(\varphi_n(t) - \mathbb{1}_{[\alpha,\beta]}(t))\right)^2 d[S]_t\right] \to 0\,,\end{aligned}$$

and

$$\int_{t_0}^{t_1} |F_\phi(t)||\varphi_n(t) - \mathbb{1}_{[\alpha,\beta]}(t)|\,d[F_\phi, S]_t \to 0\,.$$

Therefore, we have

$$\begin{aligned}0 &= \int_{t_0}^{t_1} F_\phi(t)\varphi_n(t)\,dS_t + \frac{1}{2}\int_{t_0}^{t_1} F_\phi(t)\varphi_n(t)\,d[F_\phi, S]_t \\ &\to \int_{t_0}^{t_1} F_\phi(t)\mathbb{1}_{[\alpha,\beta]}(t)\,dS_t + \frac{1}{2}\int_{t_0}^{t_1} F_\phi(t)\mathbb{1}_{[\alpha,\beta]}(t)\,d[F_\phi, S]_t = \int_{\alpha}^{\beta} F_\phi(t) \circ dS_t\,.\end{aligned}$$

The stochastic Fubini theorem then gives

$$\begin{aligned}0 &= \int_\alpha^\beta F_\phi(t) \circ dS_t = \int_\alpha^\beta \int_{\mathcal{D}} f(t,x)\phi(x)\,\mu(dx) \circ dS_t \\ &= \int_{\mathcal{D}} \int_\alpha^\beta f(t,x)\phi(x) \circ dS_t\,\mu(dx) = \int_{\mathcal{D}} \left(\int_\alpha^\beta f(t,x) \circ dS_t \right) \phi(x)\,\mu(dx)\,,\end{aligned}$$

and the total set property then gives our result. $\mathscr{Q}.\mathscr{E}.\mathscr{D}.$

Remark 2.33. Should we wish to consider equation componentwise, then we will need to place additional assumptions on the driving semimartingale S_t. If we were to do this, from Hamilton's principle we could obtain relations $\delta\ell^i/\delta v = 0$, $\forall i$. Whilst this may feel more analogous to the deterministic case, in our examples we will see that, in practice, the components of ℓ may have little to no physical meaning.

Lagrange multipliers. In the stochastic case, care must be taken when formulating constraints imposed by a Lagrange multiplier. Suppose we have some variable from our collection of physical variables, $c \in \{v_i\}$, by the compatibility of v with the driving semimartingale we know that $\mathrm{d}c = C \circ dS_t = \sum_i C_i \circ dS_t^i$, for some continuous C. Suppose we wish to *impose* the form of C as a constraint, then a Lagrange multiplier may be done to achieve this by writing the action in the form

$$\int_{t_0}^{t_1} \int_{\mathcal{D}} \big(l(v) \circ dS_t + \lambda(\mathrm{d}c - C \circ dS_t) \big) \mu(dx)\,,$$

where each component of C is not necessarily nonzero.

Suppose that, rather than a dynamical constraint, we wish to impose a relationship between our variables which does not feature time integration and is true for each time t. We represent this by $f(v) = 0$, for some function f.

This constraint can be imposed by a Lagrange multiplier in a manner which is compatible with the driving semimartingale, $\mathrm{d}\lambda = \Lambda \circ dS_t$. The action then takes the form

$$\int_{t_0}^{t_1} \int_{\mathcal{D}} \big(l(v) \circ dS_t + (\mathrm{d}\lambda) f(v)\big) \mu(dx) = \int_{t_0}^{t_1} \int_{\mathcal{D}} \big(l(v) + \Lambda f(v)\big) \mu(dx) \circ dS_t \,.$$

Crucially, when imposing such constraints, integration must be performed with respect to the entire driving semimartingale, as we will illustrate with the pressure term in incompressible stochastic fluid dynamics.

Remark 2.34. We must be careful with semantics in the stochastic case. Each component of the object Λ is a Lagrange multiplier, even though it may be tempting to refer to λ or even $\mathrm{d}\lambda$ as such. Furthermore, when we say that a Lagrange multiplier is compatible with a semimartingale S_t, we mean that λ is compatible with S_t. This is despite the fact that, formally, the Lagrange multipliers Λ_i may not be compatible with S_t.

2.2.1 Incompressible stochastic fluid dynamics

In Section 2.1.2, we illustrated the deterministic case with the example of the incompressible Euler equations, and introduced the stochastic Euler-Poincaré theorem (Theorem 2.23) corresponding to the approach of modelling continuum dynamics using stochastic advection by Lie transport. In the example of the Euler equations, a Lagrange multiplier is used to enforce that the volume element is constant. The framework of semimartingale driven variational principles has highlighted that the incompressibility constraint cannot, in the stochastic case, be thought of as part of the deterministic Lagrangian and must instead be formulated as a *stochastic* constraint. If we consider fluid dynamics in the framework of stochastic advection by Lie transport, the volume element, D, satisfies the following equation

$$\mathrm{d}D + \nabla \cdot (D\boldsymbol{u}\,dt + \sum_i D\boldsymbol{\xi}_i \circ dW_t^i) = 0\,, \tag{2.61}$$

and this, together with the constraint $D = 1$, implies

$$\int_0^t \nabla \cdot \boldsymbol{u}_s\,ds + \sum_i \int_0^t \nabla \cdot \boldsymbol{\xi}_i \circ dW_t^i = 0\,. \tag{2.62}$$

The uniqueness of the Doob-Meyer decomposition then implies that $\nabla\cdot\boldsymbol{\xi}_i = 0$ for each i and $\nabla \cdot \boldsymbol{u} = 0$. This incompressibility constraint, $D = 1$, can be imposed into the stochastic Euler-Poincaré theorem by using a Lagrange multiplier, as follows.

For the following theorem we maintain the notation which we have been using throughout this paper, in particular a_t denotes the set of all advected quantities of which the volume element, D, is one.

Theorem 2.35 (A stochastic Euler-Poincaré theorem for incompressible dynamics)**.** *With notation as in Theorem 2.23, we let a denote the collection of advected variables, to which D belongs, and we use the* pressure *as a Lagrange multiplier to enforce $D = 1$. The following are equivalent*

a) The variational principle

$$\begin{aligned}0 &= \delta \int_{t_0}^{t_1} \ell(u,a)\,dt - \langle \mathrm{d}\pi, D-1\rangle - \langle \Lambda, \mathrm{d}a + \mathcal{L}_{\mathrm{d}x_t} a\rangle \\ &= \delta \int_{t_0}^{t_1} \ell(u,a)\,dt - \langle \mathrm{d}\pi, D-1\rangle - \langle \Lambda, \mathrm{d}a + \mathcal{L}_u a\,dt\rangle \\ &\qquad\qquad + \sum_i \langle \Lambda \diamond a, \xi_i(x)\rangle \circ dW_t^i\,,\end{aligned}$$

holds on $\mathfrak{X} \times V^*$.

b) The stochastic incompressible Euler-Poincaré equations

$$
\begin{aligned}
\mathrm{d}\frac{\delta \ell}{\delta u} + \mathcal{L}_{\mathrm{d}x_t}\frac{\delta \ell}{\delta u} &= \frac{\delta \ell}{\delta a} \diamond a\, dt - \mathrm{d}\pi \diamond D\,, \\
\mathrm{d}a + \mathcal{L}_{\mathrm{d}x_t} a &= 0\,, \\
D &= 1\,,
\end{aligned}
\tag{2.63}
$$

hold on $\mathfrak{X} \times V^*$.

The proof of this is analogous to that of Theorem 2.23.

Remark 2.36. The stochastic Lagrange multiplier, which defines the pressure, has the structure, $\mathrm{d}\pi = P \circ dS_t$, which reflects its compatibility with the driving semimartingale. The advected quantities will use the capitalisation notation to denote compatibility, as in Definition 2.25.

The action integral corresponding to the Theorem 2.35 is of the form (2.47) with $S_t = (t, W_t^1, ..., W_t^n, ...)$ and

$$
\ell = (\ell(u, a_t) - P_0(D-1) - \Lambda A_t^0 - \Lambda \mathcal{L}_u a_t, -P_i(D-1) - \Lambda A_t^i - \Lambda \mathcal{L}_{\xi_i} a_t),
$$

where $da = A \circ dS_t$ and ℓ has infinitely many components represented by indexing them by i.

Remark 2.37. Where the action integral is written as above, where the integrand contains a Lagrange multiplier of the form $\mathrm{d}\pi$, then it is technically improper to refer to a variational derivative with respect to π. Instead we perform variational derivatives with respect to P, as is consistent with Lemma 2.31. Nonetheless, when considering the pressure terms of this form, to ease notation we will refer to variations with respect to π rather than P, and we here define these variations to be equivalent.

Example (Stochastic Euler equations for incompressible flow)**.** As discussed in the example of the deterministic Euler equations for incompressible flow, the Lagrangian corresponding to the structure discovered by Arnold [4] takes the form

$$\ell(u, D) = \int_{\mathcal{D}} \frac{D}{2} |\boldsymbol{u}|^2 \, d^n x \, .$$

Notice that the diamond terms from the Euler-Poincaré equation for incompressible dynamics take the form

$$\begin{aligned} \frac{\delta \ell}{\delta a} \diamond a \, dt &= D \nabla \left(\frac{\delta \ell}{\delta D} \right) \cdot d\boldsymbol{x} \otimes d^n x \, dt \, , \\ {\rm d}\pi \diamond D &= D \nabla ({\rm d}\pi) \cdot d\boldsymbol{x} \otimes d^n x \, , \end{aligned}$$

and the variational derivatives of the action are, similarly to the deterministic case, given by

$$\frac{\delta \ell}{\delta u} = D u^\flat \otimes d^n x \, , \qquad \frac{\delta \ell}{\delta D} = \frac{1}{2} |\boldsymbol{u}|^2 \, .$$

Note that, by comparing this to the variational derivatives of the Lagrangian in the deterministic example (2.40), we see a fundamental difference in how the pressure is treated. In particular, the incompressibility constraint is considered as a part of the Lagrangian, ℓ, in the deterministic Euler-Poincaré framework. In the stochastic case, this cannot happen since the Lagrange multiplier needs to be compatible with the entire driving semimartingale and thus cannot be defined under the deterministic Lebesgue time integral only, as is the case with the remainder of the physics encoded within the Lagrangian, ℓ, for a model featuring stochastic advection by Lie transport.

The left hand side of the Euler-Poincaré momentum equation is, for the Euler

Lagrangian,

$$\mathrm{d}(Du^\flat \otimes d^n x) + \mathcal{L}_{\mathrm{d}x_t}(Du^\flat \otimes d^n x) = D\big(\mathrm{d}u^\flat + \mathcal{L}_{\mathrm{d}x_t} u^\flat\big) \otimes d^n x\,,$$

where this is true since $Dd^n x$ is an advected variable. This implies that, after setting $D = 1$, the Euler-Poincaré momentum equation is

$$\mathrm{d}\boldsymbol{u} + (\mathrm{d}\boldsymbol{x}_t \cdot \nabla)\boldsymbol{u} + (\nabla \mathrm{d}\boldsymbol{x}_t) \cdot \boldsymbol{u} = \frac{1}{2}\nabla|\boldsymbol{u}|^2\, dt - \nabla \mathrm{d}\pi\,.$$

Notice now that the deterministic part of $(\nabla \mathrm{d}\boldsymbol{x}_t) \cdot \boldsymbol{u}$ cancels with $\frac{1}{2}\nabla|\boldsymbol{u}|^2\, dt$ to give

$$\mathrm{d}\boldsymbol{u} + (\mathrm{d}\boldsymbol{x}_t \cdot \nabla)\boldsymbol{u} + \sum_{i=1}^{\infty}(\nabla \boldsymbol{\xi}_i) \cdot \boldsymbol{u} \circ dW_t^i = -\nabla \mathrm{d}\pi\,.$$

To clarify this equation we notice that, $\nabla \boldsymbol{v}$ is a second order tensor for any vector $\boldsymbol{v}$, and thus by $(\nabla \boldsymbol{v}) \cdot \boldsymbol{u}$ we mean $\sum_j u_j \nabla v_j$. The stochastic Euler equations for incompressible fluids with stochastic advection by Lie transport are therefore given by

$$\mathrm{d}\boldsymbol{u} + \boldsymbol{u} \cdot \nabla \boldsymbol{u}\, dt + \sum_{i=1}^{\infty}\left(\boldsymbol{\xi}_i \cdot \nabla \boldsymbol{u} + \sum_{j=1}^{n} u_j \nabla \xi_i^j\right) \circ dW_t^i + \nabla \mathrm{d}\pi = 0\,, \tag{2.64}$$

$$\nabla \cdot \mathrm{d}\boldsymbol{x}_t = 0\,. \tag{2.65}$$

To illustrate the necessity for the pressure term to have this stochastic structure, and the need for the framework of semimartingale driven variational principles to see that the pressure takes this form, we will calculate explicitly the equation satisfied by the pressure. Physically, the pressure ensures that the fluid remains incompressible. Mathematically, we can therefore expect to find this equation by using the incompressibility constraint. Taking divergence of the stochastic Euler momentum equation (2.64) in its integral form, we have

$$\nabla\cdot(\boldsymbol{u}_t-\boldsymbol{u}_0)+\sum_{k=1}^{\infty}\int_0^t(\Delta\xi_k)\cdot\boldsymbol{u}+\sum_{i,j=1}^{n}\left(\frac{\partial\xi_k^i}{\partial x_j}\frac{\partial u_i}{\partial x_j}+\frac{\partial\xi_k^i}{\partial x_j}\frac{\partial u_j}{\partial x_i}\right)\circ dW_s^k \\ +\int_0^t\sum_{i,j=1}^{n}\frac{\partial u_i}{\partial x_j}\frac{\partial u_j}{\partial x_i}\,ds+\Delta\pi_t-\Delta\pi_0=0\,.$$

The incompressibility constraint then implies that pressure satisfies the following

$$\begin{aligned}\mathrm{d}\tilde{\pi}&=-\sum_{i,j=1}^{n}\frac{\partial u_i}{\partial x_j}\frac{\partial u_j}{\partial x_i}\,dt\\&\quad-\sum_{k=1}^{\infty}\left((\Delta\xi_k)\cdot\boldsymbol{u}+\sum_{i,j=1}^{n}\left(\frac{\partial\xi_k^i}{\partial x_j}\frac{\partial u_i}{\partial x_j}+\frac{\partial\xi_k^i}{\partial x_j}\frac{\partial u_j}{\partial x_i}\right)\right)\circ dW_t^k\,,\\ \pi&=\Delta^{-1}\tilde{\pi}\,.\end{aligned}$$

Remark 2.38 (The need for the stochastic pressure.). If we considered the Lagrange multiplier of $D-1$ in the derivation of the stochastic Euler equation to simply be $\pi\,dt$, or similar, even if we considered π to be a stochastic process then this would be insufficient to derive meaningful equations. This is observable in the above equation for the stochastic pressure, since a Lagrange multiplier of $\pi\,dt$ would imply that a sum of Lebesgue integrals is equal to a stochastic integral. This is despite the fact that the action integral would seem reasonable and would not hint at the nonsensical equations that it would produce, since the difficulties only make themselves known at the level of the equations. This example provides an explicit illustration of why the Lagrange multiplier needs this more general form, and why the identification of the driving semimartingale is crucial for deriving equations from stochastic variational principles of this form.

3

The dynamics of a free surface

3.1 Introduction

The mathematics behind the dynamics occurring on the free surface of a fluid has been an active field of study for the best part of two centuries. Following John Scott Russell's now famous observation of a 'wave of translation' on the Union Canal, mathematicians became interested in developing an understanding of the behaviour of such waves from the perspective of hydrodynamics. The wave observed by Scott Russell was propagating sufficiently fast to ensure that his horse could not keep up, yet the Union Canal is not known for a rushing current. It was therefore well understood that disturbances on the free surface can be distinct from material transport and have their own behaviour. These waves are perhaps more akin to a field of crops rippling in the wind, or the agitation of a large crowd of people. Indeed, Scott Russell noted in his 1845 Report on Waves [92] that these three phenomena were linked together in Book II of Homer's Iliad, following a speech by Agamemnon. Due to the limited language available to historic authors to refer to colour, this connection predates references to the sea being 'blue'. Despite the distinction between material transport of fluid and the propagation of a wave on its surface being evident, potentially for thousands of years, mathematical progress on modelling the interactions between these two concepts proved to be challenging. Indeed, a complete theory of wave-current interaction does not exist and there are phenomena on the free surface for which we still have no sound explanation, several of which are explained and illustrated in Sections 3.3.5 and 3.4.

We will begin this chapter with a summary of some mathematical preliminaries, including the evaluation of the three dimensional Euler equations onto the free surface. This will be followed by a novel stochastic perturbation of classical water wave theory for a potential flow, which preserves the Hamiltonian structure of the deterministic case. Under an additional assumption on the structure of the noise, this Hamiltonian structure will be rewritten entirely in terms of the Dirichlet to Neumann operator. Following this, a new geometric framework with which we may interpret wave current interactions is presented. Here, the dynamics of waves and currents are separated into two maps, which makes a clear distinction between horizontal and vertical motions.

3.1.1 A brief review of literature

The mathematics of a fluid with free boundary is a classical problem with a long history. As such, there are a number of different approaches present in the literature, many of which are often labelled as wave-current interaction. There are numerous approaches to the interactions between waves and currents, and the definition of a 'wave' and a 'current' can differ depending on one's perspective.

A classical approach to modelling geophysical flows with a free upper boundary is by integrating from the bathymetry to the free surface, and considering vertically averaged variables and columnar motion. This involves the shallow water approximation, and culminates in popular equations such as the shallow water and Green-Naghdi [43] models. This methodology can produce models for wave-like motion on the free surface, though to study wave-current interactions there is often an additional assumed distinction between *slow* currents and *fast* waves. The development of the generalised Lagrangian mean (GLM) approach by Andrew and McIntyre [3] was an influential advancement in this field, in which the fluid motion is decomposed into a mean part and an oscillatory part. This is achieved through a slow/fast decomposition of a Lagrangian trajectory into a Lagrangian mean trajectory and a fluctuating displacement with zero Eulerian mean. Quite differently, the Craik-Leibovich approach to wave-current interaction concerns itself with the interaction between wind driven rapidly oscillating waves and the mean fluid velocity [30]. The expression derived is known as the Stokes vortex force, which generates structures similar to Langmuir circulation in the upper oceanic boundary layer. We may think of this as a three dimensional effect related to air-sea interaction. An extensive body of literature can be found on both of these perspectives on wave-current interaction. It can be shown that both the Craik-Leibovich [48] and GLM [41, 50] approaches have geometric interpretations and variational structures, through which stochastic versions of the respective theories have been formulated [56].

Whilst we will not be exploring the aforementioned models of wave-current interaction, we will be working with the closure of the water wave problem provided by a potential flow assumption, which we will be referring to as the *classical water wave equations*. As discussed above, we will be deriving a stochastic version of this theory by exploiting the geometric structure of these equations [106] and, by doing so, will be filling in a gap in the literature.

The new approach to wave-current interaction introduced in this chapter will differ from those found in the literature in some fundamental ways. In particular, rather than defining a wave and a current by a slow/fast decomposition of the fluid motion, we will be considering a distinction between horizontal and vertical motion on the fluid surface. Since the currents will be represented by a purely two dimensional fluid theory, this approach is inherently different to those designed to explain explicit three dimensional effects such as Langmuir circulations. Instead, we will seek to design a framework in which we can study the interactions of wave like vertical motions with a two dimensional fluid flow, reminiscent of the data obtained from satellite oceanography. We will have distinct maps for the horizontal and vertical motions, and a general Euler-Poincaré theory will be presented. By designing a preliminary example Lagrangian within this modelling framework, we will illustrate the structure and potential of the equations resulting from this approach. Note that the wave component of this Lagrangian will give oscillating wave like behaviour, and does not directly give the same solution properties as more familiar models of water wave motion. We will give a discussion and stochastic perturbation of a more classical approach to surface gravity waves in Section 3.3, before deviating from this in Section 3.4.

3.2 Preliminaries and perspective

The problem of waves on the surface of a three dimensional fluid is, despite the intrinsic three dimensionality of reality, one which features natural two dimensionality [43]. This comes from the fact that the upper boundary of a fluid such as a river, sea, or lake is a two dimensional submanifold in $\mathbb{R}^3$. Furthermore, waves propagate in a two dimensional manner, and many of the observational remarks of these problems surround two dimensional transport along the surface. Therefore, when deriving models to explain certain wave phenomena, an attempt is often made to reduce the dimension of the problem such that we are left with a two dimensional momentum equation. This is usually done by some sort of vertical averaging. Instead of vertical averaging, we will instead look at the evaluation of variables on the free surface to attempt to close the system in terms of quantities which are observable and measurable from satellite imagery. We will proceed here with a discussion of the three dimensional problem, before discussing the trace or 'shadow' of these dynamics on the free surface. This is used to motivate a closed two dimensional model found by applying Hamilton's principle to an action integral. The structure of this model is then investigated.

3.2.1 The three dimensional beginnings

We will begin from the non-homogeneous three dimensional Euler equations. Assume we have a three dimensional spatial domain with coordinates

$\boldsymbol{x} = (\boldsymbol{r}, z) = (x, y, z)$, where we have given additional notation for the two dimensional horizontal coordinates and the vertical coordinate. We will denote the velocity field of the fluid inside this domain by $\boldsymbol{u} = (\boldsymbol{v}, w)$. Note that, again, we have provided the notation necessary for a decomposition of horizontal and vertical velocities, given by $\boldsymbol{v}$ and w respectively. The incompressible Euler equations for a three dimensional fluid with a gravitational forcing term may be found by applying Hamilton's principle to the following action integral

$$\begin{aligned} \int\int D\rho \left(\frac{1}{2}|\boldsymbol{u}|^2 - gz\right) - \pi(D-1) \\ + \varphi(\partial_t D + \operatorname{div}(D\boldsymbol{u})) + \gamma(\partial_t \rho + \boldsymbol{u} \cdot \nabla\rho)\, d^3x\, dt\,. \end{aligned} \tag{3.1}$$

In the above, $\boldsymbol{u}$ denotes the three dimensional velocity field, π the pressure, ρ the thermal buoyancy, and D the volume element given by the determinant of the Lagrange to Euler map. The three dimensional operator ∇ is defined by $(\partial_x, \partial_y, \partial_z)$, and as expected g represents acceleration due to gravity. We will later use the notation $\nabla_{\boldsymbol{r}} = (\partial_x, \partial_y)$ to denote the two dimensional version of this operator. Notice the presence of Lagrange multipliers φ, and γ which constrain the relevant advection constraints. Notice also that we have not used the coordinate free description of the advection constraints, instead using the equation corresponding to our choice of coordinates. The coordinate free description is still relevant, and Section 2.1.2 should be consulted in order to properly understand the mathematics behind variational principles of this form. Taking variations of the action (3.1) gives

$$
\begin{aligned}
\delta D: &\quad (\partial_t + \mathcal{L}_u)\varphi = \rho\left(\frac{1}{2}|\boldsymbol{u}|^2 - gz\right) - \pi\,,\\
\delta \rho: &\quad (\partial_t + \mathcal{L}_u)\left(\frac{\gamma}{D}\right) = \frac{1}{2}|\boldsymbol{u}|^2 - gz\,,\\
\left.\begin{aligned}
\delta \varphi: &\quad (\partial_t + \mathcal{L}_u)(D\,d^3x) = (\partial_t D + \mathrm{div}(D\boldsymbol{u}))d^3x = 0\\
\delta \pi: &\quad D - 1 = 0
\end{aligned}\right\} &\qquad \Longrightarrow \ \mathrm{div}\boldsymbol{u} = 0\,,\\
\delta \gamma: &\quad (\partial_t + \mathcal{L}_u)\rho = 0\,,\\
\delta u: &\quad \rho\,\boldsymbol{u}\cdot d\boldsymbol{x} = d\varphi - \frac{\gamma}{D}d\rho\,.
\end{aligned}
$$

Since the Euler equations are so well studied and understood, we will not extrapolate on the full geometric structure of this model here. We notice that the above relations can be assembled into the momentum equation for the system as follows

$$
\begin{aligned}
\rho(\partial_t + \mathcal{L}_u)(\boldsymbol{u}\cdot d\boldsymbol{x}) &= d(\partial_t + \mathcal{L}_u)\varphi - (\partial_t + \mathcal{L}_u)\left(\frac{\gamma}{D}\right)d\rho\\
&= d\left(\rho\left(\frac{1}{2}|\boldsymbol{u}|^2 - gz\right) - \pi\right) - \left(\frac{1}{2}|\boldsymbol{u}|^2 - gz\right)d\rho\\
&= \rho d\left(\frac{1}{2}|\boldsymbol{u}|^2 - gz\right) - d\pi\,.
\end{aligned}
$$

A *free boundary* problem may be formulated by assuming that our three dimensional spatial domain has an upper boundary, $z = \zeta(\boldsymbol{r}, t)$, which is a function of time and space. This will be considered as a boundary condition on the equations derived from the above variational principle, though it is worth noting that it is possible to embed such conditions into the variational principle itself [24]. We now introduce a convenient notation with which we will be able to cleanly represent conditions on the free boundary.

Definition 3.1 (Evaluation on a free surface)**.** The evaluation of a time dependent object $f(\boldsymbol{x},t) = f(\boldsymbol{r},z,t)$, which depends on all three spatial coordinates, on the free surface, $z = \zeta(\boldsymbol{r},t)$, is an object which is independent of the vertical coordinate, z, and is denoted by the following

$$\widehat{f}(\boldsymbol{r},t) := f(\boldsymbol{r},\zeta(\boldsymbol{r},t),t)\,. \tag{3.2}$$

Definition 3.2 (Evaluation on a free surface as a pullback)**.** The evaluation of a variable, $f(\boldsymbol{x},t)$, on the free surface defined in Definition 3.1 may be written in terms of the pullback by a time dependent function $Z_t : \mathbb{R}^3 \mapsto \mathbb{R}^3$ as

$$Z_t^* f = (f \circ Z_t)(x,y,z) = \widehat{f}\,, \tag{3.3}$$

where Z_t is defined by

$$Z_t(x,y,z) = (x,y,\zeta(t,x,y))\,. \tag{3.4}$$

Notice that the following vector, $\boldsymbol{n}$, is *normal* to the surface

$$\boldsymbol{n} = \begin{pmatrix} -\nabla_{\boldsymbol{r}}\zeta \\ 1 \end{pmatrix}, \tag{3.5}$$

where, as mentioned earlier, $\nabla_{\boldsymbol{r}} := (\partial_x, \partial_y)$ is the two dimensional gradient operator in the horizontal components. This normal vector may be used to interpret the following, commonly used, boundary condition at the free surface.

Definition 3.3 (The kinematic boundary condition)**.** The kinematic boundary condition states that a particle on the free surface remains on the free surface. This is described mathematically as

$$(\partial_t + \boldsymbol{u} \cdot \nabla)(\zeta - z) = 0 \,, \quad \text{on} \quad z = \zeta \,.$$

Using the notation from Definition 3.1, we have that

$$\partial_t \zeta + \widehat{\boldsymbol{v}} \cdot \nabla_{\boldsymbol{r}} \zeta = \widehat{w} \,. \tag{3.6}$$

Notice that this is equivalent to

$$\partial_t \zeta = \boldsymbol{u} \cdot \boldsymbol{n} \,, \quad \text{on} \quad z = \zeta(\boldsymbol{r}, t) \,,$$

which is a sensible statement on how the rate of change of the free surface relates to the velocity.

A model may be formulated where the dynamics of the interior of the three dimensional domain is governed by the homogeneous Euler equations, where buoyancy is constant, and the upper surface has a kinematic boundary condition, as follows

$$\partial_t \boldsymbol{u} + \boldsymbol{u} \cdot \nabla \boldsymbol{u} = -\boldsymbol{g} - \frac{1}{\rho_0} \nabla \pi \,, \tag{3.7}$$

$$\nabla \cdot \boldsymbol{u} = 0 \,, \tag{3.8}$$

$$\partial_t \zeta = \boldsymbol{u} \cdot \boldsymbol{n} \,, \quad \text{on} \quad z = \zeta(\boldsymbol{r}, t) \,, \tag{3.9}$$

where $\boldsymbol{g} = (0, 0, g)$. This system may be closed by the inclusion of a dynamic boundary condition on the pressure, which requires that the stress is continuous across the boundary.

3.2.2 Evaluation of the Euler equations onto a free surface

In order to determine what these equations look like on the free surface, we first look at what happens to advected variables. Since the kinematic boundary constraint ensures that the evaluation of an advected variable on the free surface does not stray from the surface, it is natural to attempt to understand how this evaluation is advected. We begin by determining the difference between evaluating the derivative of a variable on the surface, as opposed to taking a derivative of the evaluation.

Proposition 3.4. *The difference between exchanging the order of differentiation and evaluation on the free surface is as follows,*

$$\begin{aligned} \partial_t\widehat{f} - \widehat{\partial_t f} &= \widehat{\partial_z f}\partial_t\zeta\,, \\ \nabla_{\boldsymbol{r}}\widehat{f} - \widehat{\nabla_{\boldsymbol{r}} f} &= \widehat{\partial_z f}\nabla_{\boldsymbol{r}}\zeta\,, \\ \nabla_{\boldsymbol{r}}\cdot\widehat{\boldsymbol{f}} - \widehat{\nabla_{\boldsymbol{r}}\cdot\boldsymbol{f}} &= \widehat{\partial_z \boldsymbol{f}}\cdot\nabla_{\boldsymbol{r}}\zeta\,, \end{aligned}$$

where $f(\boldsymbol{x},t)$ and $\boldsymbol{f}(\boldsymbol{x},t) = (f_1(\boldsymbol{x},t), f_2(\boldsymbol{x},t))$ are used to denote arbitrary variables with values in $\mathbb{R}$ and $\mathbb{R}^2$ respectively.

Remark 3.5. The equations above are a comment on exchanging the order of differentiation and pullback, when the evaluation on the free surface is interpreted as in Definition 3.2. Using the notation from this definition, the equations are equivalent to

$$\begin{aligned} \partial_t(Z_t^* f) - Z_t^*(\partial_t f) &= (Z_t^*\partial_z f)\partial_t\zeta\,, \\ \nabla_{\boldsymbol{r}}(Z_t^* f) - Z_t^*(\nabla_{\boldsymbol{r}} f) &= (Z_t^*\partial_z f)\nabla_{\boldsymbol{r}}\zeta\,, \\ \nabla_{\boldsymbol{r}}\cdot(Z_t^*\boldsymbol{f}) - Z_t^*(\nabla_{\boldsymbol{r}}\cdot\boldsymbol{f}) &= (Z_t^*\partial_z\boldsymbol{f})\cdot\nabla_{\boldsymbol{r}}\zeta\,. \end{aligned}$$

Proof. The proof of the first identity follows immediately by the chain rule, viz.

$$
\begin{aligned}
\partial_t \widehat{f} &= \partial_t f(\boldsymbol{r}, \zeta(\boldsymbol{r}, t), t) \\
&= [\partial_t f(\boldsymbol{x}, t)]\big|_{z=\zeta} + \widehat{\partial_z f} \partial_t \zeta = \widehat{\partial_t f} + \widehat{\partial_z f} \partial_t \zeta \,.
\end{aligned}
$$

The second identity follows by the same method and the third by applying this method twice, as follows

$$
\begin{aligned}
\nabla_{\boldsymbol{r}} \cdot \widehat{\boldsymbol{f}} &= \partial_x \widehat{f}_1 + \partial_y \widehat{f}_2 \\
&= \widehat{\partial_x f_1} + \widehat{\partial_z f_1} \partial_x \zeta + \widehat{\partial_y f_2} + \widehat{\partial_z f_2} \partial_y \zeta = \widehat{\nabla_{\boldsymbol{r}} \cdot \boldsymbol{f}} + \widehat{\partial_z \boldsymbol{f}} \cdot \nabla_{\boldsymbol{r}} \zeta \,.
\end{aligned}
$$

$\mathscr{Q}.\mathscr{E}.\mathscr{D}.$

These results imply the following remarkable property of advection on a free surface.

Proposition 3.6. *The advection operator, $\mathcal{D} := \partial_t + \boldsymbol{u} \cdot \nabla$, applied to some quantity, f, satisfies the following free surface relationship*

$$
\widehat{\mathcal{D} f} = \widehat{\mathcal{D}} \widehat{f} \,. \tag{3.10}
$$

where $\widehat{\mathcal{D}} = \partial_t + \widehat{\boldsymbol{v}} \cdot \nabla_{\boldsymbol{r}}$ defines a two dimensional advection operator on the free surface.

Remark 3.7. Proposition 3.6 implies that, on the free surface, the three dimensional advection of a scalar advected quantity is equivalent to the two dimensional advection of the trace of that quantity on the free surface, by the horizontal component of velocity only.

Proof. The proof of equation (3.10) is by direct calculation, using the identities from Proposition 3.4 and the kinematic boundary condition (3.6), as follows

$$
\begin{aligned}
\widehat{\mathcal{D}f} &= \widehat{\partial_t f} + \widehat{\boldsymbol{v}} \cdot \widehat{\nabla_{\boldsymbol{r}} f} + \widehat{w}\widehat{\partial_z f} \\
&= \partial_t \widehat{f} - \widehat{\partial_z f}\partial_t \zeta + \widehat{\boldsymbol{v}} \cdot (\nabla_{\boldsymbol{r}} \widehat{f} - \widehat{\partial_z f}\nabla_{\boldsymbol{r}}\zeta) + \widehat{w}\widehat{\partial_z f} \\
&= \partial_t \widehat{f} + \widehat{\boldsymbol{v}} \cdot \nabla_{\boldsymbol{r}} \widehat{f} + \widehat{\partial_z f}\big(\widehat{w} - \partial_t\zeta - \widehat{\boldsymbol{v}} \cdot \nabla_{\boldsymbol{r}}\zeta\big) \\
\text{(by (3.6))} \qquad &= \partial_t \widehat{f} + \widehat{\boldsymbol{v}} \cdot \nabla_{\boldsymbol{r}} \widehat{f} =: \widehat{\mathcal{D}}\widehat{f}\,.
\end{aligned}
$$

$\mathscr{Q}.\mathscr{E}.\mathscr{D}.$

As a result of this, despite the following equation looking unlikely, we have

$$
\partial_t \boldsymbol{u} + \boldsymbol{u} \cdot \nabla \boldsymbol{u} = \partial_t \boldsymbol{u} + \boldsymbol{v} \cdot \nabla_{\boldsymbol{r}} \boldsymbol{u}\,, \quad \text{on} \quad z = \zeta(\boldsymbol{r}, t)\,. \tag{3.11}
$$

Evaluation of the momentum equation on the free surface. Using the notation and results proven thus far, we may evaluate the Euler momentum equation (3.7) on the free surface. We first project the first two components of Euler, the horizontal components, as follows

$$
\partial_t \widehat{\boldsymbol{v}} + \widehat{\boldsymbol{v}} \cdot \nabla_{\boldsymbol{r}} \widehat{\boldsymbol{v}} = -\frac{1}{\rho_0}\widehat{\nabla_{\boldsymbol{r}}\pi} = -\frac{1}{\rho_0}\nabla_{\boldsymbol{r}}\widehat{\pi} + \frac{\widehat{\partial_z \pi}}{\rho_0}\nabla_{\boldsymbol{r}}\zeta\,, \tag{3.12}
$$

and the same for vertical component

$$
\partial_t \widehat{w} + \widehat{\boldsymbol{v}} \cdot \nabla_{\boldsymbol{r}} \widehat{w} = -g - \frac{\widehat{\partial_z \pi}}{\rho_0}\,. \tag{3.13}
$$

Notice that several terms on the right hand size contain a z derivative, which is inconsistent with a 2D theory, and an equation for $\widehat{\pi}$ is needed to close the theory. The z derivatives also prevent us from being able to define a variational principle in two dimensions which captures the full unapproximated dynamics of the free surface of a three dimensional Euler model. Furthermore, the equations (3.12) and (3.13), together with the kinematic boundary condition (3.6) and incompressibility (3.8), are not a closed set of equations.

The terms featuring a z derivative may be eliminated by assembling the momentum equations into *Choi's relation*, which is an unapproximated relationship between variables evaluated on the free surface and is defined as follows.

Definition 3.8. By *Choi's relation*, we mean the following equation

$$\widehat{\mathcal{D}}\widehat{\boldsymbol{v}} + (\widehat{\mathcal{D}}^2\zeta + g)\nabla_{\boldsymbol{r}}\zeta = -\frac{1}{\rho_0}\nabla_{\boldsymbol{r}}\widehat{\pi}\,, \tag{3.14}$$

which may be found by combining the equations (3.12) and (3.13). *

Remark 3.9 (A comment on pressure). A sensible boundary condition for equations of this class is to assume that the trace of the pressure on the free surface is constant. This is the dynamic boundary condition which ensures that there is no jump in pressure between the incompressible fluid and the external atmospheric pressure, which is assumed to not be variable. It is worth noting that this does not imply that gradients of the pressure are zero at the surface, which is a consequence of the fact that differentiation and evaluation on the surface do not commute, as was shown in Proposition 3.4. Indeed

$$0 = \nabla_{\boldsymbol{r}}\widehat{\pi} = \widehat{\nabla_{\boldsymbol{r}}\pi} + \widehat{\partial_z\pi}\nabla_{\boldsymbol{r}}\zeta\,,$$

and thus $\widehat{\nabla_{\boldsymbol{r}}\pi}$ is not necessarily zero.

*A non-homogeneous version of Choi's relation also exists, where ρ_0 is simply replaced by $\widehat{\rho}$.

3.2.3 Impossibility of unapproximated closure

It is clear that making additional assumptions or approximations, such as potential flow, is necessary to close the system. Thus, a two dimensional framework can only hope to capture some essence of the full physics.

Finding a closed variational two dimensional theory. The equations (3.12) and (3.13) cannot be found by a two dimensional variational principle, due to the fact that they are not closed and feature a z derivative. As in Crisan et al. [33], taking variations of the action integral, where the 'hat' notation has been dropped on the advected variables for simplicity,

$$\begin{aligned} S &= \int \ell(\widehat{v}, D, \widehat{\phi}, \widehat{w}, \zeta, \lambda) dt \\ &= \int\int D\rho \left(\frac{1}{2}\big(|\widehat{\boldsymbol{v}}|^2 + \widehat{w}^2\big) - g\zeta \right) + \lambda\big(\partial_t \zeta + \widehat{\boldsymbol{v}} \cdot \nabla_{\boldsymbol{r}} \zeta - \widehat{w}\big) \\ &\qquad + \widehat{\phi}\big(\partial_t D + \mathrm{div}_{\boldsymbol{r}}(D\widehat{\boldsymbol{v}})\big) + \gamma\big(\partial_t \rho + \widehat{\boldsymbol{v}} \cdot \nabla_{\boldsymbol{r}} \rho\big) - \pi(D-1)\, d^2r\, dt\,, \end{aligned} \tag{3.15}$$

yields a closure of the equations (3.12) and (3.13). The integrand of the Lagrangian corresponding to this action is equivalent to that of the 3D action, after evaluation on the free surface and addition of the kinematic boundary condition, however the spatial integral is taken over two dimensions rather than three. Notice that this approach will not produce the terms in $\widehat{\partial_z \pi}$, since no z derivatives are possible in this setup, and furthermore it similarly does not capture the z derivative term in the incompressibility constraint $0 = \widehat{\nabla_{\boldsymbol{r}} \cdot \boldsymbol{v}} + \widehat{\partial_z w} = \nabla_{\boldsymbol{r}} \cdot \widehat{\boldsymbol{v}} + \widehat{\partial_z \boldsymbol{u} \cdot \boldsymbol{n}}$. Despite this, it achieves mathematical closure into a closed two dimensional fluid theory. This transition toward a two dimensional fluid theory, and two dimensional incompressibility, limits the ability of models of this class to capture the full picture seen in the three dimensional equations. In particular, the pressure on the free surface in this case must be allowed to vary and enforce incompressibility. As shown in Crisan et al. [34], the action integral (3.15) is, after integration by parts, equivalent to imposing the *classical water wave equations*, which we will see in Section 3.3, as constraints by Lagrange multipliers, D and λ. The mathematical closure this gives to Choi's relation makes it an appealing framework in which we can make approximations to model how wave dynamics can interact with two dimensional free surface transport flows.

An assumption on vertical pressure gradients. When formulating the variational formulation of a fluid governed by the Green-Naghdi or Great Lake equations, as in Holm and Luesink [61], including terms of the same order as the aspect ratio squared gives that the pressure deviates from hydrostatic balance by a term which is linear in the vertical coordinate. If, rather than prescribing the form of this deviation, we assume that the dependence of the pressure on the vertical coordinate is a linear perturbation around hydrostatic balance, then a harmonic oscillator appears. Indeed, if

$$\partial_z \pi = -g\rho_0 - \epsilon z \,, \tag{3.16}$$

then the wave dynamics become a harmonic oscillator

$$\begin{aligned} \partial_t \zeta + \widehat{\boldsymbol{v}} \cdot \nabla_{\boldsymbol{r}} \zeta &= \widehat{w} \,, \\ \partial_t \widehat{w} + \widehat{\boldsymbol{v}} \cdot \nabla_{\boldsymbol{r}} \widehat{w} &= -\frac{\epsilon}{\rho_0} \zeta \,. \end{aligned}$$

Moreover, this assumption on the dependence of the pressure on z implies, by integration in z, that

$$\pi = -g\rho_0 z - \frac{\epsilon}{2}z^2 + p(\boldsymbol{r}, t)\,. \tag{3.17}$$

Thus, the right hand side of the equation for $\widehat{\boldsymbol{v}}$ becomes

$$\begin{aligned}
-\frac{1}{\rho_0}\nabla_{\boldsymbol{r}}\widehat{\pi} + \frac{\widehat{\partial_z \pi}}{\rho_0}\nabla_{\boldsymbol{r}}\zeta &= -\frac{1}{\rho_0}\nabla_{\boldsymbol{r}}\big(- g\rho_0\zeta - \frac{\epsilon}{2}\zeta^2 + p(\boldsymbol{r}, t)\big) \\
&\quad + \frac{1}{\rho_0}\big(- g\rho_0 - \epsilon\zeta\big)\nabla_{\boldsymbol{r}}\zeta \\
&= -\frac{1}{\rho_0}\nabla_{\boldsymbol{r}} p(\boldsymbol{r}, t)\,.
\end{aligned} \tag{3.18}$$

Notice that this is obviously equal to $\frac{1}{\rho_0}\widehat{\nabla_{\boldsymbol{r}}\pi}$. The equations therefore become

$$\partial_t \widehat{\boldsymbol{v}} + \widehat{\boldsymbol{v}} \cdot \nabla_{\boldsymbol{r}} \widehat{\boldsymbol{v}} = -\frac{1}{\rho_0}\nabla_{\boldsymbol{r}} p\,, \tag{3.19}$$

$$\partial_t \widehat{w} + \widehat{\boldsymbol{v}} \cdot \nabla_{\boldsymbol{r}} \widehat{w} = -\frac{\epsilon}{\rho_0}\zeta\,, \tag{3.20}$$

$$\partial_t \zeta + \widehat{\boldsymbol{v}} \cdot \nabla_{\boldsymbol{r}} \zeta = \widehat{w}\,, \tag{3.21}$$

$$\nabla_{\boldsymbol{r}} \cdot \widehat{\boldsymbol{v}} + \widehat{\partial_z \boldsymbol{u}} \cdot \boldsymbol{n} = 0\,. \tag{3.22}$$

In this approximation, we have a momentum equation for the horizontal currents which does not involve the wave variables, and a pair of equations for the vertical oscillations which occur in the frame of reference moving with the horizontal flow. We do not have a complete separation between waves and currents since the incompressibility constraint, and hence the equation for the pressure, creates an interdependence between the horizontal and vertical dynamics. The three dimensional incompressibility constraint is responsible for the entanglement between the wave and current variables and, in particular, the effect of the waves on the current. However, the effects of the waves on the current variables are, in the above equations, given by the kinematic boundary condition and, in the thermal case, by variable thermal buoyancy $\rho \not\equiv \rho_0$.

3.3 The classical water wave equations (CWWE)

A well studied approach to closing the free surface problem is the so-called *classical water wave equations* (CWWE). These equations have a Hamiltonian structure, as discovered by Zakharov [106]. Following Craig and Sulem [29], the equations can be rearranged into a form which enables numerical simulation by writing them in terms of the free surface and trace of the potential on the free surface only, which is achieved using the *Dirichlet to Neumann map.*

In order to achieve this closure, the bulk flow is assumed to be governed by the Euler equations (3.7)-(3.9), where $\rho = \rho_0$ is constant, under the additional assumption that the flow is *irrotational.* This translates mathematically to an assumption that the curl of the (three dimensional) velocity field is zero, $\nabla \times \boldsymbol{u} = 0$. If we assume further that the spatial domain is simply connected, then the velocity field is *conservative.* We thus have the existence of the velocity potential, ϕ, which is defined as the potential corresponding to the velocity field, and the incompressibility constraint implies that this satisfies Laplace's equation

$$\boldsymbol{u} = \nabla\phi \quad \Longrightarrow \quad \Delta\phi = 0\,. \tag{3.23}$$

In this case, we may rewrite the Euler momentum equation (3.7) as

$$\partial_t \phi + \frac{1}{2}|\nabla \phi|^2 + gz = -\frac{1}{\rho_0}\pi\,, \tag{3.24}$$

which is known as Bernoulli's integrated form of the Euler equation. Note also that, under the potential flow approximation, the kinematic boundary condition becomes

$$\partial_t \zeta = \widehat{\partial_z \phi} - \widehat{\nabla_{\boldsymbol{r}} \phi} \cdot \nabla_{\boldsymbol{r}} \zeta = \widehat{\partial_z \phi}(1 + |\nabla_{\boldsymbol{r}} \zeta|^2) - \nabla_{\boldsymbol{r}} \widehat{\phi} \cdot \nabla_{\boldsymbol{r}} \zeta\,. \tag{3.25}$$

Here, we will derive a new *stochastic* formulation of the classical water wave assumption. This will be achieved by assuming that the velocity field in the bulk flow is governed by the incompressible Euler equations with stochastic advection by Lie transport. Recall, from the example in Section 2.2.1, that these equations are (2.64) and (2.65), with the addition of vertical gravitational forcing. Since we are beginning with a stochastic perturbation which preserves the geometric structure of the underlying flow, we may expect that this leads us to a stochastic classical water wave equation with a reasonable Hamiltonian structure. Indeed, we will see that this is the case.

Note that a stochastic perturbation of the water wave theory has recently appeared in the literature [37]. Our approach here differs in that we will begin with a structure preserving approach to the addition of noise in the underlying fluid model. Thus, rather than fitting a Hamiltonian structure to a stochastic equation, this approach will have a variational structure by design.

3.3.1 A stochastic perturbation of the water wave problem

Before we begin with the derivation of the model equations, the reader should keep in mind that the well known deterministic theory may be recovered at each point by setting the stochastic perturbation terms to zero. The stochastic terms in this case represent transport noise in the equation governing the the fluid below the free surface, which will be taken to be the stochastic Euler equation from Section 2.2.1.

A stochastic kinematic boundary. Beginning with the kinematic boundary condition, we will make the three dimensional theory stochastic. The kinematic boundary condition with stochastic advection by Lie transport can be written as

$$(\mathrm{d} + \mathrm{d}\boldsymbol{x}_t \cdot \nabla)(z - \zeta) = 0\,.$$

Decomposing the noise into two dimensional horizontal and one dimensional vertical components, we have that

$$\mathrm{d}\boldsymbol{x}_t = \begin{pmatrix} \boldsymbol{v} \\ w \end{pmatrix} dt + \sum_{i=1}^{\infty} \begin{pmatrix} \boldsymbol{\xi}_i^{(\boldsymbol{r})} \\ \xi_i^{(z)} \end{pmatrix} \circ dW_t^i =: \begin{pmatrix} \mathrm{d}\boldsymbol{r}_t \\ \mathrm{d}z_t \end{pmatrix}, \tag{3.26}$$

where we have denoted the components of the perturbations as $\boldsymbol{\xi}_i = (\boldsymbol{\xi}_i^{(\boldsymbol{r})}, \xi_i^{(z)})$. The kinematic boundary condition is therefore

$$(\mathrm{d} + \mathrm{d}\boldsymbol{r}_t \cdot \nabla)\zeta = \mathrm{d}z_t\,, \tag{3.27}$$

or, equivalently,

$$\mathrm{d}\zeta = \widehat{\boldsymbol{u}} \cdot \boldsymbol{n}\, dt + \sum_{i=1}^{\infty} \widehat{\boldsymbol{\xi}}_i \cdot \boldsymbol{n} \circ dW_t^i\,. \tag{3.28}$$

The potential flow assumption. The N dimensional stochastic Euler equations are given by equations (2.64) and (2.65), which we repeat here with the addition of the vertical gravitational forcing

$$\mathrm{d}\boldsymbol{u} + \boldsymbol{u}\cdot\nabla\boldsymbol{u}\,dt + \sum_{i=1}^{\infty}\left(\boldsymbol{\xi}_i\cdot\nabla\boldsymbol{u} + \sum_{j=1}^{N} u_j\nabla\xi_i^j\right)\circ dW_t^i + \boldsymbol{g}\,dt = -\nabla\mathrm{d}\pi\,, \tag{3.29}$$

$$\nabla\cdot\mathrm{d}\boldsymbol{x}_t = 0\,, \tag{3.30}$$

where we have denoted the components of $\boldsymbol{\xi}_i$ as $(\xi_i^1, \xi_i^2, \dots)$ and, as before, $\boldsymbol{g} = (0, 0, g)$. In the case of incompressible fluids, the perturbations, $\boldsymbol{\xi}_i$, mirror the structure of the deterministic velocity field in that its divergence is zero. In the case of irrotational fluids we will, for now, assume that the velocity field, $\boldsymbol{u} = \nabla\phi$, is irrotational and the perturbations are only incompressible. The kinematic boundary condition (3.27) becomes

$$\mathrm{d}\zeta + \widehat{\nabla_{\boldsymbol{r}}\phi}\cdot\nabla_{\boldsymbol{r}}\zeta\,dt + \sum_{i=1}^{\infty}\widehat{\boldsymbol{\xi}_i^{(\boldsymbol{r})}}\cdot\nabla_{\boldsymbol{r}}\zeta\circ dW_t^i = \widehat{\partial_z\phi}\,dt + \sum_{i=1}^{\infty}\widehat{\xi_i^{(z)}}\circ dW_t^i\,. \tag{3.31}$$

The Euler momentum equation becomes

$$\begin{aligned}\mathrm{d}\nabla\phi + (\nabla\phi\cdot\nabla)\nabla\phi\,dt + \sum_{i=1}^{\infty}\left((\boldsymbol{\xi}_i\cdot\nabla)\nabla\phi + \sum_{j=1}^{3}(\partial_j\phi)\nabla\xi_i^j\right)\circ dW_t^i \\ + \boldsymbol{g}\,dt + \nabla\mathrm{d}\pi = 0\,,\end{aligned}$$

where we have taken the dimension to be three, $N = 3$, and by the sum over the derivatives ∂_j we mean a sum over $\{\partial_x, \partial_y, \partial_z\}$. Recall that the nonlinearity simplifies into a gradient term

$$(\nabla\phi\cdot\nabla)\nabla\phi = \frac{1}{2}\nabla(|\nabla\phi|^2)\,,$$

but it is not immediately obvious that the same is true for each nonlinear stochastic term. In order to simplify these terms, we return to the coordinate free language of exterior calculus as discussed in Section 2.1.2. We notice that each stochastic term corresponds to a Lie derivative of $u^\flat = \boldsymbol{u}\cdot d\boldsymbol{x}$ with respect to the vector field $\xi_i = \boldsymbol{\xi}_i \cdot \nabla$. Using Cartan's formula, may relate this to the interior product by

$$\mathcal{L}_{\xi_i} u^\flat = \xi_i \lrcorner\, du^\flat + d(\xi_i \lrcorner\, u^\flat)\,. \tag{3.32}$$

Since we have a potential flow, the vector field $u \in \mathfrak{X}$ is related to its potential ϕ by

$$u = (d\phi)^\sharp\,.$$

The Lie derivative then becomes

$$\begin{aligned}\mathcal{L}_{\xi_i}((d\phi)^\sharp)^\flat &= \mathcal{L}_{\xi_i} d\phi\\ &= \xi_i \lrcorner\, d^2\phi + d(\xi_i \lrcorner\, d\phi)\\ &= d(\xi_i \lrcorner\, d\phi)\,, \quad \text{since} \quad d^2\phi = 0\,.\end{aligned} \tag{3.33}$$

Returning to Euclidean coordinates, we see that this corresponds to

$$(\boldsymbol{\xi}_i\cdot\nabla)\nabla\phi + \sum_{j=1}^{3}(\partial_j\phi)\nabla\xi_i^j = \nabla(\boldsymbol{\xi}_i\cdot\nabla\phi)\,, \tag{3.34}$$

where the left hand side is the Lie derivative of a 1-form and the right hand side is the exterior derivative of the interior product between a vector field, ξ_i, and a 1-form, $u^\flat$, associated to another vector field, u, through the musical isomorphism $\flat$. Whilst this calculation follows immediately from Cartan's formula in exterior calculus, it may also be performed, with some difficulty, in Euclidean coordinates using vector calculus.

As a result of this calculation, we have the following

$$\mathrm{d}\nabla\phi + \frac{1}{2}\nabla(|\nabla\phi|^2)\,dt + \sum_{i=1}^{\infty}\nabla(\boldsymbol{\xi}_i\cdot\nabla\phi)\circ dW_t^i + \boldsymbol{g}\,dt + \nabla\mathrm{d}\pi = 0\,.$$

The pressure here has the form discussed in Section 2.2.1, and can thus be written in the form $\mathrm{d}\pi = P_0\,dt + \sum P_i \circ dW_t^i$. In the deterministic case, it was assumed that the pressure was constant in time and space along the free surface. In the stochastic case, both the drift and diffusion parts of the pressure must have such an assumption. This, roughly speaking, can be interpreted as suggesting that the deterministic part of the pressure inherits its structure from the deterministic assumption, and we have assumed that there is no perturbation around this. We therefore have the 'Bernoulli form' of the stochastic Euler equation

$$\mathrm{d}\phi + \frac{1}{2}|\nabla\phi|^2\,dt + \sum_{i=1}^{\infty}\boldsymbol{\xi}_i\cdot\nabla\phi\circ dW_t^i + gz\,dt = 0\,. \tag{3.35}$$

We immediately see promise at this stage in the calculation, since the stochastic term is the Lie derivative of the scalar velocity potential along the vector field ξ_i.

As is done in the deterministic case, we wish to evaluate this onto the free surface. In the deterministic case, this is achieved simply by using the relationships from Proposition 3.4. In the stochastic case, this is more involved. We begin by considering the map, Z_t, as in Definition 3.2. Recall that this map has the property that the evaluation of the free surface is a pullback by this map

$$Z^*f = (f\circ Z)(x,y,z) = f(x,y,\zeta(t,x,y)) = \widehat{f}\,, \tag{3.36}$$

for an arbitrary function on the three dimensional spatial domain, f. We can therefore consider the evaluation onto the free surface by an application of the stochastic Kunita-Itô-Wentzell formula [11].

Proposition 3.10. *For a function, f, which is compatible with the driving semimartingale as*

$$\mathrm{d}f = F_0\, dt + \sum_i F_i \circ dW_t^i \,,$$

the evaluation of the function on the free surface, $\widehat{f}$, satisfies

$$\mathrm{d}\widehat{f} - \widehat{\partial_z f}\, \mathrm{d}\zeta = \widehat{F_0}\, dt + \sum_i \widehat{F}_i \circ dW_t^i \,. \tag{3.37}$$

Remark 3.11. This is a stochastic generalisation of the relationship from Proposition 3.4 corresponding to the derivative in the time variable.

Proof. Recall that the free surface, in the stochastic case, satisfies equation (3.28) and Z therefore satisfies

$$\mathrm{d}Z_t = G_0\, dt + \sum_i G_i \circ dW_t^i =: \begin{pmatrix} 0 \\ 0 \\ \widehat{\boldsymbol{v}} \cdot \boldsymbol{n} \end{pmatrix} dt + \sum_i \begin{pmatrix} 0 \\ 0 \\ \widehat{\boldsymbol{\xi}}_i \cdot \boldsymbol{n} \end{pmatrix} \circ dW_t^i \,,$$

where the vector field $G = \boldsymbol{G} \cdot \nabla$ is defined through the above equation. The Kunita-Itô-Wentzell formula is therefore

$$\begin{aligned} \mathrm{d}(Z^* f) = Z^* F_0\, dt &+ \sum_{i=1}^{\infty} Z^* F_i \circ dW_t^i \\ &+ Z^* \mathcal{L}_{G_0} f\, dt + \sum_{i=1}^{\infty} Z^* \mathcal{L}_{G_i} f \circ dW_t^i \,. \end{aligned} \tag{3.38}$$

In this case, the Lie derivative of a function is a directional derivative, and hence

$$\mathrm{d}(Z^* f) = Z^* F_0 \, dt + \sum_{i=1}^{\infty} Z^* F_i \circ dW_t^i \\ + Z^*(\widehat{\boldsymbol{u}} \cdot \boldsymbol{n} \, \partial_z f) \, dt + \sum_{i=1}^{\infty} Z^*(\widehat{\boldsymbol{\xi}}_i \cdot \boldsymbol{n} \, \partial_z f) \circ dW_t^i \, .$$

Returning to the hat notation, we have that

$$\mathrm{d}\widehat{f} = \widehat{F_0} \, dt + \sum_i \widehat{F}_i \circ dW_t^i + \widehat{\partial_z f} \, \mathrm{d}\zeta \, , \tag{3.39}$$

and we have proven our claim. $\mathscr{Q}.\mathscr{E}.\mathscr{D}.$

As an immediate consequence to this, we may evaluate equation (3.35) onto the free surface to obtain

$$\mathrm{d}\widehat{\phi} - \widehat{\partial_z \phi} \, \mathrm{d}\zeta + \frac{1}{2} |\widehat{\nabla \phi}|^2 \, dt + \sum_{i=1}^{\infty} \widehat{\boldsymbol{\xi}}_i \cdot \widehat{\nabla \phi} \circ dW_t^i + g\zeta \, dt = 0 \, .$$

Substituting in the kinematic boundary condition equation (3.31), written in terms of the velocity potential, we have

$$\mathrm{d}\widehat{\phi} - \widehat{\partial_z \phi} \left(\widehat{\partial_z \phi} \, dt - \widehat{\nabla_{\boldsymbol{r}} \phi} \cdot \nabla_{\boldsymbol{r}} \zeta \, dt + \sum_{i=1}^{\infty} \big(\widehat{\xi_i^{(z)}} - \widehat{\boldsymbol{\xi}_i^{(\boldsymbol{r})}} \cdot \nabla_{\boldsymbol{r}} \zeta \big) \circ dW_t^i \right) \\ + \frac{1}{2} |\widehat{\nabla \phi}|^2 \, dt + \sum_{i=1}^{\infty} \widehat{\boldsymbol{\xi}}_i \cdot \widehat{\nabla \phi} \circ dW_t^i + g\zeta \, dt = 0 \, .$$

After cancellations we have

$$\mathrm{d}\widehat{\phi} + \frac{1}{2} |\widehat{\nabla_{\boldsymbol{r}} \phi}|^2 \, dt - \frac{1}{2} \widehat{\partial_z \phi}^2 \, dt + \widehat{\partial_z \phi} (\widehat{\nabla_{\boldsymbol{r}} \phi} \cdot \nabla_{\boldsymbol{r}} \zeta) \, dt + g\zeta \, dt \\ + \sum_{i=1}^{\infty} \left(\widehat{\boldsymbol{\xi}_i^{(\boldsymbol{r})}} \cdot \widehat{\nabla_{\boldsymbol{r}} \phi} + \widehat{\partial_z \phi} \big(\widehat{\boldsymbol{\xi}_i^{(\boldsymbol{r})}} \cdot \nabla_{\boldsymbol{r}} \zeta \big) \right) \circ dW_t^i = 0 \, , \tag{3.40}$$

and we see that the deterministic part of this equation is equivalent to the classical deterministic theory [106].

3.3.2 A stochastic Hamiltonian structure for water waves

We claim that our stochastic equations, (3.31) and (3.40), have a Hamiltonian formulation in the spirit of Bismut [13]. That is, there is a family of Hamiltonians, $\{H, H_1, H_2, \dots\}$, such that the equations (3.31) and (3.40) can be expressed as

$$\mathrm{d}\zeta = \frac{\delta H}{\delta \widehat{\phi}}\, dt + \sum_{i=1}^{\infty} \frac{\delta H_i}{\delta \widehat{\phi}} \circ dW_t^i\,, \quad \text{and} \quad \mathrm{d}\widehat{\phi} = -\frac{\delta H}{\delta \zeta}\, dt - \sum_{i=1}^{\infty} \frac{\delta H_i}{\delta \zeta} \circ dW_t^i\,.$$

As we will substantiate with the proof of Theorem 3.15, these Hamiltonians are given by

$$\begin{aligned} H &= \int \int_{-\infty}^{\zeta} \frac{1}{2} |\nabla \phi|^2 \, dz \, d^2 r + \frac{1}{2} g \int \zeta^2 \, d^2 r\,, \\ H_i &= \int \int_{-\infty}^{\zeta} \boldsymbol{\xi}_i \cdot \nabla \phi \, dz \, d^2 r\,. \end{aligned}$$

To demonstrate that this is true, we must first consider how to take variations of Hamiltonians of this form.

Remark 3.12. The key feature which complicates the calculation of variations of the above Hamiltonians is the fact that a variation of the free surface elevation deforms the potential yet, for ζ and $\widehat{\phi}$ to be canonically conjugate variables, we wish to keep one constant whilst taking variations with respect to the other. By the definition of evaluation of the potential on the free surface, $\widehat{\phi} = \phi(\boldsymbol{r}, \zeta(\boldsymbol{r}, t))$, it is evident that the variation in ζ will induce a variation in $\widehat{\phi}$, and it can be proposed that the form of this variation is $(\partial\phi/\partial z)\delta\zeta$. We prove this formally in the following proposition, which illuminates a connection between this and the Lie chain rule. A corollary to this proposition will illustrate how to take independent variations of our Hamiltonians.

Proposition 3.13. *Variations of the free surface elevation, ζ, induce a variation in the potential, $\delta_\zeta \phi$, given by*

$$\delta_\zeta \widehat{\phi} = \frac{\partial \phi}{\partial z} \delta\zeta \,, \quad on \quad z = \zeta \,. \tag{3.41}$$

Proof. This relationship follows from the *Lie chain rule* pullback relation introduced in Definition 2.19. Thus, in order to approach this problem, we notice that the evaluation on the free surface may be defined as a pullback by a smooth time dependent three dimensional map, $Z_t : \mathbb{R}^3 \mapsto \mathbb{R}^3$, defined as

$$Z_t(x, y, z) = (x, y, \zeta(t, x, y)) \,. \tag{3.42}$$

This implies that for an arbitrary function on the three dimensional spatial domain, f, we have

$$Z_t^* f = (f \circ Z_t)(x, y, z) = f(x, y, \zeta(t, x, y) = \widehat{f} \,. \tag{3.43}$$

We will be considering variations of Z_t only, and will study the effect this has on a functional of the pullback of ϕ by Z_t. We are therefore interested in the first variation of Z_t, defined as

$$\delta Z_t = \frac{d}{d\epsilon}\bigg|_{\epsilon=0} Z_{t,\epsilon} \,, \quad \text{where} \quad Z_{t,\epsilon} = g_\epsilon Z_t \,, \tag{3.44}$$

which is a vector field defined on the horizontal domain. Here, g_ϵ is a infinitesimal Lie group action of a diffeomorphism on $\mathbb{R}^3$. The above definition is motivated by the Taylor expansion of $Z_{t,\epsilon}$ around $\epsilon = 0$. We then seek a variational derivative of some functional of the composition, $F(Z_t^*\phi)$, where this can be interpreted as a functional of $\widehat{\phi}$. This variation, through its relation to the Taylor series of $Z_{t,\epsilon}^*\phi$, may be written as

$$\begin{aligned}
\delta F(Z_t^*\phi) &= \frac{d}{d\epsilon}\bigg|_{\epsilon=0} F(Z_t^*\phi + \epsilon\delta(Z_t^*\phi)) \\
&= \frac{d}{d\epsilon}\bigg|_{\epsilon=0} F\left(Z_{t,0}^*\phi + \epsilon\frac{d}{d\epsilon}\bigg|_{\epsilon=0}(Z_{t,\epsilon}^*\phi)\right) = \frac{d}{d\epsilon}\bigg|_{\epsilon=0} F(Z_{t,\epsilon}^*\phi)\,.
\end{aligned}$$

We consider further the final derivative in the above equation, before evaluation at $\epsilon = 0$. To ease legibility, we will define the pullback as a map depending on ϵ as $f(\epsilon) = Z_{t,\epsilon}^*\phi$, and we have

$$\begin{aligned}
\frac{d}{d\epsilon}F(f(\epsilon)) &= \lim_{h\to 0}\frac{F(f(\epsilon+h)) - F(f(\epsilon))}{h} \\
&= \lim_{h\to 0}\frac{F\left(f(\epsilon) + h\frac{d}{dh}\Big|_{h=0} f(\epsilon+h) + \mathcal{O}(h^2)\right) - F(f(\epsilon))}{h} \\
&=: \left\langle \frac{\delta F}{\delta f(\epsilon)}, \frac{d}{dh}\bigg|_{h=0} f(\epsilon+h)\right\rangle = \left\langle \frac{\delta F}{\delta f(\epsilon)}, \frac{d}{d\epsilon}f(\epsilon)\right\rangle\,.
\end{aligned}$$

Returning to the variation, the above calculation implies

$$\delta F(Z_t^*\phi) = \left\langle \frac{\delta F}{\delta(Z_{t,\epsilon}^*\phi)}, \frac{d}{d\epsilon}(Z_{t,\epsilon}^*\phi)\right\rangle\bigg|_{\epsilon=0} = \left\langle \frac{\delta F}{\delta(Z_{t,\epsilon}^*\phi)}, Z_{t,\epsilon}^*(\mathcal{L}_X\phi)\right\rangle\bigg|_{\epsilon=0}, \tag{3.45}$$

where the final equality makes use of the *Lie chain rule* (as defined in Definition 2.19) and $X = \frac{d}{d\epsilon}\big|_{\epsilon=0} g_\epsilon$ is the vector field which generates the flow. Having evaluated the equation (3.45) at $\epsilon = 0$, we find that

$$\delta F(Z_t^*\phi) =: \left\langle \frac{\delta F}{\delta(Z_t^*\phi)}, \delta_\zeta(Z_t^*\phi)\right\rangle = \left\langle \frac{\delta F}{\delta(Z_t^*\phi)}, Z_t^*(\mathcal{L}_X\phi)\right\rangle\,, \tag{3.46}$$

where the notation δ_ζ is used to reflect the fact that this has resulted from the variation of the free surface, and that the variation of the potential directly has not been considered. Noticing that $\delta Z_t = \frac{d}{d\epsilon}\big|_{\epsilon=0} Z_{t,\epsilon} = Z_t^* X$, and that the Lie derivative is the directional derivative in local coordinates, we have that

$$\delta_\zeta(Z_t^*\phi) = Z_t^*(X \cdot \nabla\phi) = (Z_t^* X) \cdot Z_t^*(\nabla\phi) = \delta\zeta Z_t^*(\partial_z \phi)\,, \tag{3.47}$$

and we have proven our claim. $\mathscr{Q.E.D.}$

Corollary 3.14. *When considering variations of the water wave Hamiltonians with respect to the free surface ζ, we must also vary the potential according to*

$$\delta\phi = -\frac{\partial\phi}{\partial z}\delta\zeta\,, \quad \textit{on} \quad z = \zeta\,. \tag{3.48}$$

in order to ensure that the canonically conjugate variable, $\widehat{\phi}$, is untouched by the variation in ζ.

Proof. The aim is to vary ϕ and ζ in such a manner that $Z_t^*\phi = \widehat{\phi}$ is constant. Considering $\phi_\epsilon(\boldsymbol{r}, z) = \phi(\boldsymbol{r}, z) + \epsilon\delta\phi(\boldsymbol{r}, z)$ and $\zeta_\epsilon(\boldsymbol{r}) = \zeta(\boldsymbol{r}) + \epsilon\delta\zeta(\boldsymbol{r})$, where the time dependence is not explicitly notated for brevity, then the composition $\phi_\epsilon(\boldsymbol{r}, \zeta_\epsilon) = Z_{t,\epsilon}^*\phi_\epsilon$ must be such that

$$\phi_\epsilon(\boldsymbol{r}, \zeta_\epsilon) = \phi(\boldsymbol{r}, \zeta) = \widehat{\phi}\,, \quad \text{for each} \quad \epsilon\,. \tag{3.49}$$

This is equivalent to the assertion that the canonically conjugate variable, $\widehat{\phi}$, should not be altered by the variation of ζ. Mathematically, the first variation of $\widehat{\phi} = Z_t^*\phi$, when both ϕ and ζ are varied, has a contribution from the variation in ϕ directly as well as a contribution from the variation in ζ which is given in Proposition 3.13. That is,

$$\phi(\boldsymbol{r}, \zeta) + \epsilon\left(\frac{\partial\phi}{\partial z}\right)(\boldsymbol{r}, \zeta)\delta\zeta + \epsilon\delta\phi(\boldsymbol{r}, \zeta) + \mathcal{O}(\epsilon^2) = \widehat{\phi}\,, \tag{3.50}$$

which implies equation (3.48) and we have proven our claim.

Note that equation (3.50) may also be found by considering the Taylor expansions of $\phi(\boldsymbol{r}, \zeta_\epsilon)$ and $\delta\phi(\boldsymbol{r}, \zeta_\epsilon)$ and substituting these into $\phi_\epsilon(\boldsymbol{r}, \zeta_\epsilon)$. $\mathscr{Q}.\mathscr{E}.\mathscr{D}.$

This corollary may be used to prove that our stochastic classical water wave system is Hamiltonian.

Theorem 3.15. *The equations* (3.31) *and* (3.40) *have a Hamiltonian structure*

$$\mathrm{d}\zeta = \frac{\delta H}{\delta \widehat{\phi}}\,dt + \sum_{i=1}^{\infty} \frac{\delta H_i}{\delta \widehat{\phi}} \circ dW_t^i\,, \tag{3.51}$$

$$\mathrm{d}\widehat{\phi} = -\frac{\delta H}{\delta \zeta}\,dt - \sum_{i=1}^{\infty} \frac{\delta H_i}{\delta \zeta} \circ dW_t^i\,, \tag{3.52}$$

where the family of Hamiltonians are given by

$$H = \int\int_{-\infty}^{\zeta} \frac{1}{2}|\nabla\phi|^2\,dz\,d^2r + \frac{1}{2}g\int \zeta^2\,d^2r\,, \tag{3.53}$$

$$H_i = \int\int_{-\infty}^{\zeta} \boldsymbol{\xi}_i \cdot \nabla\phi\,dz\,d^2r\,. \tag{3.54}$$

Proof. For this to be true, we will need to demonstrate that the variational derivatives of these Hamiltonians are as follows

$$\frac{\delta H}{\delta \widehat{\phi}} = \boldsymbol{n} \cdot \widehat{\nabla\phi}\,, \tag{3.55}$$

$$\frac{\delta H}{\delta \zeta} = \frac{1}{2}|\widehat{\nabla_{\boldsymbol{r}}\phi}|^2 - \frac{1}{2}\widehat{\partial_z\phi}^2 + \widehat{\partial_z\phi}\widehat{\nabla_{\boldsymbol{r}}\phi} \cdot \nabla_{\boldsymbol{r}}\zeta + g\zeta\,, \tag{3.56}$$

$$\frac{\delta H_i}{\delta \widehat{\phi}} = \boldsymbol{n} \cdot \widehat{\boldsymbol{\xi}}_i\,, \tag{3.57}$$

$$\frac{\delta H_i}{\delta \zeta} = \widehat{\boldsymbol{\xi}_i^{(r)}} \cdot \widehat{\nabla_{\boldsymbol{r}}\phi} + \widehat{\partial_z\phi}\big(\widehat{\boldsymbol{\xi}_i^{(r)}} \cdot \nabla_{\boldsymbol{r}}\zeta\big)\,. \tag{3.58}$$

Notice that the first two variational derivatives are akin to those found by Zakahrov [106], we will use the same method here. We begin with the variation of H with respect to $\widehat{\phi}$. Since the velocity potential, ϕ, is a harmonic function, we may use Green's first identity on the kinetic energy term

$$\frac{1}{2}\int\int_{-\infty}^{\zeta}|\nabla\phi|^2\,dz\,d^2r=\frac{1}{2}\int\phi(\nabla\phi\cdot\boldsymbol{n})\,d^2r\,. \tag{3.59}$$

Note that the normal, $\boldsymbol{n}$, given by equation (3.5) is not a unit normal, but the factor through which it may be transformed into a unit normal also appears in the following expression for an infinitesimal region of the free surface, $ds=\sqrt{1+|\nabla_{\boldsymbol{r}}\zeta|^2}\,d^2r$. The integral on the right hand side is taken to be over the free surface since the normal component of velocity is assumed to vanish on all other boundaries. The existence of a symmetric Green's function relating $\widehat{\phi}$ and $\widehat{\nabla\phi}\cdot\boldsymbol{n}$ follows from the Dirichlet to Neumann map and, as in Zakharov [106], this implies the variational derivative (3.55).

The variational derivative of H with respect to ζ is trivial for the potential energy, and for the kinetic energy follows from the approach discussed in Corollary 3.14. A variation of the kinetic energy gives

$$\begin{aligned}
\frac{1}{2}\int\int_{-\infty}^{\zeta+\delta\zeta}|\nabla(\phi+\delta\phi)|^2 &= \frac{1}{2}\int\left[|\nabla\phi|^2\right]\delta\zeta\,d^2r+\int\int_{-\infty}^{\zeta}\nabla\phi\cdot\nabla\delta\phi\,dz\,d^2r\\
\text{(by Green's second identity)} &= \frac{1}{2}\int\left[|\nabla\phi|^2\right]_{z=\zeta}\delta\zeta\,d^2r+\int\left[\delta\phi(\nabla\phi\cdot\boldsymbol{n})\right]_{z=\zeta}\,d^2r\\
\text{(by Corollary 3.14)} &= \frac{1}{2}\int\left[|\nabla\phi|^2\right]_{z=\zeta}\delta\zeta\,d^2r\\
&\quad-\int\widehat{\partial_z\phi}\,\delta\zeta(\widehat{\partial_z\phi}-\widehat{\nabla_{\boldsymbol{r}}\phi}\cdot\nabla_{\boldsymbol{r}}\zeta)\,d^2r\\
&= \int\left(\frac{1}{2}|\widehat{\nabla_{\boldsymbol{r}}\phi}|^2-\frac{1}{2}(\widehat{\partial_z\phi})^2+\widehat{\partial_z\phi}\widehat{\nabla_{\boldsymbol{r}}\phi}\cdot\nabla_{\boldsymbol{r}}\zeta\right)\delta\zeta\,d^2r\,.
\end{aligned}$$

This implies the required variational derivative (3.56).

The variational derivatives of the stochastic Hamiltonians, H_i, are performed similarly. Beginning with the variational derivative of H_i with respect to $\widehat{\phi}$. Rather than Green's identity, we use the divergence theorem. Noting that

$$\nabla \cdot (\phi \boldsymbol{\xi}_i) = \boldsymbol{\xi}_i \cdot \nabla \phi + \phi \nabla \cdot \boldsymbol{\xi}_i = \boldsymbol{\xi}_i \cdot \nabla \phi \,,$$

where we have used the fact that $\boldsymbol{\xi}_i$ are divergence free. The divergence theorem implies that

$$\begin{aligned} H_i = \int \int_{-\infty}^{\zeta} \boldsymbol{\xi}_i \cdot \nabla \phi \, dz \, d^2 r &= \oint_{z=\zeta} \phi(\boldsymbol{\xi}_i \cdot \boldsymbol{n}) \frac{1}{\sqrt{1 + |\nabla_{\boldsymbol{r}} \zeta|^2}} \, ds \\ &= \int \phi(\boldsymbol{\xi}_i \cdot \boldsymbol{n}) \, d^2 r \,. \end{aligned} \tag{3.60}$$

The justification of this is the same as for the variation of H and, since $\boldsymbol{\xi}_i$ are independent of ϕ, this immediately implies the variational derivative (3.57).

It only remains to calculate the variational derivative of H_i with respect to ζ. This again invokes Corollary 3.14 and closely follows the deterministic case, indeed

$$\begin{aligned} \int \int_{-\infty}^{\zeta + \delta \zeta} \boldsymbol{\xi}_i \cdot \nabla (\phi + \delta \phi) \, dz \, d^2 r &= \int \widehat{\boldsymbol{\xi}}_i \cdot \widehat{\nabla \phi} \, \delta \zeta \, d^2 r \\ &\qquad + \int \int_{\zeta}^{\zeta + \delta \zeta} \boldsymbol{\xi}_i \cdot \nabla \delta \phi \, dz \, d^2 r \\ &= \int \widehat{\boldsymbol{\xi}}_i \cdot \widehat{\nabla \phi} \, \delta \zeta \, d^2 r + \int \left[\delta \phi (\boldsymbol{\xi}_i \cdot \boldsymbol{n}) \right]_{z=\zeta} \, d^2 r \,. \end{aligned}$$

The final line of this calculation follows again from divergence theorem, since the divergence of $\boldsymbol{\xi}_i$ is zero. Continuing the calculation, we see that

$$\begin{aligned}\int\int_{-\infty}^{\zeta+\delta\zeta} \boldsymbol{\xi}_i \cdot \nabla(\phi+\delta\phi)\,dz\,d^2r &= \int \widehat{\boldsymbol{\xi}}_i \cdot \widehat{\nabla\phi}\,\delta\zeta\,d^2r \\ &\quad - \int \widehat{\partial_z\phi}\,\delta\zeta(\widehat{\xi_i^{(z)}} - \widehat{\boldsymbol{\xi}_i^{(\boldsymbol{r})}} \cdot \nabla_{\boldsymbol{r}}\zeta)\,d^2r \\ &= \int \left(\widehat{\boldsymbol{\xi}_i^{(\boldsymbol{r})}} \cdot \widehat{\nabla_{\boldsymbol{r}}\phi} + \widehat{\partial_z\phi}\big(\widehat{\boldsymbol{\xi}_i^{(\boldsymbol{r})}} \cdot \nabla_{\boldsymbol{r}}\zeta\big)\right)\delta\zeta\,d^2r\,,\end{aligned}$$

which gives our result. $\mathscr{Q}.\mathscr{E}.\mathscr{D}.$

Remark 3.16. Notice that if we set $\boldsymbol{\xi}_i$ to be zero, this recovers the deterministic theory exactly.

3.3.3 The Dirichlet to Neumann map

We will rearrange the equations such that they are written in terms of the free surface and trace of the potential on the free surface only. To do so, we use relationships from Proposition 3.4 to rewrite the deterministic part of equation (3.40) as

$$\begin{aligned}\frac{1}{2}|\widehat{\nabla_{\boldsymbol{r}}\phi}|^2 - \frac{1}{2}\widehat{\partial_z\phi}^2 + \widehat{\partial_z\phi}(\widehat{\nabla_{\boldsymbol{r}}\phi}\cdot\nabla_{\boldsymbol{r}}\zeta) + g\zeta &= \frac{1}{2}|\nabla_{\boldsymbol{r}}\widehat{\phi} - \widehat{\partial_z\phi}\nabla_{\boldsymbol{r}}\zeta|^2 - \frac{1}{2}\widehat{\partial_z\phi}^2 \\ &\quad + \widehat{\partial_z\phi}(\nabla_{\boldsymbol{r}}\widehat{\phi} - \widehat{\partial_z\phi}\nabla_{\boldsymbol{r}}\zeta)\cdot\nabla_{\boldsymbol{r}}\zeta + g\zeta \\ &= g\zeta + \frac{1}{2}|\nabla_{\boldsymbol{r}}\widehat{\phi}|^2 - \frac{1}{2}\widehat{\partial_z\phi}^2(1 + |\nabla_{\boldsymbol{r}}\zeta|^2)\end{aligned}$$

The stochastic part of equation (3.40) may be rearranged as

$$\begin{aligned}\widehat{\boldsymbol{\xi}_i^{(\boldsymbol{r})}} \cdot \widehat{\nabla_{\boldsymbol{r}}\phi} + \widehat{\partial_z\phi}\big(\widehat{\boldsymbol{\xi}_i^{(\boldsymbol{r})}}\cdot\nabla_{\boldsymbol{r}}\zeta\big) &= \widehat{\boldsymbol{\xi}_i^{(\boldsymbol{r})}}\cdot(\nabla_{\boldsymbol{r}}\widehat{\phi} - \widehat{\partial_z\phi}\nabla_{\boldsymbol{r}}\zeta) + \widehat{\partial_z\phi}\big(\widehat{\boldsymbol{\xi}_i^{(\boldsymbol{r})}}\cdot\nabla_{\boldsymbol{r}}\zeta\big) \\ &= \widehat{\boldsymbol{\xi}_i^{(\boldsymbol{r})}}\cdot\nabla_{\boldsymbol{r}}\widehat{\phi}\,.\end{aligned}$$

As in Craig and Sulem [29], both the deterministic part of equation (3.40) and the kinematic boundary condition can be written in terms of the *Dirichlet to Neumann map*. This is convenient since it enables numerical integration, as well as allowing the consideration of an asymptotic expansion of the map. Given that the potential satisfies Laplace's equation in the bulk of the fluid, the map takes the Dirichlet boundary data and returns the Neumann boundary condition which corresponds to the same solution. This map therefore takes the trace of the potential, $\widehat{\phi}$, and returns the velocity in the normal direction at the surface, $\boldsymbol{n}\cdot\widehat{\boldsymbol{u}}$. The map can be written in multiple equivalent forms as

$$\begin{aligned} G(\zeta)\widehat{\phi} &:= (-\nabla_{\boldsymbol{r}}\zeta, 1)\cdot\widehat{\nabla\phi} = -\nabla_{\boldsymbol{r}}\zeta\cdot\widehat{\nabla_{\boldsymbol{r}}\phi} + \widehat{\partial_z\phi} \\ &= -\nabla_{\boldsymbol{r}}\zeta\cdot\nabla_{\boldsymbol{r}}\widehat{\phi} + \widehat{\partial_z\phi}|\nabla_{\boldsymbol{r}}\zeta|^2 + \widehat{\partial_z\phi}\,. \end{aligned} \tag{3.61}$$

The stochastic kinematic boundary condition (3.31) may be rewritten as

$$\mathrm{d}\zeta = G(\zeta)\widehat{\phi}\,dt + \sum_{i=1}^{\infty}\big(\widehat{\xi_i^{(z)}} - \widehat{\boldsymbol{\xi}_i^{(\boldsymbol{r})}}\cdot\nabla_{\boldsymbol{r}}\zeta\big)\circ dW_t^i\,. \tag{3.62}$$

The stochastic Bernoulli boundary equation (3.40) becomes

$$\begin{aligned} \mathrm{d}\widehat{\phi} + g\zeta\,dt + \frac{1}{2}|\nabla_{\boldsymbol{r}}\widehat{\phi}|^2\,dt - \frac{1}{2\big(1+|\nabla_{\boldsymbol{r}}\zeta|^2\big)}\big(G(\zeta)\widehat{\phi} + \nabla_{\boldsymbol{r}}\zeta\cdot\nabla_{\boldsymbol{r}}\widehat{\phi}\,\big)^2\,dt& \\ + \sum_{i=1}^{\infty}\left(\widehat{\boldsymbol{\xi}_i^{(\boldsymbol{r})}}\cdot\nabla_{\boldsymbol{r}}\widehat{\phi}\right)\circ dW_t^i = 0\,.& \end{aligned} \tag{3.63}$$

The pair of equations (3.62) and (3.63) are a closed system of SPDEs for the water wave problem, which is a stochastic generalisation of that found by Craig and Sulem [29]. As has been noted in the deterministic case, the Hamiltonian (3.53) found by Zakharov may be rewritten in terms of the

Dirichlet to Neumann map as

$$H = \frac{1}{2}\int \widehat{\phi}\, G(\zeta)\widehat{\phi} + g\zeta^2 \, d^2r \,. \tag{3.64}$$

The equivalence of these Hamiltonians follows from applying Green's first identity to the kinetic energy term. Indeed, since ϕ is a harmonic function, we have

$$\frac{1}{2}\int\int_{-\infty}^{\zeta} |\nabla\phi|^2 \, dx\, d^2r = \frac{1}{2}\int \widehat{\phi}(\widehat{\nabla\phi}\cdot \boldsymbol{n})\, d^2r \,,$$

noting the relationship between the normal vector, $\boldsymbol{n}$, its associated unit normal, and the infinitesimal surface element, $ds = \sqrt{1+|\nabla_{\boldsymbol{r}}\zeta|^2}\, d^2r$, as discussed in the proof of Theorem 3.15.

For the Hamiltonians corresponding to the stochastic terms, H_i, we have that

$$H_i = \int \widehat{\phi}(\widehat{\boldsymbol{\xi}}_i \cdot \boldsymbol{n})\, d^2r \,. \tag{3.65}$$

This follows from the divergence theorem, and can also be found in the proof of Theorem 3.15.

We have therefore found that our stochastic extension to the classical water wave equations can be written purely in terms of the canonically conjugate variables, $\widehat{\phi}$ and ζ. Furthermore, its Hamiltonians may also be expressed in this way.

3.3.4 On the structure of the noise

Thus far, we have been working under the assumption that the deterministic part of the transport, $\boldsymbol{u}$, is irrotational. We have made no further comment on the structure of the stochastic perturbations, $\boldsymbol{\xi}_i$, which are assumed to have the same divergence-free form as in the Euler equations. Whilst this means that the large scale flow is irrotational, the whole dynamical portrait encompasses small scale stochastic motions which may have nonzero vorticity. This is desirable, since the lack of vorticity in the deterministic picture is a significant limiting factor.

If we make a further assumption that the noise terms are also irrotational, and each can be written in terms of a potential as

$$\boldsymbol{\xi}_i = \nabla \varphi_i \,,$$

then we see that the stochastic terms can, too, be written in terms of the Dirichlet to Neumann operator. Indeed, the stochastic terms in equation (3.63) become

$$\begin{aligned}\widehat{\boldsymbol{\xi}_i^{(\boldsymbol{r})}} \cdot \nabla_{\boldsymbol{r}} \widehat{\phi} &= \widehat{\nabla_{\boldsymbol{r}} \varphi_i} \cdot \nabla_{\boldsymbol{r}} \widehat{\phi} \\ &= \nabla_{\boldsymbol{r}} \widehat{\phi} \cdot \left(\nabla_{\boldsymbol{r}} \widehat{\varphi}_i - \widehat{\partial_z \varphi_i} \nabla_{\boldsymbol{r}} \zeta \right) \\ &= \nabla_{\boldsymbol{r}} \widehat{\phi} \cdot \nabla_{\boldsymbol{r}} \widehat{\varphi}_i - \frac{\nabla_{\boldsymbol{r}} \widehat{\phi} \cdot \nabla_{\boldsymbol{r}} \zeta}{1 + |\nabla_{\boldsymbol{r}} \zeta|^2} \big(G(\zeta) \widehat{\varphi}_i - \nabla_{\boldsymbol{r}} \widehat{\varphi}_i \cdot \nabla_{\boldsymbol{r}} \zeta \big) \,.\end{aligned}$$

The stochastic classical water wave equation (3.63) can therefore be rewritten, fully in terms of the Dirichlet to Neumann map, as

$$\begin{aligned}\mathrm{d}\widehat{\phi} + g\zeta \, dt + \frac{1}{2} |\nabla_{\boldsymbol{r}} \widehat{\phi}|^2 \, dt - \frac{1}{2\big(1 + |\nabla_{\boldsymbol{r}} \zeta|^2\big)} \big(G(\zeta) \widehat{\phi} + \nabla_{\boldsymbol{r}} \zeta \cdot \nabla_{\boldsymbol{r}} \widehat{\phi} \,\big)^2 \, dt \\ + \sum_{i=1}^{\infty} \nabla_{\boldsymbol{r}} \widehat{\phi} \cdot \nabla_{\boldsymbol{r}} \widehat{\varphi}_i - \frac{\nabla_{\boldsymbol{r}} \widehat{\phi} \cdot \nabla_{\boldsymbol{r}} \zeta}{1 + |\nabla_{\boldsymbol{r}} \zeta|^2} \big(G(\zeta) \widehat{\varphi}_i - \nabla_{\boldsymbol{r}} \widehat{\varphi}_i \cdot \nabla_{\boldsymbol{r}} \zeta \big) \circ dW_t^i = 0 \,.\end{aligned} \tag{3.66}$$

Similarly, the kinematic boundary condition (3.28) is

$$\mathrm{d}\zeta + G(\zeta)\widehat{\phi}\, dt + \sum_{i=1}^{\infty} G(\zeta)\widehat{\varphi}_i \circ dW_t^i \,. \tag{3.67}$$

To further illustrate that this stochastic perturbation of the water wave problem preserves the geometric structure of the deterministic case, we note that the Hamiltonians, H_i, defined in equation (3.54) can be rewritten in terms of the Dirichlet to Neumann map in the same manner as the deterministic Hamiltonian (3.53) was transformed into an equivalent form (3.64). Indeed, again using Green's first identity we have

$$H_i = \int \widehat{\phi}\, G(\zeta)\widehat{\varphi}_i \, d^2r \,. \tag{3.68}$$

Due to the properties of the Dirichlet-to-Neumann map, this may be beneficial in some cases. It should be noted that this further assumption on the structure of the noise is not required for a Hamiltonian structure to exist, and making this assumption destroys the hope of vorticity within the stochastic terms. This should be considered carefully since if we are calibrating the stochastic terms using data, then it is unlikely that these will be irrotational.

3.3.5 Limitations of the classical water wave system

The classical water wave equations are adequate for a number of applications. By considering expansions of the Dirichlet to Neumann map, a broad class of wave equations can be found. However, they are not a perfect closure of the system and are unsuited to many problems considered by mathematicians and physicists today. In particular, the potential flow assumption is often a poor approximation in reality and the equations cannot support thermal gradients and inhomogeneity. These limitations have been noted and attempts have been made to circumvent them, in particular Castro and Lannes [19] proposed an extension to the classical theory to allow for vorticity in the bulk flow. The inability to apply these modelling approaches to fluid problems with thermal gradients is severely limiting when considering geophysical flows. Contemporary satellite oceanographic data contains a plethora of such features which must be interpreted, however a number of these features remain unexplained. Indeed, improvements in the quality and availability of imagery of the ocean have revealed that the submesoscale ocean is a rich dynamical landscape, containing complex features that were not imagined until photographed. Among these are the famous 'spirals on the sea', studied by Munk et al. [79], which were observed by such photographs throughout the second half of the 20th century. Despite being referred to as 'perhaps the most fundamental entity in ocean dynamics' at their scale by Scully-Power [93], the spirals on the sea remain elusive, like many other such interactions between currents and surface disturbances on the sea surface. Understanding these complex phenomena would be aided by models which feature a more complete set of variables than the classical water wave equations.

3.4 A variational principle for free surface dynamics

As motivated above, we seek a new modelling framework in which we can incorporate thermal gradients and understand the relationships between waves and currents observed on the free surface. In order to do this, the model we seek will be written only in terms of *observable* quantities. We will be deriving and presenting the model introduced in Holm, Hu, and Street [57], before illustrating its geometric properties. The dynamics will take place on a two dimensional spatial domain, $\mathcal{D} \in \mathbb{R}^2$.

The observable quantities in which we are interested are the velocity evaluated on the free surface, $(\widehat{\boldsymbol{v}}, \widehat{w})$, the free surface elevation, ζ, and the advected buoyancy, ρ, associated with horizontal thermal gradients. As discussed earlier in this section, the vertical components of the dynamics, $\widehat{w}$ and ζ, will be considered as *wave variables*, and the *currents* will be represented by the horizontal velocity field, $\widehat{v} = \widehat{\boldsymbol{v}} \cdot \nabla_{\boldsymbol{r}} \in \mathfrak{X}(\mathcal{D})$, and the variables associated with its evolution. Through this approach, currents are understood to be flows which transport physical properties, and waves are disturbances on the surface which propagate due to a restoring force.

3.4.1 A composition of maps

The characterisation of the evaluation of a variable onto the free surface as a pullback by the map Z_t, as defined in Proposition 3.13, introduces the idea of decomposing the flow into two maps. Namely, we consider a decomposition into a flow map $\Phi_t : \mathbb{R}^2 \mapsto \mathbb{R}^2$, for the two dimensional currents and a composition of maps, $\zeta_t \Phi_t : \mathbb{R}^2 \mapsto \mathbb{R}$, for the vertical free surface elevation. We have $Z_t = (\Phi_t \boldsymbol{r}_0, \zeta_t \Phi_t \boldsymbol{r}_0)$ or, in components,

$$\boldsymbol{r}_t = \Phi_t \boldsymbol{r}_0 \,, \quad \text{and} \quad z_t = \zeta_t \Phi_t \boldsymbol{r}_0 = \zeta_t \boldsymbol{r}_t \,. \tag{3.69}$$

For a horizontal spatial domain, $\mathcal{D} \subset \mathbb{R}^2$, we have that the flow map is a diffeomorphism, $\Phi_t \in \mathrm{Diff}(\mathcal{D})$, and the free surface elevation is a function, $\zeta \in \mathcal{F}(\mathcal{D})$. The horizontal flow map, Φ_t, acts as usual for a two dimensional fluid, where $\boldsymbol{r}_0$ is a Lagrangian coordinate. The vertical map, ζ_t, acts on the curve $\Phi_t \boldsymbol{r}_0$ in the two dimensional plane and the two dimensional dynamics acts as the Lagrangian coordinate for the vertical motion. Taking time derivatives, we have

$$
\begin{aligned}
\frac{d\boldsymbol{r}_t}{dt} &= \frac{d}{dt}(\Phi_t \boldsymbol{r}_0) = \widehat{v}_t(\Phi_t \boldsymbol{r}_0) =: \widehat{v}_t(\boldsymbol{r}_t)\,, \qquad \Longrightarrow \qquad \widehat{v}_t = \frac{d\Phi_t}{dt}\Phi_t^{-1}\,,\\
\frac{dz_t}{dt} &=: \widehat{w}_t(\boldsymbol{r}_t) = \frac{d}{dt}\big(\zeta_t(\Phi_t \boldsymbol{r}_0)\big) = \partial_t \zeta_t(\boldsymbol{r}_t) + \nabla_{\boldsymbol{r}}\zeta_t(\boldsymbol{r}_t)\cdot \widehat{\boldsymbol{v}}_t(\boldsymbol{r}_t)\,.
\end{aligned}
$$

Notice that the first line of this is analogous to equation (2.23), meaning that the horizontal flow has the standard geometric structure of a fluid motion. The second line recovers the kinematic boundary condition (3.6). From a modelling perspective, these equations say that *if* we have a suitable two dimensional model which models the two dimensional flow on the free surface, *then* the composition of maps approach allows us to model how such a flow affects the disturbances propagating on its surface. It is worth noting that, without the inclusion of additional terms, this approach will not be able to determine how those waves will affect the currents. In ocean scale flows, it is not unreasonable to postulate that the small and fast oscillations on the surface are affected more by the large and comparatively slow currents than vice versa. In general, however, the existence of a two dimensional model which fits the data available for flows on the free surface of the ocean is an open problem. In some regions of the ocean, this flow is approximately incompressible whereas in others, for example near strong thermal fronts, the ocean currents are convergent or divergent due to the effects of downwelling. Nonetheless, in what follows the two dimensional flow will be taken to be incompressible and, despite the fact we will have horizontal gradients of thermal buoyancy, will still enable us to understand the mechanisms at play when waves interact with strong thermal fronts. Due to this, it is unlikely that the model simulations will reproduce perfect replications of ocean flows since the currents at thermal gradients will, in general, have a different structure. This approach will instead attempt to reveal the interactions of the wave dynamics with the temperature gradients, from a mathematical perspective.

Approaches similar to this one have been employed in a variational setting to other problems with multiple interacting dynamical features, including superfluids [58], complex fluids [49], and wave current interaction [55]. The same structure may be applied to the dynamics of a swinging spring, as will be seen in an upcoming work.

THE GEOMETRY OF TWO MAPS. Given a Lagrangian, L, which depends on both the flow map, $\Phi^{\dagger}$, and the composition of maps, $\zeta \circ \Phi$, it reduces under right invariance as in the traditional setting. A variable which is advected by the currents is given by a, where, $a_t = a_0 \Phi_t^{-1}$. Note that this differs from the introduction of a reduced Lagrangian for semidirect product Lie algebras in Section 2.1.2, as in equation (2.20), in that the action of the inverse map is a right action. We may define the reduced action, ℓ, by using the right invariance of the Lagrangian as follows

$$
\begin{aligned}
L\left(\Phi, \frac{\partial \Phi}{\partial t}, \zeta \circ \Phi, \frac{\partial(\zeta \circ \Phi)}{\partial t}, a_0\right) &= L\left(e, \frac{\partial \Phi}{\partial t}\Phi^{-1}, (\zeta \circ \Phi)\Phi^{-1}, \frac{\partial(\zeta \circ \Phi)}{\partial t}\Phi^{-1}, a_t\right) \\
&= L(e, \widehat{v}, \zeta, (\partial_t + \mathcal{L}_{\widehat{v}})\zeta, a) =: \ell(\widehat{v}, \zeta, \widehat{w}, a)\,.
\end{aligned}
$$

We therefore have a reduced Lagrangian to which we may apply Hamilton's principle, noting that in this case the variations must have the following

†The capital greek letter, Φ, is used for the flow map to distinguish it from the velocity potential, ϕ, in previous sections.

forms

$$\delta\widehat{v} = \partial_t \eta - \mathrm{ad}_{\widehat{v}}\eta\,, \quad \delta\widehat{w} = (\partial_t + \mathcal{L}_{\widehat{v}})\delta\zeta + \mathcal{L}_{\delta\widehat{v}}\zeta\,, \quad \text{and} \quad \delta a = -\mathcal{L}_\eta a\,,$$

where $\eta \in \mathfrak{X}(\mathcal{D})$ vanishes at the endpoints, and the form of the variation in $\widehat{w}$ is inherited from the kinematic boundary condition. We may derive Euler-Poincaré equations by taking variations as follows

$$\begin{aligned}
0 &= \int \left\langle \frac{\delta \ell}{\delta \widehat{v}}, \delta\widehat{v} \right\rangle + \left\langle \frac{\delta \ell}{\delta \zeta}, \delta\zeta \right\rangle + \left\langle \frac{\delta \ell}{\delta \widehat{w}}, \delta\widehat{w} \right\rangle + \left\langle \frac{\delta \ell}{\delta a}, \delta a \right\rangle dt \\
&= \int \left\langle \frac{\delta \ell}{\delta \widehat{v}}, \partial_t \eta - \mathrm{ad}_{\widehat{v}}\eta \right\rangle + \left\langle \frac{\delta \ell}{\delta \zeta}, \delta\zeta \right\rangle \\
&\qquad + \left\langle \frac{\delta \ell}{\delta \widehat{w}}, (\partial_t + \mathcal{L}_{\widehat{v}})\delta\zeta + \mathcal{L}_{\delta\widehat{v}}\zeta \right\rangle + \left\langle \frac{\delta \ell}{\delta a}, -\mathcal{L}_\eta a \right\rangle dt \\
&= \int \left\langle -(\partial_t + \mathrm{ad}^*_{\widehat{v}})\frac{\delta \ell}{\delta \widehat{v}} + \frac{\delta \ell}{\delta a} \diamond a, \eta \right\rangle \\
&\qquad + \left\langle (-\partial_t + \mathcal{L}^T_{\widehat{v}})\frac{\delta \ell}{\delta \widehat{w}} + \frac{\delta \ell}{\delta \zeta}, \delta\zeta \right\rangle + \left\langle -\frac{\delta \ell}{\delta \widehat{w}} \diamond \zeta, \partial_t \eta - \mathrm{ad}_{\widehat{v}}\eta \right\rangle dt \\
&= \int \left\langle -(\partial_t + \mathrm{ad}^*_{\widehat{v}})\left(\frac{\delta \ell}{\delta \widehat{v}} - \frac{\delta \ell}{\delta \widehat{w}} \diamond \zeta \right) + \frac{\delta \ell}{\delta a} \diamond a, \eta \right\rangle \\
&\qquad + \left\langle (-\partial_t + \mathcal{L}^T_{\widehat{v}})\frac{\delta \ell}{\delta \widehat{w}} + \frac{\delta \ell}{\delta \zeta}, \delta\zeta \right\rangle dt\,.
\end{aligned}$$

The Euler-Poincaré equations corresponding to the decomposition of maps approach to wave-current interaction are therefore given by

$$(\partial_t + \mathrm{ad}^*_{\widehat{v}})\left(\frac{\delta \ell}{\delta \widehat{v}} - \frac{\delta \ell}{\delta \widehat{w}} \diamond \zeta \right) = \frac{\delta \ell}{\delta a} \diamond a\,, \tag{3.70}$$

$$(-\partial_t + \mathcal{L}^T_{\widehat{v}})\frac{\delta \ell}{\delta \widehat{w}} + \frac{\delta \ell}{\delta \zeta} = 0\,, \tag{3.71}$$

$$(\partial_t + \mathcal{L}_{\widehat{v}})\zeta = \widehat{w}\,, \tag{3.72}$$

$$(\partial_t + \mathcal{L}_{\widehat{v}})a = 0\,, \tag{3.73}$$

where the quantities, a, are advected by the two dimensional flow of currents. The transpose of the Lie derivative, $\mathcal{L}_{\widehat{v}}^T$, appears applied to the variational derivative $\delta\ell/\delta\widehat{w}$. Notice that this is obtained by noting that $\mathcal{L}_{\widehat{v}}\delta\zeta = \widehat{\boldsymbol{v}}\cdot\nabla_{\boldsymbol{r}}\delta\zeta$ and performing integration by parts, from which we see that $\mathcal{L}_{\widehat{v}}^T(\delta\ell/\delta\widehat{w}) = -\mathcal{L}_{\widehat{v}}(\delta\ell/\delta\widehat{w})$. Note that equation (3.70) is a modified version of the standard Euler-Poincaré momentum equation (2.28) for fluid dynamics with advected quantities, and in general the wave variables do not decouple from this. Revisiting Theorem 2.21, the abstract Kelvin-Noether circulation theorem, we see that for the decomposition of maps approach we have

$$\frac{d}{dt}\oint_{c(t)}\frac{1}{D}\left(\frac{\delta\ell}{\delta\widehat{v}}-\frac{\delta\ell}{\delta\widehat{w}}\diamond\zeta\right)=\oint_{c(t)}\frac{1}{D}\frac{\delta\ell}{\delta a}\diamond a\,,\tag{3.74}$$

where D is the advected mass density and $c(t)\in\mathcal{C}$ (see the preamble to Theorem 2.21 for a discussion of these). From here, it is evident that we have *gained* a contribution to the total momentum from the wave variables. The structure of the momentum for this problem is more evident on the Hamiltonian side, which is obtained through a Legendre transform

$$h(m,\lambda,\zeta,a) := \langle m,\widehat{v}\rangle + \langle\lambda,\widehat{w}\rangle - \ell(\widehat{v},\zeta,\widehat{w},a)\,,\tag{3.75}$$

The equation (3.70) can be rewritten as

$$\begin{aligned}(\partial_t+\mathrm{ad}^*_{\widehat{v}})\frac{\delta\ell}{\delta\widehat{v}} &= \frac{\delta\ell}{\delta a}\diamond a + (\partial_t+\mathrm{ad}^*_{\widehat{v}})\left(\frac{\delta\ell}{\delta\widehat{w}}\diamond\zeta\right)\\ &= \frac{\delta\ell}{\delta a}\diamond a + \frac{\delta\ell}{\delta\zeta}\diamond\zeta + \frac{\delta\ell}{\delta\widehat{w}}\diamond\widehat{w}\,.\end{aligned}\tag{3.76}$$

We may then write the Lie-Poisson equations in matrix form as

$$\partial_t \begin{pmatrix} m \\ a \\ \lambda \\ \zeta \end{pmatrix} = - \begin{pmatrix} \mathrm{ad}^*_\square m & \square \diamond a & -\lambda \diamond \square & \square \diamond \zeta \\ \mathcal{L}_\square a & 0 & 0 & 0 \\ -\mathcal{L}^T_\square \lambda & 0 & 0 & 1 \\ \mathcal{L}_\square \zeta & 0 & -1 & 0 \end{pmatrix} \begin{pmatrix} \frac{\delta h}{\delta m} = \widehat{v} \\ \frac{\delta h}{\delta a} = -\frac{\delta \ell}{\delta a} \\ \frac{\delta h}{\delta \lambda} = \widehat{w} \\ \frac{\delta h}{\delta \zeta} = -\frac{\delta \ell}{\delta \zeta} \end{pmatrix} . \tag{3.77}$$

Rather than writing the Lie-Poisson equations in terms of the momentum of the currents, m, we may write them in terms of the *total momentum* corresponding to the integrand of the Kelvin-Noether theorem, $M := m - \lambda \diamond \zeta$. This gives the untangled form of the Lie-Poisson equations in matrix form

$$\partial_t \begin{pmatrix} M \\ a \\ \lambda \\ \zeta \end{pmatrix} = - \begin{pmatrix} \mathrm{ad}^*_\square M & \square \diamond a & 0 & 0 \\ \mathcal{L}_\square a & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & -1 & 0 \end{pmatrix} \begin{pmatrix} \frac{\delta h}{\delta M} = \widehat{v} \\ \frac{\delta h}{\delta a} = -\frac{\delta \ell}{\delta a} \\ \frac{\delta h}{\delta \lambda} = \widehat{w} - \mathcal{L}_{\widehat{v}} \zeta \\ \frac{\delta h}{\delta \zeta} = -\frac{\delta \ell}{\delta \zeta} - \mathcal{L}^T_{\widehat{v}} \lambda \end{pmatrix} , \tag{3.78}$$

where the variational derivatives of the Hamiltonian reflect this change since the Hamiltonian given in equation (3.75) may be written in terms of M as

$$\begin{aligned} h &= \langle M + \lambda \diamond \zeta, \widehat{v} \rangle + \langle \lambda, \widehat{w} \rangle - \ell(\widehat{v}, \zeta, \widehat{w}, a) \\ &= \langle M, \widehat{v} \rangle - \langle \lambda, \mathcal{L}_{\widehat{v}} \zeta \rangle + \langle \lambda, \widehat{w} \rangle - \ell(\widehat{v}, \zeta, \widehat{w}, a) \, . \end{aligned} \tag{3.79}$$

3.4.2 The interaction of waves with a two dimensional flow

Following this, we may seek an action integral of the form of equation (3.15), as proposed by Crisan et al. [33], where the potential energy is chosen to give a harmonic oscillator structure to the wave dynamics

$$
\begin{aligned}
0 &= \delta \int \ell(\widehat{v}, \zeta, \widehat{w}, D) \\
&= \delta \int \int_{\mathcal{D}} \frac{D}{2}\left(|\widehat{\boldsymbol{v}}|^2 + \sigma^2 \widehat{w}^2 - \frac{\zeta^2}{Fr^2} \right) - p(D-1)\, d^2r\, dt\,.
\end{aligned}
\tag{3.80}
$$

Here, we have introduced the ratio, σ, between horizontal and vertical length scales and velocities, and the Froude number to represent the ratio of inertial and gravitational forces. For the model corresponding to this action, the vertical motion can be interpreted as *advected dynamics*, since we will see that it will not influence the dynamics associated with the flow map, Φ_t, and therefore extends the notion of an advected quantity[‡]. We calculate the variational derivatives of this action as

$$
\frac{\delta \ell}{\delta \widehat{v}} = D\widehat{\boldsymbol{v}} \cdot d\boldsymbol{r} \otimes d^2r\,, \tag{3.81}
$$

$$
\frac{\delta \ell}{\delta \widehat{w}} = D\sigma^2 \widehat{w}\,, \tag{3.82}
$$

$$
\frac{\delta \ell}{\delta \zeta} = -\frac{D}{Fr^2}\zeta\,, \tag{3.83}
$$

$$
\frac{\delta \ell}{\delta D} = \frac{1}{2}\left(|\widehat{\boldsymbol{v}}|^2 + \sigma^2 \widehat{w}^2 - \frac{\zeta^2}{Fr^2} \right) - p =: \varpi - p\,, \tag{3.84}
$$

In order to assemble these into the Euler-Poincaré equations (3.70)-(3.73), we first note that the equation (3.70) is written in terms of the diamond operator. Following Remark 2.22, we note the following

$$
\begin{aligned}
\frac{\delta \ell}{\delta \widehat{w}} \diamond \zeta &= -D\sigma^2 \widehat{w} \diamond \zeta = -D\sigma^2 \widehat{w} d\zeta \otimes d^2r\,, \\
\frac{\delta \ell}{\delta D} \diamond D &= Dd\left(\frac{\delta \ell}{\delta D} \right) \otimes d^2r = Dd(\varpi - p) \otimes d^2r\,.
\end{aligned}
$$

[‡]In general, the composition of maps approach does not have the property that the vertical motion does not influence the horizontal dynamics, as can be seen from the Euler-Poincaré equations. This is a property of this particular Lagrangian.

The Euler-Poincaré equations are therefore given by

$$(\partial_t + \mathcal{L}_{\widehat{v}})\left((D\widehat{\boldsymbol{v}}\cdot d\boldsymbol{r} + D\sigma^2\widehat{w}d\zeta)\otimes d^2r\right) = Dd(\varpi - p)\otimes d^2r\,, \tag{3.85}$$

$$(\partial_t + \mathcal{L}_{\widehat{v}})(D\sigma^2\widehat{w}) = -\frac{D}{Fr^2}\zeta\,, \tag{3.86}$$

$$(\partial_t + \mathcal{L}_{\widehat{v}})\zeta = \widehat{w}\,, \tag{3.87}$$

$$(\partial_t + \mathcal{L}_{\widehat{v}})(D\,d^2r) = 0\,, \tag{3.88}$$

$$D = 1 \quad \Longrightarrow \quad \mathrm{div}_{\boldsymbol{r}}\widehat{\boldsymbol{v}} = 0\,. \tag{3.89}$$

After dividing equation (3.85) through by the advected volume element, we have

$$(\partial_t + \mathcal{L}_{\widehat{v}})\left(\widehat{\boldsymbol{v}}\cdot d\boldsymbol{r} + \sigma^2\widehat{w}d\zeta\right) = d(\varpi - p)\,, \tag{3.90}$$

and we therefore have the Kelvin circulation theorem

$$\frac{d}{dt}\oint_{c(t)}\widehat{\boldsymbol{v}}\cdot d\boldsymbol{r} + \sigma^2\widehat{w}d\zeta = \oint_{c(t)} d(\varpi - p) = 0\,, \tag{3.91}$$

where $c(t)$ is a closed loop advected by the horizontal flow.

Taking a closer look at the momentum equation, it may be simplified by instead considering the equation (3.76), which is obtained by combining our Euler-Poincaré equations (3.70)-(3.73). To determine the form of this equation, we note that the terms featuring the diamond operator are

$$\frac{\delta\ell}{\delta\zeta}\diamond\zeta = -\frac{\delta\ell}{\delta\zeta}d\zeta\otimes d^2r = \frac{D}{Fr^2}\zeta d\zeta\otimes d^2r\,,$$
$$\frac{\delta\ell}{\delta\widehat{w}}\diamond\widehat{w} = -\frac{\delta\ell}{\delta\widehat{w}}d\widehat{w}\otimes d^2r = -D\sigma^2\widehat{w}d\widehat{w}\otimes d^2r\,.$$

Substituting into equation (3.76) gives

$$(\partial_t + \boldsymbol{\mathcal{L}}_{\widehat{v}})(D\widehat{\boldsymbol{v}}\cdot d\boldsymbol{r}\otimes d^2r) = Dd(\varpi - p)\otimes d^2r + \frac{D}{Fr^2}\zeta d\zeta \otimes d^2r - D\sigma^2\widehat{w}d\widehat{w}\otimes d^2r\,,$$

and dividing by the advected volume form gives

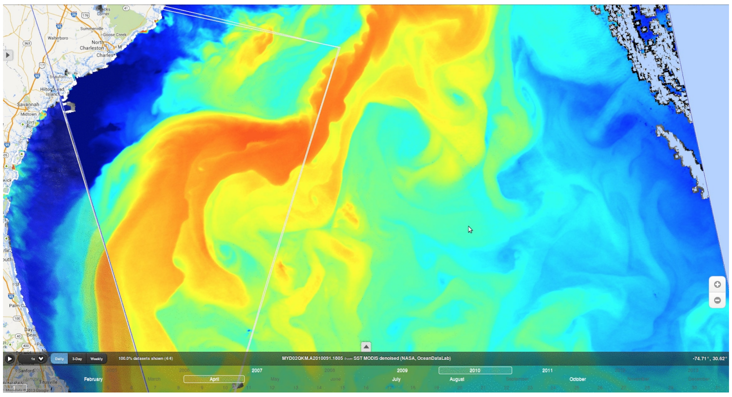
$$\begin{aligned}(\partial_t + \boldsymbol{\mathcal{L}}_{\widehat{v}})(\widehat{\boldsymbol{v}}\cdot d\boldsymbol{r}) &= d(\varpi - p) + \frac{\zeta}{Fr^2}d\zeta - \sigma^2\widehat{w}d\widehat{w}\\ &= d\left(\frac{1}{2}\left(|\widehat{\boldsymbol{v}}|^2 + \sigma^2\widehat{w}^2 - \frac{\zeta^2}{Fr^2}\right) - p\right) + \frac{\zeta}{Fr^2}d\zeta - \sigma^2\widehat{w}d\widehat{w}\\ &= d\left(\frac{1}{2}|\widehat{\boldsymbol{v}}|^2\right) - dp\,.\end{aligned}$$

The above equation may be recognised as the Euler-Poincaré equation for the two dimensional incompressible Euler equations, and it is therefore apparent that the dynamics decouples into a system of equations where no wave variables are present in the equation for the currents. The equations may be expressed in local coordinates as

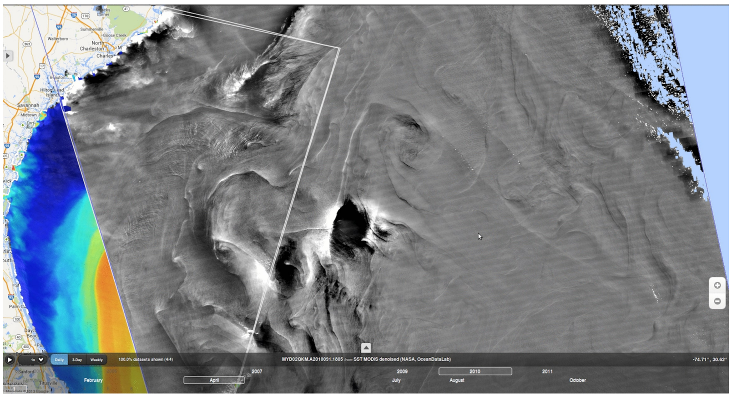
$$\begin{aligned}\partial_t\widehat{\boldsymbol{v}} + \widehat{\boldsymbol{v}}\cdot\nabla_{\boldsymbol{r}}\widehat{\boldsymbol{v}} &= -\nabla_{\boldsymbol{r}}p\,,\\ \partial_t\widehat{w} + \widehat{\boldsymbol{v}}\cdot\nabla_{\boldsymbol{r}}\widehat{w} &= -\frac{1}{\sigma^2Fr^2}\zeta\,,\\ \partial_t\zeta + \widehat{\boldsymbol{v}}\cdot\nabla_{\boldsymbol{r}}\zeta &= \widehat{w}\,,\\ \nabla_{\boldsymbol{r}}\cdot\widehat{\boldsymbol{v}} &= 0\,.\end{aligned} \tag{3.92}$$

This system corresponds to a horizontal flow given by a two dimensional Euler equation, and wave dynamics given by a harmonic oscillation in the frame of reference moving with the two dimensional flow. Whilst this is detached from classical models of water wave motion, we have a set of equations which illustrates how the composition of maps approach allows us to study the interactions between wave-like and current-like motions from a new perspective.

A potential vorticity formulation of this system may be obtained by taking the exterior derivative of equation (3.90). Without the inclusion of thermal effects, this potential vorticity is advected since the right hand side of (3.90) vanishes when we take its exterior derivative. As we will see, including thermal inhomogeneity creates a more complex potential vorticity structure.



(a) Sea surface temperature on April 1st 2010 near the Gulf Stream, from the Envisat AATSR measurements.



(b) Sea surface glitter contrasts on April 1st 2010 near the Gulf Stream, from the Envisar MERIS observations.

Figure 3.1: Visible here is the emergent coherence between sea surface temperature and the glitter patterns. It is apparent that sea surface roughness is most dramatic along the strongest thermal fronts. For an interpretation of sun glitter measurements, see Chapron et al. [21], Rascle et al. [89], or Yurovskaya et al. [105]. Images courtesy of B. Chapron.

Thermal effects. In the above system of equations, fluid elements are all oscillating in phase with the same frequency across the domain. Satellite imagery, such as in Figure 3.1, demonstrates that wave activity, as detected by sea surface sun glitter, is closely correlated with thermal effects. This motivates an approach for which wave activity is occurring in the frame of reference moving with the currents, as we have with our composition of maps. Coherence between wave activity and thermal gradients can be achieved through a thermal desynchronisation of the kinetic and potential energies. Indeed, consider the action given by

$$\begin{aligned} S &= \int_\ell (\widehat{\boldsymbol{v}}, \zeta, D, \rho)\, dt \\ &= \int\int_{\mathcal{D}} \frac{D\rho}{2}\left(|\widehat{\boldsymbol{v}}|^2 + \sigma^2 \widehat{w}^2 - \frac{\rho_{ref}}{\rho}\frac{\zeta^2}{Fr^2}\right) - p(D-1)\, d^2r\, dt\,, \end{aligned} \tag{3.93}$$

where ρ is the variable thermal buoyancy. Notice that in the coefficient of the potential energy, the dependence on the buoyancy is removed through the introduction of a reference buoyancy, ρ_{ref}. This approximation states that thermal variations away from the reference density have a more significant effect on the kinetic energy than the potential energy, and is discussed further in Holm et al. [57]. The variational derivatives of the action (3.93) are

$$\frac{\delta \ell}{\delta \widehat{v}} = D\rho \widehat{\boldsymbol{v}} \cdot d\boldsymbol{r} \otimes d^2 r \,, \tag{3.94}$$

$$\frac{\delta \ell}{\delta \widehat{w}} = D\rho \sigma^2 \widehat{w} \,, \tag{3.95}$$

$$\frac{\delta \ell}{\delta \zeta} = -\frac{D\rho_{ref}}{Fr^2}\zeta \,, \tag{3.96}$$

$$\frac{\delta \ell}{\delta \rho} = \frac{D}{2}\Big(|\widehat{\boldsymbol{v}}|^2 + \sigma^2 \widehat{w}^2\Big) = D\tilde{\varpi} \,, \tag{3.97}$$

$$\begin{aligned} \frac{\delta \ell}{\delta D} &= \frac{\rho}{2}\Big(|\widehat{\boldsymbol{v}}|^2 + \sigma^2 \widehat{w}^2 - \frac{\rho_{ref}\zeta^2}{Fr^2}\Big) - p =: \rho\tilde{\varpi} - \tilde{p} \,, \\ &\text{where} \quad \tilde{p} = p + \frac{\rho_{ref}\zeta^2}{2Fr^2} \,, \end{aligned} \tag{3.98}$$

and the Euler-Poincaré equations (3.70)-(3.73) give

$$\begin{aligned} (\partial_t + \mathcal{L}_{\widehat{v}})\big((D\rho \widehat{\boldsymbol{v}} \cdot d\boldsymbol{r} + D\rho\sigma^2 \widehat{w} d\zeta) \otimes d^2 r\big) &= Dd(\rho\tilde{\varpi} - \tilde{p}) \otimes d^2 r \\ &\quad - D\tilde{\varpi} d\rho \otimes d^2 r \,, \end{aligned} \tag{3.99}$$

$$(\partial_t + \mathcal{L}_{\widehat{v}})(D\rho\sigma^2 \widehat{w}) = -\frac{D\rho_{ref}}{Fr^2}\zeta \,, \tag{3.100}$$

$$(\partial_t + \mathcal{L}_{\widehat{v}})\zeta = \widehat{w} \,, \tag{3.101}$$

$$(\partial_t + \mathcal{L}_{\widehat{v}})\rho = 0 \,, \tag{3.102}$$

$$(\partial_t + \mathcal{L}_{\widehat{v}})(D\, d^2 r) = 0 \,, \tag{3.103}$$

$$D = 1 \quad \Longrightarrow \quad \mathrm{div}_{\boldsymbol{r}} \widehat{\boldsymbol{v}} = 0 \,. \tag{3.104}$$

From these equations, we can see that the wave-like dynamics now depend on the thermal density. A comparison between equations (3.100) and (3.20) reveals a similarity in the effect of density variations on the vertical velocity. Similarly to the case without thermal buoyancy, and using notation consistent with previous examples, this system of equations has a Kelvin-Noether circulation theorem

$$\frac{d}{dt}\oint_{c(t)} \big(\widehat{\boldsymbol{v}}\cdot d\boldsymbol{r} + \sigma^2 \widehat{w}\, d\zeta\big) = -\oint_{c(t)} \frac{1}{\rho} d\tilde{p}\,, \tag{3.105}$$

where $c(t)$ is a closed loop moving with the horizontal flow. This follows from dividing equation (3.99) through by the advected quantities Dd^2r and ρ, noting that

$$\frac{d(\rho\tilde{\varpi} - \tilde{p})}{\rho} - \frac{\tilde{\varpi}d\rho}{\rho} = -\frac{1}{\rho}d\tilde{p} + d\tilde{\varpi}\,.$$

Proposition 3.17. *The Kelvin-Noether theorem corresponding to equation* (3.105) *decouples into one corresponding to the currents,*

$$\frac{d}{dt}\oint_{c(t)} \widehat{\boldsymbol{v}}\cdot d\boldsymbol{r} = -\oint_{c(t)} \frac{1}{\rho}dp - d\frac{|\widehat{\boldsymbol{v}}|^2}{2}\,, \tag{3.106}$$

and one corresponding to the waves

$$\frac{d}{dt}\oint_{c(t)} \sigma^2\widehat{w}\, d\zeta = -\oint_{c(t)} \frac{\rho_{ref}}{\rho Fr^2}\zeta d\zeta\,. \tag{3.107}$$

Proof. Beginning with first equation (3.106), this may be shown by subtracting the wave momentum terms to the right hand side of equation (3.99) as

$$\begin{aligned}
\frac{d}{dt}\oint_{c(t)} \widehat{\boldsymbol{v}}\cdot d\boldsymbol{r} &= -\oint_{c(t)} \frac{1}{\rho}d\tilde{p} + \sigma^2(\partial_t + \mathcal{L}_{\widehat{\boldsymbol{v}}})\big(\widehat{w}\, d\zeta\big) - d\tilde{\varpi} \\
&= -\oint_{c(t)} \frac{1}{\rho}d\tilde{p} + \sigma^2\big((\partial_t + \mathcal{L}_{\widehat{v}})\widehat{w}\big)d\zeta + \sigma^2\widehat{w}d\widehat{w} - d\tilde{\varpi} \\
&= -\oint_{c(\widehat{\boldsymbol{v}})} \frac{1}{\rho}d\widetilde{p} - \frac{\rho_{ref}}{Fr^2\rho}\zeta d\zeta + \sigma^2\widehat{w}d\widehat{w} - d\tilde{\varpi} = -\oint_{c(\widehat{\boldsymbol{v}})} \frac{1}{\rho}dp - d\frac{|\widehat{\boldsymbol{v}}|^2}{2}\,.
\end{aligned}$$

The Kelvin-Noether theorem for the wave variables, equation (3.107), follows by combining equations (3.105) and (3.106) as

$$\begin{aligned}
\frac{d}{dt}\oint_{c(t)} \sigma^2 \widehat{w}\, d\zeta &= -\oint_{c(t)} \frac{1}{\rho}d\tilde{p} - d\tilde{\varpi} - (\partial_t + \mathcal{L}_{\hat{v}})(\boldsymbol{v}\cdot d\boldsymbol{r}) \\
&= -\oint_{c(t)} \frac{1}{\rho}d\tilde{p} - d\tilde{\varpi} + \frac{1}{\rho}dp - d\frac{|\widehat{\boldsymbol{v}}|^2}{2} \\
&= -\oint_{c(t)} \frac{1}{\rho}(\tilde{p} - p) = -\oint_{c(t)} \frac{\rho_{ref}}{\rho Fr^2}\zeta d\zeta\,.
\end{aligned}$$

$\mathscr{Q.E.D.}$

Remark 3.18. Notice that Proposition 3.17 implies that $\widehat{\boldsymbol{v}}$ satisfies the thermal Euler equations. It also implies that waves do note create circulation in the currents, and vice versa, whilst the thermal buoyancy acts to create circulation in both the waves and currents.

A POTENTIAL VORTICITY FORMULATION. A potential vorticity formulation is possible here, since the equation (3.105) representing the Kelvin-Noether circulation theorem has a nonzero right hand side, and thus the exterior derivative of the right hand side of (3.99) is nonzero. As we will demonstrate, the system closes for a *thermally weighted* potential vorticity. The thermal weighting is expected due to the fact that the abstract Kelvin-Noether theorem (Theorem 2.21) features a division by D through the Euler-Poincaré equation analogous to (3.99) which, in this example, leaves behind the thermal buoyancy multiplying the integrand of equation (3.105).

Noting the relationship between the exterior derivative and curl of a vector field, we thus define weighted potential vorticity by

$$
\begin{aligned}
q\,d^2r &:= d\Big(\rho\big(\widehat{\boldsymbol{v}}\cdot d\boldsymbol{r} + \sigma^2\widehat{w}d\zeta\big)\Big)\,d^2r \\
&= d\rho\wedge\big(\widehat{\boldsymbol{v}}\cdot d\boldsymbol{r} + \sigma^2\widehat{w}d\zeta\big) + \rho\big(\widehat{\boldsymbol{z}}\cdot \mathrm{curl}\widehat{\boldsymbol{v}} + \sigma^2 d\widehat{w}\wedge d\zeta\big)\,d^2r\,,
\end{aligned}
\tag{3.108}
$$

where $\widehat{\boldsymbol{z}}$ is the unit normal in the vertical direction. In the above, to transition from the first line to the second we note that the multiplication of a scalar function by a one form is the exterior (or *wedge*) product, and the exterior derivative of such a product obeys a Leibniz rule. Upon the introduction of a stream function for the currents, $\widehat{\boldsymbol{v}} = \nabla^{\perp}\psi$, this can be decomposed into potential vorticity for the currents, q_c, and for the waves, q_w, in the following way

$$
q\,d^2r = \mathrm{div}(\rho\nabla\psi)\,d^2r + \sigma^2 J(\rho\widehat{w},\zeta)\,d^2r =: (q_c + q_w)\,d^2r\,. \tag{3.109}
$$

The equations in Proposition 3.17 then imply a pair of decoupled potential vorticity equations

$$
\begin{aligned}
(\partial_t + \mathcal{L}_{\widehat{v}})(q_c\,d^2r) &= \frac{1}{2}d\rho\wedge d|\widehat{\boldsymbol{v}}|^2 = \frac{1}{2}J\big(\rho,|\nabla\psi|^2)\big)\,d^2r\,, \\
(\partial_t + \mathcal{L}_{\widehat{v}})(q_w\,d^2r) &= \sigma^2\frac{1}{2}d\rho\wedge d\widehat{w}^2 = \frac{1}{2}J\Big(\rho\,,\sigma^2\widehat{w}^2\Big)\,d^2r\,.
\end{aligned}
\tag{3.110}
$$

The decoupling of the potential vorticity formulation is not surprising, given the decoupling of the Kelvin-Noether circulation theorem. Noting that $D=1$ and the advection operator applied to the variables ρ, ζ, and $\widehat{w}$ becomes $\partial_t\,\square + J(\psi,\square)$, the system of equations is evident. The combination of equations (3.110) gives an equation for q

$$
\partial_t q + J(\psi,q) = \frac{1}{2}J(\rho,\tilde{\varpi})\,. \tag{3.111}
$$

3.4.3 The slowly varying envelope approximation

There is a distinct difference in time scales between wave oscillations and the flow of currents in the upper ocean. The vertical motion is, usually, significantly quicker than the rate of change of horizontal current features. We will use the Wentzel–Kramers–Brillouin (WKB) approximation, common in mathematical physics, to account for this difference in time scale, a methodology which follows from Bretherton and Garrett [17] and, later, Gjaja and Holm [42]. Note that there is a difference between the application of this approximation here and how it is found elsewhere in the literature. Namely, we define wave activity to be the dynamics occurring in the vertical direction, rather than a rapid fluctuation around a mean flow. We therefore assume that the surface elevation is of the form

$$\zeta(\boldsymbol{r},t) = \Re\left(a(\boldsymbol{r},t)\exp\left(\frac{i\theta(\boldsymbol{r},t)}{\epsilon}\right)\right)\,, \quad \text{with} \quad \epsilon \ll 1\,, \tag{3.112}$$

where $\Re$ denotes taking the real part. We have here introduced a pair of new variables which are both permitted to vary slowly in space and time, the complex amplitude, a, and the phase factor, θ. Notice that equation (3.112) is equivalent to

$$\zeta = \frac{1}{2}a\exp\left(\frac{i\theta}{\epsilon}\right) + \frac{1}{2}a^*\exp\left(-\frac{i\theta}{\epsilon}\right) \tag{3.113}$$

We will be substituting this expression into the action given in equation (3.93), together with the kinematic boundary condition replacing the dependence of the action on $\widehat{w}$. Noting that the derivative of ζ is

$$\frac{d\zeta}{dt} = \frac{1}{2}\frac{da}{dt}\exp\left(\frac{i\theta}{\epsilon}\right)+\frac{ia}{2\epsilon}\frac{d\theta}{dt}\exp\left(\frac{i\theta}{\epsilon}\right)+\frac{1}{2}\frac{da^*}{dt}\exp\left(-\frac{i\theta}{\epsilon}\right)-\frac{ia^*}{2\epsilon}\frac{d\theta}{dt}\exp\left(-\frac{i\theta}{\epsilon}\right),$$

the kinetic energy becomes

$$
\begin{aligned}
\int\int_{\mathcal{D}} \frac{D\rho\sigma^2}{2}\Big(\frac{d\zeta}{dt}\Big)^2\, d^2r\, dt = \int\int_{\mathcal{D}} \frac{D\rho\sigma^2}{2}\bigg[&\frac{1}{2}\Big|\frac{da}{dt}\Big|^2 + \frac{|a|^2}{2\epsilon^2}\Big(\frac{d\theta}{dt}\Big)^2 \\
&- \frac{ia^*}{2\epsilon}\frac{da}{dt}\frac{d\theta}{dt} + \frac{ia}{2\epsilon}\frac{da^*}{dt}\frac{d\theta}{dt}\bigg]\, d^2r\, dt \\
= \int\int_{\mathcal{D}} \frac{D\rho\sigma^2}{2}\bigg[&\frac{1}{2}\Big|\frac{da}{dt}\Big|^2 + \frac{|a|^2}{2\epsilon^2}\Big(\frac{d\theta}{dt}\Big)^2 + \frac{1}{\epsilon}\frac{d\theta}{dt}\Im\Big(a^*\frac{da}{dt}\Big)\bigg],
\end{aligned}
$$

where $\Im$ denotes taking the imaginary part, and exponential terms are neglected since $\epsilon \ll 1$. The potential energy term is simpler

$$
\int\int_{\mathcal{D}} -\frac{D\rho_{ref}}{2Fr^2}\zeta^2\, d^2r\, dt = \int\int_{\mathcal{D}} -\frac{D\rho_{ref}}{2Fr^2}\frac{|a|^2}{2}\, d^2r\, dt\,. \tag{3.114}
$$

The action integral for the WKB approximation is therefore

$$
\begin{aligned}
S = \int\int_{\mathcal{D}} &\frac{D\rho|\widehat{\boldsymbol{v}}|^2}{2} - p(D-1) \\
&+ \frac{D\rho\sigma^2}{4}\bigg[\Big|\frac{da}{dt}\Big|^2 + \frac{2}{\epsilon}\frac{d\theta}{dt}\Im\Big(a^*\frac{da}{dt}\Big) + \frac{|a|^2}{\epsilon^2}\bigg(\Big(\frac{d\theta}{dt}\Big)^2 - \frac{\rho_{ref}\epsilon^2}{\rho\sigma^2 Fr^2}\bigg)\bigg] d^2r\, dt \\
= \int\int_{\mathcal{D}} &\frac{D\rho|\widehat{\boldsymbol{v}}|^2}{2} - p(D-1) \\
&+ \frac{D\rho\sigma^2|a|^2}{4\epsilon^2}\bigg(\big(\partial_t\theta + \widehat{\boldsymbol{v}}\cdot\nabla_{\boldsymbol{r}}\theta\big)^2 - \frac{\rho_{ref}\epsilon^2}{\rho\sigma^2 Fr^2}\bigg)\, d^2r\, dt\,,
\end{aligned} \tag{3.115}
$$

where only the leading order term in ϵ remains. This is a valid assumption when $\epsilon \ll 1$ and $\epsilon^2/(\sigma^2 Fr^2) = \mathcal{O}(1)$. According to satellite oceanography data, these approximations are reasonable in many regions of interest in the upper ocean [57]. As was shown in Holm, Hu, and Street [57], taking variations of the action (3.115) yields

$$
\begin{aligned}
\frac{\delta \ell}{\delta \widehat{v}} &= D\rho\Big(\widehat{\boldsymbol{v}}\cdot d\boldsymbol{r} + \mathcal{N} d\frac{d\theta}{dt}\Big)\otimes d^2r \quad \text{with} \quad \mathcal{N} := \frac{\sigma^2|a|^2}{2\epsilon^2}\,,\\
\frac{\delta \ell}{\delta D} &= \frac{\rho}{2}|\widehat{\boldsymbol{v}}|^2 - p\,,\\
\frac{\delta \ell}{\delta \rho} &= \frac{D}{2}|\widehat{\boldsymbol{v}}|^2\,,\\
\frac{\delta \ell}{\delta |a|^2} &= \frac{D\rho}{4Fr^2}\bigg(\Big(\frac{d\theta}{dt}\Big)^2 - \frac{\rho_{ref}}{\rho}\bigg) = 0\,, \quad \text{since} \quad \frac{\epsilon^2}{\sigma^2 Fr^2} = \mathcal{O}(1)\\
&\Longrightarrow \quad \frac{d\theta}{dt} = \pm\frac{\sqrt{\rho\rho_{ref}}}{\rho}\,,\\
\frac{\delta \ell}{\delta \theta} &= 0\,, \quad \Longrightarrow \quad \partial_t\mathcal{A} + \mathrm{div}(\mathcal{A}\widehat{\boldsymbol{v}}) = 0\,, \quad \text{with} \quad \mathcal{A} := D\rho\mathcal{N}\frac{d\theta}{dt}\,d^2r\,,\\
D = 1\,, &\quad \Longrightarrow \quad \partial_t|a|^2 + \widehat{\boldsymbol{v}}\cdot\nabla_{\boldsymbol{r}}|a|^2 = 0\,.
\end{aligned}
\tag{3.116}
$$

Despite the approximations made, we still have a Kelvin-Noether theorem where

$$
\frac{d}{dt}\oint_{c(t)}\Big(\widehat{\boldsymbol{v}}\cdot d\boldsymbol{r} + \mathcal{N}d\frac{d\theta}{dt}\Big) = -\oint_{c(t)}\frac{1}{\rho}dp\,. \tag{3.117}
$$

The calculation for this follows similarly to the other examples, noting that $\mathcal{N}$ and $\frac{d\theta}{dt}$ are advected and the diamond terms emerging from $\delta\ell/\delta D$ and $\delta\ell/\delta\rho$ are as for the inhomogeneous Euler equations,. Indeed, the Euler-Poincaré momentum equation is

$$
(\partial_t + \mathcal{L}_{\widehat{v}})\frac{\delta \ell}{\delta \widehat{v}} = D\nabla_{\boldsymbol{r}}\frac{\delta \ell}{\delta D} - \frac{\delta \ell}{\delta \rho}\nabla_{\boldsymbol{r}}\rho\,,
$$

and dividing by the advected variables D and ρ gives

$$
(\partial_t + \mathcal{L}_{\widehat{v}})\Big(\widehat{\boldsymbol{v}}\cdot d\boldsymbol{r} + \mathcal{N}d\frac{d\theta}{dt}\Big) = -\rho^{-1}dp + d\Big(\frac{1}{2}|\widehat{\boldsymbol{v}}|^2\Big)\,.
$$

Remark 3.19. It is immediately obvious from equation (3.117) that the decoupling between wave and current degrees of freedom observed before the slowly varying envelope approximation is present also here. This follows from the advection of the second term in the integrand of Kelvin's theorem

$$(\partial_t + \mathcal{L}_{\widehat{v}}) \left(\mathcal{N} d \frac{d\theta}{dt} \right) = 0\,. \tag{3.118}$$

Remark 3.20 (Alignment of gradients.). As was noted in [57], we have the following advected quantity

$$(\partial_t + \mathcal{L}_{\widehat{v}})(d|a|^2 \wedge d\rho) = 0\,,$$

meaning that the alignment of gradients of $|a|^2$ and ρ is preserved by the flow.

Properties of the solution. Solutions to these equations were examined numerically in [57]. The approximated equations can be written into a potential vorticity formulation in the same manner as the unapproximated equations. The potential vorticity of the currents, q_c, has the same structure as in equation (3.110), however the potential vorticity for the waves is given by

$$q_w = \frac{\mathcal{N}}{|a|^2} J(\sqrt{\rho \rho_{ref}}, |a|^2)\,,$$

and is advected. The numerical integration of this demonstrates that there is a synchronisation between thermal fronts and wave activity present within the model [57]. This gives hope to the application of this modelling approach to understand the phenomena illustrated in Figure 3.1. This illustrates the potential for more sophisticated models of this class to reveal further information about the interaction of wave activity with horizontal thermal gradients.

3.4.4 A stochastic approach to decomposition of maps

Recalling the Euler-Poincaré approach to the decomposition of maps, presented in Section 3.4.1, we may introduce a stochastic perturbation of this system again using stochastic advection by Lie transport. Since the modelling assumption here is that the flow is decomposed into a horizontal part and a vertical part, and the model is a continuum theory on a two dimensional domain, we will be introducing a two dimensional noise into the currents. Recall that in our stochastic formulation of the classical water wave equations, there is a vertical component of the noise appearing in the kinematic boundary condition (3.27). To maximise generality, we will include such a term in our reasoning here. We will therefore be considering the two dimensional stochastic vector field, ${\rm d}r_t$, and we will have vertical noise represented by ${\rm d}z_t$, as defined in Section 3.3.1.

Working with the same reduced Lagrangian to which we applied a deterministic version of Hamilton's principle in Section 3.4.1, we note that in the stochastic case the variations have the following forms

$$\delta\widehat{v}\,dt = {\rm d}\eta - {\rm ad}_{{\rm d}r_t}\eta\,, \quad \delta\widehat{w}\,dt = ({\rm d} + \mathcal{L}_{{\rm d}r_t})\delta\zeta + \mathcal{L}_{\delta\widehat{v}}\zeta\,, \quad \text{and} \quad \delta a = -\mathcal{L}_\eta a\,,$$

where $\eta \in \mathfrak{X}(\mathcal{D})$ vanishes at the endpoints. For a discussion of the form of these in the case of a standard fluid theory, see Luesink [72]. We may derive the stochastic Euler-Poincaré equations by taking variations as follows

$$
\begin{aligned}
0 &= \int \left\langle \frac{\delta \ell}{\delta \widehat{v}}, \delta \widehat{v} \right\rangle + \left\langle \frac{\delta \ell}{\delta \zeta}, \delta \zeta \right\rangle + \left\langle \frac{\delta \ell}{\delta \widehat{w}}, \delta \widehat{w} \right\rangle + \left\langle \frac{\delta \ell}{\delta a}, \delta a \right\rangle dt \\
&= \int \left\langle \frac{\delta \ell}{\delta \widehat{v}}, {\rm d}\eta - {\rm ad}_{{\rm d}r_t}\eta \right\rangle + \left\langle \frac{\delta \ell}{\delta \zeta}, \delta \zeta \right\rangle dt \\
&\qquad + \left\langle \frac{\delta \ell}{\delta \widehat{w}}, ({\rm d} + \mathcal{L}_{{\rm d}r_t})\delta\zeta + \mathcal{L}_{\delta \widehat{v}}\zeta \right\rangle + \left\langle \frac{\delta \ell}{\delta a}, -\mathcal{L}_\eta a \right\rangle dt \\
&= \int \left\langle -({\rm d} + {\rm ad}^*_{{\rm d}r_t})\frac{\delta \ell}{\delta \widehat{v}} + \frac{\delta \ell}{\delta a} \diamond a\, dt, \eta \right\rangle \\
&\qquad + \left\langle (-{\rm d} + \mathcal{L}^T_{{\rm d}r_t})\frac{\delta \ell}{\delta \widehat{w}} + \frac{\delta \ell}{\delta \zeta}\, dt, \delta\zeta \right\rangle + \left\langle -\frac{\delta \ell}{\delta \widehat{w}} \diamond \zeta, {\rm d}\eta - {\rm ad}_{{\rm d}r_t}\eta \right\rangle \\
&= \int \left\langle -({\rm d} + {\rm ad}^*_{{\rm d}r_t})\left(\frac{\delta \ell}{\delta \widehat{v}} - \frac{\delta \ell}{\delta \widehat{w}} \diamond \zeta \right) + \frac{\delta \ell}{\delta a} \diamond a\, dt, \eta \right\rangle \\
&\qquad + \left\langle (-{\rm d} + \mathcal{L}^T_{{\rm d}r_t})\frac{\delta \ell}{\delta \widehat{w}} + \frac{\delta \ell}{\delta \zeta}\, dt, \delta\zeta \right\rangle .
\end{aligned}
$$

The Euler-Poincaré equations corresponding to the composition of maps approach to wave-current interaction are therefore given by

$$
({\rm d} + {\rm ad}^*_{{\rm d}r_t})\left(\frac{\delta \ell}{\delta \widehat{v}} - \frac{\delta \ell}{\delta \widehat{w}} \diamond \zeta \right) = \frac{\delta \ell}{\delta a} \diamond a\, dt\,, \tag{3.119}
$$

$$
(-{\rm d} + \mathcal{L}^T_{{\rm d}r_t})\frac{\delta \ell}{\delta \widehat{w}} + \frac{\delta \ell}{\delta \zeta}\, dt = 0\,, \tag{3.120}
$$

$$
({\rm d} + \mathcal{L}_{{\rm d}r_t})\zeta = {\rm d}z_t\,, \tag{3.121}
$$

$$
({\rm d} + \mathcal{L}_{{\rm d}r_t})a = 0\,, \tag{3.122}
$$

where the quantities, a, are advected by the two dimensional stochastic flow of currents. Note that the equation for ζ here agrees with the kinematic boundary condition for the classical water wave theory given by equation (3.27). The inclusion of the vertical component of the noise in equation (3.27) is due to the relationship between the kinematic boundary condition and three dimensional advection, as discussed in Definition 3.3. The vertical noise component, found in $\mathrm{d}z_t$, does not permeate the rest of the model since advected quantities are carried by the two dimensional flow of currents, which is a two dimensional fluid theory.

We see that for the decomposition of maps approach we have a stochastic Kelvin-Noether theorem represented by the equation

$$\mathrm{d}\oint_{c(t)} \frac{1}{D}\left(\frac{\delta \ell}{\delta \widehat{v}} - \frac{\delta \ell}{\delta \widehat{w}} \diamond \zeta\right) = \oint_{c(t)} \frac{1}{D}\frac{\delta \ell}{\delta a} \diamond a \, dt \,, \tag{3.123}$$

where the notation is largely inherited from the deterministic case.

We gain a new perspective on the structure of the noise by considering the Hamiltonian formulation. In the stochastic case, we can observe this by considering a Legendre transform to a phase space variational principle

$$0 = \delta \int\int_{\mathcal{D}} \langle m, \mathrm{d}r_t \rangle + \langle \lambda, \mathrm{d}z_t \rangle - h(m,a)\, dt - \sum_i h_i(m,a) \circ dW_t^i \,,$$

where m and λ are as in the deterministic case, and the stochastic vector field $\mathrm{d}r_t$ is written in terms of the Hamiltonians as

$$\mathrm{d}r_t = \frac{\delta h}{\delta m}\, dt + \sum_i \frac{\delta h_i}{\delta m} \circ dW_t^i \,.$$

We may the stochastic equations for the decomposition of maps approach may therefore be expressed as Lie-Poisson equation in matrix form as

$$\mathrm{d}\begin{pmatrix} m \\ a \\ \lambda \\ \zeta \end{pmatrix} = -\begin{pmatrix} \mathrm{ad}^*_{\square} m & \square \diamond a & -\lambda \diamond \square & \square \diamond \zeta \\ \mathcal{L}_{\square} a & 0 & 0 & 0 \\ -\mathcal{L}^T_{\square} \lambda & 0 & 0 & 1 \\ \mathcal{L}_{\square} \zeta & 0 & -1 & 0 \end{pmatrix} \begin{pmatrix} \mathrm{d}r_t \\ -\frac{\delta \ell}{\delta a}\, dt \\ \mathrm{d}z_t \\ -\frac{\delta \ell}{\delta \zeta}\, dt \end{pmatrix} . \tag{3.124}$$

From here, it is evident that we may use this methodology to incorporate a structure preserving noise into equations of this class, including those presented in this section.

3.5 Concluding remarks

Despite being a topic which has captured the imagination of the scientific community for an extended period of time, the dynamics of a free surface are by no means fully described by existing modelling techniques. In this chapter, we have derived a structure preserving stochastic perturbation of the classical theory. The perturbation is inspired by one which has been shown to be compatible with a data driven approach to numerically modelling fluid problems, as found in Cotter et al. [25, 26, 27]. Thus we can, in theory, use this approach to account for uncertainty in the water wave problem. Moreover, the stochastic perturbations of the velocity field need not be irrotational, introducing the possibility of using such a model to account for the effects of vorticity. We have furthermore introduced a new approach, making use of the composition of two maps, which can support additional variables. This allows us to include variables such as thermal buoyancy, which play an important role in large scale ocean dynamics. We also demonstrated that models of this class can be made stochastic through their variational structure.

We have developed several new concepts in this section and each of these opens new research questions. In particular, the decomposition of maps approach enables the study of waves interacting with a two dimensional fluid equation. Whilst we used an Euler equation with harmonic oscillations, there is opportunity to allow the two maps to represent other models for the wave and current dynamics. Separately, one can find many papers about the classical water wave equations in the literature. Their analytical properties are understood, their solutions numerically studied, and asymptotic expansions have been considered. All of these topics are potential future research directions for the stochastic classical water wave equations we have derived in this chapter.

4

The influence of memory on inertial particle motion

There are pressing environmental reasons which give importance to understanding the motion of inertial particles on the free surface of a fluid. In the language of the previous chapter, the currents are responsible for the transport of material. This is related to the remarks of John Scott Russell included in the introduction to Chapter 3. An object with mass will be carried by this flow, although this is not straightforward due to the inertia of the object. The inertia modifies the trajectory from that of a Lagrangian trajectory, and the object's procession alters the flow of the fluid in which it resides.

As will become apparent, the governing equation for this will be a fractional order differential equation. These equations do not have a known variational structure, unlike those studied thus far in this thesis. We will therefore not be considering stochastic parameterisations. It is worth noting that, borrowing language from discretised equations, stochastic equations change how we move forward from one timestep to the next. This increment gains a random component. Likewise, the equations we see here will differ from the standard theory of differential equations in terms of how we move in time from one step to the next. This is the second class of equations seen in this thesis which changes the process of stepping through time, relative to standard fluid theories. In this example, the equations have this structure since they feature *memory*.

4.1 Summary of the chapter

Within the study of the trajectory of an inertial particle through a fluid domain, the Maxey-Riley equation [77] is widely used. Since its initial publication, many analytical properties of this equation have remained unknown and are nontrivial due to the nonlinear and fractional in time nature of the model. In this section, we will feature the findings of a recent work [34] where a comprehensive analysis of the Maxey-Riley equation is given.

The Maxey-Riley equation will be considered as a fractional order equation, and results from fractional calculus will be applied to resolve difficulties posed by the fractional derivative 'Basset-Boussinesq' term [97]. Despite the fact that fractional calculus has an active research community, the Maxey-Riley equation is a relatively rare example of a fractional order equation which is adopted by the applied mathematics community. We will show that many of the classical properties of ordinary differential equations (ODEs) also apply to this equation. In particular, the following will be covered.

- The notion of *weak* and *strong* solutions to the Maxey-Riley equation is introduced, and the definition of a *maximal solution* is given.
- Given some conditions on the velocity field of the surrounding fluid, the existence of a unique maximal weak solution is shown.
- We prove that a unique global in time weak solution exists, using a fractional version of a Grönwall-type inequality.
- Under an additional assumption on the initial conditions, the global solution is proven to be sufficiently regular to classify the solution as *strong*.
- By considering the derivative of the solution with respect to its initial conditions, we show that inertial particle trajectories given by the Maxey-Riley equation do not collide, and that two trajectories can be chosen to be arbitrarily close together at time t by choosing suitable initial conditions.

4.2 Motivation and historical overview

There is a vast quantity of data available demonstrating the dynamics of the upper ocean. In particular, data can be obtained from drifter buoys floating on the ocean's surface, as seen in Figure 4.1. This data has been collected since the early 1980s, and provides a rich dataset of surface ocean currents over a relatively long time period. Moreover, as demonstrated in Figure 4.2, modern satellite imagery can identify the trajectories of small inertial particles, which may be used to provide quantitative data on the

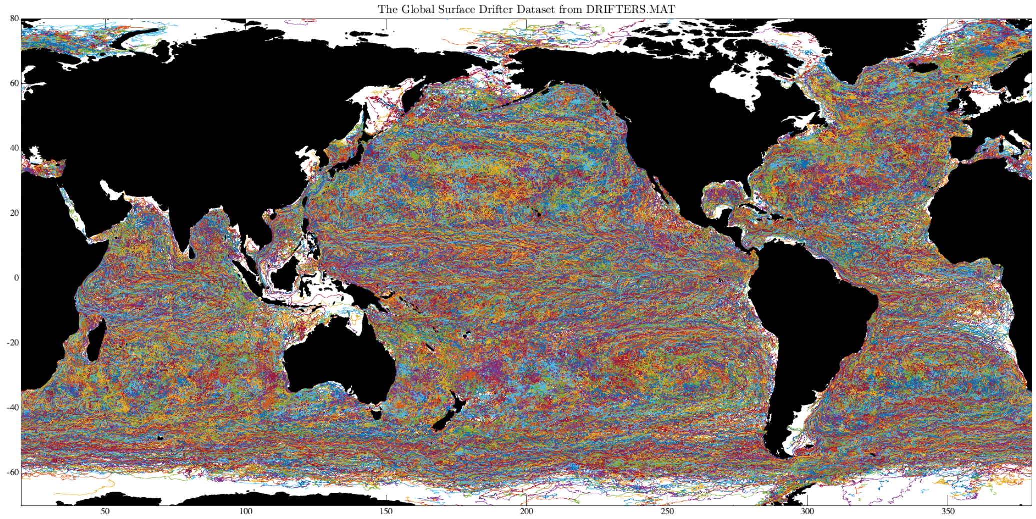
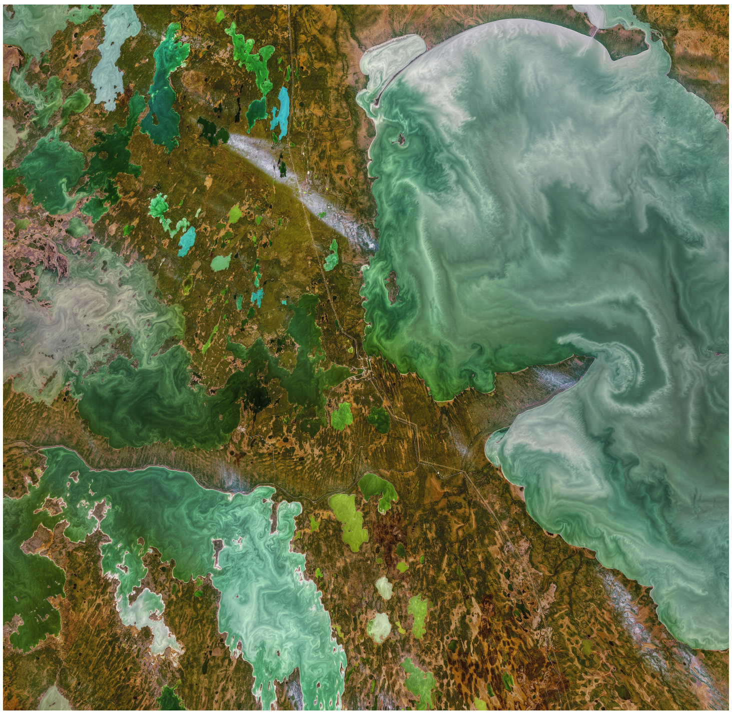


Figure 4.1: Visible here are trajectories of drifter buoys on the ocean surface, using data from the Global Drifter Program of the National Oceanic and Atmospheric Administration's Atlantic Oceanographic and Meteorological Laboratory. The data is displayed such that the path of each drifter is represented by a coloured line [70]. Thanks to the authors of [31], where this image was also featured, for bringing this to my attention.

ocean's dynamics. Due to the expensive nature of collecting this data, its accurate interpretation is of importance. When considering how to interpret it with accuracy, notice that both the satellite imagery and the drifter buoy data both represent the path of an object with mass moving through the ocean. Neither data set captures a pure Lagrangian trajectory. The effort and expense of collecting all of this data is justified by the growing scientific interest in the upper ocean, partially due to its role in weather and climate systems. From this, it is apparent that understanding how objects with mass move through the ocean is a matter of great scientific interest. For a review of models of inertial particle motion and their history, see [78].

Often, when modelling the movement of inertial particles in the ocean, the assumption is made that the object's mass does not influence its trajectory with no thorough justification of whether this is a good assumption or not. A detailed summary of current methods in this field can be found in [101]. Our

(a) Imagery of the Solway Firth, or Tràchd Romhra, between Dumfries and Galloway and Cumbria, showing the dynamics of the firth in October 2019.



(b) A snapshot of Lake Winnipeg, Lake Winnipegosis, and Cedar Lake, in Manitoba, Canada. Captured on 3rd November 2021.

Figure 4.2: The two images above are from the Landsat 8 satellite and are courtesy of the U.S. Geological Survey. Both images capture inertial particles, in this case sediment, algae, and other flotsam, which may be visually enhanced and used to illustrate the dynamics occurring in the region of the ocean where the particles are present. The ability of satellites to identify these particles provides a methodology for collecting data on ocean currents.

present understanding of models which feature inertial effects is sufficiently poor that purely data driven approaches, without any physical modelling, can provide meaningful insight into environmental issues connected to this topic, including ocean plastics [99]. Given the enormous environmental impact of small plastic pollutants (see e.g. [100]), a better mathematical understanding of models related to this phenomenon are crucial to give a more accurate analysis of these issues. In this chapter, we will investigate a commonly used mathematical model for the trajectory of an object with mass through a fluid domain, and the results proven have implications on a wider class of equations which share the same form.

In spite of the fact that inertial particles are increasingly relevant in contemporary applications of mathematics, their interest as a field of study is nothing new. Indeed, the study of interactions between an inertial particle and its surrounding flow predates the publication of the Navier-Stokes equations. As one can see upon observing the titles of many of the historical works cited here, much of the early work in this direction relates to the forces experienced by pendula. Interest in this direction was heavily motivated by the desire to manufacture more accurate clocks. This was such a pressing issue at the time that an act of Parliament, known as the Longitude Act [1], was passed by the Parliament of Great Britain in 1714, which promised significant financial rewards to scientists who advanced the ability to measure longitude at sea. A solution to the 'longitude problem' was of international interest; the 'Board of Longitude', founded in Britain to monitor progress in this direction, awarded prizes to individuals working in a number of different countries. Similar initiatives were established in other countries. Leading thinkers developed interesting mathematics as a result of the consideration of the forces acting on the sphere attached to the end of a pendulum. This work found wider application, since it applies to any sphere moving relative to the fluid in which it is suspended. The idea of the ‘added mass term’, which models the inertia resulting from the deflection of fluid around an accelerating (or decelerating) body immersed in fluid, was proposed by Bessel in 1828 [10]. The term itself was later formulated first, for a rigid sphere, by Poisson [87], and shortly after for a rigid ellipse in potential flow by Green [44]. This work was built upon by Clebsch [23], who considered the effect of rotation of the ellipse. Observations on viscous effects for flow around a sphere were made by Stokes [94], who considered this problem by solving what are now widely known as the Stokes equations for *creeping* flow, also known as *Stokes flow*. Thus, from our contemporary viewpoint, his work is valid in the limit of zero Reynolds number, defined using the relative velocity between the sphere and the fluid away from the sphere. These viscous effects will be known to many readers as *Stokes drag*.

Several decades after the formulation of the added mass term and Stokes drag, the influence of a particles history on its future was found. Again in the context of creeping flow, Boussinesq [16], in 1885, formulated an equation where the velocity of the sphere relative to the fluid is a function of time which is permitted to be arbitrary. In the previously mentioned study, this function was considered to be exogenous and usually fixed as sinusoidal. The reclassification of this velocity as an independent variable is a shift in modelling approach and generalises the problem away from periodic motions and pendula. The force experienced by the sphere as a result of its motion through the fluid was found to consist of an added mass term, a drag term, and a viscous term which depends on the history of the sphere’s trajectory. The viscous term appears as a time integral from the initial time through to the present, and is due to the force corresponding to the formation of a lagging boundary layer. Shortly afterwards in 1888, Basset [6] published, in his book on hydrodynamics, a calculation under the same assumptions which reached the same conclusion. Whilst it is unclear who derived the term first, it is the later author whose name commonly appears in the literature and the term is known as the *Basset history integral*, or similar, and corresponds to the *Basset force*. Alternatively, the term may be named the *Basset-Boussinesq* term, which will be used in this section.

After the derivation of the main forces experienced by an inertial body as it moves through a fluid medium, attempts were made to relax the assumptions under which the equations are valid. The creeping flow assumption is restrictive and often unphysical in the applications of interest, thus there was a desire to extend this work to the case where the Reynolds number is greater than zero. The need for this was identified in the late 19th century by Whitehead [104], although his solution was not valid far from the sphere. This problem was solved by an improvement due to Oseen in 1910 [82]. Oseen's correction was shown to be valid asymptotically by Lamb in 1911 [67], who used the same method to resolve the problem of two dimensional flow past a cylinder which had eluded Stokes decades previously. For small positive Reynolds number, the drag coefficient was improved upon, also by Oseen [83], and Faxén [40] offered a further improvement by including the effects of the nonuniformity of the fluid flow around the sphere. This contribution is commonly known as Faxén's law, and the corresponding terms as the Faxén corrections. A move towards more contemporary models was made Tchen's thesis [98], where the inertial particle was assumed to be within a time dependent velocity field, as opposed to within a steady flow, and the resulting model was used to investigate particles suspended in a turbulent fluid.

4.3 The Maxey-Riley equation

Work on inertial spherical particles culminated in a publication by Maxey and Riley [77] in 1983, which clarified these previous contributions and presented an equation of motion which has dominated the literature since its introduction [78]. After a modification, by Maxey [75, 76], allowing the particle's initial velocity to be distinct from that of the fluid, and a correction to the added mass term by Auton et al. [5], the equation reached the state in which it is commonly used in the literature today. This equation is known as the *Maxey-Riley* equation and has proven popular to researchers as an 'off-the-shelf' solution. Despite the seminal paper [77] shining a light on the assumptions under which the equation is valid, thus raising awareness to the community of its limitations, the Maxey-Riley equation has been applied to a wide number of physical problems. Whilst the direct applicability of the Maxey-Riley equation to problems such as the advection of drifter buoys is questionable due to assumptions on the Reynolds number, its study is still relevant since it has a distinctive mathematical structure inherited from the Basset-Boussinesq term and it is well used in the wider research community.

Unfortunately, the equation is difficult to implement in practice. This is in part due to the Basset-Boussinesq memory term, which presents severe storage complications if attempting to numerically solve the equation for a large scale operational fluid problem. The equation also poses mathematical problems, it is nonlinear and features a nonlocal integral term. Since these complications revolve around the Basset-Boussinesq term, some studies simplify the equation by omitting the memory term to create a more usable equation [78]. The Faxén corrections have also been omitted in the literature to aid usability [46]. Whilst this is sometimes argued with scaling assumptions, a growing body of work in support of the Basset-Boussinesq term [35, 36, 45] has made these simplifications look questionable. On the other hand, the Basset-Boussinesq term implies that particles always retained a dependence on their initial velocity, regardless of the viscosity of the fluid, which may be argued to be unphysical since one expects viscous dissipation to eradicate this memory [90]. It's worth noting that making crude simplifications to the Maxey-Riley equation may appear reasonable on the surface, however throwing away any term in the Maxey-Riley equation is equivalent to ignoring the corresponding result from the historic literature. Nonetheless, returning again to motivation from Figures 4.1 and 4.2, authors have attempted to use a simplified version of Maxey-Riley for ocean transport applications [7]. Attempts have been made to circumvent the need for the Basset history term by using a simplified Maxey-Riley equation and including a stochastic noise to match the equation to experimental data [91]. Successful attempts have been made to use the Maxey-Riley equation as inspiration to create a framework more tailored to oceanography applications [8, 9, 81]. As a separate issue, the suitability of the Maxey-Riley equation for particles in the ocean remains contentious due to the assumption on the size of the Reynolds number in the seminal paper [77]. A new approach [88] has given hope to the idea of numerically solving the 'full' Maxey-Riley equation with memory, thus interest in the Basset-Boussinesq term is not fading.

We now concern ourselves with the mathematical analysis of the Maxey-Riley equations and, more broadly, with nonlinear inertial particles featuring the Basset-Boussinesq term. Indeed, little mention of this topic has appeared in the literature. Recently, Farazmand and Haller [39] show the local existence and uniqueness of weak solutions of the Maxey-Riley equation. This was the only known result in this direction, prior to the publication corresponding to this chapter [34]. In the aforementioned study, [39], it is also shown that, if the solution is differentiable at its initial time, the equations of motion can be re-written into a form which does permit strong solutions. This chapter will extend this work.

The equations of motion. The results we prove here extend those found by Farazmand and Haller [39]. To ensure clarity we will use notation consistent with theirs, which is equivalent to the model suggested in the seminal work [77] with the commonly accepted corrections made by Auton [5] and Maxey [75, 76].

Note that, unlike in previous sections, we will not be using bold characters to denote vector quantities. This is because we will have vectors of dimension $2n$ as well as n, and the equations are more legible as they appear. It is clear from the context which variables are vector valued, and this is stated where necessary.

For a fluid moving in a domain $\mathscr{D} \subseteq \mathbb{R}^n$ with velocity field $u : \mathscr{D} \times [0, \infty) \to \mathbb{R}^n$, we denote the trajectory of a inertial particle with mass released at time t_0 by $y : [t_0, \infty) \to \mathbb{R}^n$, and its velocity by $v : [t_0, \infty) \to \mathbb{R}^n$. Since physical applications have dimension three, we will nondimensionalise with the assumption that our sphere and domain are three dimensional, which simplifies the form of the constants. Note that the equations which result from this have a higher dimensional generalisation of the same form. We nondimensionalise our problem by length scale L, time scale T, and velocity U which are characteristic to the ambient flow u. For this flow, we have a Reynolds number (Re) and, for a particle of radius a, the problem corresponds to a Stokes number (St) where these quantities are defined via the kinematic viscosity ν

by

$$\mathrm{Re} = \frac{\mathrm{UL}}{\nu}, \quad \mathrm{St} = \frac{2}{9}\frac{\mathrm{a}^2}{\nu \mathrm{T}} . \tag{4.1}$$

In a frame of reference moving with the particle, the Maxey-Riley equation may be written in the following form

$$\begin{aligned} \dot{y} &= v \\ \dot{v} &= R\frac{Du}{Dt} + \left(1 - \frac{3R}{2}\right) g + \frac{R}{2}\frac{D}{Dt}\left(u + \frac{\gamma}{10}\mu^{-1}\Delta u\right) \\ &\quad - \mu\left(v - u - \frac{\gamma}{6}\mu^{-1}\Delta u\right) - \kappa\mu^{1/2}\frac{d}{dt}\int_{t_0}^{t}\frac{w(s)}{\sqrt{t-s}}\,ds , \end{aligned} \tag{4.2}$$

where $u = u(y(t), t)$, and $w(t)$ is defined by

$$w(t) = \dot{y}(t) - u(y(t), t) - \frac{\gamma}{6}\mu^{-1}\Delta u(y(t), t) . \tag{4.3}$$

The additional parameters are defined by

$$R = \frac{2\rho_f}{\rho_f + 2\rho_p}, \quad \mu = \frac{R}{\mathrm{St}}, \quad \kappa = \sqrt{\frac{9R}{2\pi}}, \quad \gamma = \frac{9R}{2\mathrm{Re}} , \tag{4.4}$$

where ρ_f and ρ_p denote the densities of the fluid and particle respectively. As in [39], we write equation (4.2) in the following form

$$\begin{aligned} \dot{y} &= w + A_u(y, t) , \\ \dot{w} &= -\mu w - M_u(y, t)w - \kappa\mu^{1/2}\frac{d}{dt}\int_{t_0}^{t}\frac{w(s)}{\sqrt{t-s}}\,ds + B_u(y, t) , \end{aligned} \tag{4.5}$$

where $A_u, B_u : \mathscr{D} \times [t_0, \infty) \to \mathbb{R}^n$ and $M_u : \mathscr{D} \times [t_0, \infty) \to \mathbb{R}^{n\times n}$ are defined by

$$
\begin{aligned}
A_u &= u + \frac{\gamma}{6}\mu^{-1}\Delta u\,, \\
B_u &= \left(\frac{3R}{2} - 1\right)\left(\frac{Du}{Dt} - g\right) + \left(\frac{R}{20} - \frac{1}{6}\right)\gamma\mu^{-1}\frac{D}{Dt}\Delta u \\
&\quad - \frac{\gamma}{6}\mu^{-1}\left(\nabla u + \frac{\gamma}{6}\mu^{-1}\nabla\Delta u\right)\Delta u\,, \\
M_u &= \nabla u + \frac{\gamma}{6}\mu^{-1}\nabla\Delta u\,.
\end{aligned}
\tag{4.6}
$$

Well-posedness properties. In order to study the desired analytical properties of the equation, we first define what it means for a solution to be sufficiently regular to be a *strong* solution. To do this, we consider the integrated version of equation (4.5)

$$
\begin{aligned}
y(t) &= y_0 + \int_{t_0}^{t} w(s) + A_u(y(s), s)\, ds\,, \\
w(t) &= w_0 + \int_{t_0}^{t} \Big(- \mu w(s) - M_u(y(s), s) w(s) \\
&\qquad\qquad - \kappa\mu^{1/2}\frac{w(s)}{\sqrt{t-s}} + B_u(y(s), s)\Big)\, ds\,.
\end{aligned}
\tag{4.7}
$$

Remark 4.1. The equation (4.7) is not a standard ordinary differential equation, since the integrand of the equation in w has t as an argument. The classical theory of ordinary differential equations therefore cannot be applied, and in order to study this equation we need to develop all notions and results from the classical theory in the new context.

Definition 4.2. A solution of (4.5) is called *weak* if it satisfies the integrated formulation (4.7). A solution of (4.5) is called *strong* if it satisfies (4.7) and also is also differentiable with respect to the time variable.

Remark 4.3. One cannot determine whether a solution exists for the system of equations (4.5) from any general theorems known to the authors from the literature of both ordinary and fractional order differential equations (see e.g. [85]). This is due to the specific nature of the nonlinearity of the system.

The analytical properties discussed here will consist of existence and uniqueness of both weak and strong solutions, and the conditions under which these solutions exist. We will summarise the existing literature results, before building on them [34]. Before this, however, we will explore the mathematical structure of the integral term which diverts this equation away from the theory of ordinary differential equations.

4.4 The Basset-Boussinesq fractional integral term

As first observed by Tatom [97] and eluded to in Remark 4.3, the Basset-Boussinesq history integral is a 'semiderivative', meaning a fractional derivative of order 1/2. Thus the Maxey-Riley equation, and similar inertial particle models such as the Basset-Boussinesq-Oseen equation, are fractional differential equations. The notion behind the definition of a derivative of fractional order dates back to a 17th century letter between L'Hôpital and Leibniz, who speculated that 'useful consequences' would come from its study. Despite this, most research and development of fractional calculus has been performed relatively recently and the formulation of the history integral term by Boussinesq and Basset is a very early example of fractional calculus is the applied sciences, though they did not know this. The remark that the Basset-Boussinesq term is a fractional derivative has been made by a number of authors whilst writing about the Maxey-Riley equation [35, 36, 39], though its implications have been largely neglected and the equation is not currently well known in fractional calculus research communities. To illustrate this point, we recall a definition of the fractional derivative [85].

Definition 4.4. For a real number $p \in \mathbb{R}$, define the integer $n \in \mathbb{Z}$ to be such that $n - 1 \leqslant p < n$. We may then define the *left Riemann-Liouville fractional derivative of order* p by

$$ {}_aD^p f(t) = \frac{1}{\Gamma(n-p)} \frac{d^n}{dt^n} \int_a^t (t-s)^{n-p-1} f(s)\, ds \,. \tag{4.8} $$

By comparing (4.8) with (4.5), we can immediately see that the history integral is a Riemann-Liouville fractional derivative of order 1/2

$$ {}_{t_0}D^{1/2} w(t) = \frac{1}{\sqrt{\pi}} \frac{d}{dt} \int_{t_0}^t \frac{w(s)}{\sqrt{t-s}}\, ds \,. \tag{4.9} $$

We may rewrite the Maxey-Riley equations (4.5) as a fractional order equation explicitly as follows

$$ \begin{aligned} {}_{t_0}D^1 y(t) &= w + A_u(y,t)\,, \\ {}_{t_0}D^1 w(t) &= -\mu w - M_u(y,t)w - \kappa\mu^{1/2}\sqrt{\pi}\,{}_{t_0}D^{1/2} w(t) + B_u(y,t)\,. \end{aligned} \tag{4.10} $$

To give a good definition of a fractional derivative, one hopes that the definition would extend the notion of regular calculus in a reasonable way. There are multiple ways of defining such a derivative, each with their own desirable properties, and here we will introduce another which has application to inertial particle motion. An approach to fractional calculus was developed by Caputo [18] which allows for initial value problems to be formulated to involve only the values of integer derivatives of the variables at t_0 (see e.g. [85]). This means that, in general, initial value problems for Caputo-type fractional differential equations feature physically interpretable initial conditions. The definition of this derivative is as follows.

Definition 4.5. For a non-integer real number $p \in \mathbb{R}\backslash\mathbb{Z}$, define the integer

$n \in \mathbb{Z}$ to be such that $n-1 < p < n$. We may then define the *Caputo fractional derivative of order* p by

$$ {}_a^C D_t^p f(t) = \frac{1}{\Gamma(n-p)} \int_a^t \frac{f^{(n)}(s)}{(t-s)^{p+1-n}}\, ds\,. \tag{4.11} $$

As shown in [85], this derivative is a true interpolation between derivatives of integer order, since

$$ \lim_{p\to n} {}_a^C D_t^p f(t) = f^{(n)}(t)\,. \tag{4.12} $$

As we will see, there is a connection between Riemann-Liouville and Caputo fractional derivatives and whether the solution of the Maxey-Riley equation which exists is weak or strong.

4.5 Existing results and their extension

A summary of existing results. As mentioned above, we will be building on a previous study [39] which contained the following as its main result.

Theorem 4.6. *If $u(x,t)$ is three times continuously differentiable in both x and t and all of its partial derivatives are uniformly bounded and Lipschitz continuous up to order three then, for any initial condition $(y_0, w_0) \in \mathscr{D} \times \mathbb{R}^n$, there exists some $T > t_0$ such that over the time interval $[t_0, T)$ the equations* (4.5) *have a unique weak solution $(y(t), w(t))$ with $(y(t_0), w(t_0)) = (y_0, w_0)$.*

Proof. See [39] for full details, a sketch of the proof is given here. Recall that, by the definition of a weak solution, the theorem is equivalent to proving that the integral equation (4.7) has a unique solution. This may be done by a fixed point argument by reformulating the problem into into a fixed point theorem for the map

$$(P\Phi)(t) = \begin{pmatrix} y_{t_0} + \int_{t_0}^{t} \eta(s) + A_u(\xi(s), s)\, ds \\ w_{t_0} + \int_{t_0}^{t} -\left(\mu + \frac{\kappa\mu^{1/2}}{\sqrt{t-s}} + M_u(\xi(s), s)\right)\eta(s) + B_u(\xi(s), s)\, ds \end{pmatrix} \tag{4.13}$$

where $\Phi := (\xi, \eta)$. It can be shown that this map is a contraction mapping on a complete metric space. $\mathscr{Q.E.D.}$

Remark 4.7. Within the proof of this result, it is also shown that y, w have continuous paths on the interval $[t_0, t_0 + T]$, and are hence bounded.

It has been observed [39] that a weak solution of the Maxey-Riley equation exists locally in time but, under additional assumptions, these solutions may be strong. A further observation was made, though not proven, that solutions to (4.7) are not necessarily (continuously) differentiable and hence are not, in general, also solutions to (4.10). This relates to the fractional structure of the model since, if continuously differentiable solutions to the differential form of the equation (4.10) exist, then under the special initial condition $w(t_0) = 0$ the Basset history term takes the form

$$\frac{d}{dt}\int_{t_0}^{t} \frac{w(s)}{\sqrt{t-s}}\, ds = \int_{t_0}^{t} \frac{\dot{w}(s)}{\sqrt{t-s}}\, ds\,. \tag{4.14}$$

Reviewing Definitions 4.4 and 4.5, we see that this is equivalent to saying that, in this special case, the Riemann-Liouville fractional derivative takes the form of a Caputo fractional derivative of the same order. Thus, if continuously differentiable solutions to (4.10) exist *and* $w(t_0) = 0$, then (4.10) can be written as

$$
\begin{aligned}
{}_{t_0}D_t^1 y(t) &= w + A_u(y,t)\,, \\
{}_{t_0}D_t^1 w(t) &= -\mu w - M_u(y,t)w - \kappa\mu^{1/2}\sqrt{\pi}\,{}_{t_0}^{C}D_t^{1/2} w(t) + B_u(y,t)\,.
\end{aligned}
\tag{4.15}
$$

It has been shown, by using a similar fixed point argument, that the equations (4.15) have a unique local in time solution.

Extending the known results. The analytical remarks on the Maxey-Riley equation summarised above may be extended. Progress in this direction has been made [34] during this project, which will constitute the remainder of this section. In particular, the local in time weak solution corresponding to Theorem 4.6 will be shown to exist *globally* in time. Furthermore, the conditions under which this solution is differentiable with respect to the time variable and is thus a *strong* solution will be rigorously proven.

4.6 Global existence and uniqueness of solutions

First notice that the existence of the solution in the previous section was stated on a closed time interval. The result also holds on the half open interval $[t_0, T)$. Indeed, if we were to include the endpoint T we would have complications defining the derivative at T due to the inability to define the limit from above. Furthermore, the inclusion of the half open interval will allow us to introduce the notion of a maximal solution. We will be extending Theorem 4.6 to a global in time existence and uniqueness result, which will be done by extending notions from the classical theory of ordinary differential equations to the fractional case, as well as using a tailored Grönwall lemma for fractional differential equations (see Appendix B). We will be working under the following assumptions.

> The velocity field, u, and its derivatives are sufficiently smooth to ensure that the first derivatives in time and space of A_u and B_u are continuous and uniformly bounded in time and space. $\quad (*)$

In the assumptions $(*)$, the operators being 'uniformly bounded' means that their supremum norms are bounded by some constant. Thus there exists some constant L_b such that

$$\|\partial_t A_u\|_\infty, \|\nabla A_u\|_\infty, \|\partial_t B_u\|_\infty, \|\nabla B_u\|_\infty < L_b\,.$$

Notice also that the assumptions $(*)$ are sufficient to ensure that A_u and B_u are Lipschitz in space uniformly in time. Therefore there exists some $L_c > 0$ such that for any $t \in [t_0, t_0 + T)$ and $y_1, y_2 \in \mathscr{D}$, we have

$$\begin{aligned}|A_u(t, y_1) - A_u(t, y_2)| &\leqslant L_c|y_1 - y_2|\,,\\ |B_u(t, y_1) - B_u(t, y_2)| &\leqslant L_c|y_1 - y_2|\,.\end{aligned}$$

Note that A_u and B_u are similarly Lipschitz in time. Notice that these assumptions are more relaxed than those used in the proof of Theorem 4.6 [39], since we do not assume boundedness of A_u and B_u. It can be shown that the proofs of the lemmata in [39] can be modified to circumvent the need for this boundedness, indeed the modifications needed are present in the methodology we will use to prove the lemmata in this section. We will now develop some notions and prove some results from ordinary differential equations.

4.6.1 The existence of a maximal solution

Definition 4.8. A solution $(\tilde{y}, \tilde{w})$ of (4.7) with domain $[t_0, \tilde{T})$ is called an *extension* of the solution (y, w) with domain $[t_0, T)$ if $t_0 < T < \tilde{T}$ and the solutions are identical on $[t_0, T)$. The solution (y, w) is called *maximal* if there exists no such extension.

We will now prove that a maximal solution exists for the Maxey-Riley equation, which will require proving an initial proposition and lemma.

Proposition 4.9. *Suppose (y_1, w_1) and (y_2, w_2) are two solutions to (4.7) with domains $[t_0, T_1)$ and $[t_0, T_2)$ respectively, corresponding to the same initial condition (y_0, w_0), then the two solutions coincide on $[t_0, \min\{T_1, T_2\})$.*

Proof. For t in $[t_0, \min\{T_1, T_2\})$ we have

$$\begin{aligned} y_i(t) &= y_0 + \int_{t_0}^{t} w_i(s) + A_u(y_i(s), s)\, ds, \quad i = 1, 2\,, \\ w_i(t) &= w_0 + \int_{t_0}^{t} \bigg(-\mu w_i(s) - M_u(y_i(s), s) w_i(s) \\ &\qquad\qquad - \kappa \mu^{1/2} \frac{w_i(s)}{\sqrt{t-s}} + B_u(y_i(s), s) \bigg)\, ds\,, \quad i = 1, 2\,. \end{aligned}$$

Now consider the Euclidean norm of the differences $\|y_1 - y_2\|$ and $\|w_1 - w_2\|$, and we find a bound on these as follows:

$$\begin{aligned} \|y_1(t) - y_2(t)\| &= \left\| \int_{t_0}^{t} w_1(s) - w_2(s) + A_u(y_1(s), s) - A_u(y_2(s), s)\, ds \right\| \\ &\leqslant \int_{t_0}^{t} \|w_1(s) - w_2(s)\| + L_c \|y_1(s) - y_2(s)\|\, ds\,, \end{aligned}$$

and

$$
\begin{aligned}
\|w_1(t) - w_2(t)\| &= \bigg\| \int_{t_0}^{t} -\mu(w_1(s) - w_2(s)) - M_u(y_1(s), s)w_1(s) \\
&\qquad + M_u(y_2(s), s)w_2(s) - \kappa\mu^{1/2}\left(\frac{w_1(s)}{\sqrt{t-s}} - \frac{w_2(s)}{\sqrt{t-s}}\right) \\
&\qquad + B_u(y_1(s), s) - B_u(y_2(s), s)\, ds \bigg\| \\
&\leqslant \int_{t_0}^{t} \mu\|w_1(s) - w_2(s)\| + L_b\|w_1(s) - w_2(s)\| \\
&\qquad - \frac{\kappa\mu^{1/2}}{\sqrt{t-s}}\|w_1(s) - w_2(s)\| + L_c\|y_1(s) - y_2(s)\|\, ds\,.
\end{aligned}
$$

We have uniqueness from an application of the Grönwall result from Appendix B with '$u(t)$' equal to $\|y_1(t) - y_2(t)\| + \|w_1(t) + w_2(t)\|$, noting that $a(t) = 0$ in the case of the above bounds. $\mathscr{Q}.\mathscr{E}.\mathscr{D}.$

In the following lemma, we will show that the above result will allow us to combine multiple local in time solutions into another solution. This will enable us to prove that a maximal solution exists.

Lemma 4.10. *Let* $\{(y_\alpha(t), w_\alpha(t))\}_{\alpha\in A}$ *be a family of solutions to* (4.7) *with initial condition* (y_0, w_0)*, where* A *is an arbitrary index set. Let the domain of* (y_α, w_α) *be* $[t_0, T_\alpha)$*. We define* T *such that* $[t_0, T) = \bigcup_{\alpha\in A}[t_0, T_\alpha)$*, and define a function on* $[t_0, T)$ *by*

$$(y(t), w(t)) = (y_\alpha(t), w_\alpha(t)), \quad \textit{if} \quad t \in [t_0, T_\alpha)\,. \tag{4.16}$$

Then $(y(t), w(t))$ *is also a solution to* (4.7) *with the same initial condition.*

Proof. We must justify first that (4.16) gives a consistent definition of (y, w), i.e. that $(y(t), w(t))$ does not depend on the choice of α. For $t \in [t_0, T_{\alpha_1})$, (4.16) gives that $(y(t), w(t)) = (y_{\alpha_1}(t), w_{\alpha_1}(t))$. If t also belongs to $[t_0, T_{\alpha_2})$, then $t \in [t_0, \min\{T_{\alpha_1}, T_{\alpha_2}\})$ and therefore our uniqueness result Proposition 4.9 implies that $(y_{\alpha_1}(t), w_{\alpha_1}(t)) = (y_{\alpha_2}(t), w_{\alpha_2}(t))$ for this value of t.

Now we prove that (y, w) defined by (4.16) defines a solution to (4.7) on $[t_0, T)$. We know that $t_0 \in [t_0, T_\alpha)$ for any α and therefore

$$(y(t_0), w(t_0)) = (y_\alpha(t_0), w_\alpha(t_0)) = (y_0, w_0)\,, \tag{4.17}$$

since (y_α, w_α) is a solution to (4.7) with initial condition (y_0, w_0). Furthermore, for any $t \in [t_0, T)$ there exists α such that $t \in [t_0, T_\alpha)$. We know that (y_α, w_α) solves (4.7) on $[t_0, T_\alpha)$, $(y, w) = (y_\alpha, w_\alpha)$ on $[t_0, T_\alpha)$, and therefore (y, w) solves (4.7) at any $t \in [t_0, T)$. $\mathscr{Q.E.D.}$

Proposition 4.11. *Assume that u satisfies the conditions in $(*)$, then we have a unique maximal solution to (4.7) with initial condition (y_0, w_0).*

Proof. We need only prove that the solution identified in Lemma 4.10 corresponding to the family of all possible solutions to (4.7) with initial condition (y_0, w_0) is the unique maximal solution. We know that (y, w) is indeed a solution of (4.7) and it is maximal since its domain contains the domains of all other possible solutions. It only remains to prove uniqueness.

Let $(\tilde{y}, \tilde{w})$ be another such maximal solution. Similar to Lemma 4.10, the union of our two maximal solutions is a solution of (4.7) with initial condition (y_0, w_0) and extends (y, w) and $(\tilde{y}, \tilde{w})$. By the definition of maximality, this union must be identical to both (y, w) and $(\tilde{y}, \tilde{w})$ and hence we have uniqueness. $\mathscr{Q.E.D.}$

4.6.2 Existence of a global in time solution

Now that we have established the existence of a maximal solution, we will now apply this to prove that a global solution exists. This will begin with the following result.

Theorem 4.12. *Assume that* u *satisfies the conditions* $(*)$*, then if* $(y(t), w(t))$ *is a maximal solution with domain* $[t_0, T)$ *and* T *is finite, then* (y, w) *leaves any compact set* $S \subset \mathscr{D} \times \mathbb{R}^n$ *as* t *approaches* T*.*

Proof. Let (y, w) be a maximal solution of (y, w) to (4.7) with domain $[t_0, T)$ corresponding to an initial condition (y_0, w_0). Assume further that there exists a compact set $S \subset \mathscr{D} \times \mathbb{R}^n$ such that the solution remains inside S, i.e. $\forall \tau \in (t_0, T)$, $\exists t_1 \in (\tau, T)$ s.t. $(y(t_1), w(t_1)) \in S$. We will find a contradiction and hence conclude that no such S exists.

Take a sequence $\{(y_n, w_n)\}_{n \in \mathbb{N}}$ defined by $(y_n, w_n) := (y(t_n), w(t_n))$ for a sequence $t_n \to T$. Furthermore, we assume that $(y_n, w_n) \in S$ for all n. Since S is compact, there exists a converging subsequence $\{(y_{n_k}, w_{n_k})\}_{k \in \mathbb{N}}$ where $t_{n_k} \to T$. We call the limit of this sequence (y_T, w_T):

$$(y_{n_k}, w_{n_k}) \xrightarrow[k \to \infty]{} (y_T, w_T) \in S\,.$$

We may take an element of the sequence which is 'arbitrarily close' to (y_T, w_T) in the following way: $\forall \varepsilon > 0\ \exists (y_{t_1}, w_{t_1}) \in S$ s.t. $|T - t_1| < \varepsilon$. We will pick t_1 close to T, and use this as an initial condition (y_{t_1}, w_{t_1}) for a Maxey-Riley equation with memory starting at a time before t_1, at t_0. In the setup, we have that (y, w) is given on $[t_0, t_1)$ (and indeed beyond this to T) and by construction our Maxey-Riley equation starting at t_1 will be shown to extend our solution beyond T, hence contradicting maximality.

We have that w at t_1 is given by

$$\begin{aligned} w(t_1) = w_0 + \int_{t_0}^{t_1} -\mu w(s) - M_u(y(s),s)w(s) + B_u(y(s),s)\, ds \\ - \kappa\mu^{1/2}\int_{t_0}^{t_1} \frac{w(s)}{\sqrt{t_1 - s}}\, ds\,. \end{aligned} \tag{4.18}$$

If w is extendable beyond T, then for $t > T$ we would have

$$\begin{aligned} w(t) = w_0 + \int_{t_0}^{t} -\mu w(s) - M_u(y(s),s)w(s) + B_u(y(s),s)\, ds \\ - \kappa\mu^{1/2}\int_{t_0}^{t} \frac{w(s)}{\sqrt{t - s}}\, ds\,, \end{aligned} \tag{4.19}$$

and thus

$$\begin{aligned} w(t) - w(t_1) &= \int_{t_1}^{t} -\mu w(s) - M_u(y(s),s)w(s) + B_u(y(s),s)\, ds \\ &\qquad - \kappa\mu^{1/2}\int_{t_0}^{t} \frac{w(s)}{\sqrt{t - s}}\, ds + \kappa\mu^{1/2}\int_{t_0}^{t_1} \frac{w(s)}{\sqrt{t_1 - s}}\, ds \\ &= \int_{t_1}^{t} -\mu w(s) - M_u(y(s),s)w(s) + B_u(y(s),s) - \kappa\mu^{1/2}\frac{w(s)}{\sqrt{t - s}}\, ds \\ &\quad + \kappa\mu^{1/2}\int_{t_0}^{t_1} \frac{w(s)}{\sqrt{t_1 - s}} - \frac{w(s)}{\sqrt{t - s}}\, ds\,. \end{aligned} \tag{4.20}$$

We will consider equation (4.20) together with

$$y(t) = y(t_1) + \int_{t_1}^{t} w(s) + A_u(y(s),s)\, ds\,. \tag{4.21}$$

We want to prove that this system has solutions on an interval of length δ depending only the compact set S and not on t_1. We define the map

$$(P\Phi)(t) = \begin{pmatrix} (P\Phi)_1(t) \\ (P\Phi)_2(t) \end{pmatrix}, \tag{4.22}$$

where

$$
\begin{aligned}
(P\Phi)_1(t) &= y_{t_1} + \int_{t_1}^{t} \eta(s) + A_u(\xi(s), s)\, ds\,, \\
(P\Phi)_2(t) &= w_{t_1} + \int_{t_1}^{t} -\left(\mu + \frac{\kappa\mu^{1/2}}{\sqrt{t-s}} + M_u(\xi(s), s)\right)\eta(s) + B_u(\xi(s), s)\, ds \\
&\qquad + \kappa\mu^{1/2} \int_{t_0}^{t_1} \frac{w(s)}{\sqrt{t_1 - s}} - \frac{w(s)}{\sqrt{t-s}}\, ds\,.
\end{aligned}
$$

Note that a solution to equations (4.20) and (4.21) corresponds to a fixed point of the map P. We define R to be such that $S \subseteq \bar{B}_0(R)$, then lemmata C.1 and C.2 from Appendix C give that, for $K = 4\max\{R, 2R\sqrt{T - t_0}\}$ and any δ chosen such that

$$
\begin{aligned}
\delta + \mu\delta + 2\kappa\mu^{1/2}\sqrt{\delta} + L_b\delta < \delta + \mu\delta + 2\kappa\mu^{1/2}\sqrt{\delta} + 3L_b\delta &< 1/5\,, \\
(2 + K)L_c\delta &< 1/4\,, \\
(2L_b\delta + A_u(0, t_0) + B_u(0, t_0))\delta &< K/4\,,
\end{aligned}
$$

the map P has a unique fixed point.

To complete our proof, notice that δ here depends only on the Euclidean norm of the initial conditions, i.e. on the compact set S. Hence we may choose t_1 to be within a distance δ from T and thus we have extended our solution beyond the supposedly maximal domain. We have found the required contradiction and proven our theorem. $\mathscr{Q}.\mathscr{E}.\mathscr{D}.$

Theorem 4.13. *Assume that u satisfies the conditions $(*)$, then for any initial condition $(y_0, w_0) \in \mathscr{D} \times \mathbb{R}^n$, there exists a unique global solution $(y(t), w(t))$ (i.e. a solution on $[t_0, \infty)$) to the integral equation (4.7) with $(y(t_0), w(t_0)) = (y_0, w_0)$.*

To prove Theorem 4.13, we must first introduce a lemma which finds an appropriate bound on the solution to (4.7).

Lemma 4.14. *If u satisfies the conditions $(*)$ and (y,w) satisfies (4.7), on the interval $[t_0,T)$, there exists some (C_Y,C_W) depending on T,y_0,w_0,κ,μ and L_b such that*

$$\sup_{t\in[t_0,T)} |y(t)| \leqslant C_Y\,, \tag{4.23}$$

$$\sup_{t\in[t_0,T)} |w(t)| \leqslant C_W\,. \tag{4.24}$$

Proof (of Lemma 4.14). We seek to apply a bound on the solution using the integrated form of the equation. We first notice that, as in Appendix C, we may bound the integral of $B_u(y(s),s)$ using the Lipschitz property. That is, for any $x_1,x_2\in\mathbb{R}^n$ and $\tau_1,\tau_2\in\mathbb{R}$, we have

$$\begin{aligned}|B_u(x_1,\tau_1)-B_u(x_2,\tau_2)| &\leqslant |B_u(x_1,\tau_1)-B_u(x_2,\tau_1)| + |B_u(x_2,\tau_1)-B_u(x_2,\tau_2)|\\ &\leqslant \|\nabla B_u\|_\infty|x_1-x_2| + \|\partial_t B_u\|_\infty|\tau_1-\tau_2|\\ &\leqslant L_b(|x_1-x_2|+|\tau_1-\tau_2|)\,.\end{aligned}$$

Therefore, choosing $x_1=y(s)$, $\tau_1=s$, and $x_2=\tau_2=0$, we have that

$$\begin{aligned}\left|\int_{t_0}^t B_u(y(s),s)\,ds\right| &\leqslant \int_{t_0}^t L_b(|y(s)|+|s|)+|B_u(0,0)|\,ds\\ &\leqslant L_b\int_{t_0}^t |y(s)|\,ds + L_b\int_{t_0}^t |s|\,ds + |B_u(0,0)|(t-t_0)\\ &\leqslant L_b\int_{t_0}^t |y(s)|\,ds + \frac{L_b(t^2-t_0^2)}{2} + |B_u(0,0)|(t-t_0)\,.\end{aligned}$$

Beginning with the equation for w, we seek a bound on the solution as follows

$$
\begin{aligned}
|w(t)| &\leqslant |w_0| + \int_{t_0}^{t} \Big| -\mu w(s) - M_u(y(s),s)w(s) - \kappa\mu^{1/2}\frac{w(s)}{\sqrt{t-s}} + B_u(y(s),s)\Big|\, ds \\
&\leqslant |w_0| + \int_{t_0}^{t} |\mu + L_b||w(s)| + \Big|\kappa\mu^{1/2}\frac{w(s)}{\sqrt{t-s}}\Big| + |B_u(y(s),s)|\, ds \\
&\leqslant |w_0| + |\mu + L_b| \int_{t_0}^{t} |w(s)|\, ds + |\kappa\mu^{1/2}| \int_{t_0}^{t} (t-s)^{-1/2}|w(s)|\, ds \\
&\qquad + L_b \int_{t_0}^{t} |y(s)|\, ds + \frac{L_b(t^2 - t_0^2)}{2} + |B_u(0,0)|(t - t_0)\,.
\end{aligned}
$$

We now proceed with the equation for y

$$
\begin{aligned}
|y(t)| &\leqslant |y_0| + \int_{t_0}^{t} |w(s)| + |A_u(y(s),s)|\, ds \\
&\leqslant \int_{t_0}^{t} |w(s)|\, ds + L_b \int_{t_0}^{t} |y(s)|\, ds + \frac{L_b(t^2 - t_0^2)}{2} + |A_u(0,0)|(t - t_0)\,.
\end{aligned}
$$

We then must consider the above inequalities for $|y(s)|$ and $|w(s)|$ as a pair. In particular, we define $a(s)$ by

$$
a(s) = |y(s)| + |w(s)|\,,
$$

and then we have the following inequality

$$
\begin{aligned}
a(s) &\leqslant a(t_0) + L_b(t^2 - t_0^2) + (|A_u(0,0)| + |B_u(0,0)|)(t - t_0) \\
&\qquad + \max\{1, \mu + L_b\} \int_{t_0}^{t} a(s)\, ds + |\kappa\mu^{1/2}| \int_{t_0}^{t} (t-s)^{-1/2} a(s)\, ds\,.
\end{aligned}
\tag{4.25}
$$

A Grönwall-style result of S. Y. Lin [71], gives a bound on the solution as required (see Appendix B). $\mathscr{Q.E.D.}$

We are now prepared to prove Theorem 4.13.

Proof (of Theorem 4.13). Let (y, w) be the unique maximal solution from Proposition 4.11 and $[t_0, T)$ its domain. We aim to show that $[t_0, T) = [t_0, \infty)$. Assume the contrary is true, then T is finite. By Theorem 4.12, (y, w) leaves any compact set $S \Subset \mathscr{D} \times \mathbb{R}^n$ as $t \to T$. Take a specific compact set

$$S = \bar{B}_0(r_1) \times \bar{B}_0(r_2)\,.$$

For t sufficiently close to T, we know $\|y\| > r_1$ and $\|w\| > r_2$. Since r_1, r_2 were chosen arbitrarily, we may deduce that

$$\|y\|, \|w\| \to \infty, \quad \text{as } t \to T\,.$$

On the contrary, we have boundedness of y and w from Lemma 4.14. Thus we have reached a contradiction and proven our theorem. $\mathscr{Q.E.D.}$

4.7 Conditions under which a strong solution exists

The well-posedness results presented thus far have concerned themselves with weak solutions. In general, the Maxey-Riley equation permits weak solutions only. Under additional assumptions, however, we will show that the solution is sufficiently regular to be classified as strong. Recall that, when we assume the particular initial condition $w(t_0) = 0$, the Basset-Boussinesq history integral takes a particular form as in equation (4.14), repeated here for convenience.

$$\frac{d}{dt}\int_{t_0}^{t} \frac{w(s)}{\sqrt{t-s}}\,ds = \int_{t_0}^{t} \frac{\dot{w}(s)}{\sqrt{t-s}}\,ds\,.$$

Thus, the Maxey-Riley equations take the form of a fractional differential equation of Caputo type (4.15). It is known [39] that this form of the equations permits a strong solution with $y(t_0) = y_0$ and $w(t_0) = 0$.

It remains to prove under which conditions the solution to (4.7) is differentiable at t_0. This will involve proving the necessary conditions for the fractional integral

$$\int_{t_0}^{t} \frac{w(s)}{\sqrt{t-s}}\, ds\,, \tag{4.26}$$

to be differentiable at t_0. Without proving this, the expression (4.14) does not make sense at $t = t_0$ and thus we cannot say that continuously differentiable solutions to (4.15) are also continuously differentiable solutions to (4.10). This integral is of interest since difficulties which arise when considering the differentiability of solutions to (4.7) stem from the history term (4.26). With the following results we clarify under what assumptions this term is differentiable at t_0.

We first study the smoothness properties of the integral (4.26) which may assist in proving differentiability. Recall that since w is a solution to (4.7), it satisfies the continuity and boundedness properties in Theorem 4.6. We can certainly prove that the integral (4.26) is continuous at t_0. Indeed, for every $\epsilon > 0$ there exists $\delta > 0$ such that $2K\sqrt{\delta} < \epsilon$. For this ϵ, δ we have that for $|t - t_0| < \delta$:

$$\begin{aligned} \left| \int_{t_0}^{t} \frac{w(s)}{\sqrt{t-s}}\, ds - \int_{t_0}^{t_0} \frac{w(s)}{\sqrt{t_0-s}}\, ds \right| &= \left| \int_{t_0}^{t} \frac{w(s)}{\sqrt{t-s}}\, ds \right| \\ &\leqslant K \int_{t_0}^{t} \frac{1}{\sqrt{t-s}}\, ds \\ &\leqslant 2K\sqrt{t-t_0} < \epsilon\,. \end{aligned} \tag{4.27}$$

From this, one can deduce that the integral is 1/2-Hölder continuous.

Lemma 4.15. *Assuming* (y, w) *is a solution of* (4.7), *we have that*

(i) *the integral* (4.26) *is 1/2-Hölder, i.e. there exists some constant* $C > 0$

such that for any t_1, t_2 with $t_0 < t_1 < t_2$ we have

$$\left|\int_{t_0}^{t_2} \frac{w(s)}{\sqrt{t_2 - s}}\, ds - \int_{t_0}^{t_1} \frac{w(s)}{\sqrt{t_1 - s}}\, ds\right| \leqslant C|t_2 - t_1|^{1/2}\,, \tag{4.28}$$

and

(ii) *if $w(t_0) = 0$ then the following limit exists and is equal to zero*

$$\lim_{t\to t_0} \frac{w(t)}{\sqrt{t - t_0}} = 0\,. \tag{4.29}$$

Proof. We prove the two parts separately, beginning with part (i).

Part (i): We bound the left hand side of (4.28) as follows

$$\begin{aligned}\left|\int_{t_0}^{t_2} \frac{w(s)}{\sqrt{t_2 - s}}\, ds - \int_{t_0}^{t_1} \frac{w(s)}{\sqrt{t_1 - s}}\, ds\right| \leqslant &\left|\int_{t_1}^{t_2} \frac{w(s)}{\sqrt{t_2 - s}}\, ds\right| \\ &+ \left|\int_{t_0}^{t_1} \frac{w(s)}{\sqrt{t_2 - s}} - \frac{w(s)}{\sqrt{t_1 - s}}\, ds\right|\,,\end{aligned}$$

and it remains to prove that both terms on the right hand side are indeed 1/2-Hölder continuous. By the same argument as (4.27), the first term is 1/2-Hölder and so is the second term since

$$\begin{aligned}\left|\int_{t_0}^{t_1} \frac{w(s)}{\sqrt{t_2 - s}} - \frac{w(s)}{\sqrt{t_1 - s}}\, ds\right| &\leqslant K\left|\int_{t_0}^{t_1} \frac{1}{\sqrt{t_2 - s}} - \frac{1}{\sqrt{t_1 - s}}\, ds\right| \\ &\leqslant \left|-2K\sqrt{t_2 - t_1} + 2K\sqrt{t_2 - t_0} - 2K\sqrt{t_1 - t_0}\right| \\ &\leqslant 2K\sqrt{t_2 - t_1} + 2K\left|\sqrt{t_2 - t_0} - \sqrt{t_1 - t_0}\right| \\ &\leqslant 2K\sqrt{t_2 - t_1} + 2K\frac{|t_2 - t_1|}{\sqrt{t_2 - t_0} + \sqrt{t_1 - t_0}} \\ &\leqslant 2K\sqrt{t_2 - t_1} + 2K\frac{|t_2 - t_1|}{\sqrt{t_2 - t_1} + \sqrt{t_1 - t_1}} \\ &\leqslant 4K\sqrt{t_2 - t_1}\,,\end{aligned}$$

where in the first line we have made use of the boundedness property of solutions to the Maxey-Riley equation, see Theorem 4.6.

Part (ii): Recall that w satisfies

$$w(t) = w(t_0) + \int_{t_0}^{t} \Big(- \mu w(s) - M_u(y(s), s)w(s) \\ - \kappa\mu^{1/2} \frac{w(s)}{\sqrt{t-s}} + B_u(y(s), s) \Big)\, ds\,. \tag{4.30}$$

Dividing through by $\sqrt{t-t_0}$ and considering $w(t_0) = 0$, we have

$$\frac{w(t)}{\sqrt{t-t_0}} = \frac{1}{\sqrt{t-t_0}} \int_{t_0}^{t} \Big(- \mu w(s) - M_u(y(s), s)w(s) + B_u(y(s), s) \Big)\, ds \\ - \frac{\kappa\mu^{1/2}}{\sqrt{t-t_0}} \int_{t_0}^{t} \frac{w(s)}{\sqrt{t-s}}\, ds\,. \tag{4.31}$$

We know that w is locally bounded and, under assumptions $(*)$, M_u is uniformly bounded and B_u is sufficiently smooth to ensure that it is locally bounded on each interval $[t_0, t]$. Hence, there exists some c_1 which (locally) bounds the integrand of the first integral on the right hand side and hence

$$\frac{1}{\sqrt{t-t_0}} \int_{t_0}^{t} \left(-\mu w(s) - M_u(y(s), s)w(s) + B_u(y(s), s)\right)\, ds \leqslant c_1\sqrt{t-t_0}\,. \tag{4.32}$$

We may therefore deduce that

$$\lim_{t\to t_0} \frac{1}{\sqrt{t-t_0}} \int_{t_0}^{t} \left(-\mu w(s) - M_u(y(s), s)w(s) + B_u(y(s), s)\right)\, ds = 0\,. \tag{4.33}$$

It remains to show the existence of the limit

$$\begin{aligned}
&\lim_{t\to t_0} \frac{1}{\sqrt{t-t_0}} \int_{t_0}^{t} \frac{w(s)}{\sqrt{t-s}}\, ds = \\
&\lim_{t\to t_0} \left[\frac{1}{\sqrt{t-t_0}} \int_{t_0}^{t} \frac{1}{\sqrt{t-s}} \int_{t_0}^{s} -\mu w(r) - M_u(y(r),r)w(r) + B_u(y(r),r)\, dr\, ds \right] \\
&\qquad - \lim_{t\to t_0} \left[\frac{1}{\sqrt{t-t_0}} \int_{t_0}^{t} \frac{\kappa\mu^{1/2}}{\sqrt{t-s}} \int_{t_0}^{s} \frac{w(r)}{\sqrt{s-r}}\, dr\, ds \right].
\end{aligned} \tag{4.34}$$

where to show the equality we have used the equation (4.7) for $w(s)$. We may bound the middle line of (4.34) by

$$\begin{aligned}
&\frac{1}{\sqrt{t-t_0}} \int_{t_0}^{t} \frac{1}{\sqrt{t-s}} \int_{t_0}^{s} - \mu w(r) - M_u(y(r),r)w(r) + B_u(y(r),r)\, dr\, ds \\
&\qquad \leqslant \frac{c_1}{\sqrt{t-t_0}} \int_{t_0}^{t} \frac{s-t_0}{\sqrt{t-s}}\, ds\,.
\end{aligned}$$

The integral on the right hand side may be calculated by making a suitable substitution

$$\begin{aligned}
\int_{t_0}^{t} \frac{s-t_0}{\sqrt{t-s}}\, ds &= -\int_{\sqrt{t-t_0}}^{0} \frac{t-t_0-u^2}{u} 2u\, du = \int_{0}^{\sqrt{t-t_0}} (t-t_0) - u^2\, du \\
&= 2\left((t-t_0)^{3/2} - \frac{1}{3}(t-t_0)^{3/2} \right) = \frac{4}{3}(t-t_0)^{3/2}\,.
\end{aligned}$$

Therefore, the limit in the second line of (4.34) is equal to 0. We now observe the limit in the final line of (4.34), i.e.

$$\lim_{t\to t_0} \left[\frac{1}{\sqrt{t-t_0}} \int_{t_0}^{t} \frac{1}{\sqrt{t-s}} \int_{t_0}^{s} \frac{w(r)}{\sqrt{s-r}}\, dr\, ds \right]. \tag{4.35}$$

First recall that w is bounded by K, then we may bound and again calculate by making a series of suitable substitutions

$$\begin{aligned}
\frac{1}{\sqrt{t-t_0}}\int_{t_0}^{t}\frac{1}{\sqrt{t-s}}\int_{t_0}^{s}\frac{w(r)}{\sqrt{s-r}}\,dr\,ds &\leqslant \frac{K}{\sqrt{t-t_0}}\int_{t_0}^{t}\frac{1}{\sqrt{t-s}}\int_{t_0}^{s}\frac{1}{\sqrt{s-r}}\,dr\,ds\\
&\leqslant \frac{2K}{\sqrt{t-t_0}}\int_{t_0}^{t}\frac{\sqrt{s-t_0}}{\sqrt{t-s}}\,ds\\
&\leqslant \frac{2K}{\sqrt{t-t_0}}\int_{0}^{\sqrt{t-t_0}}\frac{2u^2}{\sqrt{t-u^2-t_0}}\,du\\
&\leqslant 4K\sqrt{t-t_0}\int_{0}^{\pi/2}\sin^2\theta\,d\theta\\
&\leqslant K\pi\sqrt{t-t_0}\,,
\end{aligned}$$

and thus

$$\lim_{t\to t_0}\left[\frac{1}{\sqrt{t-t_0}}\int_{t_0}^{t}\frac{1}{\sqrt{t-s}}\int_{t_0}^{s}\frac{w(r)}{\sqrt{s-r}}\,dr\,ds\right]=0\,.$$

Putting these calculations together, we have proven our claim. $\mathscr{Q}.\mathscr{E}.\mathscr{D}.$

Lemma 4.16. *A solution $w(t)$ of* (4.7) *is* 1/2*-Hölder on the interval $[t_0, t_0+\epsilon)$ and locally Lipschitz on $[t_0+\epsilon,\infty)$, for any $\epsilon>0$.*

Proof. By Lemma 4.15 (i), there exists some constant $C>0$ such that for any t_1, t_2 with $t_0<t_1<t_2$ we have that the memory term in (4.7) is 1/2-Hölder on the interval $[t_0, t_0+\epsilon)$ (see (4.28)). As all other terms in (4.7) are Lipschitz continuous, we deduce the 1/2-Hölder continuity of w.

It remains to prove that $w(t)$ is Lipschitz on any interval $[t_0+\epsilon, T]$, for any $\epsilon>0$ and $T>t_0+\epsilon$. Without loss of generality we will assume that $t_0=0$. The definition of $w(t)$ is then

$$\begin{aligned}
w(t)=w(0)+\int_0^t\Bigg(&-\mu w(s)-M_u(y(s),s)w(s)\\
&-\kappa\mu^{1/2}\frac{w(s)}{\sqrt{t-s}}+B_u(y(s),s)\Bigg)\,ds\,.
\end{aligned}$$

For $t, s \in [\epsilon, \infty)$, we aim to bound the difference between $w(t)$ and $w(s)$ by iterating an argument by which we multiply the time parameter by a number in $[0,1]$ before taking the supremum. Indeed, for $r_0 \in [0,1]$ and $t, s \in [\epsilon, \infty)$ we have

$$\begin{aligned} w(r_0 t) - w(r_0 s) = &\int_{r_0 s}^{r_0 t} (-\mu w(q) - M_u(y(q), q) w(q) + B_u(y(q), q)) \\ &- \kappa \mu^{1/2} \left(\int_0^{r_0 t} \frac{w(q)}{\sqrt{r_0 t - q}}\, dq - \int_0^{r_0 s} \frac{w(q)}{\sqrt{r_0 s - q}}\, dq \right) . \end{aligned}$$

Making substitutions $q = r_1 r_0 t$ and $q = r_1 r_0 s$ in the penultimate and final integrals in the above equation respectively, we have

$$\begin{aligned} w(r_0 t) - w(r_0 s) &= \int_{r_0 s}^{r_0 t} (-\mu w(q) - M_u(y(q), q) w(q) + B_u(y(q), q)) \\ &\quad - \kappa \mu^{1/2} \left[\sqrt{r_0 t} \int_0^1 \frac{w(r_1 r_0 t)}{\sqrt{1 - r_1}}\, dr_1 - \sqrt{r_0 s} \int_0^1 \frac{w(r_1 r_0 s)}{\sqrt{1 - r_1}}\, dr_1 \right] \\ &= \int_{r_0 s}^{r_0 t} (-\mu w(q) - M_u(y(q), q) w(q) + B_u(y(q), q)) \\ &\quad - \kappa \mu^{1/2} \Bigg[(\sqrt{r_0 t} - \sqrt{r_0 s}) \int_0^1 \frac{w(r_1 r_0 t)}{\sqrt{1 - r_1}}\, dr_1 \\ &\qquad + \sqrt{r_0 s} \int_0^1 \frac{w(r_1 r_0 t) - w(r_1 r_0 s)}{\sqrt{1 - r_1}}\, dr_1 \Bigg] , \end{aligned}$$

and from local boundedness of w and the assumptions $(*)$, this implies

$$|w(r_0 t) - w(r_0 s)| \leqslant K_0 |t - s| + \kappa \mu^{1/2} \sqrt{r_0 s} \int_0^1 \frac{|w(r_1 r_0 t) - w(r_1 r_0 s|)}{\sqrt{1 - r_1}}\, dr_1 .$$

To obtain the above estimate, we used the fact that (recall that $r_0 \in [0,1]$)

$$|\sqrt{r_0 t} - \sqrt{r_0 s}| \leqslant \sqrt{r_0} \frac{t - s}{\sqrt{t} + \sqrt{s}} \leqslant \frac{1}{2\sqrt{\epsilon}} |t - s| .$$

Within the integrand of the above inequality, we may iterate the argument by substituting in the definition of w to evaluate $w(r_1r_0t) - w(r_1r_0s)$. We claim that after iterating k times we have

$$
\begin{aligned}
&|w(r_0t) - w(r_0s)| \leqslant K_k|t-s| \\
&\quad + (\kappa\mu^{1/2}\sqrt{r_0s})^{k+1}\int_0^1\sqrt{\frac{r_1^k}{1-r_1}}\int_0^1\sqrt{\frac{r_2^{k-1}}{1-r_2}}\int_0^1\cdots \\
&\quad \cdots\int_0^1\frac{|w(r_{k+1}r_k\ldots r_0t) - w(r_{k+1}r_k\ldots r_0s)|}{\sqrt{1-r_{k+1}}}dr_{k+1}\ldots dr_1\,,
\end{aligned}
\tag{4.36}
$$

where $\{K_i, i \in \mathbb{N}\}$, κ is an appropriately chosen collection of constants. We will prove this by induction, where the above calculation acts as a base case. Suppose our inductive hypothesis (4.36) is true, then we iterate once more by substituting in the difference

$$
\begin{aligned}
&\big|w(r_{k+1}r_k\ldots r_0t) - w(r_{k+1}r_k\ldots r_0s)\big| \leqslant K_k|t-s| \\
&\quad + \kappa\mu^{1/2}\left|\left(\sqrt{r_{k+1}\ldots r_0t} - \sqrt{r_{k+1}\ldots r_0s}\right)\int_0^1\frac{w(r_{k+2}\ldots r_0t)}{\sqrt{1-r_{k+2}}}\,dr_{k+2}\right| \\
&\quad + \kappa\mu^{1/2}\left|\sqrt{r_{k+1}\ldots r_0s}\int_0^1\frac{w(r_{k+2}\ldots r_0t) - w(r_{k+2}\ldots r_0s)}{\sqrt{1-r_{k+2}}}\,dr_{k+2}\right| \\
&\leqslant K_{k+1}|t-s| \\
&\quad + \kappa\mu^{1/2}\sqrt{r_{k+1}\ldots r_0s}\int_0^1\frac{|w(r_{k+2}\ldots r_0t) - w(r_{k+2}\ldots r_0s)|}{\sqrt{1-r_{k+2}}}\,dr_{k+2}\,.
\end{aligned}
$$

Using our inductive hypothesis (4.36), we get

$$
\begin{aligned}
&|w(r_0t) - w(r_0s)| \leqslant K_{k+1}|t-s| \\
&\quad + (\kappa\mu^{1/2}\sqrt{r_0s})^{k+2}\int_0^1\sqrt{\frac{r_1^{k+1}}{1-r_1}}\int_0^1\sqrt{\frac{r_2^{k}}{1-r_2}}\int_0^1\cdots \\
&\quad \cdots\int_0^1\frac{|w(r_{k+2}\ldots r_0t) - w(r_{k+2}\ldots r_0s)|}{\sqrt{1-r_{k+2}}}dr_{k+2}\ldots dr_1\,,
\end{aligned}
\tag{4.37}
$$

and hence we have proven our claim by induction.

Noting that $r_i \in [0,1]$ for each $i \in \mathbb{N}$, we take supremum over $\{r_i, i \in \mathbb{N}\}$ to deduce that

$$\sup_{\alpha\in[0,1]} |w(\alpha t) - w(\alpha s)| \leqslant K_k|t-s| + \mathcal{I}_k \sup_{\alpha\in[0,1]} |w(\alpha t) - w(\alpha s)|\,, \tag{4.38}$$

where

$$\mathcal{I}_k = (\kappa\mu^{1/2}\sqrt{r_0 s})^{k+1} \int_0^1 \sqrt{\frac{r_1^k}{1-r_1}} \int_0^1 \sqrt{\frac{r_2^{k-1}}{1-r_2}} \int_0^1 \cdots \int_0^1 \frac{dr_{k+1}\ldots dr_1}{\sqrt{1-r_{k+1}}}\,. \tag{4.39}$$

We now observe the integral

$$a_k = \int_0^1 \sqrt{\frac{x^k}{1-x}}\, dx = 2\int_0^{\pi/2} \sin^{k+1}\theta\, d\theta\,. \tag{4.40}$$

Note that the sequence $(a_k)_k$ converges to zero as k tends to infinity, and does so at the same rate as $1/\sqrt{k}$. Indeed, note that

$$a_k = 2\int_0^{\pi/2} \sin^{k-1}\theta(1-\cos^2\theta)\, d\theta = a_{k-2} - \frac{a_k}{k}\,,$$

where we have used that

$$\begin{aligned} 0 = \frac{2\sin^k\theta}{k}\cos\theta\bigg|_0^{\pi/2} &= \int_0^{\pi/2} 2\sin^{k-1}\theta\cos^2\theta - \frac{2\sin^{k+1}\theta}{k}\, d\theta \\ &= \int_0^{\pi/2} 2\sin^{k-1}\theta\cos^2\theta\, d\theta - \frac{a_k}{k}\,. \end{aligned}$$

Therefore we have that

$$a_k = \frac{k}{k+1}a_{k-2} = \left(1 - \frac{1}{k+1}\right)a_{k-2}\,.$$

Since $1 - y \leqslant \exp(-y)$ for positive y, we have, for k odd

$$a_k \leqslant \frac{\pi}{2} \exp\left(-\sum_{n=2}^{\frac{k+1}{2}} \frac{1}{2n} \right),$$

and the case where k is even is similar. Note that we have used the fact that $a_1 = \pi/2$, in the case where k is even we would instead use $a_0 = 2$. Therefore

$$\sqrt{k} a_k \leqslant \frac{\pi}{2} \exp\left(\frac{1}{2} \log k - \frac{1}{2} \sum_{n=2}^{\frac{k+1}{2}} \frac{1}{n} \right).$$

In the following we will use the fact that the sequence $(b_k)_k$ defined as

$$b_k = \log k - \sum_{n=1}^{k} \frac{1}{n}$$

converges to the Euler-Mascheroni constant γ. We have that

$$\sqrt{k} a_k \leqslant \frac{\pi}{2} \exp\left(\frac{1}{2} \log k - \frac{1}{2} \left(\log(\frac{k+1}{2}) - 1 - b_{\frac{k+1}{2}} \right) \right).$$

Observe that the limit on the right hand side converges to $\frac{\pi}{2} \exp\left(\frac{\log(2)}{2} + \frac{1}{2} + \frac{\gamma}{2} \right)$. In particular, the sequence on the right hand side is bounded, since it converges, and hence the sequence $\sqrt{k} a_k$ is bounded above. The analysis for the even terms is similar. We have found that there exists M such that

$$\sqrt{k} a_k \leqslant M \quad \Longrightarrow \quad a_k \leqslant \frac{M}{\sqrt{k}},$$

thus $\mathcal{I}_k \to 0$ as

$$\mathcal{I}_k \leqslant (M \kappa \mu^{1/2} \sqrt{r_0 s})^{k+1} \frac{1}{\sqrt{k!}} \to 0 \quad \text{as} \quad k \to \infty .$$

Hence we choose $k \in \mathbb{N}$ sufficiently large to ensure that

$$\mathcal{I}_k < 1\,, \tag{4.41}$$

and thus (4.36) implies

$$(1 - \mathcal{I}_k) \sup_{\alpha \in [0,1]} |w(\alpha t) - w(\alpha s)| \leqslant K_k |t - s|\,. \tag{4.42}$$

Rearranging gives

$$|w(t) - w(s)| \leqslant \sup_{\alpha \in [0,1]} |w(\alpha t) - w(\alpha s)| \leqslant \frac{K_k}{1 - \mathcal{I}_k} |t - s|\,, \tag{4.43}$$

and we have thus proven our claim. $\mathscr{Q}.\mathscr{E}.\mathscr{D}.$

We may now approach the main result of the section.

Theorem 4.17. *Under the assumptions* $(*)$, *there exists a strong solution of the Maxey-Riley equation* (4.5) *if and only if* $w(t_0) = 0$.

Proof. To prove this theorem, note that the existence of classical solutions is equivalent to the differentiability of the integral (4.26) at t_0, since the differentiability of the remaining terms in (4.7) is trivial.

Firstly, assume that $w(t_0) \neq 0$. Then by adding and subtracting $w(t_0)$ to the numerator of the integrand of (4.26), we have

$$\begin{aligned} \int_{t_0}^{t} \frac{w(s)}{\sqrt{t-s}}\, ds &= \int_{t_0}^{t} \frac{w(s) - w(t_0)}{\sqrt{t-s}}\, ds + \int_{t_0}^{t} \frac{w(t_0)}{\sqrt{t-s}}\, ds \\ &= \int_{t_0}^{t} \frac{w(s) - w(t_0)}{\sqrt{t-s}}\, ds + 2w(t_0)\sqrt{t - t_0}\,. \end{aligned}$$

One may observe that the second term is not differentiable at t_0, indeed

$$\lim_{t\to t_0} \frac{2w(t_0)\sqrt{t-t_0}}{t-t_0} = \lim_{t\to t_0} \frac{2w(t_0)}{\sqrt{t-t_0}} = \infty\,.$$

Since, by Lemma 4.16, w is 1/2-Hölder as t approaches t_0, we know that for suitably small $t-t_0$ there exists some c_2 such that

$$\left|\int_{t_0}^{t} \frac{w(s)-w(t_0)}{\sqrt{t-s}}\right| ds \leqslant \int_{t_0}^{t} \frac{|w(s)-w(t_0)|}{\sqrt{t-s}} \leqslant c_2 \int_{t_0}^{t} \frac{\sqrt{s-t_0}}{\sqrt{t-s}}\, ds\,. \tag{4.44}$$

We have calculated the integral on the right hand side in the proof of Lemma 4.15 (ii), and thus

$$\frac{1}{t-t_0}\left|\int_{t_0}^{t} \frac{w(s)-w(t_0)}{\sqrt{t-s}}\, ds\right| \leqslant \frac{c_2\pi}{2}\,. \tag{4.45}$$

We may conclude that if $w(t_0) \neq 0$, then the integral (4.26) is not differentiable at t_0 and thus solutions of (4.7) are not differentiable at t_0 and are not classical solutions. Hence the contrapositive is true, and if solutions of (4.7) are classical solutions, then $w(t_0) = 0$.

We next prove the reverse implication, by assuming that $w(t_0) = 0$. By Lemma 4.15 (i), for $w(t_0) = 0$ the following function is bounded

$$q(s) = \frac{w(s)}{\sqrt{s-t_0}}\,,$$

since

$$\frac{|w(s)|}{\sqrt{s-t_0}} = \frac{|w(s)-w(t_0)|}{\sqrt{s-t_0}} \leqslant C\,,$$

by the definition of Hölder continuity. In this case we have

$$\left|\frac{1}{t-t_0}\int_{t_0}^{t}\frac{\sqrt{s-t_0}\,q(s)}{\sqrt{t-s}}\,ds\right| \leqslant \frac{C}{t-t_0}\left|\int_{t_0}^{t}\frac{\sqrt{s-t_0}}{\sqrt{t-s}}\,ds\right| \\ \leqslant \frac{C}{t-t_0}\left|\frac{\pi}{2}(t-t_0)\right| = \frac{C\pi}{2}\,. \tag{4.46}$$

By Lemma 4.15 (ii) the following limit exists

$$\lim_{s\to t_0} q(s) = 0\,, \tag{4.47}$$

and therefore for t suitably close to t_0 we have

$$\frac{1}{t-t_0}\int_{t_0}^{t}\frac{\sqrt{s-t_0}(0-\epsilon)}{\sqrt{t-s}}\,ds \leqslant \frac{1}{t-t_0}\int_{t_0}^{t}\frac{\sqrt{s-t_0}\,q(s)}{\sqrt{t-s}}\,ds \\ \leqslant \frac{1}{t-t_0}\int_{t_0}^{t}\frac{\sqrt{s-t_0}(0+\epsilon)}{\sqrt{t-s}}\,ds\,,$$

and thus

$$-\frac{C\pi}{2}\epsilon \leqslant \frac{1}{t-t_0}\int_{t_0}^{t}\frac{\sqrt{s-t_0}\,q(s)}{\sqrt{t-s}}\,ds \leqslant \frac{C\pi}{2}\epsilon\,.$$

Hence the integral (4.26) is differentiable at $t = t_0$, with value zero.

It remains to prove differentiability away from the initial time. By Lemma 4.16, $w(t)$ is Lipschitz and thus absolutely continuous, hence there exists a locally bounded measurable function $a : [t_0, \infty) \mapsto \mathbb{R}$ such that

$$w(s) = w(s) - w(t_0) = \int_{t_0}^{s} a(r)\,dr\,.$$

By integrating by parts, we have the identity

$$0 = \left(\sqrt{t-s}\int_{t_0}^{s} a(r)\,dr\right)\bigg|_{t_0}^{t} = \int_{t_0}^{t} \frac{d}{ds}\left(\sqrt{t-s}\big(w(t)-w(t_0)\big)\right) ds$$
$$= \int_{t_0}^{t} \frac{w(s)-w(t_0)}{2\sqrt{t-s}}\,ds + \int_{t_0}^{t} a(s)\sqrt{t-s}\,ds\,,$$

and we thus it suffices to prove the differentiability of

$$\begin{aligned} f(t) &= \int_{t_0}^{t} a(s)\sqrt{t-s}\,ds \\ &= -\int_{t_0}^{t} \frac{w(s)-w(t_0)}{2\sqrt{t-s}}\,ds = -\int_{t_0}^{t} \frac{w(s)}{2\sqrt{t-s}}\,ds\,. \end{aligned} \tag{4.48}$$

Notice that if we prove that f is differentiable for any $t > t_0$, then the Basset history term is differentiable and hence so is $w(t)$. The proof is from first principles, we have that

$$\begin{aligned} f'(t) &= \lim_{\epsilon\to 0} \frac{f(t+\epsilon)-f(t)}{\epsilon} \\ &= \lim_{\epsilon\to 0}\left[\frac{1}{\epsilon}\int_{t_0}^{t+\epsilon} a(s)\sqrt{t+\epsilon-s}\,ds - \frac{1}{\epsilon}\int_{t_0}^{t} a(s)\sqrt{t-s}\,ds\right] \\ &= \lim_{\epsilon\to 0}\left[\frac{1}{\epsilon}\int_{t}^{t+\epsilon} a(s)\sqrt{t+\epsilon-s}\,ds\right] \\ &\quad + \lim_{\epsilon\to 0}\int_{t_0}^{t} a(s)\left(\frac{\sqrt{t+\epsilon-s}-\sqrt{t-s}}{\epsilon}\right) ds\,. \end{aligned} \tag{4.49}$$

It remains to prove that the limit can be exchanged with the integral. For $\epsilon > 0$, note that

$$\left|\frac{1}{\epsilon}\int_{t}^{t+\epsilon} a(s)\sqrt{t+\epsilon-s}\,ds\right| \leqslant \left|\frac{1}{\epsilon}\int_{t}^{t+\epsilon} a(s)\sqrt{\epsilon}\,ds\right| \leqslant \big(\sup_{s\in[t,t+\epsilon]}|a(s)|\big)\sqrt{\epsilon}\,,$$

the boundedness of $a(s)$ gives

$$\lim_{\epsilon\to 0}\left[\frac{1}{\epsilon}\int_t^{t+\epsilon} a(s)\sqrt{t+\epsilon-s}\,ds\right] = 0\,,$$

In the last term in (4.49) we can switch between the integration and the limit with respect to ϵ by observing that, for any $\epsilon > 0$, we have

$$\left|a(s)\frac{\sqrt{t+\epsilon-s}-\sqrt{t-s}}{\epsilon}\right| = \left|a(s)\frac{t+\epsilon-s-(t-s)}{\epsilon(\sqrt{t+\epsilon-s}+\sqrt{t-s})}\right| \leqslant \frac{\sup_{s\in[t_0,t]}|a(s)|}{2\sqrt{t-s}}\,,$$

and the above upper bound is integrable on the interval $[t_0, t]$.

For $\epsilon < 0$ sufficiently small so that $t + \epsilon > t_0$ one shows in a similar manner that

$$\lim_{\epsilon\to 0}\left[\frac{1}{\epsilon}\int_{t+\epsilon}^{t} a(s)\sqrt{t-s}\,ds\right] = 0\,,$$

Also

$$\int_{t_0}^{t+\epsilon} a(s)\left[\frac{\sqrt{t+\epsilon-s}-\sqrt{t-s}}{\epsilon}\right] ds = \int_{t_0}^{t} a(s)q(s,\epsilon)ds\,,$$

where $q(s, \epsilon) = 0$ for $s \in [t + \epsilon, t]$ and

$$0 \leqslant q(s,\epsilon) = \frac{\sqrt{t+\epsilon-s}-\sqrt{t-s}}{\epsilon} = \frac{t+\epsilon-s-(t-s)}{\epsilon(\sqrt{t+\epsilon-s}+\sqrt{t-s})} \leqslant \frac{1}{\sqrt{t-s}}\,,$$

and we have, for any $s \in [t_0, t]$,

$$\left|a(s)\frac{\sqrt{t+\epsilon-s}-\sqrt{t-s}}{\epsilon}\right| \leqslant \frac{\sup_{s\in[t_0,t]}|a(s)|}{\sqrt{t-s}}\,,$$

and the above upper bound is integrable on the interval $[t_0, t]$.

Hence we have the required differentiability by the dominated convergence theorem, that is, we have explicitly that

$$f'(t) = \int_{t_0}^{t} \frac{a(s)}{2\sqrt{t-s}}\, ds\,.$$

$\mathscr{Q}.\mathscr{E}.\mathscr{D}.$

4.8 Properties of the solution as a function of the initial conditions

Suppose $(y(t), w(t))$ denotes a solution of (4.5) corresponding to an initial condition $(y_0, w_0) \in \mathbb{R}^{2n}$. We denote the derivatives of y and w with respect to (y_0, w_0) by Dy and Dw respectively. Note that these derivatives are matrix valued and may be considered as a map $[t_0, \infty) \to \mathbb{R}^{n\times 2n}$.

The solution of the Maxey-Riley equation can be interpreted as a function of its initial condition. Suppose $(y(t), w(t))$ denotes a solution of (4.5) corresponding to an initial condition $(y_0, w_0) \in \mathbb{R}^{2n}$, then we denote the derivatives of y and w with respect to (y_0, w_0) by Dy and Dw respectively. These derivatives satisfy the following equation [39]

$$\begin{aligned} Dy(t) &= (I_n|O_n) + \int_{t_0}^{t} \Big(Dw(s) + \nabla A_u(y(s), s) Dy(s) \Big)\, ds\,, \\ Dw(t) &= (O_n|I_n) + \int_{t_0}^{t} \bigg(-\mu Dw(s) - \mathcal{L}(y(s), w(s), s) Dy(s) \\ &\qquad - M_u(y(s), s) Dw(s) - \kappa\mu^{1/2} \frac{Dw(s)}{\sqrt{t-s}} \\ &\qquad\qquad + \nabla B_u(y(s), s) Dy(s) \bigg)\, ds\,, \end{aligned} \tag{4.50}$$

where $\mathcal{L}$ is an n-dimensional square matrix with components defined by

$$\mathcal{L}_{ij}(y(s), w(s), s) = \sum_k \frac{\partial M_{ik}}{\partial y_j}(y(s), s) w_k(s)\,,$$

and I_n and O_n denote the n-dimensional identity and null matrices respectively.

This interpretation will enable us to learn qualitative information about the solution. In particular, we will prove results about the proximity of ‘nearby’ trajectories, a non-collision result, and give a discussion on time reversibility. In this section, we will work under the following assumptions.

$$\text{The velocity field, } u\text{, is four times continuously differentiable and its partial derivatives are Lipschitz continuous up to order four.} \tag{$**$}$$

The Maxey-Riley equation is formulated to take a fluid velocity field and return a single particle trajectory corresponding to a particle’s given initial position and velocity. This raises the question of how two such particles would interact if their initial conditions are similar.

Proposition 4.18. *Under the conditions* $(**)$*, the distance between two trajectories at any time* t *is controlled by the difference between their initial conditions. Indeed, for two initial conditions* $x_1, x_2 \in \mathbb{R}^{2n}$*, there exists some constant* M *such that*

$$\|y(t, x_2) - y(t, x_1)\| \leqslant M\|x_2 - x_1\|\,,$$
$$\|w(t, x_2) - w(t, x_1)\| \leqslant M\|x_2 - x_1\|\,,$$

where the notation $y(t, x_i)$ *and* $w(t, x_i)$ *is used to reflect the dependence of the solution on its initial conditions.*

Remark 4.19. An interpretation of this proposition is that that two trajectories can be chosen to be arbitrarily close at time t by selecting suitably close initial conditions for them.

Proof. The equation (4.50) for the derivatives with respect to the initial conditions has solutions under assumptions $(**)$. With a similar methodology to that developed in Section 4.6, this result may be shown to hold globally in time. Under assumptions $(**)$, we have that ∇A_u, $\mathcal{L}$, M_u, and B_u are all bounded and we thus have sufficient conditions to apply a fractional Grönwall argument as in Appendix B. Let

$$f(t) = \sup_{s \in [t_0,t]} \sum_{i=1}^{n} |Dy^i(s)| + |Dw^i(s)| \,, \tag{4.51}$$

then there exist constants C_1, C_2 such that

$$f(t) \leqslant 2n + \int_{t_0}^{t} C_1 f(s)\, ds + \int_{t_0}^{t} C_2 (t-s)^{-1/2} f(s)\, ds \,. \tag{4.52}$$

We may once again apply the Grönwall inequality from Appendix B. Hence Dy and Dw are bounded above on intervals $[t_0, T)$ for all $T > t_0$. Suppose M is such that Dy and Dw are bounded above by M. Then, for two initial conditions $x_1, x_2 \in \mathbb{R}^{2n}$ we have

$$\begin{aligned}
\|y(t,x_2) - y(t,x_1)\| &= \int_{x_1}^{x_2} Dy(t,z)\, dz \leqslant M \|x_2 - x_1\| \,, \\
\|w(t,x_2) - w(t,x_1)\| &= \int_{x_1}^{x_2} Dy(t,z)\, dz \leqslant M \|x_2 - x_1\| \,.
\end{aligned}$$

$\mathscr{Q}.\mathscr{E}.\mathscr{D}.$

Whilst the previous result implies that two particles can be chosen to be arbitrarily close to one another, the next may be interpreted as a *non-collision* property that two particles cannot collide. To prove this result, we will first need to comment on the evolution of an inverse matrix.

Remark 4.20 (Non-fractional evolution of an inverse matrix)*.* If an $n \times n$ matrix, M_t, evolves according to

$$M_t = I_n + \int_{t_0}^{t} A_s M_s \, ds \,,$$

and another matrix, N_t, according to

$$N_t = I_n - \int_{t_0}^{t} N_s A_s \, ds \,.$$

Then we have

$$\begin{aligned} N_t M_t &= I_n + \int_{t_0}^{t} N_s \frac{dM_s}{ds} \, ds + \int_{t_0}^{t} \frac{dN_s}{ds} M_s \, ds \\ &= I_n + \int_{t_0}^{t} N_s A_s M_s \, ds - \int_{t_0}^{t} N_s A_s M_s \, ds = I_n \,. \end{aligned}$$

Thus, $\det(N_t M_t) = \det(N_t)\det(M_t) = 1$ and therefore $\det(M_t) \neq 0$ and M_t is invertible. Moreover, N_t is the inverse of M_t for all t.

In the fractional case, this is less straightforward and requires the usage of *left* and *right* fractional derivatives [85].

Lemma 4.21 (Fractional evolution of an inverse matrix)**.** *Suppose that M_t evolves according to the fractional equation*

$$M_t = I_n + \frac{1}{\sqrt{\pi}} \int_{t_0}^{t} \frac{A_s M_s}{\sqrt{t-s}} \, ds \,. \tag{4.53}$$

Then its inverse satisfies

$$M_s^{-1} = I_n + \frac{1}{\sqrt{\pi}} \int_s^t \frac{M_r^{-1} A_r}{\sqrt{r-s}}\, dr\,. \tag{4.54}$$

Proof. Recall that, from Definition 4.4, the evolution of the matrix M_t is given in differential form by

$$\frac{dM_t}{dt} =_{t_0} D^{1/2}(A_t M_t)\,. \tag{4.55}$$

Recall that the *right* Riemann-Liouville fractional derivative may be defined similarly to the standard, or *left* Riemann-Liouville derivative in Definition 4.4, by instead changing the equation (4.8) to

$$D_b^p f(t) = \frac{1}{\Gamma(n-p)} \left(-\frac{d}{dt}\right)^n \int_t^b (s-t)^{n-p-1} f(s)\, ds\,. \tag{4.56}$$

We then note that, for continuous functions f and g, we have the following result

$$\int_a^b f(t)_a D^\alpha g(t)\, ds = \int_a^b g(s) D_b^\alpha f(s)\, ds\,. \tag{4.57}$$

Suppose now that the matrix N_s evolves by

$$N_s = I_n + \frac{1}{\sqrt{\pi}} \int_s^t \frac{N_r A_r}{\sqrt{r-s}}\, dr\,, \tag{4.58}$$

where this evolution depends on time. In differential form this is

$$\frac{dN_s}{ds} = -D_t^{1/2}(N_s A_s)\,. \tag{4.59}$$

Now suppose that $A_t = A$ is constant, then we have

$$\begin{aligned} N_t M_t &= I_n + \int_{t_0}^{t} N_s \frac{dM_s}{ds}\, ds + \int_{t_0}^{t} \frac{dN_s}{ds} M_s \, ds \,, \\ &= I_n + \int_{t_0}^{t} N_{st_0} D^{1/2}(AM_s)\, ds - \int_{t_0}^{t} D_t^{1/2}(N_s A) M_s \, ds \,. \end{aligned} \tag{4.60}$$

Now by the equation (4.57), we have

$$N_t M_t = I_n \,, \tag{4.61}$$

and thus we have determined the equation for the inverse when M_t is a matrix evolving according to a fractional differential equation. $\mathscr{Q}.\mathscr{E}.\mathscr{D}.$

Proposition 4.22 (Non-collision of inertial particles)**.** *Under conditions* $(**)$*, the distance between two trajectories is always strictly positive if their initial conditions are distinct.*

Proof. Define the matrix $D\varphi$ by

$$D\varphi(t) := \begin{pmatrix} Dy(t) \\ Dw(t) \end{pmatrix},$$

then this matrix evolves according to the equation

$$\begin{aligned} D\varphi(t) = I_{2n} - \int_{t_0}^{t} \begin{pmatrix} O_n & O_n \\ O_n & -\kappa\mu^{1/2} I_n \end{pmatrix} D\varphi(s) \frac{ds}{\sqrt{t-s}} \\ + \int_{t_0}^{t} \begin{pmatrix} \nabla A_u(y(s),s) I_n & I_n \\ \big(\nabla B_u(y(s),s) - \mathcal{L}(y(s),w(s),s)\big) I_n & -\big(\mu + M_u(y(s),s)\big) I_n \end{pmatrix} D\varphi(s)\, ds \,. \end{aligned}$$

Immediately following Lemma 4.21, the inverse of $D\varphi(t)$ evolves according to

$$D\varphi^{-1}(s) = I_{2n} - \int_s^t D\varphi^{-1}(r) \begin{pmatrix} O_n & O_n \\ O_n & -\kappa\mu^{1/2} I_n \end{pmatrix} \frac{dr}{\sqrt{r-s}}$$

$$- \int_{t_0}^s D\varphi^{-1}(r) \begin{pmatrix} \nabla A_u(y(r),r) I_n & I_n \\ \big(\nabla B_u(y(r),r) - \mathcal{L}(y(r),w(r),r)\big) I_n & -\big(\mu + M_u(y(r),r)\big) I_n \end{pmatrix} dr\,.$$

Since left and right Riemann-Liouville derivatives are equivalent up to time reversal, the same Grönwall argument from Appendix B may be applied. This is valid since u is sufficiently smooth and bounded under the assumptions $(**)$ for the coefficients to satisfy the required conditions for the Grönwall theorem to hold. Hence there exists some $\widetilde{M}$ which is an upper bound for Dy^{-1} and Dw^{-1}, and hence

$$0 < \frac{1}{\widetilde{M}} \leqslant Dy, Dw\,. \tag{4.62}$$

$\mathscr{Q}.\mathscr{E}.\mathscr{D}.$

4.9 Concluding remarks

Revisiting the promises made in Section 4.1, we have presented a thorough analysis of the Maxey-Riley equation. This has built on the work of Farazmand and Haller [39], and made use of the Grönwall theorem in Appendix B. Most significantly, in this chapter we have shown that the solution of the equation exists globally in time and we have rigorously proven the conditions under which a strong solution exists. Therefore, despite the fact that the Maxey-Riley equation is a fractional order differential equation, and is nonlinear, this does not inhibit the solutions from being global in time or differentiable.

Whilst we have answered many of the key questions surrounding the analysis of this model, there are still questions that remain. It remains to study how the concept of time reversibility applies to the system, since the solution depends on its history. This is key for determining the origin of an object with mass, given its current location. It may also be possible to relax the assumptions on the underlying velocity field.

5

Conclusion

Within this thesis, we have introduced and developed a broad selection of mathematical tools for the purposes of understanding problems in fluid dynamics. Along the way, we have seen and used concepts from stochastic analysis, fractional calculus, exterior calculus, differential geometry, Lie group theory, and more. New methodologies have been introduced, and existing concepts expanded on. We have revealed new properties of well established methods, as well as proposed new approaches to overcome modelling challenges. In Chapter 2, we summarised some ideas from geometric mechanics and how this may apply to fluids, as well as gave an in depth discussion of how to formulate fluid problems from a stochastic variational principle. We made use of this in Chapter 3, where we discussed both new and existing approaches to free surface dynamics from a variational perspective.

This involved a novel stochastic version of classical water wave theory, as well as a new deterministic and stochastic approach to wave current interaction using a composition of two maps. This approach allows the free surface problem to be closed with a larger set of physical variables compared to the standard theories, which can aid in understanding complex wave current interactions which have thus far remained elusive. Finally, in Chapter 4, we considered the behaviour of an object which is carried by a flow of currents. This involved a rigorous analysis of the Maxey-Riley equation. Since the complications involved in this analysis surrounded the fractional order term, the results found have implications on the solution properties of other such inertial particles with memory.

Future research directions

The work presented in this thesis has raised a number of further questions. These include topics which correspond to each chapter, questions of a mathematical nature, and new challenges in physical modelling. Some particularly interesting topics are the following.

- As found in Chapter 2, the definition of a semimartingale driven variational principle provides a framework for correctly interpreting and formulating a given action integral, when the integral is a stochastic integral. However it is unclear how many of the deterministic results from geometric mechanics generalised to an arbitrary stochastic integral. For example, it is not immediately obvious what assumptions are necessary on the driving semimartingale for an Euler-Poincaré reduction to be possible. The connection to an arbitrary stochastic Hamiltonian formulation is also unexplored. Stochastic advection by Lie transport gives a particular case of a semimartingale driven variational principle where these notions may be shown to extend from the deterministic case, as we illustrated with the water wave example in Chapter 3, but it is unclear what are the minimal assumptions necessary for this to be true in general.

- The new stochastic extension of classical water wave theory, found in Section 3.3, is yet to be fully explored. It is not known whether these equations are well posed, in the same way as the deterministic case is known to be. If it is well posed, does the noise improve or reduce the regularity of the solutions?

- The deterministic classical water wave equations have been numerically integrated, and asymptotically expanded. This remains to be done with the stochastic case, where there is an opportunity to use the noise coefficients to explore vorticity effects and model uncertainty.

- Introduced in Chapter 3 (and following on from [33] and [57]), the composition of maps approach to wave current interaction merits further study. In particular, the identification of a more physical two dimensional equation to represent the flow of currents observed from satellite imagery is an important problem. This is a particular issue at thermal fronts, where upwelling and downwelling effects can cause convergence or divergence of the observed flow.

- The composition of maps approach is also attractive mathematically. It has an interesting geometric structure, which may be applied to a range of other systems. A detailed study of its structure would aid in evaluating its potential for further application.

- The use of stochastic advection by Lie transport in the composition of maps approach supports the matching of a model with satellite observations. This presents a significant data assimilation challenge.

- The well posedness of the Maxey-Riley equation was, in Chapter 4, proven for some assumptions on the velocity field of the ambient fluid. These assumptions are not claimed to be minimal, and the relaxing of these assumptions would be desirable. It is not known whether these equations are well posed in the case where the ambient fluid is given as a solution to a stochastic fluid equation. It also not known whether the solution of the Maxey-Riley equation gains (or loses) regularity when the ambient fluid is assumed to be smoother (or rougher).

- The Maxey-Riley equation is only one example of an equation for modelling an inertial particle with memory effects. It remains to prove that the result extends to a more general class of equations which possess a similar structure.

Bibliography

[1] Great Britain. *An Act for Providing a Publick Reward for such Person or Persons as shall Discover the Longitude at Sea: Anne. Chapter 13.* (1714) London. Available at https://archives.parliament.uk/collections/getrecord/GB61_HL_PO_PU_1_1713_13An35

[2] Arnaudon, M., Chen, X., and Cruzeiro, A. B. (2014) Stochastic Euler-Poincaré reduction. *Journal of Mathematical Physics.* **55**, 081507. https://doi.org/10.1063/1.4893357

[3] Andrews, D. G., and McIntyre, M. E. (1978). An exact theory of nonlinear waves on a Lagrangian-mean flow. *Journal of Fluid Mechanics.* **89** (4), 609–646. https://doi.org/10.1017/S0022112078002773

[4] Arnold, V. I. (1966) Sur la géométrie differentielle des groupes de Lie de dimenson infinie et ses applications à l'hydrodynamique des fluids parfaits. *Annales de l'institut Fourier.* **16**, 319–361.

[5] Auton, T., Hunt, J., and Prud'Homme, M. (1988). The force exerted on a body in inviscid unsteady non-uniform rotational flow. *Journal of Fluid Mechanics.* **197**, 241–257. https://doi.org/10.1017/S0022112088003246

[6] Basset, A. B. (1888) *A treatise on hydrodynamics: With numerous examples.* Cambridge, Deighton, Bell and Co.

[7] Beron-Vera, F. J. et al. (2015) Dissipative inertial transport patterns near coherent Lagrangian eddies in the ocean. *Chaos*. **25** (8), 087412. https://doi.org/10.1063/1.4928693

[8] Beron-Vera, F. J. and Miron, P. (2020) A minimal Maxey–Riley model for the drift of Sargassum rafts. *Journal of Fluid Mechanics*. **904**, A8. https://doi.org/10.1017/jfm.2020.666

[9] Beron-Vera, F. J., Olascoaga, M. J., and Miron, P. (2019) Building a Maxey-Riley framework for surface ocean inertial particle dynamics. *Physics of Fluids*. **31** (9), 096602. https://doi.org/10.1063/1.5110731

[10] Bessel, F. W. (1828) *Untersuchungen über die Länge des einfachen Secundendpendels*. Berlin, Königlichen Akademie der Wissenschaften. Available at https://catalog.hathitrust.org/Record/008629879

[11] Bethencourt de Léon, A., Holm, D. D., Luesink, E., and Takao, S. (2020) Implications of Kunita–Itô–Wentzell Formula for k-Forms in Stochastic Fluid Dynamics. *Journal of Nonlinear Science*. **30** (4), 1421–1454. https://doi.org/10.1007/s00332-020-09613-0

[12] Billingsley, P. (1995) *Probability and Measure*. Wiley Series in Probability and Statistics. 3rd edition. New York, Wiley.

[13] Bismut, J. -M. (1981) *Mécanique aléatoire*. Lecture Notes in Mathematics 866. Berlin, Springer. https://link.springer.com/book/10.1007/BFb0088591.

[14] Bloch, A., Krishnaprasad, P. S., Marsden, J. E., and Ratiu, T. S. (1996) The Euler-Poincaré Equations and Double Bracket Dissipation. *Commu-*

nications in Mathematical Physics. **175** (1), 1–41. https://doi.org/10.1007/BF02101622

[15] Bou-Rabee, N., and Owhadi, H. (2009) Stochastic variational integrators. *IMA Journal of Numerical Analysis.* **29** (2), 421–443. https://doi.org/10.1093/imanum/drn018

[16] Boussinesq, J. (1885) Sur la résistance qu'oppose un fluide indéfini au repos, sans pesanteur, aumouvement varié d'une sphère solide qu'il mouille sur toute sa surface, quand les vitesses restent bien continues et assez faibles pour que leurs carrés et produits soient néglige. *Comptes Rendusde l'Académie des Sciences.* **100**, 935–937.

[17] Bretherton, F. P. and Garrett, C. J. R. (1968) Wavetrains in inhomogeneous moving media. *Proceedings of the Royal Society A.* **302**, 529–554. http://doi.org/10.1098/rspa.1968.0034

[18] Caputo, M. (1967) Linear Models of Dissipation whose Q is almost Frequency Independent – II. *Geophysical Journal International.* **13** (5), 529–539. https://doi.org/10.1111/j.1365-246X.1967.tb02303.x

[19] Castro, A., and Lannes, D. (2015). Well-Posedness and Shallow-Water Stability for a New Hamiltonian Formulation of the Water Waves Equations with Vorticity. *Indiana University Mathematics Journal.* **64** (4), 1169–1270.

[20] Cendra, H., and Marsden, J. E. (1987) Lin constraints, Clebsch potentials and variational principles. *Physica D: Nonlinear Phenomena.* **27**, 63–89. https://doi.org/10.1016/0167-2789(87)90005-4

[21] Chapron, B., Kudryavtsev, V. N., Collard, F., Rascle, N., Kubryakov, A. A. and Stanichny, S. V. (2020) Studies of Sub-Mesoscale Variability of

the Ocean Upper Layer Based on Satellite Observations Data. *Physical Oceanography*. **27** (6), PAGESS619–630. http://doi.org/10.22449/1573-160X-2020-6-619-630

[22] Chen, X., Cruzeiro, A. B., and Ratiu, T. S. (2015) Stochastic variational principles for dissipative equations with advected quantities. *arXiv*. [Preprint] https://arxiv.org/abs/1506.05024. [Accessed 27 July 2022]

[23] Clebsch, R. F. A. (1856) Über die Bewegung eines Ellipsoids in einer tropfbaren Flüssigkeit. *Journal für die reine und angewandte Mathematik (Crelle)*. **52**, 103–132. Available at https://eudml.org/doc/147642

[24] Cotter, C. J., and Bokhove, O. (2010) Variational water-wave model with accurate dispersion and vertical vorticity. JournalofEngineeringMathematics. **67**, 33–54. https://doi.org/10.1007/s10665-009-9346-3

[25] Cotter, C. J., Crisan, D., Holm, D. D., Pan, W., and Shevchenko, I. (2019) Numerically Modeling Stochastic Lie Transport in Fluid Dynamics. *Multiscale Modeling & Simulation*. **17** (1), 192–232. https://doi.org/10.1137/18M1167929

[26] Cotter, C. J., Crisan, D., Holm, D. D., Pan, W., and Shevchenko, I. (2020) Modelling uncertainty using stochastic transport noise in a 2-layer quasi-geostrophic model. *Foundations of Data Science*. **2** (2), 173–205. https://doi.org/10.3934/fods.2020010

[27] Cotter, C. J., Crisan, D., Holm, D. D., Pan, W., and Shevchenko, I. (2020) A Particle Filter for Stochastic Advection by Lie Trans-

port: A Case Study for the Damped and Forced Incompressible Two-Dimensional Euler Equation. *SIAM/ASA Journal on Uncertainty Quantification.* **8** (4), 1446–1492. https://doi.org/10.1137/19M1277606

[28] Cotter, C. J., and Holm, D. D. (2009) Continuous and Discrete Clebsch Variational Principles. *Foundations of Computational Mathematics.* **9**, 221–242. https://doi.org/10.1007/s10208-007-9022-9

[29] Craig, W. and Sulem, C. (1993) Numerical Simulation of Gravity Waves. *Journal of Computational Physics.* **108** (1), 73–83. https://doi.org/10.1006/jcph.1993.1164

[30] Craik, A. D. D., and Leibovich, S. (1976) A rational model for Langmuir circulations. *Journal of Fluid Mechanics.* **73** (3), 401–426. https://doi.org/10.1017/S0022112076001420

[31] Crisan, D., Flandoli, F., and Holm, D. D. (2019) Solution Properties of a 3D Stochastic Euler Fluid Equation. *Journal of Nonlinear Science.* **29** (3), 813–870. https://doi.org/10.1007/s00332-018-9506-6

[32] Crisan, D., Darryl, D. D., Leahy, J. -M., and Nilssen, T. (2020) Variational principles for fluid dynamics on rough paths. *arXiv.* [Preprint] https://arxiv.org/abs/2004.07829. [Accessed 13th April 2022]

[33] Crisan, D., Holm, D. D., and Street, O. D. (2021) Wave-current interaction on a free surface. *Studies in Applied Mathematics.* **147** (4), 1277–1338. https://doi.org/10.1111/sapm.12425

[34] Crisan, D. and Street, O. D. (2021) On the analytical aspects of inertial particle motion. *arXiv.* [Preprint] https://arxiv.org/abs/2111.08470 [Accessed 17th February 2022].

[35] Daitche, A. (2015) On the role of the history force for inertial particles in turbulence. *Journal of Fluid Mechanics*. **782**, 567–593. https://doi.org/10.1017/jfm.2015.551

[36] Daitche, A. and Tél, T. (2014) Memory effects in chaotic advection of inertial particles. *New Journal of Physics*. **16**, 073008. https://doi.org/10.1088/1367-2630/16/7/073008

[37] Dinvay, E. and Memin, E. (2022) Hamiltonian formulation of the stochastic surface wave problem. *arXiv*. [Preprint] https://arxiv.org/abs/2201.07764. [Accessed 25th May 2022]

[38] Ebin, D. G. and Marsden, J. E. (1970) Groups of diffeomorphisms and the motion of an incompressible fluid. *Annals of Mathematics*. **92** (1), 102–163. https://doi.org/10.2307/1970699

[39] Farazmand, M., and Haller, G. (2015) The Maxey–Riley Equation: Existence, Uniqueness and Regularity of Solutions. *Nonlinear Analysis: Real World Applications*. **22**, 98–106. https://doi.org/10.1016/j.nonrwa.2014.08.002

[40] Faxén, H. (1922) Der Widerstand gegen die Bewegung einer starren Kugel in einer zähen Flüssigkeit, die zwischen zwei parallelen ebenen Wänden eingeschlossen ist. *Annalen der Physik*. **373** (10), 89–119. https://doi.org/10.1002/andp.19223731003

[41] Gilbert, A., and Vanneste, J. (2018) Geometric generalised Lagrangian-mean theories. *Journal of Fluid Mechanics*. **839**, 95–134. https://doi.org/10.1017/jfm.2017.913

[42] Gjaja, I., and Holm, D. D. (1996) Self-consistent Hamiltonian dynamics of wave mean-flow interaction for a rotating stratified incompressible

fluid. *Physica D: Nonlinear Phenomena.* **98**, 343–378. https://doi.org/10.1016/0167-2789(96)00104-2

[43] Green, A. E., and Naghdi, P. M. (1976). A derivation of equations for wave propagation in water of variable depth. *Journal of Fluid Mechanics.* **78** (2), 237–246. https://doi.org/10.1017/S0022112076002425

[44] Green, G. (1836) Researches on the Vibrations of Pendulums in Fluid Media. *Transactions of the Royal Society of Edinburgh* **13**, 54–62. Available at https://www.biodiversitylibrary.org/bibliography/2290

[45] Guseva, K., Daitche, A., Feudel, U., and Tél, T. (2016) History effects in the sedimentation of light aerosols in turbulence: the case of marine snow. *Physical Review Fluids.* **1**, 074203. https://doi.org/10.1103/PhysRevFluids.1.074203

[46] Haller, G. and Sapsis, T. (2008) Where do inertial particles go in fluid flows? *Physica D: Nonlinear Phenomena.* **237** (5), 573–583. https://doi.org/10.1016/j.physd.2007.09.027

[47] Hamilton, E. P. and Nashed, M. Z. (1982) Global and Local Variational Derivatives and Integral Representations of Gâteaux Differentials. *Journal of Functional Analysis.* **49** (1). 128–144. http://dx.doi.org/10.1016/0022-1236(82)90088-X

[48] Holm, D. D. (1996) The ideal Craik-Leibovich equations. *Physica D: Nonlinear Phenomena.* **98**, 415–441. https://doi.org/10.1016/0167-2789(96)00105-4

[49] Holm, D. D. (2002). Euler-Poincaré Dynamics of Perfect Complex Fluids. In: Newton, P., Holmes, P., Weinstein, A. (eds) *Geometry, Me-*

chanics, and Dynamics. Springer, New York, NY. https://doi.org/10.1007/0-387-21791-6_4

[50] Holm, D. D. (2002) Lagrangian averages, averaged Lagrangians, and the mean effects of fluctuations in fluid dynamics. *Chaos*. **12**, 518–530. https://doi.org/10.1063/1.1460941

[51] Holm, D. D. (2011) *Geometric Mechanics - Part I: Dynamics and Symmetry*, 2nd edition. Imperial College Press.

[52] Holm, D. D. (2011) *Geometric Mechanics - Part II: Rotating, Translating and Rolling*, 2nd edition. Imperial College Press.

[53] Holm, D. D. (2015) Variational principles for stochastic fluid dynamics. *Proceedings of the Royal Society A*. **471**, 20140963. https://doi.org/10.1098/rspa.2014.0963

[54] Holm, D. D. (2019) Stochastic Parametrization of the Richardson Triple. *Journal of Nonlinear Science*. **29** (1). 89–113. http://dx.doi.org/10.1007/s00332-018-9478-6

[55] Holm, D. D. (2019) Stochastic Closures for Wave–Current Interaction Dynamics. *Journal of Nonlinear Science*. **29** (6). 2987–3031. https://doi.org/10.1007/s00332-019-09565-0

[56] Holm, D.D. (2021) Stochastic Variational Formulations of Fluid Wave–Current Interaction. *Journal of Nonlinear Science*. **31**, 4. https://doi.org/10.1007/s00332-020-09665-2

[57] Holm, D. D., Hu, R., and Street, O. D. (2022) Ponderomotive coupling of waves to sea surface currents via horizontal density gradients.

arXiv. [Preprint] https://arxiv.org/abs/2202.04446. [Accessed 4th May 2022]

[58] Holm, D. D. and Kupershmidt, B. A. (1982) Poisson structures of superfluids. *Physics Letters A*. **91** (9), 425–430. https://doi.org/10.1016/0375-9601(82)90740-X

[59] Holm, D. D. and Kupershmidt, B. A. (1983) Poisson brackets and Clebsch representations for magnetohydrodynamics, multifluid plasmas, and elasticity. *Physica D: Nonlinear Phenomena*. **6** (3), 347–363. https://doi.org/10.1016/0167-2789(83)90017-9

[60] Holm, D. D. and Luesink, E. (2021) Stochastic Geometric Mechanics with Diffeomorphisms. In: Ugolini, S. et al. (eds.) *Geometry and Invariance in Stochastic Dynamics*. Springer Proceedings in Mathematics & Statistics, 378. Switzerland, Springer, pp.169–185. https://doi.org/10.1007/978-3-030-87432-2

[61] Holm, D. D. and Luesink, E. (2021) Stochastic Wave–Current Interaction in Thermal Shallow Water Dynamics. *Journal of Nonlinear Science*. **31**, 29. https://doi.org/10.1007/s00332-021-09682-9

[62] Holm, D. D., Marsden, J. E., and Ratiu, T. S. (1998) The Euler-Poincaré Equations and Semidirect Products with Applications to Continuum Theories. *Advances in Mathematics*. **137** (1), 1–81. https://doi.org/10.1006/aima.1998.1721

[63] Holm, D. D., Schmah, T., and Stoica, C. (2009) *Geometric Mechanics and Symmetry: From Finite to Infinite Dimensions*. Oxford Texts in Applied and Engineering Mathematics. Oxford University Press.

[64] Hsu, E. P. (2002) *Stochastic Analysis on Manifolds*. Graduate Studies in Mathematics, 38. Providence, American Mathematical Society.

[65] Kallenberg, O. (2017) *Random Measures, Theory and Applications*. Probability Theory and Stochastic Modelling, 77. Switzerland, Springer. https://doi.org/10.1007/978-3-319-41598-7.

[66] Karatzas, I. and Shreve , S. E. (1998) *Brownian Motion and Stochastic Calculus*. Graduate Texts in Mathematics, 113. 2nd edition. New York, Springer. https://doi.org/10.1007/978-1-4612-0949-2

[67] Lamb, H. (1911) XV. On the uniform motion of a sphere through a viscous fluid. *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science*. **21**. 112–121. https://doi.org/10.1080/14786440108637012

[68] Lázaro-Camí, J. -A. and Ortega, J. -P. (2008) Stochastic hamiltonian dynamical systems. *Reports on Mathematical Physics*, **61** (1), 65–122. https://doi.org/10.1016/S0034-4877(08)80003-1

[69] Lee, J. M. (2013) *Introduction to Smooth Manifolds*. Graduate Texts in Mathematics, 218. 2nd edition. New York, Springer. https://doi.org/10.1007/978-1-4419-9982-5

[70] Lilly, J. M. (2021) jLab: A data analysis package for Matlab, v.1.7.1 https://doi.org/10.5281/zenodo.4547006 http://www.jmlilly.net/software

[71] Lin, S. Y. (2013) Generalized Gronwall inequalities and their applications to fractional differential equations. *Journal of Inequalities and Applications*. 549. https://doi.org/10.1186/1029-242X-2013-549

[72] Luesink, E. (2021) *Stochastic geometric mechanics of thermal ocean dynamics*. PhD thesis, Imperial College London.

[73] Marsden, J. E. and Ratiu, T. S. (1999) *Introduction to Mechanics and Symmetry: A Basic Exposition of Classical Mechanical Systems*. Texts in Applied Mathematics, 17. 2nd edition. New York, Springer.

[74] Marsden, J. E. and Scheurle, J. (1993) The Reduced Euler-Lagrange Equations. *Fields Institute Communications*. **1**. 139–164.

[75] Maxey, M. R. (1987) The motion of small spherical particles in a cellular flow field. *The Physics of Fluids*. **30** (7), 1915–1928. https://doi.org/10.1063/1.866206

[76] Maxey, M. R. (1993) The Equation of Motion for a Small Rigid Sphere in a Nonuniform or Unsteady flow. In: Stock, D. E. et al. (eds.) *Gas-solid flows – 1993*. American Society of Mechanical Engineers, United States. pp. 57–62.

[77] Maxey, M. R. and Riley, J. J. (1983) Equation of motion for a small rigid sphere in a nonuniform flow. *The Physics of Fluids*. **26** (4), 883–889. https://doi.org/10.1063/1.864230

[78] Michaelides, E. E. (1997) Review — The Transient Equation of Motion for Particles, Bubbles, and Droplets. *Journal of Fluids Engineering*. **119** (2), 233–247. https://doi.org/10.1115/1.2819127

[79] Munk, W., Armi, L., Fischer, K., and Zachariasen, F. (2000) Spirals on the sea. *Proceedings of the Royal Society A*.**456** (1997), 1217–1280. http://doi.org/10.1098/rspa.2000.0560

[80] Noether, E. (1918) Invariante Variationsprobleme. *Nachrichten von der Gesellschaft der Wissenschaften zu Göttingen, Mathematisch-Physikalische Klasse 1918*. 235–257.

[81] Olascoaga, M. J. et al. (2020) Observation and quantification of inertial effects on the drift of floating objects at the ocean surface. *Physics of Fluids*. **32** (2), 026601. https://doi.org/10.1063/1.5139045

[82] Oseen, C. W. (1910) Über die Stokes'sche Formel und über eine verwandte Aufgabe in der Hydrodynamik. *Arkiv för matematik, astronomi och fysik*. **6** (29).

[83] Oseen, C. W. (1913) Über den Gültigkeitsbereich Stokesschen Widerstandsformel. *Arkiv för matematik, astronomi och fysik*. **9** (19).

[84] Oseen, C. W. (1927) *Hydrodynamik*. Leipzig, Akademische Verlagsgesellschaft.

[85] Podlubny, I. (1999) *Fractional Differential Equations*. Mathematics in Science and Engineering, **198**. San Diego, Academic Press.

[86] Poincaré, H. (1901) Sur une forme nouvelle des équations de la méchanique. *Comptes rendus de l'Académie des Sciences*. **132**, 369–371.

[87] Poisson, S. D. (1832) Mémoire sur les mouvements simultanés d'un pendule et de l'air environnant. *Mémoires de l'Académie royale des sciences de l'Institut de France.*. **11**, 521–581. Available at https://www.biodiversitylibrary.org/bibliography/4002

[88] Prasath, S. G., Vasan, V., and Govindarajan, R. (2019) Accurate solution method for the Maxey-Riley equation, and the effects of Basset

history. *Journal of Fluid Mechanics*. **868**, 428–460. https://doi.org/10.1017/jfm.2019.194

[89] Rascle, N., Molemaker, J., Marié, L., Nouguier, F., Chapron, B., Lund, B. and Mouche, A. (2017) Intense deformation field at oceanic front inferred from directional sea surface roughness observations. *Geophysical Research Letters*. **44** (11), 5599–5608. https://doi.org/10.1002/2017GL073473

[90] Reeks, M. W. and McKee, S. (1984) The dispersive effects of Basset history forces on particle motion in a turbulent flow. *The Physics of Fluids*. **27** (7), 1573–1582. https://doi.org/10.1063/1.864812

[91] Sapsis, T. P., Ouellette, N. T., Gollub, J. P., and Haller, G. (2011) Neutrally buoyant particle dynamics in fluid flows: Comparison of experiments with Lagrangian stochastic models. *Physics of Fluids*. **23** (9), 093304. https://doi.org/10.1063/1.3632100

[92] Scott Russell, J. (1845) *Report on Waves*. Report of the fourteenth meeting of the British Association for the Advancement of Science; held at York in September 1844. 311–390. London, John Murray.

[93] Scully-Power, P. (1986) *Navy Oceanographer Shuttle observations, STS 41-G Mission Report*. Naval Underwater Systems Center, NUSC Technical Document 7611.

[94] Stokes, G. G. (1851) On the Effect of the Internal Friction of Fluids on the Motion of Pendulums. *Transactions of the Cambridge Philosophical Society*. **9** (2), 8–106. Available at https://www.biodiversitylibrary.org/bibliography/2348

[95] Street, O. D. and Crisan, D. (2021) Semi-martingale driven variational principles. *Proceedings of the Royal Society A*. **477**, 20200957. https://doi.org/10.1098/rspa.2020.0957

[96] Takao, S. (2020) *Stochastic Geometric Mechanics for Fluid Modelling and MCMC*. PhD thesis, Imperial College London.

[97] Tatom, F. B. (1988) The Basset term as a semiderivative. *Applied Scientific Research*. **45** (3), 283–285. https://doi.org/10.1007/bf00384691

[98] Tchen, C. (1947) *Mean value and correlation problems connected with the motion of small particles suspended in a turbulent fluid*. PhD thesis, Technische Universiteit Delft.

[99] van Sebille, E., England, M. H., and Froyland, G. (2012) Origin, dynamics and evolution of ocean garbage patches from observed surface drifters. *Environmental Research Letters*. **7** (4), 044040. https://doi.org/10.1088/1748-9326/7/4/044040

[100] van Sebille, E. et al. (2015) A global inventory of small floating plastic debris. *Environmental Research Letters*. **10** (12), 124006. https://doi.org/10.1088/1748-9326/10/12/124006

[101] van Sebille, E. et al. (2018) Lagrangian ocean analysis: Fundamentals and practices. *Ocean Modelling*. **121**, 49–75. https://doi.org/10.1016/j.ocemod.2017.11.008

[102] van Sebille, E. et al. (2020) The physical oceanography of the transport of floating marine debris. *Environmental Research Letters*. **15** (2), 023003. https://doi.org/10.1088/1748-9326/ab6d7d

[103] Wang, L., Hong, J., Scherer, R., and Bai, F. (2009) Dynamics and variational integrators of stochastic hamiltonian systems. *International Journal of Numerical Analysis and Modeling*. **6** (4), 586–602.

[104] Whitehead, A. N. (1889) Second approximations to viscous fluid motion. A sphere moving steadily in a straight line. *The Quarterly Journal of Pure and Applied Mathematics*. **23**, 143–152. Available at https://catalog.hathitrust.org/Record/006024259

[105] Yurovskaya, M., Rascle, N., Kudryavtsev, V., Chapron, B., Marié, L. and Molemaker, J.. (2018) Wave spectrum retrieval from airborne sunglitter images. *Remote Sensing of Environment*. **217**, 61–71. https://doi.org/10.1016/j.rse.2018.07.026

[106] Zakharov, V. E. (1968) Stability of periodic waves of finite amplitude on the surface of a deep fluid. *Journal of Applied Mechanics and Technical Physics*. **9** (2),190–194. https://doi.org/10.1007/BF00913182

A

Further discussions on Geometric Mechanics

In this appendix, we will be expanding on some of the concepts introduced in the review of geometric mechanics in Section 2.1. This will involve proving some of the elementary results and elaborating on the Euler-Poincaré theorem. We will also include some examples to illustrate the application of a reduced variational principle, as well as the requirement for the semidirect product Euler-Poincaré theorem (Theorem 2.18). We begin by proving Theorems 2.8 and 2.13.

When considering how to apply Hamilton's principle, $\delta S = 0$, to an action given by

$$S := \int_{t_0}^{t_1} L(q(t), \dot{q}(t), t)\, dt\,, \tag{A.1}$$

a complication emerges when considering how to take variations of $q(t)$ and $\dot{q}(t)$ independently. This may be overcome by the Hamilton-Pontryagin approach, where variations are taken of the Lagrangian depending on $(p, v) \in TQ$ and v is identified with the tangent lift vector via a Lagrange multiplier, p, which we will later define to be the canonically conjugate momentum.

Theorem 2.8. *For any differentiable $L(q, \dot{q})$, we have that Hamilton's principle implies the Euler-Lagrange equations*

$$\frac{d}{dt}\frac{\partial L}{\partial \dot{q}_j} = \frac{\partial L}{\partial q_j}\,, \quad \textit{for} \quad j = 1, \ldots, n\,, \tag{2.3 revisited}$$

where $q \in Q$ is assumed to be n dimensional.

Proof. In this proof, we will drop the notation which makes explicit the derivative with respect to each component of q, thus the proof will read as if it is one dimensional. We will prove this using the Hamilton–Pontryagin approach. In the following, we write the Lagrangian in terms of the coordinates $(q, v) \in TQ$, where we use a Lagrange multiplier to constrain that $v = \dot{q}$. We take variations as follows

$$\begin{aligned}
\delta S &= \delta \int_{t_0}^{t_1} L(q, v) + \left\langle p, \frac{dq}{dt} - v \right\rangle dt \\
&= \int_{t_0}^{t_1} \left\langle \frac{\partial L}{\partial q} - \frac{dp}{dt}, \delta q \right\rangle + \left\langle \frac{\partial L}{\partial v} - p, \delta v \right\rangle + \left\langle \delta p, \frac{dq}{dt} - v \right\rangle dt + \left\langle p, \delta q \right\rangle \Big|_{t_0}^{t_1}.
\end{aligned}$$

The fundamental lemma of the calculus of variations then gives the following three relationships

$$\frac{\partial L}{\partial q} = \frac{dp}{dt}\,,$$
$$p = \frac{\partial L}{\partial v}\,,$$
$$v = \frac{dq}{dt}\,.$$

These may be assembled into the Euler-Lagrange equations,

$$\frac{d}{dt}\frac{\partial L}{\partial \dot q} = \frac{\partial L}{\partial q}\,,$$

thus completing our proof. $\mathscr{Q}.\mathscr{E}.\mathscr{D}.$

Theorem 2.13. *Applying Hamilton's principle to the action written on the phase space in terms of the Hamiltonian, as follows,*

$$0 = \delta\int_{t_0}^{t_1} \langle p, \dot q\rangle - H(q,p)\,dt\,, \qquad \text{(2.8 revisited)}$$

implies Hamilton's canonical equations

$$\dot q = \frac{\partial H}{\partial p}\,, \quad \textit{and} \quad \dot p = -\frac{\partial H}{\partial q}\,. \qquad \text{(2.9 revisited)}$$

Proof. The proof is by direct computation. We take variations of the action, as follows

$$\begin{aligned} 0 &= \delta\int_{t_0}^{t_1} \langle p, \dot q\rangle - H(q,p)\,dt \\ &= \int_{t_0}^{t_1} \Big\langle \dot q - \frac{\partial H}{\partial p}, \delta p\Big\rangle + \Big\langle -\dot p - \frac{\partial H}{\partial q}, \delta q\Big\rangle\,dt + \Big\langle p, \delta q\Big\rangle\Big|_{t_0}^{t_1}\,. \end{aligned}$$

This directly implies Hamilton's canonical equations, recalling Theorem 2.10.

$\mathscr{Q.E.D.}$

The Hamiltonian dynamics of a phase space function, $F : T^*Q \mapsto \mathbb{R}$, defines the canonical Poisson bracket as follows.

Definition A.1 (Canonical Poisson bracket)**.** The *canonical Poisson bracket*, $\{\cdot,\cdot\}$, is a relationship which maps two smooth phase space functions to another, which is defined for $F : T^*Q \mapsto \mathbb{R}$ as follows

$$\begin{aligned}\frac{dF}{dt} &= \frac{\partial F}{\partial q}\dot{q} + \frac{\partial F}{\partial p}\dot{p} \\ &= \frac{\partial F}{\partial q}\frac{\partial H}{\partial p} - \frac{\partial F}{\partial p}\frac{\partial H}{\partial q} := \{F, H\}\,,\end{aligned} \tag{A.2}$$

where we have used Hamilton's canonical equations to compute the second line.

Remark A.2. Hamilton's canonical equations may be rewritten in terms of the canonical Poisson bracket as

$$\begin{aligned}\dot{q} &= \{q, H\}\,, \\ \dot{p} &= \{p, H\}\,.\end{aligned}$$

As mentioned prior to the statement of Theorem 2.10, the endpoint term arising during the derivation of Hamilton's canonical equations is related to Noether's theorem in the same way as in the Lagrangian case. When Noether's theorem is formulated on phase space, it has different implications. In particular, we have that

$$\frac{dJ_\xi}{dt} = \{J_\xi, H\} = \frac{\partial L_\xi}{\partial q}\frac{\partial H}{\partial p} - \frac{\partial L_\xi}{\partial p}\frac{\partial L_\xi}{\partial q} = -\delta H\,,$$

which implies that Noether's theorem for the Hamiltonian formulation follows from the Lie symmetry of the Hamiltonian, H, under $\delta H = \{H, J_\xi\}$. See Holm [51] for further details.

The Euler-Poincaré equations and semidirect products

In the left invariant case, the Euler-Poincaré equations are found by applying Hamilton's principle to the action corresponding to the *reduced* Lagrangian, $\ell(\xi)$, where variations are taken to be of the form

$$\delta\xi = \dot{\eta} + [\xi, \eta] =: \dot{\eta} + \mathrm{ad}_\xi \eta\,, \quad \text{for} \quad \eta = g^{-1}\delta g\,, \tag{A.3}$$

where $[\cdot,\cdot]$ is the commutator. See Bloch et al. [14] for a further details and proof of the form of these variations. Hamilton's principle then gives

$$\begin{aligned}
0 &= \delta \int_{t_0}^{t_1} \ell(\xi)\, dt \\
&= \int_{t_0}^{t_1} \frac{\delta \ell}{\delta \xi} \delta\xi\, dt = \int_{t_0}^{t_1} \frac{\delta \ell}{\delta \xi} \big(\dot{\eta} + \mathrm{ad}_\xi \eta\big)\, dt \\
&= \int_{t_0}^{t_1} \bigg(- \frac{d}{dt}\bigg(\frac{\delta \ell}{\delta \xi}\bigg) + \mathrm{ad}^*_\xi \frac{\delta \ell}{\delta \xi}\bigg) \eta\, dt
\end{aligned}$$

This implies the *Euler-Poincaré equations* for the *left* invariant case

$$\frac{d}{dt}\bigg(\frac{\delta \ell}{\delta \xi}\bigg) = \mathrm{ad}^*_\xi \frac{\delta \ell}{\delta \xi}\,. \tag{2.17 revisited}$$

This equation can be similarly derived from a reduced version of the Hamilton-Pontryagin principle found in the proof of Theorem 2.8, which takes the form

$$0 = \delta \int_{t_0}^{t_1} \ell(\xi) + \langle \mu, g^{-1}\dot{g} - \xi \rangle\, dt\,.$$

This calculation is one part of the following theorem.

Theorem A.3 (Euler-Poincaré theorem [14, 62, 74]). *Let G be a Lie group, $L : TQ \mapsto \mathbb{R}$ a left invariant Lagrangian, and $\ell(\xi) : \mathfrak{g} \mapsto \mathbb{R}$ its restriction to the tangent space of G at the identity. For a curve $g_t \in G$, let $\xi_t = g_t^{-1}\dot{g}_t$, then the following are equivalent:*

a. *Hamilton's principle*
$$0 = \delta \int_{t_0}^{t_1} L(g, \dot{g})\, dt\,,$$
holds for variations of g vanishing at the endpoints.

b. *The curve g_t satisfies the Euler-Lagrange equations for L on G.*

c. *The variational principle on $\mathfrak{g}$*
$$0 = \delta \int_{t_0}^{t_1} \ell(\xi)\, dt\,,$$
holds for variations of the form
$$\delta\xi = \dot{\eta} + [\xi, \eta]\,,$$
where η vanishes at the endpoints.

d. *The Euler-Poincaré equations* (2.17) *hold.*

For a proof of this, see Bloch et al. [14].

Remark A.4. The *right invariant* case is similar, with the following changes:

$$
\begin{aligned}
\xi = g^{-1}\dot{g} \quad &\rightarrow \quad \xi = \dot{g}g^{-1}\,, \\
\delta\xi = \dot{\eta} + [\xi,\eta] \quad &\rightarrow \quad \delta\xi = \dot{\eta} - [\xi,\eta]\,, \\
\frac{d}{dt}\left(\frac{\delta\ell}{\delta\xi}\right) = \mathrm{ad}^*_\xi\frac{\delta\ell}{\delta\xi} \quad &\rightarrow \quad \frac{d}{dt}\left(\frac{\delta\ell}{\delta\xi}\right) = -\mathrm{ad}^*_\xi\frac{\delta\ell}{\delta\xi}\,.
\end{aligned}
$$

To illustrate the power of this system, we introduce the following classic example.

Example (The rigid body)**.** Rigid body motion is a classical example of a system which lends itself to the reduced form. The configuration manifold for this problem is the rotation group, $G = SO(3)$. For this group, we have an isometry for the Lie algebra, $\mathfrak{so}(3) \cong \mathbb{R}^3$. This is through the isomorphism known as the *hat map*, which identifies the vector cross product of two elements of $\mathbb{R}^3$ with the commutator on $\mathfrak{so}(3)$. In matrix form, for a vector $\boldsymbol{\Omega} \in \mathbb{R}^3$ we have that

$$
\hat{\Omega}\boldsymbol{v} = \boldsymbol{\Omega} \times \boldsymbol{v}\,, \quad \text{for all} \quad \boldsymbol{v} \in \mathbb{R}^3\,. \tag{A.4}
$$

For this reason, we can characterise rigid body dynamics by finding their Euler-Poincaré form on the Lie algebra $\mathbb{R}^3$ using the Lagrangian

$$
\ell(\boldsymbol{\Omega}) = \frac{1}{2}\langle \mathbb{I}\boldsymbol{\Omega}, \boldsymbol{\Omega}\rangle = \frac{1}{2}\big(\mathcal{I}_1\Omega_1^2 + \mathcal{I}_2\Omega_2^2 + \mathcal{I}_3\Omega_3^2\big)
$$

where $\mathbb{I}$ is the moment of inertia tensor and $\boldsymbol{\Omega}$ is the body angular velocity. Variations of $\boldsymbol{\Omega}$ are restricted to be of the form

$$
\delta\boldsymbol{\Omega} = \dot{\boldsymbol{\Xi}} + \boldsymbol{\Omega} \times \boldsymbol{\Xi}\,, \tag{A.5}
$$

where $\boldsymbol{\Xi}(t)$ is a curve in $\mathbb{R}^3$ that vanishes at the endpoints in time [62]. Taking variations, we have

$$
\begin{aligned}
0 &= \int_{t_0}^{t_1} \langle \mathbb{I}\boldsymbol{\Omega}, \delta\boldsymbol{\Omega} \rangle \, dt = \int_{t_0}^{t_1} \langle \mathbb{I}\boldsymbol{\Omega}, \dot{\boldsymbol{\Xi}} + \boldsymbol{\Omega} \times \boldsymbol{\Xi} \rangle \, dt \\
&= \int_{t_0}^{t_1} \langle -\frac{d}{dt}\mathbb{I}\boldsymbol{\Omega}, \boldsymbol{\Xi} \rangle + \langle \mathbb{I}\boldsymbol{\Omega}, \boldsymbol{\Omega} \times \boldsymbol{\Xi} \rangle \, dt \\
&= \int_{t_0}^{t_1} \langle -\frac{d}{dt}\mathbb{I}\boldsymbol{\Omega} + \mathbb{I}\boldsymbol{\Omega} \times \boldsymbol{\Omega}, \boldsymbol{\Xi} \rangle \, dt + \langle \mathbb{I}\boldsymbol{\Omega}, \boldsymbol{\Xi} \rangle \Big|_{t_0}^{t_1} .
\end{aligned}
$$

This implies Euler's equation for the rigid body

$$
\mathbb{I}\dot{\boldsymbol{\Omega}} = \mathbb{I}\boldsymbol{\Omega} \times \boldsymbol{\Omega} \, . \tag{A.6}
$$

For a more thorough discussion of this example, together with the explicit reduction from the Lagrangian on $TSO(3)$, see Holm, Marsden, and Ratiu [62].

As was commented on in Section 2.1.2, when the symmetry of the Lagrangian is broken by some variable, a semidirect product formulation emerges. This culminated in the Euler-Poincaré theorem (Theorem 2.18 in Section 2.1.2). The following example illustrates why this is natural.

Example (The heavy top)**.** The heavy top is a classical system related to the rigid body. It is defined as a rigid body of mass m rotating around a fixed point of support, where the system feels a gravitational acceleration pointing downwards, $-g\hat{\boldsymbol{z}}$. The dynamics of a rigid body rotating around a point is nontrivial and, as seen in Figure A.1, is sufficiently interesting to occupy even the greatest of minds. This system introduces a slight complication which does not arise during the rigid body example. Namely, the presence of the gravitational field breaks the symmetry and the system is no longer $SO(3)$

invariant. One needs to keep track of the direction which gravity acts in relative to the body. We will see that the heavy top is a semidirect product system on $\mathfrak{se}^*(3) = \mathfrak{so}^*(3) \ltimes \mathbb{R}^3$, but does not correspond to geodesic motion on $SE(3)$. It can, however, be considered as geodesic motion on $SE(3) \times \mathbb{R}^3$ [52].

Indeed, the unit vector in the direction of gravity breaks the symmetry of the Lagrangian. The kinetic energy of the problem is the same as for the rigid body, which can be thought of as a reduction of a Lagrangian written in terms of a curve $\boldsymbol{R}_t \in SO(3)$. Indeed, the body angular velocity $\boldsymbol{\Omega}$ is related to $\boldsymbol{R}$ by $\boldsymbol{R}^{-1}\dot{\boldsymbol{R}}\boldsymbol{v} = \boldsymbol{\Omega} \times \boldsymbol{v}$ for all $\boldsymbol{v} \in \mathbb{R}^3$. The reduced kinetic energy is written in terms of elements of $\mathbb{R}^3$, which we recall may be interpreted as elements of $\mathfrak{so}(3)$ through the hat map.

The potential energy can be formulated as a map $TSO(3) \times \mathbb{R}^3 \mapsto \mathbb{R}^3$. We denote by g the acceleration due to gravity, by M the mass of the body and by l its length. The vector $\boldsymbol{\chi}$ denotes the unit vector pointing from the point of rotation to the body's centre of mass. The potential energy is $Mgl\boldsymbol{R}^{-1}\boldsymbol{v} \cdot \boldsymbol{\chi}$, where $\boldsymbol{v} \in \mathbb{R}^3$. This is left invariant, indeed for $\tilde{\boldsymbol{R}}$ we have

$$Mgl\boldsymbol{R}^{-1}\boldsymbol{v} \cdot \boldsymbol{\chi} = Mgl(\tilde{\boldsymbol{R}}\boldsymbol{R})^{-1}\tilde{\boldsymbol{R}}\boldsymbol{v} \cdot \boldsymbol{\chi}\,.$$

We define a vector $\boldsymbol{\Gamma} = \boldsymbol{R}^{-1}\boldsymbol{v}$ which, for $\boldsymbol{v} = \widehat{\boldsymbol{z}}$, may be interpreted as the orientation of the body relative to the vertical axis at time t. An application of the Euler-Poincaré theorem (Theorem 2.18) for $\ell : \mathfrak{so}(3) \times \mathbb{R}^3 \mapsto \mathbb{R}$ defined by

$$\ell(\boldsymbol{\Omega}, \boldsymbol{\Gamma}) = \mathbb{I}\boldsymbol{\Omega} \cdot \boldsymbol{\Omega} - Mgl\boldsymbol{\Gamma} \cdot \boldsymbol{\chi}\,,$$

gives the following equations for the motion of the heavy top

$$\mathbb{I}\dot{\mathbf{\Omega}} = \mathbb{I}\mathbf{\Omega} \times \mathbf{\Omega} + Mgl\mathbf{\Gamma} \times \boldsymbol{\chi}\,, \tag{A.7}$$

$$\dot{\mathbf{\Gamma}} = \mathbf{\Gamma} \times \mathbf{\Omega}\,. \tag{A.8}$$

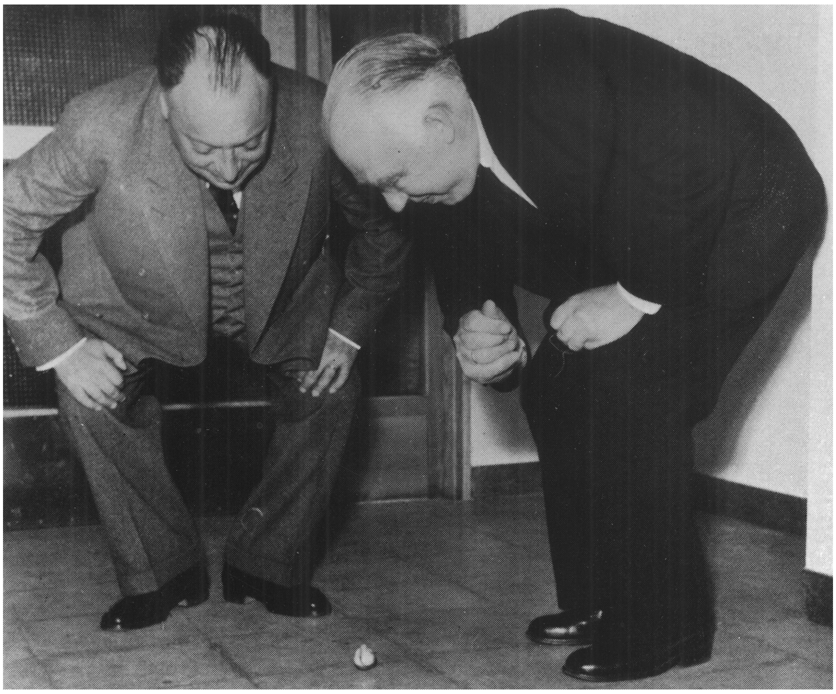


Figure A.1: Wolfgang Pauli and Niels Bohr with a 'tippe top'. Photograph by Erik Gustafson. Courtesy of the Niels Bohr Archive.

B

A Grönwall lemma for fractional differential equations

The following version of Grönwall's lemma may be found as Theorem 1.4 in [71].

Theorem B.1. *If, for any $t \in [0, T)$, we have*

$$u(t) \leqslant a(t) + \sum_{i=1}^{n} b_i(t) \int_0^t (t-s)^{\beta_i - 1} u(s)\, ds, \tag{B.1}$$

where all the functions are nonnegative and continuous, the constants β_i are positive, and $b_i \quad (i = 1, 2, \ldots, n)$ are the bounded and monotonic increasing *functions on* $[0,T]$*. Then, for any* $t \in [0,T)$*, we have*

$$u(t) \leqslant \sup_{t\in[0,T]} \Bigg\{ a(t) + \sum_{k=1}^{\infty} \Bigg[\sum_{1',2',,\ldots,k'=1}^{n} \frac{\prod_{i=1}^{k} [b_{i'}(t)\Gamma(\beta_{i'})]}{\Gamma(\sum_{i=1}^{k} \beta_{i'})} \int_0^t (t-s)^{\sum_{i=1}^{k} \beta_{i'}-1} a(s)\, ds \Bigg] \Bigg\}. \tag{B.2}$$

Remark B.2. For bounded b_i and a, the infinite sum in (B.2) converges. To show this, we assume that $a(t) \leqslant A$ and $b_i(t) \leqslant B$ for all $i = 1, \ldots, n$, furthermore we may assume without loss of generality that β_i are ordered $\beta_1 \leqslant \beta_2 \leqslant \cdots \leqslant \beta_n$. We label the terms of this series a_k, and for $k > 2/\beta_1$ we have

$$\begin{aligned} a_k &\leqslant A \sum_{1',2',,\ldots,k'=1}^{n} \frac{B^k (\max_i \Gamma(\beta_i))^k}{\Gamma(\sum_i^k \beta_{i'})} \frac{1}{k\beta_1} \max\{t,1\}^{k\beta_n} \\ &\leqslant Ank \frac{x^k}{\Gamma(k\beta_1)k\beta_1} \qquad \text{where } x := B\max\{T,1\}^{\beta_n} \max_i \Gamma(\beta_i) \\ &\leqslant A' \frac{x^k}{\Gamma(k\beta_1)}, \end{aligned}$$

for a constant A'. Note that the inequality in the second line above is only true for $k > 2/\beta_1$ since the gamma function is increasing on the interval $[2,\infty)$ however is decreasing nearer to 0. We split k into the following subsets $S_m := \{k : m \leqslant k\beta_1 \leqslant m+1\}$, and notice that on S_m we have $\Gamma(k\beta_1) > (m-1)!$. Thus we may bound a_k by the terms of the following series

$$\sum_{m=2}^{\infty} \sum_{k\in S_m} A' \frac{x^k}{(m-1)!}.$$

We have

$$\sum_{k\in S_m} x^k \leqslant \frac{m+1}{\beta_1} y^{\frac{m+1}{\beta_1}}, \qquad \text{where } y := \max\{x,1\},$$

and the sum defined by

$$\sum_{m=2}^{\infty} A' \frac{m+1}{\beta_1} \frac{y^{\frac{m+1}{\beta_1}}}{(m-1)!}$$

obviously converges. Hence the infinite sum in (B.2) converges as claimed.

C

Lemmata for the map P

Proofs of the following lemmata, which are analogous to lemmata 1 and 2 from [39], are required to complete the proof of Theorem 4.12. We define the following space of functions

$$X_{\delta,K} := \{f \in C([t_1, t_1 + \delta); \mathbb{R}^m) : \|f\| \leqslant K\}$$

where m can be either n or $2n$ as required. Note that in the following we will be dealing with the map P defined by (4.22), in the context of which Φ is $2n$-dimensional and η, ξ are n-dimensional.

Lemma C.1. *For P as defined by equation* (4.22)*, there exists a $K > 0$ large enough and $\delta = \delta(K) > 0$ small enough and independent of the initial condition such that P maps functions from $X_{\delta,K}$ to $X_{\delta,K}$.*

Proof. We must first prove that $P\Phi$ is continuous for continuous Φ, given assumption $(*)$. This continuity is obvious with the exception of the continuity of the integral

$$\int_{t_1}^{t} \frac{\eta(s)}{\sqrt{t-s}}\, ds, \tag{C.1}$$

for $\eta \in X_{\delta,K}$, as well as the integral

$$\int_{t_0}^{t_1} \frac{w(s)}{\sqrt{t-s}}\, ds, \tag{C.2}$$

for $w \in X_{\delta,K}$. Following same argument as in the proof of Lemma 4.15 (i), we may see that (C.1) is continuous. It remains only to prove that (C.2) has the required continuity. Recalling that R is such that $w \in S \subseteq \bar{B}_0(R)$, we have that (C.2) is continuous at τ since for all $\varepsilon > 0$, if we have $|t-\tau| < \frac{\varepsilon^2}{16R^2}$, then

$$\begin{aligned}
\left| \int_{t_0}^{t_1} \frac{w(s)}{\sqrt{t-s}}\, ds - \int_{t_0}^{t_1} \frac{w(s)}{\sqrt{\tau-s}}\, ds \right| &= \left| \int_{t_0}^{t_1} \frac{w(s)}{\sqrt{t-s}} - \frac{w(s)}{\sqrt{\tau-s}}\, ds \right| \\
&\leqslant R \left| \int_{t_0}^{t_1} \frac{1}{\sqrt{t-s}} - \frac{1}{\sqrt{\tau-s}}\, ds \right| \\
&= 2R\left| -\sqrt{t-t_1} + \sqrt{t-t_0} + \sqrt{\tau-t_1} - \sqrt{\tau-t_0} \right| \\
&\leqslant 2R\left|\sqrt{\tau-t_1} - \sqrt{t-t_1}\right| + 2R\left|\sqrt{t-t_0} - \sqrt{\tau-t_0}\right| \\
&= 2R\frac{|\tau-t|}{\sqrt{\tau-t_1}+\sqrt{t-t_1}} + 2R\frac{|t-\tau|}{\sqrt{t-t_0}+\sqrt{\tau-t_0}} \\
&\leqslant 2R\frac{|\tau-t|}{\sqrt{|\tau-t|}} + 2R\frac{|t-\tau|}{\sqrt{|t-\tau|}} \leqslant 4R\sqrt{|\tau-t|} \\
&< \varepsilon.
\end{aligned}$$

We find a bound on P as follows

$$
\begin{aligned}
|(P\Phi)(t)| \leqslant \|y_0 + \int_{t_1}^{t} \eta(s) + A_u(\xi(s), s)\, ds\|_\infty \\
+ \|w_0 + \int_{t_1}^{t} \left(\mu + \frac{\kappa\mu^{1/2}}{\sqrt{t-s}} + M_u(\xi(s), s)\right) \eta(s) + B_u(\xi(s), s)\, ds\|_\infty \\
+ \|\kappa\mu^{1/2} \int_{t_0}^{t_1} \frac{w(s)}{\sqrt{t_1 - s}} - \frac{w(s)}{\sqrt{t-s}}\, ds\|_\infty .
\end{aligned}
$$

Let us examine the integral in the final term as follows

$$
\begin{aligned}
\int_{t_0}^{t_1} \frac{w(s)}{\sqrt{t_1 - s}} - \frac{w(s)}{\sqrt{t-s}}\, ds &= \int_{t_0}^{t_1} \frac{w(s)}{\sqrt{t-s}\sqrt{t_1-s}} (\sqrt{t-s} - \sqrt{t_1 - s})\, ds \\
&= \int_{t_0}^{t_1} \frac{w(s)(t-t_1)}{\sqrt{t-s}\sqrt{t_1-s}(\sqrt{t-s} + \sqrt{t_1-s})},
\end{aligned}
$$

and we can bound this using $|t - t_1| < \delta$, $\sqrt{t-s} \geqslant \sqrt{t-t_1}$ and $\sqrt{t-s} + \sqrt{t_1 - s} \geqslant \sqrt{t-t_1} > \sqrt{\delta}$

$$
\begin{aligned}
\left\| \int_{t_0}^{t_1} \frac{w(s)}{\sqrt{t_1-s}} - \frac{w(s)}{\sqrt{t-s}}\, ds \right\|_\infty &\leqslant \left\| \int_{t_0}^{t_1} \frac{w(s)\delta}{\sqrt{t-s}\sqrt{t_1-s}\sqrt{\delta}}\, ds \right\|_\infty \\
&\leqslant \left\| \int_{t_0}^{t_1} \frac{w(s)}{\sqrt{t_1-s}}\, ds \right\|_\infty \leqslant R \left\| \int_{t_0}^{t_1} \frac{1}{\sqrt{t_1-s}}\, ds \right\|_\infty \\
&\leqslant 2R\sqrt{T-t_0}.
\end{aligned}
$$

Recall that, from assumption $(*)$, we have for any $x_1, x_2 \in \mathbb{R}^n$ and $\tau_1, \tau_2 \in \mathbb{R}$

$$
\begin{aligned}
|A_u(x_1, \tau_1) - A_u(x_2, \tau_2)| &\leqslant |A_u(x_1, \tau_1) - A_u(x_2, \tau_1)| + |A_u(x_2, \tau_1) - A_u(x_2, \tau_2)| \\
&\leqslant \|\nabla A_u\|_\infty |x_1 - x_2| + \|\partial_t A_u\|_\infty |\tau_1 - \tau_2| \\
&\leqslant L_b \left(|x_1 - x_2| + |\tau_1 - \tau_2|\right) , \\
|B_u(x_1, \tau_1) - B_u(x_2, \tau_2)| &\leqslant L_b \left(|x_1 - x_2| + |\tau_1 - \tau_2|\right) ,
\end{aligned}
$$

by the mean value theorem. By integrating from t_1 to t,

$$\left\| \int_{t_1}^{t} A_u(\xi(s), s) \right\|_\infty \leqslant \int_{t_1}^{t} L_b \left(|\xi(s)| + |s - t_0|\right) + |A_u(0, t_0)| \, ds$$
$$\leqslant L_b \|\xi\|_\infty (t - t_1) + L_b \delta (t - t_1) + |A_u(0, t_0)|(t - t_1)$$
$$\left\| \int_{t_1}^{t} B_u(\xi(s), s) \right\|_\infty \leqslant L_b \|\xi\|_\infty (t - t_1) + L_b \delta (t - t_1) + |B_u(0, t_0)|(t - t_1)$$

Hence we may improve our bound

$$|(P\Phi)(t)| \leqslant |y_{t_1}| + |w_{t_1}| + \|\Phi\|_\infty \Big[(t - t_1) + \mu(t - t_1) + 2\kappa\mu^{1/2}\sqrt{t - t_1}$$
$$+ L_b(t - t_1) + 2L_b(t - t_1) \Big] + 2R\sqrt{T - t_0}$$
$$+ (t - t_1)\big[2L_b\delta + A_u(0, t_0) + B_u(0, t_0)\big] .$$

Setting $K = 4\max\{R, 2R\sqrt{T - t_0}\}$ and choosing δ such that

$$\delta + \mu\delta + 2\kappa\mu^{1/2}\sqrt{\delta} + 3L_b\delta < 1/4, \quad (2L_b\delta + A_u(0, t_0) + B_u(0, t_0))\delta < K/4 \,,$$

we have our result and hence our lemma holds. $\mathscr{Q}.\mathscr{E}.\mathscr{D}.$

Lemma C.2. *For P as defined by equation* (4.22)*, there exists δ such that for any $\Phi_1, \Phi_2 \in X_{\delta,K}$ we have*

$$\|P\Phi_1 - P\Phi_2\|_\infty \leqslant \frac{1}{2}\|\Phi_1 - \Phi_2\|_\infty. \tag{C.3}$$

Proof. The proof of this is as in Lemma 2 in [39], since in $P\Phi_1$ and $P\Phi_2$ the integral from t_0 to t_1 is the same and thus cancels. Thus the proof exactly follows that of the standard Maxey-Riley system without additional memory, with no modifications necessary since the boundedness of A_u and B_u is not used. Thus this lemma holds for δ sufficiently small to ensure that

$$\delta + \mu\delta + 2\kappa\mu^{1/2}\sqrt{\delta} + L_b\delta < 1/4, \quad (2+K)L_c\delta < 1/4\,. \tag{C.4}$$

$\mathscr{Q}.\mathscr{E}.\mathscr{D}.$

Index